Universität Augsburg

Deep learning approaches for the improvement of commercial microwave link and weather radar derived precipitation information

Kumulative Dissertation zur Erlangung des Doktorgrades
an der Fakultät für Angewandte Informatik

Vorgelegt von
Julius Polz

Augsburg, February 22, 2024

Erstes Gutachten: Prof. Dr. Harald Kunstmann
Zweites Gutachten: Prof. Dr. Sabine Timpf
Drittes Gutachten: Prof. Dr. Remko Uijlenhoet
Tag der mündlichen Prüfung: 2. Dezember 2024

Paulownia Trees at Akasaka in the Evening Rain
Utagawa Hiroshige II, Japan, Edo Period (1615–1868)[1]

[1]The Metropolitan Museum of Art, 1970, www.metmuseum.org/art/collection/search/36708

Contents

Abstract

Climate observations are crucial for societies around the globe to adapt to natural hazards in a changing climate. However, large parts of the world, especially developing countries, do not have sufficient access to climate information. Rainfall is the major driver of hydrological processes that cause flooding or droughts which are responsible for the majority of natural disasters. Precipitation is especially hard to estimate and forecast due to its high spatial and temporal variability. There is also a large heterogeneity in the abundance of conventional rainfall sensors such as rain gauges and weather radars and their setup is cost and maintenance-intensive. To close the observational gap for rainfall sensors commercial microwave links (CMLs) to measure path-averaged rainfall are a promising alternative since more than 90% of the global population lives in areas where they are deployed. However, due to their opportunistic nature and indirect measurement, they are prone to systematic and random errors that require quantification, attribution to causes, and correction in order to provide high-quality quantitative precipitation estimates (QPE). The same holds for systematic errors in weather radar QPE. The objective of this thesis is the improvement of CML and weather radar QPE by mitigating systematic errors. The main innovation is the application of deep learning techniques which have proven to provide high-performing solutions to model atmospheric processes. Convolutional neural networks (CNNs) are applied to improve the detection of rain events in commercial microwave link data in order to reduce the impact of attenuation falsely attributed to rainfall by more than 50%. Another application is the simultaneous increase of the temporal resolution, ground-adjustment, and advection-correction of radar QPE to reduce biases by 20% and mitigate a sampling error. Additional studies to investigate and disentangle the complex error structure of commercial microwave links have been conducted: First, the performance of state-of-the-art CML processing techniques and the resulting CML QPE were compared using one year of country-wide rainfall observations identifying processing steps with the highest impact on QPE quality. Second, missing rainfall extremes due to a complete loss of signal in heavy rain (blackouts) have been investigated showing that they occur more frequently than radar-derived climatology suggests. Third, signal fluctuations that are not due to rainfall (anomalies) have been detected using manual data flagging and their impact on CML QPE has been investigated. While there was ambiguity in the flagging, removing anomalies significantly improved the quality of rainfall estimates. In summary, the presented results show that systematic errors in CML and weather radar QPE can be quantified and corrected using a data-driven approach, but attribution to causes remains difficult. Trained artificial neural networks prove to be a robust tool to provide high-quality QPE that can be easily transferred to new locations and future time periods within the same climatic region. CML QPE is shown to have a remarkably high quality when compared to gauge-adjusted weather radar QPE and the results presented will foster a successful deployment of CMLs to close the observational gap in climate science.

Zusammenfassung

Klimabeobachtungen sind von großer Bedeutung, damit sich menschliche Gesellschaften an Naturgefahren in einem sich verändernden Klima anpassen können. Große Teile der Welt, insbesondere Entwicklungsländer, haben jedoch keinen ausreichenden Zugang zu Klimainformationen. Besonders Niederschlag ist, als Hauptfaktor hydrologischer Prozesse, für einen Großteil der auftretenden Naturkatastrophen, wie Überschwemmungen oder Dürren, verantwortlich und deshalb als Variable für Klimaanpassungsstrategien unerlässlich. Aufgrund seiner hohen räumlichen und zeitlichen Variabilität ist es zudem besonders schwierig Niederschlag präzise zu messen und vorherzusagen, weshalb ein besonders dichtes Netz an Observationen nötig ist. Die globale Verteilung herkömmlicher Niederschlagssensoren, wie Regenmesser oder Wetterradare, ist jedoch sehr inhomogen und ihr Betrieb ist sehr kosten- und wartungsintensiv. Um die Lücken in der Verfügbarkeit präziser Niederschlagsobservationen zeitnah zu schließen ist deshalb die Nutzung bestehender Infrastruktur besonders vielversprechend. Kommerzielle Richtfunkstrecken (CMLs), die zur Messung des pfadgemittelten Niederschlags verwendet werden können sind eine vielversprechende Alternative zu herkömmlichen Messnetzen, da mehr als 90% der Weltbevölkerung in Gebieten mit Mobilfunknetzwerken lebt, in denen CMLs betrieben werden. Aufgrund der opportunisten Nutzung von CML Daten und der indirekten Messung sind CMLs jedoch anfällig für systematische und zufällige Messfehler. Diese müssen quantifiziert, möglichen Ursachen zugeordnet und anschließend korrigiert werden, um hochwertige quantitative Niederschlagsmessungen (QPE) zu liefern. Das Gleiche gilt für systematische Fehler in Wetterradarmessungen. Das Ziel dieser Arbeit ist die Verbesserung von CML und Wetterradar QPE durch die Kompensation von systematischen Fehlern. Die wichtigste Neuerung in dieser Dissertation ist die Anwendung von Deep-Learning-Techniken, die geeignete Lösungen für die Modellierung von Niederschlagsprozessen bieten. Im speziellen werden in der vorgestellten Arbeit Convolutional Neural Networks (CNNs) eingesetzt, um die Detektion von Regenereignissen in Daten kommerzieller Richtfunkstrecken zu verbessern und den Beitrag von fälschlicherweise dem Regen zugeschriebener Dämpfung um mehr als 50% zu reduzieren. Eine weitere Anwendung von CNNs bestand in der gleichzeitigen Erhöhung der zeitlichen Auflösung, einer Anpassung an bodennahe Messungen und einer Advektionskorrektur von Radarniederschlagsmessungen. Dies erlaubte die Reduktion des mittleren Fehlers um 20% und die deutliche Abmilderung von Sampling-Fehlern. Um die Korrektur von Messfehlern zu unterstützen wurden weitere Untersuchungen durchgeführt, um das Verständnis über die komplexe Fehlerstruktur kommerzieller Richtfunkstrecken zu verbessern. Zuerst wurde die Effektivität im Forschungsfeld etablierter CML-Datenverarbeitungstechniken und die daraus resultierende CML-QPE anhand eines Jahres landesweiter Messungen verglichen. Anschließend wurden fehlende Niederschlagsextreme aufgrund eines vollständigen Signalausfalls bei Starkregen (Blackouts) untersucht, wobei sich herausstellte, dass diese häufiger auftreten, als eine von Wetterradarmessungen abgeleitete Klimatologie vermuten lies. Zusätzlich wurden Sig-

nalschwankungen, die nicht auf Regenfälle zurückzuführen sind (Anomalien), durch manuelle Qualitätskontrolle der Daten identifiziert und ihre Auswirkungen auf die CML-QPE untersucht. Trotz Mehrdeutigkeiten bei der Kennzeichnung konnte durch die Entfernung der Anomalien die Qualität der Niederschlagsschätzungen deutlich verbessert werden. Zusammenfassend zeigen die Ergebnisse dieser Dissertation, dass systematische Fehler in CML und Wetterradardaten mit einem datengetriebenen Ansatz quantifiziert und korrigiert werden können, wobei eine Zuordnung zu potentiellen Ursachen schwierig bleibt. Trainierte künstliche neuronale Netze erweisen sich als robustes Instrument zur Bereitstellung qualitativ hochwertiger QPE, die leicht auf neue Standorte und künftige Zeiträume innerhalb derselben Klimaregion übertragen werden können. Es konnte außerdem gezeigt werden, dass CML-QPE eine bemerkenswert hohe Qualität im Vergleich mit stationsangeeichten Wetteradardaten aufweist. Die präsentierten Ergebnisse sind ein wichtiger Schritt hin zur Operationalisierung von CML Niederschlagsmessungen und damit einer Verbesserung der globalen Verfügbarkeit von wichtigen Klimainformationen.

Chapter 1

Introduction

1.1 Motivation

The Observational Gap in Climate Science

Climate change mitigation and adaptation are among the most important challenges humanity is facing today (IPCC, 2022a,b). A large proportion of the earth's population is experiencing limited freshwater supply, a development which has been known for decades and which will be an increasing issue due to population growth and climate change (Watson et al., 1996; Vörösmarty et al., 2000). At the same time, the International Federation of Red Cross and Red Crescent Societies (IFRC) reports that 83% of all-natural disasters in the 2010s were caused by weather and climate extremes killing more than 410,000 people (IFRC, 2021). Rainfall is the major driver of the hydrologic cycle and its spatial and temporal distribution strongly affects human societies with at least half of all natural disasters caused by rainfall, like floods and landslides (IFRC, 2021). Climate model simulations and observations support the physical theory that, globally, both mean and extreme precipitation increase with higher average temperatures (Lenderink et al., 2011; Berg et al., 2013; IPCC, 2021). Some studies even suggest a potential underestimation of this relative increase in climate projections (Allan and Soden, 2008). At the same time, the contrast between wet and dry periods will likely increase and lead to more intense droughts with a faster onset (Trenberth et al., 2014). Due to this development, the relevance of finding measures to mitigate the impact of natural disasters is extremely high.

Systematic observations of the earth's climate help mitigate losses caused by climate extremes. Both, drought monitoring and flood forecasting rely on accurate initial conditions provided by rainfall observations. Novel early warning systems build their strengths on underlying seamless prediction pipelines that combine observations, nowcasting, and numerical weather prediction to provide optimal forecasts as potentially hazardous events approach (Trömel et al., 2021). The assimilation of observations has become a key ingredient for accurate short to medium-range forecasting models (Potthast et al., 2022). Additionally, accurate reference data is beneficial for earth system model development.

Yet, climate observations remain sparse and are not homogeneously distributed around the globe. According to the United Nations Framework Convention on Climate Change,

"one third of the world, including sixty percent of Africa, does not have access to early warning and climate information services" (UNFCCC, 2022). They particularly emphasize "the need to address existing gaps in the global climate observing system, particularly in developing countries". To close this large observational gap in due time not only the time and cost-intensive installation of conventional sensors should be considered. The opportunistic usage of available infrastructure is a promising alternative.

Rainfall Remote Sensing

This thesis focuses on rainfall. It is especially challenging to estimate and predict, since it is characterized by high spatial and temporal variability, exhibiting fluctuations on almost all spatial and temporal scales (Berg et al., 2013). Accurate quantitative precipitation estimation (QPE), therefore, needs dense observation networks. Weather radars provide rainfall observations with high spatial representativeness and country-wide coverage. Gauge-adjusted weather radar QPE is considered to have the highest accuracy among rainfall measurement techniques. However, the lack of climate information in parts of the globe also applies to rainfall sensors. For example, the World Meteorological Organization does not report on any operational weather radars in West Africa (WMO). Additionally, precipitation estimates from rain gauges only observe at most 1% of the earth's surface using the strong assumption of spatial representativeness of the measurement in a 5 km radius around each rain gauge (Kidd et al., 2017).

To improve coverage, one of the most promising existing infrastructures to use is Commercial Microwave Links (CML). Used as backhaul of mobile networks they use one-way microwave transmitters and receivers to pass information over distances of hundreds of meters to several kilometers. Due to the transmission frequency in the range of 1 to 100 GHz their signal is attenuated by raindrops which makes them suitable rainfall sensors (Atlas and Ulbrich, 1977; Messer et al., 2006; Leijnse et al., 2007).

The potential for observing rainfall with CMLs is huge, since, according to the Groupe Speciale Mobile (GSM) Association, 94% of the human population already lived in regions with broadband telecommunication access in 2021 (Gsma, 2022). This means that the largest part of the earth's population lives in areas where CMLs are deployed. Unfortunately, CML data access is still sparse and challenging (Chwala and Kunstmann, 2019). CML rainfall estimation is an active research field on the verge of operational application (Uijlenhoet et al., 2018). CML data has been acquired for research purposes in a number of countries including the Netherlands (Overeem et al., 2016b), Germany (Chwala et al., 2016), Israel (Goldshtein et al., 2009), Burkina Faso (Doumounia et al., 2014), Sri Lanka (Overeem et al., 2021) and Sweden (Andersson et al., 2022).

Successful applications of CML QPE include river runoff simulations in a pre-alpine catchment in Germany (Smiatek et al., 2017) or urban drainage modeling in the Czech Republic (Pastorek et al., 2019). CMLs have also been used to estimate other observables like atmospheric water vapor (Rubin et al., 2022; Fencl et al., 2021) or atmospheric temperature inversion (David and Gao, 2016).

The physical foundations of relating specific attenuation to rainfall rate are well understood. The challenge lies in the attribution of total path loss to rainfall since multiple other atmospheric factors contribute to signal fluctuations (van Leth et al., 2018). Quality control and data processing are still the most important challenges to be solved (Chwala and Kunstmann, 2019). Even if vast amounts of CML data were available today, processing it would not be straightforward due to the heterogeneous data quality.

The Need for Automated Quality Control

In the 21st century the term "Big Data" has become prevalent. Yet, its definition remains vague. It not only refers to the amount of data but also to the diversity and complexity (Agapiou, 2017). With an increasing amount of climate observations from different types of sensors the demand for robust data quality control and flexible frameworks for merging observations is rising. A common approach to data quality control is manual data flagging by experts which is both too labor-intensive to scale with the amount of data and too subjective due to human decisions without objective rules. Standard statistical tests need parameterization and due to failing simple rules, this also amounts to labor-intensive manual tasks. Considering the vast amount of data that can not be checked by a human, but only through automated algorithms, a new way of tackling quality assurance and knowledge discovery in climate observations is needed. This especially holds for CML data, but also for weather radars (Chwala and Kunstmann, 2019; Villarini and Krajewski, 2010).

Deep Learning for High-Dimensional Physical Problems

Bronstein et al. (2021) argue that "modern data analysis is synonymous with high dimensional learning." Indeed, geoscientific data often poses extremely high dimensional problems, e.g. asking for river runoff at a single gauge given a time-evolving rainfall field over the catchment. In its most general form, the solution to such a problem is a function in a space as high dimensional as the number of cells on a grid times the number of time steps. The so-called curse of dimensionality hampers finding approximate solutions to high dimensional problems (Bellman, 1984; Zimek et al., 2012). Ways to reduce the dimensionality of the problem, i.e. finding a lower dimensional embedding of the data, can help, but requires knowledge about the structure of the data. The most radical and often applied solution to the runoff example is to integrate over time and space (i.e. the upstream catchment of the gauge) which assumes that the problem is invariant to the spatial and temporal disaggregation of the input data. On the one hand, it is easily seen that such an assumption can not hold given the spatiotemporal dynamics of rainfall and runoff in complex terrain. On the other hand, the assumption leads to an easy solution that approximates the problem surprisingly well in some cases.

In general, physical problems present structured data that underlie principles of symmetry, i.e. there is a certain invariance to transformations of the data. Transformations that reduce the dimensionality of a problem ideally drop redundant information and achieve a compressed representation of the important information. Deep learning delivers

a blueprint for exploiting symmetries when dealing with high dimensional learning problems (Bronstein et al., 2021). However, deep learning methods were initially developed for language processing and image recognition tasks, where a clear understanding of the world and the processes generating the data is present. The geosciences underlie clear rules as well, yet the processes are more complex and less understood, which is why they are interesting objects of study in the first place. This is both a challenge and an opportunity for state-of-the-art machine learning models that can be used to find hidden representations of processes generating data (Reichstein et al., 2019).
Deep learning has already proven to yield flexible high-performance solutions for many geoscientific problems. Famous applications are the modeling of rainfall-runoff using long short-term memory networks (Kratzert et al., 2018), precipitation nowcasting and downscaling using generative adversarial networks (Ravuri et al., 2021; Leinonen et al., 2021) or improving medium-range weather forecasts and seasonal predictions (Lam et al., 2022; Schneider et al., 2022). Another benefit of neural networks is that, once they are trained, they can process data at low latency and high energy efficiency. For example, Pathak et al. (2022) demonstrate that a neural network approach can produce forecasts comparable to the high-resolution Integrated Forecasting System (IFS) model of the European Center for Medium-Range Weather Forecasting (ECMWF) at 0.008% of its energy consumption and $4.5 \cdot 10^5$ times faster. Today, all these achievements seem like beacons illuminating parts of the vast space of possible geoscientific applications of deep learning. However, the true potential is still only being discovered. It is a well-known phenomenon that theory and application foster each other's development. The adaptation of existing algorithms to new problems yields new knowledge about the methods themselves. Thus both fields, the computer- and geosciences, are advancing while highly interdisciplinary research is conducted.

Thesis Objective

This thesis strives to explore deep learning approaches for the improvement of CML and weather radar-derived precipitation information. Facing the climate observation gap requires new approaches for assuring the quality of opportunistic and conventional rainfall sensors. By an application of deep neural networks for rainfall estimation we[1] hope to gain knowledge about the specific challenges that arise and to improve the quality of rainfall estimates.

[1]Note that the first person plural will be used in the theoretical introduction of this thesis to emphasize the participatory nature of the scientific endeavor. Adapting this active style of writing from modern mathematics is a personal choice of the author. In the main chapters, it is used to refer to the shared effort and consensus among the co-authors of the respective works.

1.2 Rainfall Estimation

1.2.1 A Definition of Rainfall

Precipitation is a term describing hydrometeors formed by condensation or deposition of atmospheric water vapor that are large enough to fall to the earth's surface which also includes frozen particles like snow or hail (Ams, 2021b). Rain is precipitation in the form of liquid water droplets with a diameter larger than 0.5 mm (Ams, 2021a). In the literature, the term rainfall is often found to be either synonymous with precipitation or with rain. This work is mostly concerned with rain and we will refer to rainfall as liquid precipitation. The cause for precipitation is condensation as a consequence of the saturation of air with water vapor either by processes cooling the air mass (e.g. adiabatic cooling of ascending and expanding air parcels) or by processes increasing the amount of water vapor (e.g. evaporation) (Wallace and Hobbs, 2006). The temperature where saturation occurs while cooling the air is called the dew point. It depends on the humidity and air pressure. The condensation of water droplets to form larger raindrops is strongly dependent on the presence of condensation nuclei.
Quantitative precipitation estimation (QPE) describes all means of measuring precipitation that falls to the earth's surface. The unit for QPE is millimeters (mm) of accumulated liquid water. A rainfall rate or intensity is measured in mm per time interval and the unit is mmh^{-1}. For non-liquid precipitation, the liquid water equivalent is used. Falling rain is characterized by the frequency distribution of drop sizes in a given volume called the drop-size distribution (DSD). As Fig. 1.1 illustrates, the shape and the resulting dielectric properties of water droplets depend on their size (Ekelund et al., 2020).

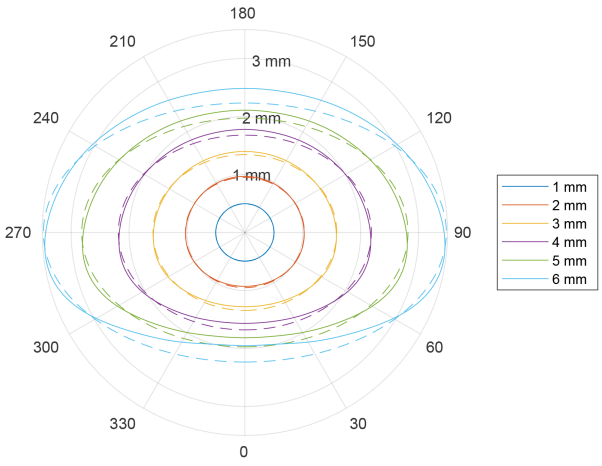

Figure 1.1: The dielectric properties of individual drops depend on their shape, which can be modeled as a function of drop diameter (Figure adapted from Ekelund et al. (2020))

The DSD uniquely defines the rain rate R in $\mathrm{mm\,h^{-1}}$ as an integral of the number of drops $N(D)$ with diameter D in cm multiplied by their volume $6 \cdot 10^{-3}\pi D^3$ and fall speed $\nu(D)$ in $\mathrm{m\,s^{-1}}$ over the drop diameter (Ekelund et al., 2020):

$$R = 6 \cdot 10^{-3}\pi \int_0^\infty N(D)\nu(D)D^3 dD \tag{1.1}$$

The fall speed power law description $\nu(D) = 1767D^{0.67}$ holds for drop sizes between 0.05cm and 0.5cm. Equation 1.1 assumes that no influences like vertical winds affect the fall speed and that the DSD does not change while raindrops are falling.
However, falling hydrometeors are affected by a number of physical processes including horizontal and vertical advection, melting and freezing processes, evaporation, condensation, collision, and coalescence. Rainfall estimation techniques are usually defined using the idealized Eq. 1.1. Therefore, the above-mentioned effects can create random as well as systematic measurement errors.
Rainfall has a high spatial and temporal variability with fluctuations across scales (Berg et al., 2013). The extreme value distribution of rainfall is heavily skewed and highly intermittent. For example, in Germany, it is, on average, not raining for more than 95% of the time (Adler et al., 2001). These characteristics make the estimation of rainfall very difficult because there are large differences between point and area measurements. In the next section, we will describe the rainfall estimation techniques considered in this thesis.

1.2.2 Measurement Techniques

There are a number of devices designed for quantitative precipitation estimation. In the following, we will describe the most important measurement techniques that are considered in this thesis as well as the main sources of systematic and random measurement errors.
Systematic measurement errors are deviations from a ground truth that do not occur randomly. Compared to random errors, they provide an opportunity for correction with respect to the ground truth if their cause is known and if individual occurrences can be detected.

1.2.2.1 Rain gauges

Rain gauges are the oldest device to estimate rainfall with known references dating back to 400 BC (Strangeways, 2010). In the 17th century, rainfall began to be measured in a scientific manner and the modern description of rainfall estimates in millimeters (mm) originates from the use of rain gauges.
The simplest form of rain gauge is an accumulation gauge, i.e. a vessel with an orifice at the top which is used to collect falling raindrops (see Fig. 1.2). Considering a vessel where the volume in liters increases linearly with the water level in millimeters within the vessel the rainfall intensity R can be directly measured as an increase of liters or mm per time interval. Today, more advanced and automatic devices like weighing, tipping bucket, capacitance, or optical gauges are used (Nystuen et al., 1996).

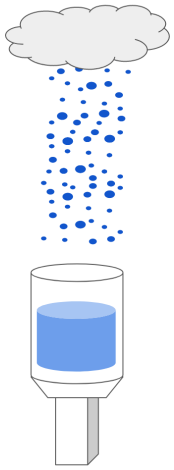

Figure 1.2: The accumulation gauge is a simple form of rain gauge with an orifice at the top which is used to collect falling raindrops.

The main source of systematic errors that affects all types of rain gauges is the influence of wind at the gauge orifice that causes turbulent airflow (Ciach, 2003). There is a potential undercatch of precipitation extremes when using tipping bucket rain gauges since the tipping mechanics usually have an upper detection limit or loose precision in heavy rainfall. Additional sources of error are wetting, evaporation, and splashing as pointed out in Chen et al. (2008) as well as a series of errors for solid or mixed-phase precipitation. Despite a high local accuracy around the small surface area that is used to collect rainfall, rain gauges lack spatial representativeness when compared to area covering precipitation remote sensing techniques such as weather radars or satellites. A single rain gauge measurement has to be considered point-like and can only represent average areal rain rates in a small radius, less than a few kilometers (Kidd et al., 2017). Spatial representativeness also depends on the temporal accumulation interval as shown by Villarini et al. (2008). They investigated how many rain gauges are needed to estimate spatial averages for varying grid spacing with a given error margin. For example, when hourly rainfall estimates are compared to satellite rainfall products with a grid cell area of 200km^2 around 25 rain gauges are needed to achieve a normalized mean absolute error below 20%. However, rain gauges are sparsely distributed with, for example, one rain gauge per 330km^2 in Germany. Globally, the density of rain gauge observations is very heterogeneous and the number of maintained weather stations is declining in many regions (Lorenz and Kunstmann, 2012). Random errors in rain gauge measurements mainly occur only due to the quantization of the rain gauge measurement and can be considered as a uniformly distributed random error (Ciach, 2003). In this thesis, rain gauges are used for the validation of improved CML and weather radar QPE.

1.2.2.2 Disdrometers

Disdrometers are measurement devices designed to measure the size, count, and velocity of falling rain drops which are used to describe the rain rate (Eq. 1.1). Early efforts to quantify the size and count used dying filter paper in a tedious manual process (Marshall et al., 1947). The Joss-Waldvogel disdrometer was the first automated measurement device of its kind. It measures the momentum of raindrops that impact a membrane (Joss and Waldvogel, 1967). This type of measurement has the drawback that fall speed and drop size can not be measured separately and that a fixed fall speed law $\nu(D)$ has to be assumed.
Particle Size Velocity (PARSIVEL) type disdrometers are laser-optical devices that measure the energy reduction caused by extinction when raindrops fall through a horizontal band of light (Löffler-Mang and Joss, 2000). By using both, a model of the size and shape of falling raindrops and the time that drops need to pass through the light band it is possible to separately estimate fall speed and drop size. It would be beneficial to always estimate the size, count, and velocity of falling raindrops, however, disdrometers are cost and maintenance-intensive devices that are mostly used for the calibration of more cost-effective measurement techniques, such as weather radars. Disdrometer data is not directly used in this thesis, but important since the attenuation-rain rate (Eq. 1.3 below) and reflectivity-rain rate (Eq. 1.7 below) relationships rely on DSD data.

1.2.2.3 Microwave links

Attenuation of microwave radiation through atmospheric phenomena like rainfall has been discovered long ago (Ryde, 1946). In a series of publications Atlas and Ulbrich (1977) derive the physical foundations and practical configurations for measuring path-integrated attenuation and deriving its relation to rainfall intensity. They show that there is a close-to-linear relationship between attenuation and rainfall rate at frequencies around 30GHz. The reason for this is demonstrated in their initial work from 1977:
The physical descriptions of rainfall rate R in mm h^{-1} (see Eq. 1.1) and specific attenuation k, in dBkm^{-1}, through a medium of scatterers are very similar, since

$$k = 0.4343 \int_0^\infty N(D)\sigma_{ext}(D)dD, \tag{1.2}$$

where $\sigma_{ext}(D)$ is the total extinction cross section in cm^2. Using the power law description $\sigma_{ext}(D) = aD^b$, we get a linear relationship between R (see Eq. 1.1) and k if we find a frequency, where $b = 3.67$. It turns out that a close to linear relationship holds for frequencies close 30GHz where scatterer diameter and wavelength have the same magnitude (Olsen et al., 1978). Comparing Eq. 1.1 and Eq. 1.2 then yields the coefficients $\alpha = 0.24$ and $\beta = 1.0$ of the power law relation

$$k = \alpha R^\beta. \tag{1.3}$$

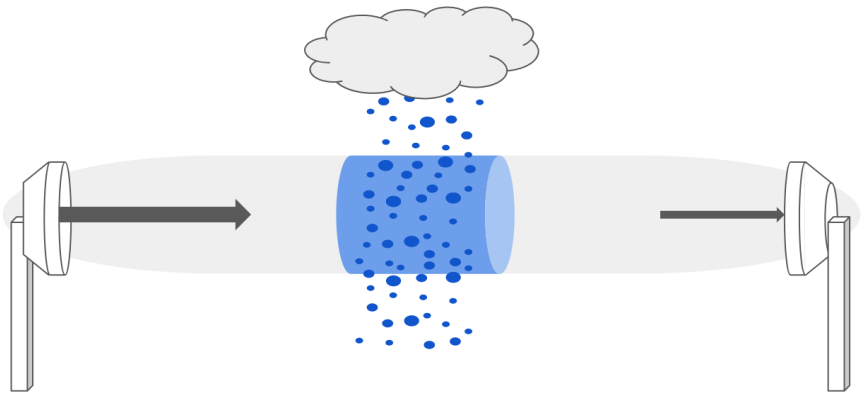

Figure 1.3: Commercial microwave link with a transmitting (left) and receiving (right) antenna. The falling raindrops affect the total path loss (TL) of the CML by scattering and absorption of microwave radiation. The amount of specific attenuation depends on the drop size distribution, frequency, and polarization.

The close to linear relationship is important because non-linearity introduces an error when aggregating attenuation which is, for example, the case when comparing path-integrated attenuation and path-averaged rainfall intensity. Additionally, the relation is mostly insensitive to the drop size distribution and the temperature dependence is negligible. Based on scattering simulations the International Telecommunication Union (ITU) recommends k-R power law coefficients α and β for vertical or horizontal polarization and frequencies f from 1 to 1000 GHz (ITU-R, 2005).
Atlas and Ulbrich (1977) concluded that, to derive path-integrated attenuation, directional one-way transmission using a focused beam was most suitable which is equivalent to the setup of microwave links (see Fig. 1.3). Microwave transmission near the ground needs to avoid obstacles in the transmission path. Not only the line-of-sight connection, but a larger ellipsoidal region called the Fresnel zone needs to be free of obstacles. Partial or complete beam blockage will lead to significant interference in the received signal level.
Although the foundation of rainfall estimation with microwave links was established already in the 1970s the concept was not applied broadly until Commercial Microwave Links (CMLs) were widely deployed as backhaul for telecommunication networks. While rainfall-induced attenuation remained a disturbance for network providers and was studied as such (Hogg, 1968), the true potential for rainfall estimation was finally discovered in the early 2000s (Messer et al., 2006; Leijnse et al., 2007). The main advantage of CMLs is that no additional infrastructure is needed to be able to use them for rainfall estimation. Uijlenhoet et al. (2018) estimate that, based on numbers provided by the GSMA, the global amount of CMLs in 2017 was 4 million and still continues to increase. The typical length ranges from hundreds of meters to several kilometers. Transmission frequencies can range from 1GHz to 100GHz but are most common between 15GHz and 40GHz (IEEE K_u-, K- and K_a-band) which is useful for rainfall estimation or between 71GHz and 86GHz (E-band), where the sensitivity to water vapor is increased (Fencl et al., 2020, 2021).

Higher frequency CMLs provide a larger bandwidth but are also more sensitive to rainfall (i.e. higher specific attenuation) (Chwala and Kunstmann, 2019). To avoid high path-integrated attenuation, lower frequencies are used with increasing CML length.
The main challenge in acquiring CML data for research purposes is that a formal agreement with network providers, who typically operate on a regional to country-wide level, is necessary. Network operators usually do not archive signal level recordings which they use for monitoring the network performance. Chwala et al. (2016) demonstrate how a data acquisition system can be used for both real-time and long-term observation of signal levels.

The sampling strategy and signal quantization of transmitted (TSL) and received (RSL) signal levels (in dB) of a CML are defined by the data acquisition. The signal level sampling on the hardware level is typically instantaneous and quantized and it is technically feasible to sample at a rate of several Hertz. A potential limit for the sampling rate is the required bandwidth within the CML network and the data acquisition. Two common sampling strategies for archiving the data that emerged are either an instantaneous sampling of signal levels at a lower rate, e.g. 1 minute (Chwala et al., 2016), or the acquisition of a 15-minute minimum and maximum of signal levels sampled at a higher rate of several Hz. An instantaneous sampling at a high sampling rate would be optimal for rainfall estimation but has to be negotiated with network providers.
Due to the opportunistic nature, that is, using the devices for purposes that were unintended during construction, of CML rainfall estimation, dedicated data processing is necessary. A commonly used processing routine to derive path-averaged rainfall estimates from TSL and RSL is as follows:
The total loss (TL), formerly also called transmitted minus received signal level ($TRSL$), is computed as a difference of raw RSL and TSL. Depending on the data acquisition, default values or missing time steps can occur and have to be treated. TL is assumed to be a compound quantity since it is a sum of the baseline attenuation (A_b) due to slow-evolving properties of the atmosphere unrelated to rainfall, such as gas concentration and water vapor, attenuation due to rainfall (A_R), excess attenuation due to the wet antenna effect (A_w) (Moroder et al., 2019), other, unknown, sources of attenuation (A_u) and finally a random error (ϵ) due to the signal quantization of the transmitter and receiver which is in the range of 0.1dB to 1dB (Zinevich et al., 2010).

$$TL = A_b + A_R + A_w + A_u + \epsilon \tag{1.4}$$

Rain event detection, the temporal separation of rainy (wet) and non-rainy (dry) periods in the time-series, is used to attribute attenuation to rainfall. It is one of the most important processing steps. A static signal level baseline to derive rainfall-induced attenuation from TL has quickly proven to be ineffective due to daily or annual cycles and unexpected jumps in the time-series (see Fig. 1.4). To be able to estimate a baseline attenuation of the given atmospheric condition, the onset of rain events needs to be known.

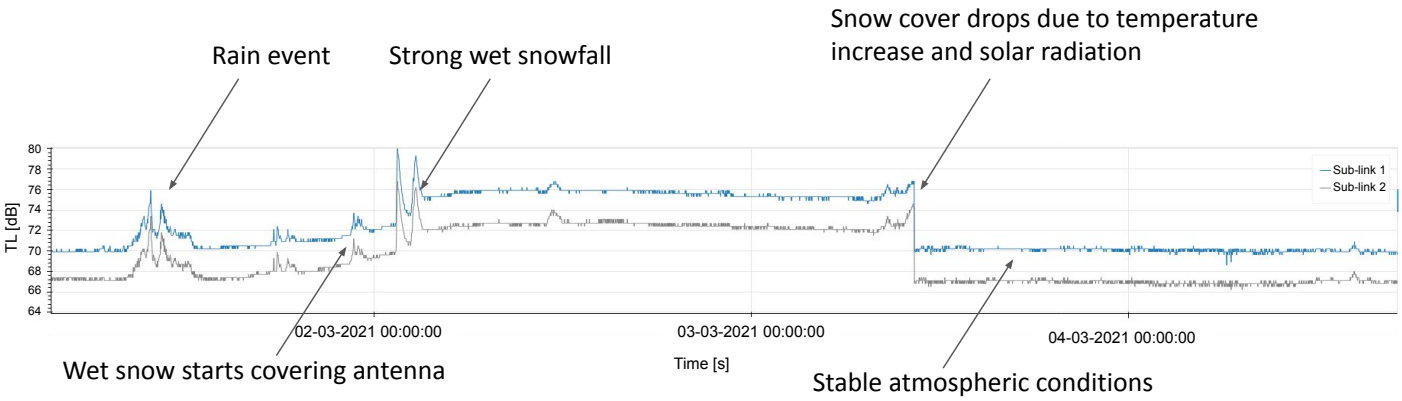

Figure 1.4: Total loss (TL) time series of an 18.2GHz CML with a path length of 12.1km from 1st to 4th of March 2019. The signal pattern shows a rain event followed by a temperature drop and wet snow that accumulates and freezes on the antenna until it eventually drops under melting conditions. The event illustrates the complexity of compound events that cause fluctuations in the CML signal level.

Schleiss and Berne (2010) established a dynamic detection method based on the rolling standard deviation (RSD) of the time-series. The assumption used to justify the algorithm was that fluctuations of a significant magnitude are only due to attenuation caused by rainfall. However, even during dry periods signal fluctuations of a similar magnitude can occur (see Fig. 1.4 or van Leth et al. (2018)). It is therefore unclear if the RSD method is suitable for country-wide rainfall estimation using a large CML dataset with heterogeneous data quality (see research question SQ2 below).

Once a rain event is detected, the baseline attenuation A_b is estimated by an average of the TL of the preceding dry period which is assumed to be an accurate estimate of the atmospheric loss. The baseline is often assumed to be constant during rain events.

Additional wet antenna attenuation A_w is estimated using either a constant or a dynamical time or rain rate dependent model (Schleiss et al., 2013; Leijnse et al., 2008; Valtr et al., 2019). More details about the compensation of A_w are given in Chapter 2.

Assuming that $A_u = 0$ during rain events, A_R can be estimated up to the random error ϵ by subtracting A_b and A_w from TL. To estimate a path-averaged rain rate, the k-R relation for specific attenuation at the CML frequency and polarization is used:

$$A_R = kL = \alpha R^\beta L \Leftrightarrow R = \frac{1}{\alpha L} A_R^{\frac{1}{\beta}}, \tag{1.5}$$

where L is the length of the CML in km.

More details about state-of-the-art CML processing routines are given in Chapters 2 and 3. CML rainfall estimates are path averages and, therefore, hard to interpret for example for stakeholders in disaster management or agriculture. Therefore, spatial interpolation as a standalone sensor solution or merging with other sources of rainfall information to obtain a gridded product is necessary (Blettner et al., 2022).

The systematic and random errors of CML measurements are as follows: As the k-R power law is based on DSD calculations, the DSD dataset needs to reflect the local rainfall climatology. Thus, systematic errors due to inadequate DSD data are possible and random errors of the power law fit are given (Chwala and Kunstmann, 2019).
A crucial aspect besides DSD uncertainties is the correct attribution of attenuation to rainfall. The assumption of Schleiss and Berne (2010) that all significant signal fluctuations are due to rainfall has proven to be too simplistic for commercially deployed microwave links. Signal anomalies, that is, fluctuations not caused by rainfall, can occur during dry periods (see Fig. 1.4). A false classification of such events as wet can lead to a large overestimation. A number of case studies have been conducted to learn about the behavior of CMLs in a real-world setting by either placing added instrumentation such as rain gauges or disdrometers along the link path (Špačková et al., 2021) or by setting up research microwave link experiments with the benefit of added observables such as phase shift or dual polarization (Chwala et al., 2014; van Leth et al., 2018; Moroder et al., 2019). However, the detection and attribution of signal anomalies without added instrumentation or the use of research microwave links remains an open question. Even if rain events are detected correctly, anomalies that occur during rain events are a problem that is not sufficiently treated in state-of-the-art processing routines of CML signal levels (see specific research question SQ4 and Chapter 5).
A limited transmission power in combination with a detection limit of the receiver of CMLs certainly limits the maximum measurable rainfall intensity. If strong attenuation leads to a complete loss of signal (blackout) no RSL reading is available and the total loss can not be computed. It is unclear to what extent this affects the correct estimation of rainfall extremes (see specific research question SQ3 and Chapter 4).
Depending on the frequency and length of a microwave link, DSD variations along the link path can cause random errors (Berne and Uijlenhoet, 2007). Similar to rain gauges, the signal quantization of CMLs leads to a random uniformly distributed error (Zinevich et al., 2010).

1.2.2.4 Weather radar

Single polarized weather radars (short for RAdio Detecting And Ranging) yield an indirect measurement of precipitation by directly measuring the reflectivity of falling hydrometeors (Doviak et al., 1994). We can describe the reflectivity Z in $\mathrm{mm}^6\,\mathrm{m}^{-3}$ as

$$Z = \frac{\lambda^4}{\pi^5 |K_W|^2} \int_0^\infty N(D)\sigma_{bsc} dD, \tag{1.6}$$

where the back-scattering cross section σ_{bsc} in cm^2 depends on the polarization of the radar and where the factor K_W summarizes the dielectric particle properties (Ekelund et al., 2020). Similar to specific attenuation k we can derive a $Z - R$ power law by comparing Equations1.6 and 1.1. However, typical radar frequencies yield highly non-linear relations

like the one derived by Chen et al. (2021) for horizontal polarization at C-band (4-8 GHz):

$$R = 0.052Z^{0.57} \Leftrightarrow Z = 178.7R^{1.75} \tag{1.7}$$

This relation is highly DSD-dependent and requires a specific choice of coefficients for different scenarios, like the relations $Z = 300R^{1.4}$ in convective and $Z = 200R^{1.6}$ in stratiform rain as used for the Weather Surveillance Radar – 1988 Doppler (WSR-88D) at S-band (3GHz) (Chen and Chandrasekar, 2021).

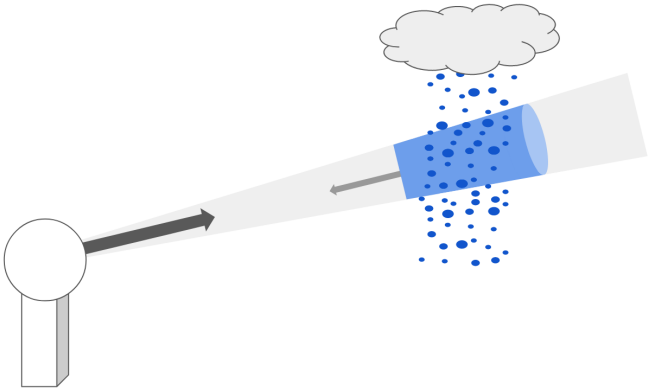

Figure 1.5: Weather radars are active sensors that measure the reflectivity of raindrops illuminated by the radar beam.

Reflectivity is derived from a measurement of back-scattered power P of a volume of raindrops at range r illuminated by the radar beam as illustrated in Fig. 1.5 (Villarini and Krajewski, 2010). Since this volume can be far from the radar, compensation for the two-way attenuation PIA_2 of atmospheric phenomena like rain, gases, and aerosols in between is necessary. Using a radar-specific constant C, P is approximated as

$$P = Z\frac{C|K_W|^2}{PIA_2^2 r^2} \tag{1.8}$$

The estimation of PIA_2 is non-trivial and often relies on an independent measurement or the use of additional sources of information, such as numerical weather models to estimate temperature gradients (Ryzhkov et al., 2014).
In addition to the indirect measurement using a DSD-dependent relationship, radar measurements need to avoid obstacles near the ground and, therefore, use an orography following beam angle. This leads to measurements high above the ground in regions far from the radar which increases the error of radar measurements compared to sensors on the ground (Pejcic et al., 2020).
Wilson and Brandes (1979) summarize the main error sources for weather radar rainfall estimation as variations in the relationship between the backscattered energy and rainfall rate, changes in the precipitation before reaching the ground, and anomalous propagation

of the beam. A more recent review of uncertainties in radar rainfall estimates by Villarini and Krajewski (2010) lists the following error components: A miscalibration of the radar constant C (see Eq. 1.8) due to deterioration of the hardware or temperature dependencies, attenuation of the beam, ground clutter and anomalous propagation, beam blockage, variability of the Z-R relation, beam broadening and overshooting, vertical variability of rainfall, vertical and horizontal air motion, and temporal sampling errors.
Especially DSD changes and precipitation drift before reaching the ground lead to systematic and time-dependent errors that hinder a direct comparison to rain gauges and other sensors on the ground. Single measurements of reflectivity are insufficient to correct such scenario-dependent errors because of the DSD uncertainty of the Z-R relation. It is, therefore, necessary to correct radar QPE using spatial and temporal dynamics as boundary conditions for correcting rainfall estimates (see research question SQ5).
Much work has been conducted to improve weather radar QPE. The main approaches are quality control algorithms like attenuation correction, adjustment of radar estimates using additional sources of information such as rain gauges (Vogl et al., 2012; Pulkkinen et al., 2016; Moraux et al., 2021), and radar polarimetry (Ryzhkov and Zrnic, 2019; Chen et al., 2021). Radar polarimetry is especially promising because measurements of differential reflectivity and specific differential phase can be used to further reduce DSD uncertainties (Chen et al., 2021). Additionally, path-integrated attenuation can be estimated by using the total span of differential reflectivity (Bringi et al., 1990; Testud et al., 2000; Chen et al., 2021).

1.2.3 Rainfall Observation Data in Germany

The following section describes rainfall data from rain gauges, CMLs and weather radars available for research purposes in Germany. All presented datasets have been considered for at least one of the studies conducted within this thesis.

1.2.3.1 Rain Gauge Data

The German meteorological service (DWD) provides multiple rain gauge datasets which are freely available on the opendata online archive (`https://opendata.dwd.de/`). One set of observations includes 1-minute automatic rainfall estimates from a network with an average station density was one rain gauge per 330km^2. The network operates weighing gauges, initially Ott Pluvio^2 sensors, which were gradually replaced by rain[e] sensors from Lambrecht meteo from 2018 to 2020. The sensor resolution is 0.01 mm (Pluvio^2) and 0.001mm (rain[e]) both quantized to 0.01mm for consistency within the dataset. The seccond set of observations are daily measurements from Hellmann rain gauges with a manual readout of the accumulated rainfall amount of the last 24 hours at 5:50 UTC with a quantization of 0.1mm. For the period from 2013 to 2021, 1066 1-minute and 2150 daily rain gauge stations were considered for studies conducted in this dissertation. Around 50% of the daily rain gauges share their location with the 1-minute rain gauges.

1.2.3.2 CML Data

The CML data accessible for research purposes in Germany is recorded from a network operated by Ericsson Germany and and counts 3904 unique CML paths distributed over entire Germany. The length of CML paths ranges between a few hundred meters to almost 30km and the frequency ranges from 10GHz to 40GHz. The hardware of the CMLs is homogeneous within the dataset using Ericsson MINI-LINK Traffic Node systems in all cases. The real-time data acquisition system used to retrieve instantaneous TSL and RSL records from two sub-links per CML path is described in Chwala et al. (2016). Data collected at Ericsson is immediately sent to and stored at a server at the Karlsruhe Institute of Technology (KIT). The temporal resolution of the recorded signal levels is 1 minute and the power resolution is 1dB for TSL and 0.3dB (occasionally 0.4dB) for RSL. Data for all of Germany is available from September 2017 onwards. The size of one year of data is approximately 100GB.

1.2.3.3 Weather Radar Data

DWD operates a network of 17 C-band weather dual-pol doppler weather radars shown as black dots in Fig. 1.6 and provides multiple precipitation products as a gridded composite of data from the individual radars. To date, all available products are based on measured reflectivity for a single polarization using a three-part Z-R relation (Bartels et al., 2004). The measurement height above ground (see Fig. 1.6) is using a beam angle of 0.8 degrees where no obstacle is blocking the beam and a terrain following angle otherwise which can lead to measurement heights up to several kilometers.
Weather radar QPE is provided in the "Radar Online Aneichung" (RADOLAN) and "Radarklimatologie" (RADKLIM) product series that apply a range of quality checks and bias correction techniques. Products selected for studies conducted within this thesis are as follows:
RADOLAN-RY is a quality-checked and attenuation-corrected composite with a spatial resolution of 1 km by 1 km and a temporal resolution of 5 minutes which is the instantaneous measurement frequency of the radar scans and therefore the highest possible resolution (Bartels et al., 2004).
RADOLAN-RW is a gauge-adjusted version of RADOLAN-RY with a temporal resolution of 1 hour using a combination of multiplicative and additive factors. The factors are derived from a comparison of the gridded radar data to the automatic rain gauges described above (Bartels et al., 2004).
RADKLIM-YW is performing a similar gauge adjustment where weights are calculated at a 1-hour resolution and remain constant for the 12 5-minute timesteps within this hour to provide a temporal resolution of 5 minutes. In addition, the product is climatologically corrected using the manual, daily rain gauge data. These climatological corrections aim to reduce static errors like beam blockage and range-dependent underestimation (Winterrath et al., 2017).

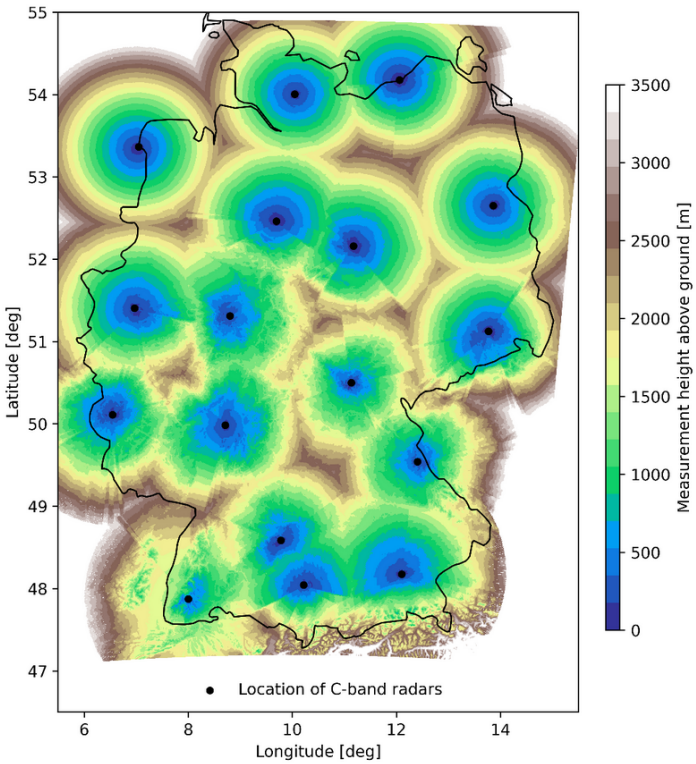

Figure 1.6: Minimal measurement height above ground derived from the German radar network composite on the RADOLAN grid. The precipitation scan uses an orography following elevation angle between 0.5 and 1.8 degrees. In regions with overlapping radar measurements, the minimum of measurement heights is shown. Figure adapted from Polz (2023b).

1.3 Deep learning

1.3.1 Statistical learning

The primary goal of data-driven modeling of geophysical processes is the approximation of an unknown function which is assumed to be generating the observed data. We call this function the data generating process. An abstract definition of machine learning is to say that instead of hard coding rules into a computational system it is given the ability to acquire its own knowledge by learning from observed data (Goodfellow et al., 2016).
Hardcoded rules in the form of physical formulas are available in the case of geophysical processes, but there are two main limitations: The first one is that, while individual processes have a concise description, a data-generating process outside a lab environment is composed of many known and even unknown processes that increase the system's complexity. It is hard to impossible to describe all of them explicitly. The second one is that modeling physical processes in the form of differential equations can be computationally expensive, which is why the computation is usually performed on larger scales. Smaller scales are parameterized. However, scale interactions and scenario-dependent choices of parameters may lead to a bad general performance. Data-driven approaches can handle smaller scales because they are less expensive to compute (Pathak et al., 2022) and they can learn which scales are sufficiently well approximated by an average (see scale separation below).
Modern machine learning draws the ability to learn from observed data from the mathematical theory of statistical learning. Essentially, statistical learning is concerned with function approximation by parameterized families of functions. The target function usually describes a prediction in the form of a regression or classification problem. The learning process is a minimization of a predefined objective function that estimates the deviation from the observed ground truth.
Classical machine learning algorithms like logistic regression or naive Bayes classifiers heavily rely on the representation of the data (Goodfellow et al., 2016). The main idea behind deep learning is to learn from raw data instead of engineered features. In order to be able to gain knowledge from raw data it is favorable to represent the data by simple concepts that are then used to describe more complex features. Deep artificial neural networks (ANNs) achieve such a hierarchy of representations by using multiple layers of artificial neurons. Before going into more detail about ANNs we will introduce the basic terms of machine learning and statistical learning. The required components for statistical learning of a given task like regression or classification are data, labels, hypothesis space, objective function and update rule.
The data is the collected information that is available for the given task. It contains the predictor variables and serves as model input. The labels are the ground truth associated with the data and serve as the expected model output. An example pair for data and labels is observed precipitation in a catchment and observed river discharge. The task of predicting discharge from precipitation is a regression problem.

The hypothesis space consists of a parameterized family of functions which are assumed to be hypothetical candidates to describe the relation between data and labels. In our case, the hypothesis space will consist of different weight configurations of a chosen neural network architecture.
The objective function is a performance measure that describes how well a considered parameterization in the hypothesis space performs, for example, the mean squared error between observed and predicted discharge.
The update rule or optimizer is an algorithm that describes the choice of a future parameterization given the current one with respect to the objective function. The goal is to successively increase the performance until an optimum is reached. For gradient-based learning, we will mostly consider variations of stochastic gradient descent. Note that if the objective function and update rule do consider the labels the learning process is called supervised learning. If they do not consider the labels it is called unsupervised learning.
Statistical learning theory aims to provide suitable hypothesis spaces, objective functions, and update rules for machine learning. We will see in the next section that the theory is able to prove that artificial neural networks are universal approximators of continuous functions. However, statistical learning theory also shows that there is a caveat to the choice of a hypothesis space, namely, the so called "no free lunch" theorem which essentially states that for a randomly chosen optimization problem, no algorithm performs better than random search (Wolpert and Macready, 1997). This leads to the conclusion that the choice of a suitable hypothesis space, objective function, and update rule is tied closely to any available knowledge about the given learning problem. We will later see which assumptions can be made about structured data from a physical world.

1.3.2 Artificial neural networks

Artificial neural networks (ANN) are a class of functions used in statistical learning theory. They are inspired by biological neurons and, therefore, are linked to the term artificial intelligence (AI). However, the term AI is very abstract and includes different approaches. One of the largest domains is machine learning which includes ANNs. Before going into details about why and when deep learning is a suitable approach to high dimensional learning problems in the geosciences we need to see that ANNs are suitable for function approximation. The artificial neuron is in fact a function

$$\begin{aligned} f: \quad & \mathbb{R}^n \to \mathbb{R} \\ & (x_i) \mapsto \rho \sum_i w_i x_i \end{aligned} \tag{1.9}$$

with weights $w_i \in \mathbb{R}$ and activation function $\rho : \mathbb{R} \to \mathbb{R}$. ANNs consist of layers of neurons. The simplest ANN is the multilayer perceptron (MLP), where the input of each neuron consists of the output of all neurons in the previous layer. Such networks that pass information uni-directional from the input to the output layer are called feed-forward neural networks. There are other concepts of information flow in ANNs, but we will omit

them in this section. An MLP is uniquely determined by a pair (W, ρ) consisting of weight matrix W and activation ρ. More complex neural network architectures such as convolutional neural networks have fewer connections between neurons than an MLP. If networks consist of more layers than one input and one output layer, they are called deep neural networks. Note that the non-linearity of the chosen activation function is an important property since a deep neural network with linear activation is equivalent to an ANN with just one layer. Common activation functions are step functions, the sigmoid function $\frac{e^x}{e^x+1}$ and the rectified linear unit (ReLU) $max(\{0, x\})$. It is easy to see how function approximation with ANNs works since it is common knowledge that any continuous function can be approximated arbitrarily well as a series of step functions (Forster, 2016).
Deep learning describes the optimization of the weights of a deep neural network to perform on a given data set. Statistical learning theory provides the framework for this optimization process. In fact, it has produced remarkable results regarding the parameterized function class of neural networks (Cybenko, 1989; Hornik et al., 1989):

Universal Approximation Theorem
Let $C(X, Y)$ denote the set of continuous functions from X to Y, then $\rho \in C(\mathbb{R}, \mathbb{R})$ is non-polynomial if and only if for all $f \in C(K, \mathbb{R}^m)$ with compact support $K \subseteq \mathbb{R}^n$ and for all $\epsilon > 0$ $\exists k \in \mathbb{N}$, $A \in \mathbb{R}^{k \times n}$, $b \in \mathbb{R}^k$, $B \in \mathbb{R}^{k \times m}$ such that

$$sup_{x \in K} \parallel f(x) - g(x)) \parallel < \epsilon,$$

where $g(x) = B(\rho(Ax + b)$ is a multi-layer perceptron with arbitrary width and bounded depth.

The theorem states that multi-layer feed-forward neural networks with a suitable activation function are universal approximators, i.e. they approximate any continuous function with compact support arbitrarily well. The compact support assumption is not restrictive in most cases.
So far we only treated the existence of an approximation to a known function with arbitrarily small error. The crucial aspect is finding the approximation when the function is unknown, but data generated by the function is available. This is the actual learning process. A consequence of the universal approximation theorem is that any inaccuracy in the approximation must be a consequence of inadequate learning. In machine learning, the learning process is very often referred to as the training process.
The most common way of training an ANN with data is by means of gradient-based learning using back-propagation (Rumelhart et al., 1986). Let $y = f(x, W)$ be the output of a feed-forward neural network $f(\cdot, W)$ with weights W that was used to evaluate the input x. Furthermore, let $\theta(y, y') = \theta(f(x, W), y')$ be the objective function evaluated at y with respect to the ground truth y'. Then, back-propagation can be used to compute the gradient $\nabla\theta(W)$ of θ with respect to the weights W. A benefit of ANNs is that partial derivatives can be computed layer-wise using the chain rule.

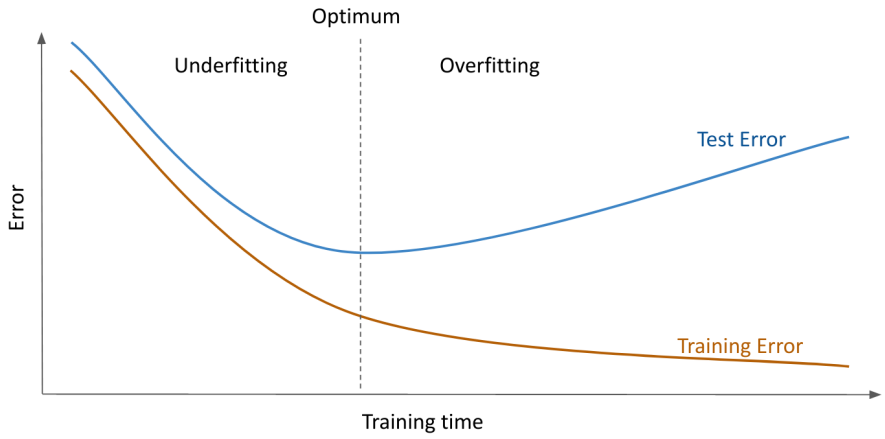

Figure 1.7: Idealized evolution of train and test error while training artificial neural networks.

Once the gradient is known, stochastic gradient descent can be used to update the weights W in order to improve on the objective function (Kiefer and Wolfowitz, 1952). We have

$$W_{n+1} = W_n - \lambda \nabla \theta(W_n), \tag{1.10}$$

where a defined learning rate $\lambda > 0$ is used to control the speed (step width) of the gradient descent to achieve $\theta(W_n) \geq \theta(W_{n+1})$ (Bottou et al., 2018). To see where an optimum is reached we need to define the training and generalization, or test, error. The training error is the objective function evaluated using the training data which, in turn, is used to optimize W. The generalization error is the objective function evaluated using a separate test set that is assumed to be independent and identically distributed (i.i.d.) with respect to the training data. To achieve a better test error it can be favorable to average gradients over a so-called "batch" of instances, or samples, of the training data. Fig. 1.7 shows an idealized picture of the evolution of train and test error during the training process. The stage where both train and test error still improve is called underfitting while overfitting refers to the stage where the training error decreases and the test error increases. The goal of the training process is to stop at the optimal test error between the two stages.

1.3.3 High-dimensional learning

The curse of dimensionality refers to a series of problems that arise in high-dimensional learning problems (Zimek et al., 2012). The main issue is the distance concentration effect which describes the phenomenon that the higher the dimensionality of a space, the more pairs of points in this space have the same distance (Aggarwal et al., 2001). Since objective functions in the training of ANNs are essentially distance metrics, this is causing problems in the computation of gradients during training.
As an example, consider an n-dimensional ball of radius 1, then half of its volume is closer than $r = 1 - \sqrt[n]{0.5}$ to its boundary. As a percentage of the sphere radius, r is 30% in

two dimensions, 16% in four dimensions, and less than 1% in 69 dimensions. This means that if data was uniformly distributed in this ball, then almost all pairs of points have the same distance in high dimensions. The consequences for the functional approximation of such data are illustrated by another example given by Bronstein et al. (2021) who show that the required number of observations to approximate a 1-Lipschitz function f is exponential in the dimensionality of f if the error is required to stay below a given threshold. However, the class of 1-Lipschitz functions is large and not clustered in its domain just like the uniformly distributed points in a ball.

Why is this important? It means that without further structure in the data, the required sample size for high-dimensional learning problems grows exponentially. However, data that is generated by physical processes is structured by underlying principles of symmetry. Symmetries are invariances under transformations of the observed space and include, for example, temporal or spatial translations or rotations. By exploiting these symmetries the dimensionality of the problem can be reduced if describing the data with less explanatory variables is possible. Most physical problems are referred to as one (time-series), two (spatial or image), or three (volume or video) dimensional. From a data-driven perspective, the dimensionality of the problem changes, since model input is not processed by physical equations that act on each point in low dimensional space (e.g. Brownian motion of a particle suspended in a medium). The process acting on a point in time is viewed as a function of all knowledge of the whole system. Each observed point in the system, e.g. a grid cell or a single point of a time series, creates its own dimension in the feature space, that is, the domain of explanatory variables of the system. As a result, the feature space is often very high dimensional. E.g. a single time-step in the radar composite used in the RADOLAN product consists of 900 by 900 grid cells which amounts to an observation in an 810000-dimensional space.

To see that dimensionality can be reduced if structural information is known let us consider a time series that we want to model as a stochastic process. The goal is to predict the future behavior of the series given its past. If the Markow condition holds, e.g. if the series describes a Brownian motion, then using only a limited number of past time steps produces equally well-performing predictions than any longer history of the series. Now what if symmetries that can be exploited to reduce the dimensionality of the problem are unknown? We will see that the extraction of useful features from raw data can be learned using suitable ANN architectures.

The term prior originally refers to a prior probability distribution in Bayesian statistics. A prior is previous knowledge about the data that can be used to increase the accuracy of a prediction. There are a number of priors that can be utilized to effectively use ANNs for rainfall estimation.

The main assumption that needs to be made to approach physical processes as function approximation problems is that the data-generating process is a continuous function in Euclidean space. In a real-world scenario, this assumption is not always verifiable. Additionally, there is noise due to random and systematic errors when observing the data.

Atmospheric science is concerned with scales where this assumption is reasonable, yet the practical applicability is always verified using an evaluation strategy that assures the transferability and robustness of the approximation.

Scale separation is a powerful prior for approaching complex dynamical systems and for effective learning in high dimensions. The underlying idea is that the behavior of the system can be described by a decomposition into different scales (Bronstein et al., 2021). A good example is Fourier analysis which assumes that a signal can be decomposed in the frequency space. Different frequencies describe the different scales. When modeling a process this is helpful, because usually the larger scales (lower frequencies) are well approximated if they are assumed to be locally constant. Scales that are small relative to the observed scale of the system can be averaged since they appear as random fluctuations. For example, this applies to fluctuations smaller than the resolution of a sensor. What is left is the dynamic process that moves fast enough to change while being observed and slow enough to not appear as random fluctuations. It makes sense to assume that, the more complex a system is the more interaction happens between scales (Peters et al., 2004). Then, an averaging of the larger and smaller scales would be less effective. But it also makes sense to focus on modeling the scales that are not well approximated by an average. A conjecture of this thesis is that scale separation is especially useful for improving rainfall estimates subject to errors that are present at different scales because the processes that cause them can be slow- or fast-moving.

ANNs apply the concept of scale separation by learning a hierarchical representation of the observed data. The reason is that their implementation uses many layers of neurons that successively increase the complexity of extracted features. It has been shown that increasing the model complexity in terms of free parameters (wider networks) is not as effective without also increasing the number of layers of the network (deeper networks) (Goodfellow et al., 2014).

Another prior that is helpful for high dimensional data from CMLs or weather radars is reducing the sparsity of useful information by feature extraction. The dimensionality of the data is defined by its spatial and temporal extent. On the one hand, longer time sequences or larger rainfall fields provide context for useful predictions. On the other hand, rainfall is characterized by a high intermittency, that is, long periods without signal alternate with shorter periods of rainfall.

Thus, not all information in a sequence of rainfall fields is useful, and trivial signals may be ignored. This fact can be used to reduce model complexity and problem dimensionality. A desired property for ANNs is to divide them into two functional parts. The first part is meant to extract a hierarchy of features from the raw data and the second part is used to make predictions based on these features. The feature extraction part is desirably an equivariant operation, meaning that a shift (in time or space) in the input sequence should result in a shift of the output sequence.

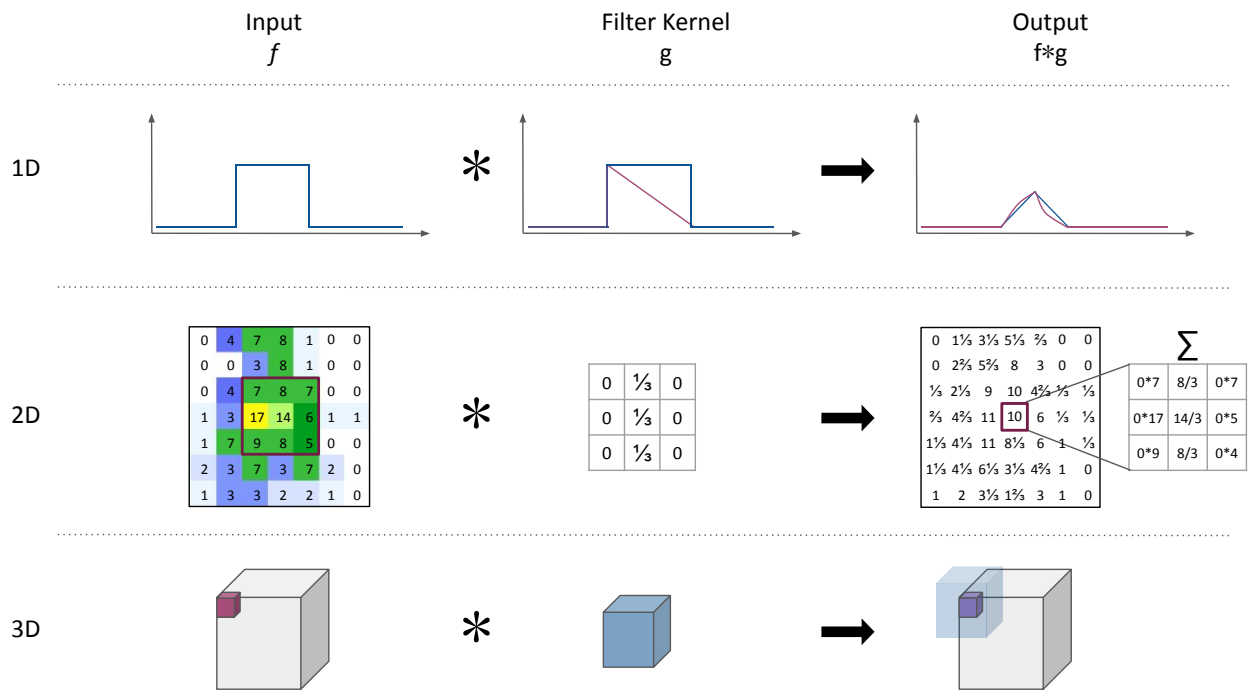

Figure 1.8: Examples of convolution in different dimensions. To be consistent with the notation of convolutional layers in neural networks, the observed signal f provided to the model is denoted as input and g is called the filter kernel. The top row shows a 1-dimensional convolution of time series data. The middle row shows a discrete example of 2-dimensional convolution where the filter kernel is equal to a column-wise moving average. The convolution can be performed in arbitrary dimensions as shown in the graphical illustration of the 3-dimensional case in the bottom row.

A suitable equivariant mathematical operation is convolution of two functions f and g:

$$f * g(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt \tag{1.11}$$

We will think of f as observed data that serves as model input and of g as a filter that is used to extract features from the data. In our case the filter kernel $g = (g_k)_{k\in[-n,n]}$ has finite support which leads to the discrete version

$$(f * g)_i = \sum_{k=-n}^{n} f_{i-k}g_k \tag{1.12}$$

Examples of convolution in 1, 2, and 3 dimensions are given in Fig. 1.8. It can be seen that the filter kernel acts locally on the input.

A convolution layer can be expressed in the same way as an MLP layer, except, that the weights are shared between neurons, i.e. instead of separate weights for all inputs from all neurons of the previous layer, a neuron receives only valid inputs from a neighborhood defined by the filter kernel and zero otherwise. A different neuron receives valid inputs

from its own, different neighborhood, but uses the same weights. ANNs that use convolution layers are called Convolutional Neural Networks (CNNs) (LeCun et al., 2015). Gradient-based training of CNNs works in the same way as for MLPs. CNNs originate from image recognition and are inspired by the visual cortex of mammals. They are designed to recognize objects or patterns, regardless of their location in images or time series (Fukushima, 1980). In addition to the location equivariant convolution, CNNs often use local and global pooling layers to select superior features and reject unimportant ones which leads to higher performance and faster convergence (Cireşan et al., 2011). It has been shown that the weight sharing and related minimization of free parameters in CNNs enhance their generalization ability (Lecun, 1989).

1.3.4 Neural network architectures

The type and order of layers of a neural network are referred to as its architecture. Here, we will elaborate on different neural network architectures that allow for effective learning in high dimensions. The main concept that is applied is convolution.
Fig. 1.9 shows a 1-dimensional CNN for time-series classification with an architecture inspired by LeNet (LeCun et al., 1998). It consists of multiple blocks with two convolution layers followed by one max pooling layer which form the feature extraction part of the network. While each pooling layer reduces the dimension of the input time series, the number of applied filter kernels is increased in each block. Once the dimensionality is reduced sufficiently, global average pooling is applied. Here, the classification part, that is, the interpretation of the features starts using fully connected layers. To avoid over-fitting to the training data dropout layers are added (Srivastava et al., 2014). In summary, this network architecture uses functional layers that are targeted to enhance the generalization, convergence, and classification performance. This architecture is used in Chapter 3. Fig. 1.10 shows a residual neural network (ResNet), an extension of the classical CNN approach (He et al., 2015). The network uses so-called residual blocks, where the convoluted input is afterward added to the raw input. In this case the raw input is additionally processed by a convolution layer with kernel size one and no activation which allows for a multiplicative scaling of the input. The larger part of the model capacity, i.e. the convolution with non-linearity, is used for the residuals rather than the quantity itself. This network design encourages the representation of features as a perturbation of the input which is beneficial for the representation of dynamical systems (Bronstein et al., 2021). Another perspective on the benefit of ResNets is that they allow for deeper networks without running into gradient problems (He et al., 2015). We will use a 3-dimensional version of this ResNet approach in Chapter 6.

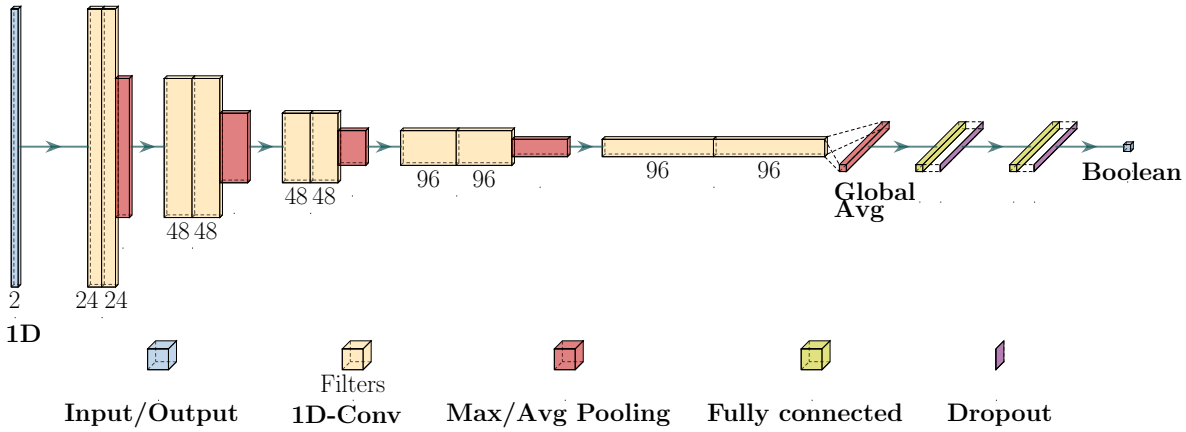

Figure 1.9: 1-D Convolutional neural network with alternating convolution and pooling operations similar to LeNet (LeCun et al., 2015). The numbers below the layers indicate the feature dimensionionality. Figure designed using LaTeX code provided by Iqbal (2018).

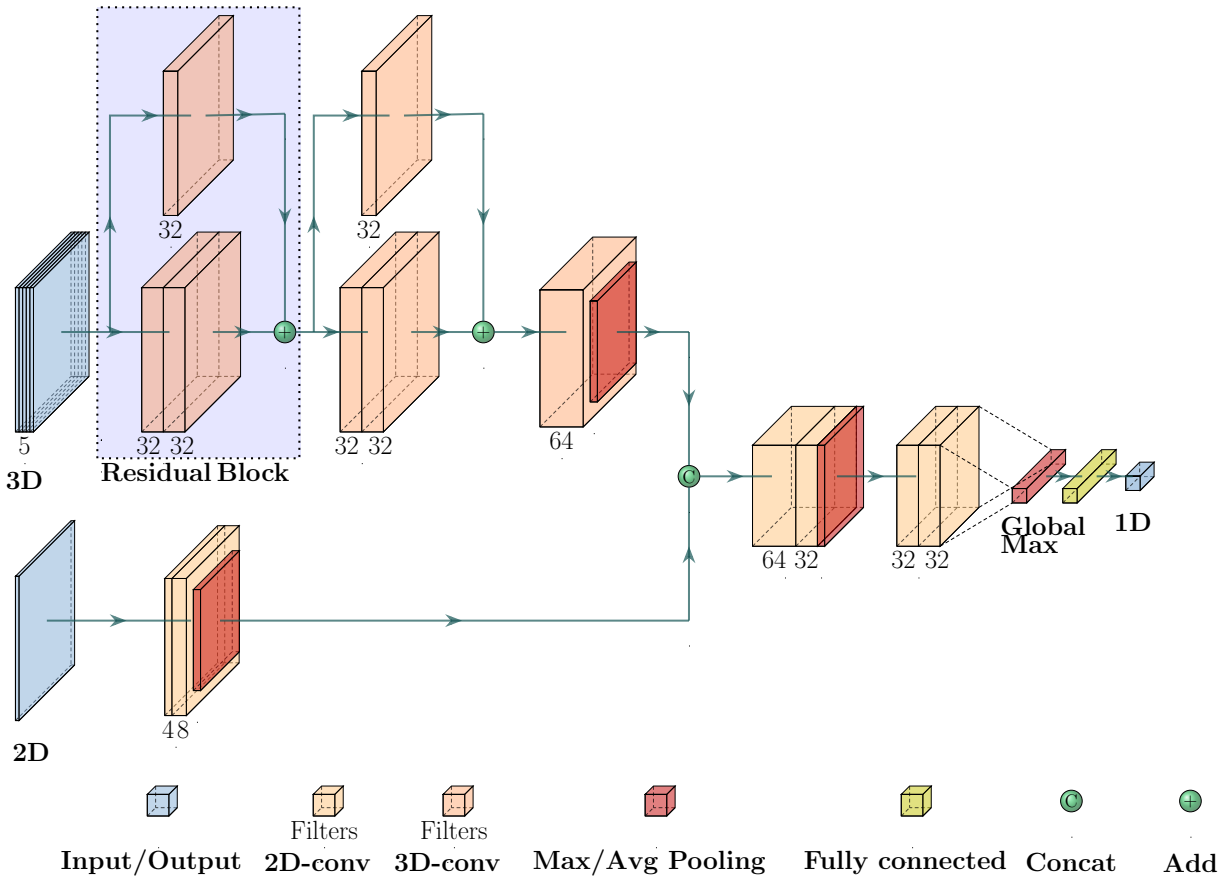

Figure 1.10: ResNet architecture with 3D and 2D input branches and skip connections to model residuals. Figure designed using LaTeX code provided by Iqbal (2018).

1.4 Research questions

The overall objective of this thesis is the improvement of CML and weather radar-derived quantitative precipitation estimates using deep learning approaches. Both types of sensors indirectly observe rainfall by measuring the microwave attenuation or reflectivity of raindrops. They are challenged by different sources of systematic and random errors that disturb the measurement. The aim is to both, analyze the systematic errors of these two types of sensors, and reduce them using deep learning approaches. The random error from DSD uncertainties in the k-R and Z-R relations, which have been treated in numerous studies, will be less of a focus than systematic errors that require sensor quality control. To achieve this goal it is necessary to define suitable objectives, identify the benchmark performance of the state-of-the-art, and find solutions to improve quantitative precipitation estimates from there. Deep learning has been shown to provide useful solutions for modeling atmospheric phenomena. Therefore, deep neural networks will be applied in this thesis to achieve high-performing signal processing routines that help mitigate systematic errors of CML and weather radar rainfall estimates. As described above, specific opportunities but also challenges that arise in the application of deep neural networks to geo-scientific problems are anticipated. Additionally, there is hope to gain more general conclusions about these challenges and opportunities. In summary, the aim is to answer the following two overarching research questions (OQ).

Overarching Research Questions

OQ1 Can systematic errors of CML and weather radar QPE be quantified, attributed to causes, and corrected?

Section 1.2.2 reviewed the potential causes for systematic errors of rainfall sensors placed in a complex natural environment. For CMLs, the identified errors were, on one hand, an incorrect attribution of attenuation to rainfall, for example, due to strong signal fluctuations during dry periods or wet antenna attenuation. On the other hand, there are missing observations due to a complete loss of signal in heavy precipitation. Random errors due to DSD variability and signal quantization need to be considered in addition to systematic errors.

For weather radars, systematic errors can be due to radar miscalibration, attenuation of the beam, ground clutter and anomalous propagation, beam blockage, scenario-dependent variability of the Z-R relation, beam broadening and overshooting, vertical variability of rainfall (compared to ground sensors), and temporal sampling errors. Random errors due to DSD uncertainty apply similarly to the case of CMLs. While the presence of such errors is known in theory, the frequency of their appearance in real-world rainfall observations is often unknown. The attribution to a cause of an observed error in the rainfall estimate is also not always possible, as is its correction. It is one objective of this thesis to achieve more clarity on the potential for quantification, attribution, and correction of systematic errors within

rainfall estimates from CMLs and weather radars.

OQ2 Are artificial neural networks a robust and transferable tool to provide quantitative precipitation estimates?

Geophysical applications of deep learning algorithms are, in general, different from classical deep learning problems like image classification. The main differences that also apply to a successful application of deep learning in rainfall estimation are as follows:

First, rainfall is characterized by a heavily skewed extreme value distribution and high intermittency. It is unclear if deep learning algorithms, that were initially developed for image recognition tasks, are *robust* tools to model such extreme value distributions and if the model skill decreases for extreme events that occur at a low frequency.

Second, the i.i.d. assumption that train and test sets are independent and identically distributed is harder to achieve. A covariate shift between training and inference is assumed to be much more likely. Mainly, the non-stationarity of observed systems may also lead to a decreased performance if the distribution of the train and test data are different. It is unclear if the *transferability* to new locations and different time periods can be achieved without a loss of skill. This especially applies to rainfall sensors where both, the local climatic conditions as well as the sensor hardware may change in time and space.

Specific Research Questions

The following specific research questions (SQ) are used to answer the overarching questions. An overview of the connections between specific and overarching questions is given in Fig. 1.11. Answers to these questions are achieved by the individual studies conducted within this thesis (see Chapters 2 to 6). The questions are ordered chronologically by the time when the studies were conducted.

SQ1 How do established signal processing techniques compare in a large-scale evaluation of one year of CML data in Germany?

CMLs are opportunistic sensing devices in the sense that they are not designed for rainfall estimation. Thus, they have an increased heterogeneity in data quality which leads to complex data processing (see Section 1.2.2.3). A series of methods for the detection of rain events and the compensation of wet antenna attenuation have been proposed by research groups in different countries. Due to a lack of openly available CML data, each method has been developed using a different dataset. To be able to measure improvements in CML signal processing routines it is essential to draw a baseline for the performance of state-of-the-art methods. Not only the application of these methods on the same data set but also the choice of a suitable framework to evaluate them is necessary.

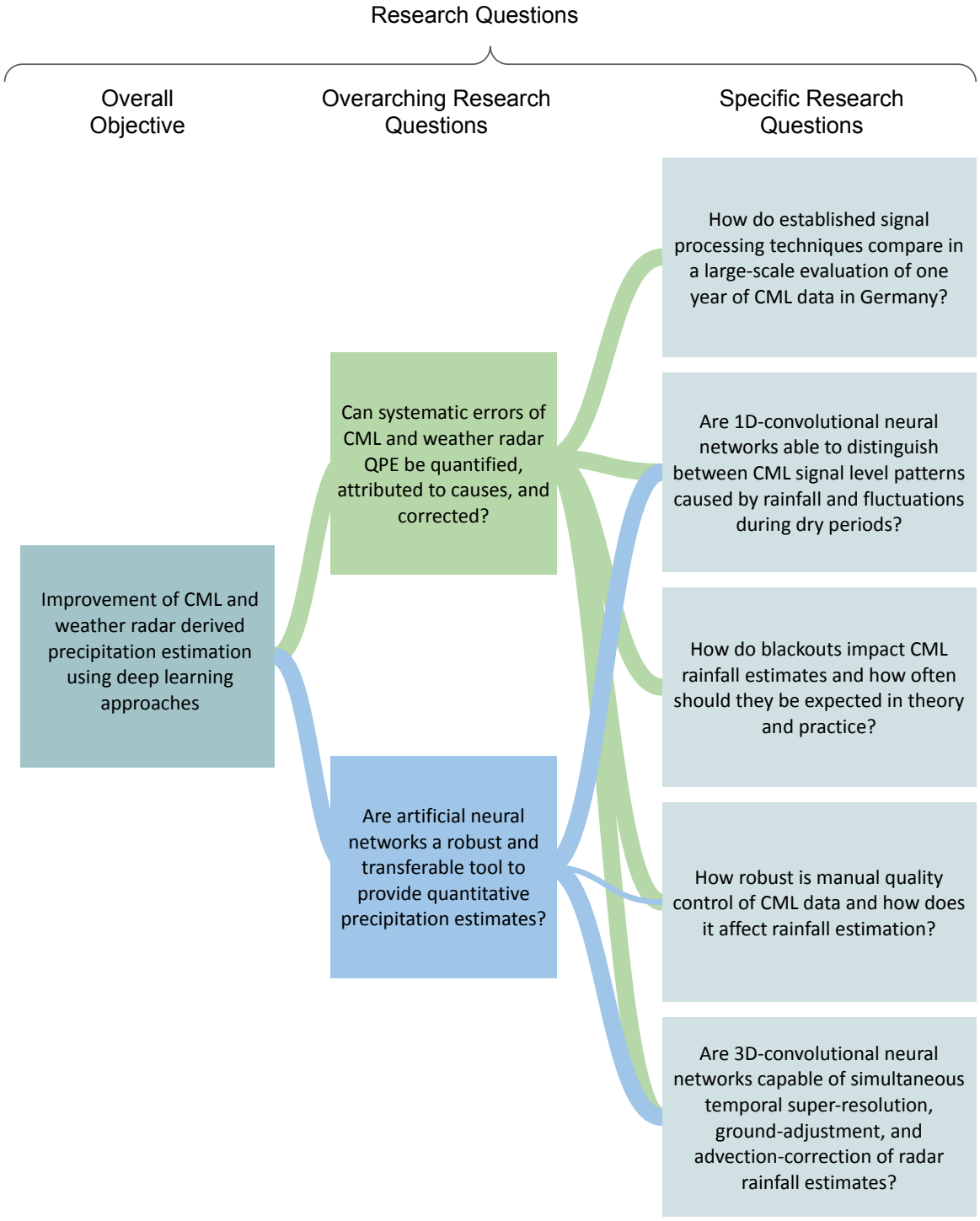

Figure 1.11: Research questions: The overall objective of improving rainfall estimates using deep learning techniques is achieved by answering the two overarching research questions. These target, on one hand, the successful mitigation of systematic QPE errors and on the other hand the robustness of the deep learning approaches. Each overarching question is related to a specific research question that was each answered in a individually designed study (see Fig. 1.12).

SQ2 Are 1D-convolutional neural networks able to distinguish between CML signal level patterns caused by rainfall and fluctuations during dry periods?

The detection of rain events in CML attenuation data by state-of-the-art methods is based on the assumption that the amount of fluctuation in the time series measured by the rolling standard deviation allows for the separation of wet and dry periods. The specific pattern of the fluctuation was not considered in previous studies. However, several phenomena that induce signal fluctuations during dry periods have been identified, but it is yet unclear if they can be distinguished from rain events based on their signal pattern. Due to the recent success of deep neural networks in pattern recognition tasks, it is now important to test their performance for rain event detection in CML attenuation data and investigate the impact that an improved rain event detection has on the resulting rainfall estimates.

SQ3 How do blackouts impact CML rainfall estimates and how often should they be expected in theory and practice?

Previous studies for the performance of CML-derived rainfall information excluded missing periods in the data assuming that they appear randomly due to disturbances in the data acquisition. However, it is known that CML hardware does not allow for the measurement of high attenuation values when the signal-to-noise ratio at the receiver is too low. The hardware, transmission power, and cut-off signal-to-noise ratio influence when these blackouts happen. Since those are largely influenced by network providers, it remains unclear how frequent blackouts occur in real-world CML data. For planning mobile networks, using high-resolution rainfall climatology is proposed by the ITU, but the suitability of such data has yet to be tested empirically. Blackouts should be considered a systematic error of CML rainfall estimates and an accurate detection of blackouts is important to develop approaches to reconstruct the missing extremes in the future.

SQ4 How robust is manual quality control of CML data and how does it affect rainfall estimation?

Understanding the cause of anomalies, that is, signal fluctuations not caused by rainfall, is useful to understand more about the potential applications of CML data. Additionally, compound events of rainfall and anomalies disturb accurate rainfall estimation even if the rain event is located correctly. However, the causes of anomalies are diverse and reference data is often scarce. There are unknown factors like the mechanical properties of the CML tower and the hardware that may or may not be exposed to influences like heat, solar radiation, or wind. Experts in the field of CML-based rainfall estimation usually develop a good understanding of the patterns in the time series caused by rainfall. Despite being able to tell which segments of the time series resemble rain events, the causes for anomalies remain less well understood. Cataloging CML anomalies using expert knowledge seems a valid solution

to overcome the scarcity of reference data, but it is unclear if there is ambiguity in the expert decision and if expert flagging of anomalies can improve CML rainfall estimates.

SQ5 Are 3D-convolutional neural networks capable of simultaneous temporal super-resolution, ground-adjustment, and advection-correction of radar rainfall estimates?

As summarized in Section 1.2.2.4, systematic errors in radar rainfall estimates when compared to on-ground sensors include a temporal undersampling of small, fast-moving rain cells and an advection-driven spatiotemporal mismatch. Additionally, biases can be caused by an inadequate choice of scenario-dependent Z-R power-law coefficients or strong attenuation and beam blockage. Typically, gauge adjustment is used to correct these errors which relies on a dense rain gauge network in the study region. On the one hand, statistical methods have been proposed to bias correct radar rainfall with respect to a rain gauge reference. On the other hand, neural network approaches have been used to learn a reflectivity-rain rate relationship and thus to improve the radar rainfall estimate. To date, a large-scale, post-processing approach is missing to correct the temporal undersampling and advection-driven errors.

1.5 Study Outline

The outline of this study follows the five main chapters composed of published journal articles that aim to answer the specific research questions (see Figures 1.11 and 1.12). The main chapters are accompanied by an introduction to the general topics of rainfall estimation and deep learning and a synthesis that connects the specific and overarching research questions in order to reflect on the main objective.
The basis for the conducted studies is comprised of country-wide CML, weather radar, and rain gauge observations in Germany (see Sect. 1.2.3). The CML dataset was collected at the Karlsruhe Institute of Technology from 2017 to 2023 and beyond (Chwala et al., 2016). The weather radar and rain gauge datasets are provided by the German meteorological service (Winterrath et al., 2017). The combined density of observations from these three types of sensors that are available for research purposes in Germany is not surpassed in a global comparison. Therefore, choosing Germany as a study region for a data-driven improvement of rainfall estimates appeared to be the optimal choice. **Chapter 1** introduces the reader to the topic of the thesis and motivates the research objective of improving rainfall estimates using deep learning algorithms. The motivation is followed by more general introductions to rainfall estimation techniques and deep learning. Based on this general introduction, overarching and specific research questions are posed using identified knowledge gaps. Afterward, the study outline and innovations of the thesis are stated, as well as further related articles with the contribution of the author. These articles partly overlap with the topic of the thesis but are not at its core.

In **Chapter 2** "Rainfall estimation from a German-wide commercial microwave link network: optimized processing and validation for 1 year of data", a large-scale evaluation of state-of-the-art CML signal processing routines is conducted using one year of data from countrywide CML observations in Germany. The study uses data from 3904 CMLs and the months of September 2017 to August 2018. The temporal resolution of the CML attenuation data is one minute and the processed rainfall estimates are aggregated to hourly, monthly, and seasonal data to be compared to a high-quality, gauge-adjusted weather radar product of the German Meteorological Service. The study aims to compare optimized versions of the rain event detection technique by Schleiss and Berne (2010) and the wet antenna attenuation compensation schemes by Leijnse et al. (2008) and Schleiss et al. (2013). The CML data is compared to the radar product once at the CML locations and once on an interpolated map-based level. The results are compared in a dedicated statistical evaluation framework that considers different suitable performance metrics and different thresholding approaches to exclude dry periods from the analysis. Chapter 2 is published as Graf et al. (2020a) in Hydrology and Earth System Science.

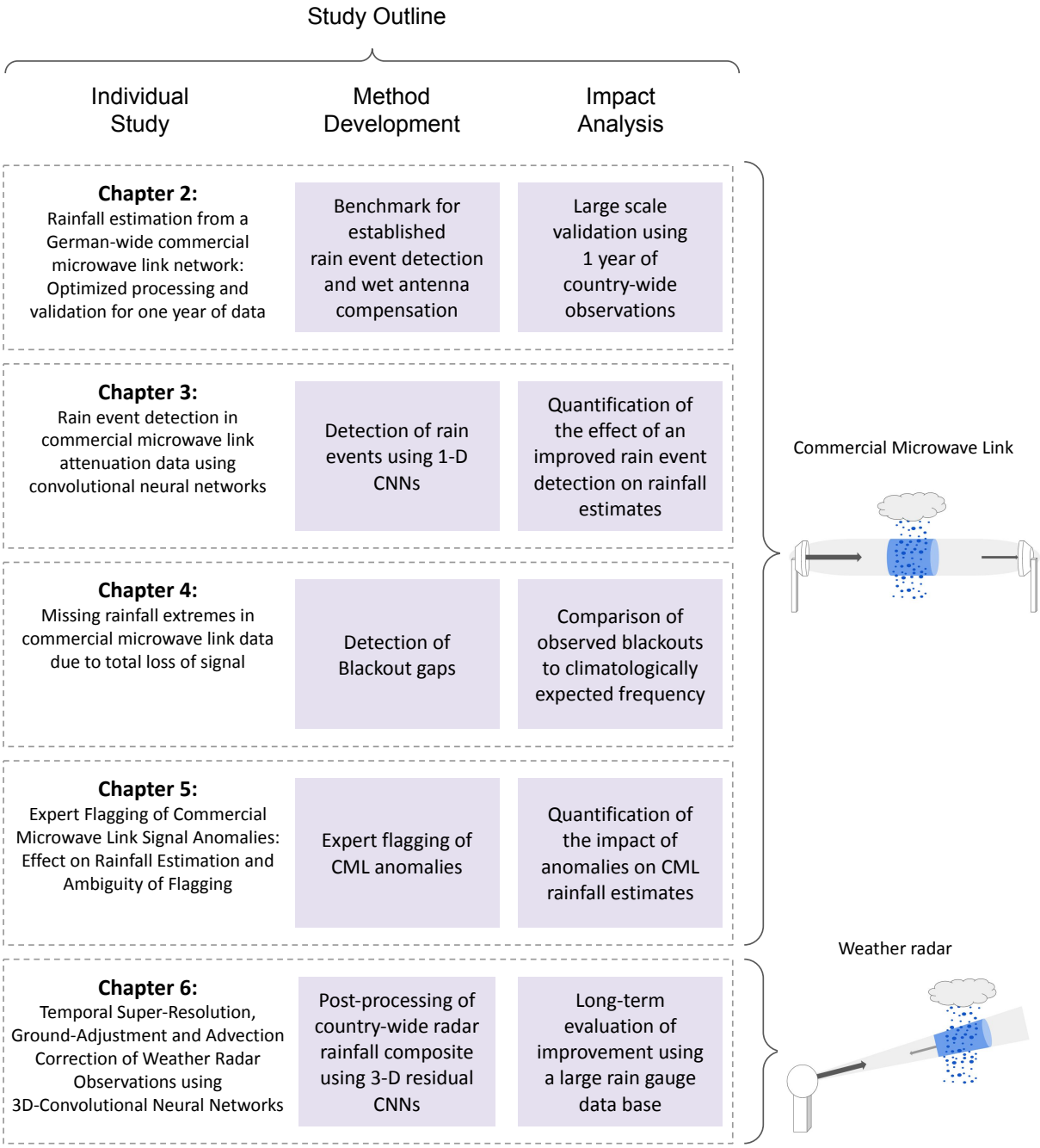

Figure 1.12: Study design: The five main chapters of the thesis describe the individual studies that are used to answer the research questions. Within each study, new methodologies to process and analyze the complex datasets have been developed and the impact of the studied phenomenon on rainfall estimates is analyzed. Chapters 2-5 focus on CML data and Chapter 6 focuses on weather radar QPE.

Chapter 3 "Rain event detection in commercial microwave link attenuation data using convolutional neural networks", is the first application of a deep learning technique to improve the processing step of rain event detection in CML signal levels. The study applies a convolutional neural network (CNN) to recognize rain events in the CML signal time series and compares this approach to the previous state-of-the-art method introduced by Schleiss and Berne (2010). The neural network is a one-dimensional version of the two-dimensional image recognition CNNs such as LeNet (LeCun et al., 2015). The network is trained on four months of data from 800 randomly selected CMLs and tested on two different months of data from more than 3000 CMLs not included in the training. This data-splitting routine is chosen to test for the transferability of the chosen method to future times and previously unseen locations. As a rain event reference for the training, hourly gauge-adjusted weather radar estimates are averaged along the link paths. The evaluation of the results is conducted using climatology invariant metrics such as the receiver operating characteristic. In addition to the rain event detection skill the improvement of the resulting rainfall estimates by the CNN in combination with the remaining optimized processing developed in Chapter 2 is analyzed. Chapter 3 is published as Polz et al. (2020) in Atmospheric Measurement Techniques.

Chapter 4 "Missing Rainfall Extremes in Commercial Microwave Link Data Due To Complete Loss of Signal", investigates the occurrence of blackouts (see Section 1.2.2.3), a systematic error in CML rainfall estimates that was neglected in previous studies. Blackouts occur when the rainfall-induced attenuation causes a complete loss of the CMLs signal. The study proposes a new algorithm to detect blackouts and uses it to analyze three years of CML data to gain knowledge about the frequency of blackouts and the probability of missing high-intensity rain rates. As a reference, a gauge-adjusted and climatology-corrected radar product with a 5-minute temporal resolution is used. Additionally, an attenuation climatology derived from 20 years of radar data is analyzed to compare the expected occurrence of blackouts to the observed frequency. Chapter 4 is published as Polz et al. (2023b) in Earth and Space Science.

Chapter 5 "Expert Flagging of Commercial Microwave Link Signal Anomalies: Effect on Rainfall Estimation and Ambiguity of Flagging", is concerned with CML signal anomalies and their detection and classification. Because of a lack of understanding of the potential causes of signal anomalies, the opportunistic nature of the sensing approach, and the lack of suitable reference data, cataloging the signal anomalies using expert knowledge is explored. To see if the manual data quality control is robust, four experts propose quality flags for data from different CMLs with a high amount of fluctuations. After a first round of flagging a set of anomaly classes is defined and discussed within the expert group. Afterward, each expert individually flags three months of data from 20 CMLs, and the results are compared. The study analyzes the degree of ambiguity between expert decisions and the impact of the detected anomalies on the quality of the rainfall estimates.

Chapter 5 is published as Polz et al. (2023a) in the proceedings of the 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW).

Chapter 6 "Temporal Super-Resolution, Ground-Adjustment and Advection Correction of Weather Radar Observations using 3D-Convolutional Neural Networks", describes the correction of systematic errors in radar rainfall estimates and their mitigation using a 3D-residual neural network approach. The focus of the study is to simultaneously adjust the radar QPE to on-ground information, correct for advection-related issues, and increase the temporal resolution. Because of the complex error structure of radar rainfall estimates a post-processing approach is chosen. The neural network is designed to predict the next five minutes of estimated rainfall from rain gauges using a time series of adjacent radar QPE. The target resolution is one minute. The model is trained using one year of data (2020) from 17 C-band radars and 581 rain gauges operated by the German meteorological service. The results are validated using eight years of independent data from 247 rain gauges with a 1-minute and 1138 rain gauges with a daily resolution. Chapter 6 is published as Polz et al. (2024) in Atmospheric Measurement Techniques.

Chapter 7 is the synthesis of all studies conducted in Chapters 2 to 6. First, the key findings of the studies are summarized and the specific research questions are answered. Then, the implications of the key findings for the overarching research questions are discussed and the overall goal is reviewed as a conclusion. Finally, an outlook on new knowledge gaps and potential research topics that emerge from this thesis is given.

1.6 Innovation

The main innovation of this thesis lies in the application of deep learning techniques for rainfall estimation with CMLs and weather radars specifically and the identification of specific challenges and opportunities of deep learning approaches for rainfall estimation more generally. Specifically, this thesis presents:

- The first-time application of convolutional neural networks for rain event detection in CML signal level time-series which is shown to outperform state-of-the-art methods
- A residual neural network approach for simultaneous temporal super-resolution, ground-adjustment, and advection correction of radar rainfall that significantly reduces the error of radar QPE at a 1-minute temporal resolution.

The main innovations are accompanied by a series of important achievements that significantly advance CML rainfall estimation:

- The first country-wide rainfall estimates from CMLs in Germany and a benchmark estimate of the relative error compared to gauge-adjusted weather radars
- The identification and quantification of missing rainfall extremes in the CML signal level time series as a serious systematic error
- A new approach to cataloging CML signal anomalies using expert knowledge and an analysis of their impact on CML rainfall estimates.

1.7 Author Contributions

Chapter 2: *Rainfall estimation from a German-wide commercial microwave link network: Optimized processing and validation for one year of data*
Graf, M., Chwala, C., **Polz, J.**, and Kunstmann, H.: Rainfall estimation from a German-wide commercial microwave link network: optimized processing and validation for 1 year of data, *Hydrol. Earth Syst. Sci.*, 24, 2931–2950, `https://doi.org/10.5194/hess-24-2931-2020`, 2020.
The main contribution of **Julius Polz** was the design and improvement of the statistical evaluation, especially the identification of the Matthews correlation coefficient as a suitable objective function for optimizing rain event detection. Julius Polz frequently discussed the study layout and the technical implementation of the analysis with the other co-authors and significantly contributed to the final outcome.
The detailed contributions of all authors are: The concept of this study was designed by Maximilian Graf, Christian Chwala, and Harald Kunstmann. Maximilian Graf and Christian Chwala designed the analysis and Maximilian Graf carried it out with contributions from Christian Chwala and **Julius Polz**. Christian Chwala set up the data acquisition to provide the raw CML data. Maximilian Graf processed and evaluated all CML data. The optimization of the individual processing steps was conducted by Maximilian Graf. The concept of the statistical analysis to set a benchmark performance for CML wet-dry classification was developed by **Julius Polz** and Maximilian Graf. The results were presented by Maximilian Graf and were discussed with all co-authors. The code both for the CML processing and the evaluation was developed by Maximilian Graf with contributions from Christian Chwala and **Julius Polz**. All figures were prepared by Maximilian Graf. The manuscript was written by Maximilian Graf with contributions from and discussions with all co-authors.

Chapter 3: *Rain event detection in commercial microwave link attenuation data using convolutional neural networks*
Polz, J., Chwala, C., Graf, M., and Kunstmann, H.: Rain event detection in commercial microwave link attenuation data using convolutional neural networks, *Atmos. Meas. Tech.*, 13, 3835–3853, `https://doi.org/10.5194/amt-13-3835-2020`, 2020
Julius Polz was leading the design of this study proposing the CNN-based methodology as well as the training and evaluation framework. He was the main contributor to the technical implementation of the methods and the lead author of the manuscript preparing all figures and writing the initial paper draft.
The detailed contributions of all authors are: **Julius Polz**, Christian Chwala, and Harald Kunstmann designed the study layout and **Julius Polz** carried it out with contributions from Christian Chwala and Maximilian Graf. Christian Chwala set up the data acquisition to provide the raw CML data. Maximilian Graf provided pre-processed CML data which was further processed by **Julius Polz**. **Julius Polz** developed the CNN-based

rain event detection model and conducted the training, validation, and model selection. **Julius Polz** evaluated the CNN-based wet-dry classification and rainfall estimates against a radar reference with contributions from Maximilian Graf. The code was developed by **Julius Polz** with the contributions of Christian Chwala and Maximilian Graf. All figures were prepared by **Julius Polz**. **Julius Polz** prepared the manuscript with contributions from all co-authors.

Chapter 4: *Missing rainfall extremes in commercial microwave link data due to total loss of signal*

Polz, J., Graf, M. and Chwala, C.: Missing rainfall extremes in commercial microwave link data due to total loss of signal, *Earth and Space Science*, vol. 10, no. 2, pp. e2022EA002456, `https://doi.org/10.1029/2022EA002456`, 2023)

The main contribution of **Julius Polz** to this study was the analysis of expected blackouts based on a large dataset of climatology-corrected and gauge-adjusted weather radar estimates while Maximilian Graf implemented the detection of blackouts in real CML data. In all other aspects, including the overall study design, he contributed equally to the other co-authors of the study.

The detailed contributions of all authors are: **Julius Polz**, Maximilian Graf, and Christian Chwala designed the study layout. **Julius Polz** and Maximilian Graf carried out the analysis in equal parts with contributions from Christian Chwala. The data was provided and prepared by all authors. In general, **Julius Polz** conducted all analyses related to the synthetic CML data based on RADKLIM-YW and Maximilian Graf conducted all analyses related to the observed CML data with respective contributions and discussion with Christian Chwala. Accordingly, Figures 1 and 2 were prepared by Maximilian Graf who provided CML-based statistics and developed a blackout gap detection algorithm while Figures 3 and 4 were prepared by **Julius Polz** who provided synthetic CML-based statistics and compared the blackout frequency in real and synthetic CML observation. Christian Chwala contributed and discussed various stages of the analysis. Accompanying example software of the blackout gap detection algorithm and example data were prepared and published by Maximilian Graf within the python software package *pycomlink*. **Julius Polz** and Maximilian Graf prepared the manuscript with the contributions of Christian Chwala. **Julius Polz** and Maximilian Graf share the authorship of this publication with equal contributions.

Chapter 5: *Expert Flagging of Commercial Microwave Link Signal Anomalies: Effect on Rainfall Estimation and Ambiguity of Flagging*

Polz, J., Glawion, L., Graf, M., Blettner, N., Lasota, E., Schmidt, L., Kunstmann, H. and Chwala, C., *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, pages 1–5, `https://doi.org/10.1109/icasspw59220.2023.10193654`, 2023

Julius Polz was the main contributor to the design of this study and implemented the

flagging tool. He also designed and conducted the evaluation of flagged data and prepared the manuscript.
The detailed contributions of all authors are: **Julius Polz**, Luca Glawion, Maximilian Graf, and Christian Chwala designed the study layout. **Julius Polz** and Luca Glawion developed the data flagging tool and discussed it with all co-authors. **Julius Polz**, Luca Glawion, Maximilian Graf, and Nico Blettner independently flagged the CML anomalies. **Julius Polz** developed and carried out the analysis of the data and evaluated the impact on the CML rainfall estimates. All figures were prepared by **Julius Polz**. **Julius Polz** prepared the manuscript with contributions from all co-authors. **Julius Polz** presented the conference publication at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

Chapter 6: *Temporal Super-Resolution, Ground-Adjustment and Advection Correction of Weather Radar Observations using 3D-Convolutional Neural Networks*
Polz, J., Glawion, L., Gebisso, H., Altenstrasser, L., Graf, M., Kunstmann, H., Vogl, S., and Chwala, C. Temporal Super-Resolution, Ground Adjustment, and Advection Correction of Radar Rainfall Using 3-D-Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 62:1–10, `https://doi.org/10.1109/TGRS.2024.3371577`, 2024
Julius Polz was the main contributor who designed the study and implemented the ResRadNet model with contributions from Luca Glawion. He also designed the evaluation of the trained model, prepared all figures, and prepared the first draft of the manuscript before discussing it with all co-authors.
The detailed contributions of all authors are: **Julius Polz**, Stephanie Vogl, and Christian Chwala designed the study layout. **Julius Polz** and Maximilian Graf pre-processed the radar and station data. **Julius Polz** developed the neural network architecture with contributions from Luca Glawion, Hiob Gebisso, Lukas Altenstrasser, and Christian Chwala. **Julius Polz** conducted the advection correction using the Lukas-Kanade Algorithm. **Julius Polz** designed the validation framework and evaluated the model performance. The code was developed by **Julius Polz** with contributions from all co-authors. **Julius Polz** prepared the manuscript with contributions from Maximilian Graf, Luca Glawion, Christian Chwala, Stephanie Vogl, and Harald Kunstmann.

1.8 Further related articles

In this section, co-authored articles related to the improvement of rainfall estimation and prediction with CMLs and weather radars that have been published or submitted for publication in the period of December 2018 to March 2023 are presented in chronological order.

Rainfall estimation using the mobile network (Regenmessung im Mobilfunknetz) Graf et al. (2021b)

Graf, M., **Polz, J.** and Chwala, C.

Physik in unserer Zeit

Abstract A most accurate estimation of precipitation is a fundamental part of climate and environmental research as well as for various applications, for example in agriculture. A new way of measuring precipitation is made possible by the mobile communications network in Germany using directional radio links (Commercial Microwave Link, CML). The great advantage of this method lies in the infrastructure, which can potentially be used worldwide. The quality of the precipitation observations from about 4000 CML in Germany shows a high agreement with the official measuring network of the German Weather Service.

Near-Realtime Quantitative Precipitation Estimation and Prediction (RealPEP) Trömel et al. (2021)

Trömel, S., Chwala, C., Furusho-Percot, C., Carbajal Henken, C., **Polz, J.**, Potthast, R., Reinoso-Rondinel, R. and Simmer, C.

Bulletin of the American Meteorological Society

Abstract On 5–7 October 2020, 250 participants discussed ideas and recent developments in the fields of quantitative precipitation estimation (QPE) based on the exploitation of measurements of polarimetric radars and microwave backhaul links, observation-based quantitative precipitation nowcasting (QPN), numerical quantitative precipitation forecasting (QPF), flash-flood prediction (FFP), and their organization into seamless prediction systems. This article describes the major findings, challenges and discussion points identified during the conference.

High-resolution rainfall maps from commercial microwave links for a data-scarce region in West Africa Djibo et al. (2023)

Djibo, M., Chwala, C., Graf, M., **Polz, J.**, Kunstmann, H., and Zougmoré, F.

Journal of Hydrometeorology

Abstract We present high-resolution rainfall maps from commercial microwave link (CML) data in the city of Ouagadougou, Burkina Faso. Rainfall was quantified based on data from

100 CMLs along unique paths and interpolated to achieve rainfall maps with a 5-minute temporal and 0.55km spatial resolution for the monsoon season of 2020. Established processing methods were combined with newly developed filtering methods, minimizing the loss of data availability. The rainfall maps were analyzed qualitatively both at a five-minute and aggregated daily scale. We observed high spatio-temporal variability on the five-minute scale which cannot be captured with any existing measurement infrastructure in West Africa. For the quantitative evaluation only one rain gauge with a daily resolution was available. Comparing the gauge data with the corresponding CML rainfall map pixel showed a high agreement with a Pearson correlation coefficient of over 0.95 and an underestimation of the CML rainfall maps of around ten percent. Because the CMLs closest to the gauge have the largest influence on the map pixel at the gauge location, we thinned out the CML network around the rain gauge synthetically in several steps and repeated the interpolation. The performance of these rainfall maps dropped only when a radius of 5 km was reached and around half of all CMLs were removed. We further compared ERA5 and GPM-IMERG data to the rain gauge and found that they show much lower correlation than data from the CML rainfall maps. This clearly highlights the large benefit that CML data can provide in the data scarce but densely populated African cities.

spateGAN: Spatio-Temporal Downscaling of Rainfall Fields using a cGAN Approach Glawion et al. (2023)

Glawion, L., **Polz, J.**, Kunstmann, H., Fersch, B. and Chwala, C.

Earth and Space Science

Abstract Climate models face limitations in their ability to accurately represent highly variable atmospheric phenomena. To resolve fine-scale physical processes, allowing for local impact assessments, downscaling techniques are essential. We propose spateGAN, a novel approach for spatio-temporal downscaling of precipitation data using conditional generative adversarial networks. Our method is based on a video super-resolution approach and trained on 10 years of country-wide radar observations for Germany. It simultaneously increases the spatial and temporal resolution of coarsened precipitation observations from 32 km to 2 km and from 1 hr to 10 min. Our experiments indicate that the ensembles of generated temporally consistent rainfall fields are in high agreement with the observational data. Spatial structures with plausible advection were accurately generated. Compared to trilinear interpolation and a classical convolutional neural network, the generative model reconstructs the resolution-dependent extreme value distribution with high skill. It showed a high fractions skill score of 0.6 (spatio-temporal scale: 32 km and 1 hr) for rainfall intensities over $15\,mm\,h^{-1}$ and a low relative bias of 3.55%. A power spectrum analysis confirmed that the probabilistic downscaling ability of our model further increased its skill. We observed that neural network predictions may be interspersed by recurrent structures not related to rainfall climatology, which should be a known issue for future studies. We were able to mitigate them by using an appropriate model architecture and model

selection process. Our findings suggest that spateGAN offers the potential to complement and further advance the development of climate model downscaling techniques, due to its performance and computational efficiency.

Potential and limitations of filling gaps in commercial microwave link data stemming from complete loss of signal during heavy rainfall Graf et al. (2023)

Graf, M., Blettner, N., **Polz, J.**, and Chwala, C.

IEEE International Conference on Acoustics, Speech, and Signal Processing ICASSP

Abstract Rainfall estimates from commercial microwave links (CML) can be impaired by the total loss of signal during heavy rainfall events. This implies that the highest rainfall intensities may not be observed by CMLs during these so-called blackouts. As CML rainfall estimation approaches an operational state, this issue has to be studied and potential mitigation strategies have to be developed. Therefore, we investigate three methods which fill blackout gaps detected by a filter established in previous work. The methods we developed consisted of a linear and a cubic interpolation as well as the infilling of the values at the detection limit of each CML. In general, the linear interpolation underestimated and the cubic interpolation overestimated rainfall intensities compared to a reference. The physically motivated infilling at the detection limit performed best. In conclusion, blackout gaps should be mitigated with one propose methods to improve the quality of rainfall estimates derived from CMLs.

Improved rain event detection in Commercial Microwave Link time series via combination with MSG SEVIRI data Graf et al. (2024)

Graf, M.*, Wagner, A.*, **Polz, J.***, Lliso, L., Lahuerta, J., Kunstmann, H., and Chwala, C.
(*equal contributions)

Atmospheric Measurement Techniques

Abstract The most reliable areal precipitation estimation is usually generated via combinations of different measurements. Path-averaged rainfall rates can be derived from commercial microwave links (CMLs), where attenuation of the emitted radiation is strongly related to rainfall rate. CMLs can be combined with data from other rainfall measurements or can be used individually. They are available almost worldwide and often represent the only opportunity for ground-based measurement in data-scarce regions. However, deriving rainfall estimates from CML data requires extensive data processing. The separation of the attenuation time series into rainy and dry periods (rain event detection) is the most important step in this processing and has a high impact on the resulting rainfall estimates. In this study, we investigate the suitability of Meteosat Second Generation Spinning Enhanced Visible and InfraRed Imager (MSG SEVIRI) satellite data as an auxiliary-data-based (ADB) rain event detection method. We compare this method with two time-series-based (TSB) rain event detection methods. We used data from 3748 CMLs

in Germany for 4 months in the summer of 2021 and data from the two SEVIRI-derived products PC and PC-Ph. We analyzed all rain event detection methods for different rainfall intensities, differences between day and night, and their influence on the performance of rainfall estimates from individual CMLs. The radar product RADKLIM-YW was used for validation. The results showed that both SEVIRI products are promising candidates for ADB rainfall detection, yielding only slightly worse results than the TSB methods, with the main advantage that the ADB method does not rely on extensive validation for different CML datasets. The main uncertainty of all methods was found for light rain. Slightly better results were obtained during the day than at night due to the reduced availability of SEVIRI channels at night. In general, the ADB methods led to improvements for CMLs performing comparatively weakly using TSB methods. Based on these results, combinations of ADB and TSB methods were developed by emphasizing their specific advantages. Compared to basic and advanced TSB methods, these combinations improved the Matthews correlation coefficient of the rain event detection from 0.49 (or 0.51) to 0.59 during the day and from 0.41 (or 0.50) to 0.55 during the night. Additionally, these combinations increased the number of true-positive classifications, especially for light rainfall compared to the TSB methods, and reduced the number of false negatives while only leading to a slight increase in false-positive classifications. Our results show that utilizing MSG SEVIRI data in CML data processing significantly increases the quality of the rain event detection step, in particular for CMLs which are challenging to process with TSB methods. While the improvement is useful even for applications in Germany, we see the main potential of using ADB methods in data-scarce regions like West Africa where extensive validation is not possible.

Chapter 2

Rainfall estimation from a German-wide commercial microwave link network: Optimized processing and validation for one year of data (Graf et al., 2020a)

Graf, M., Chwala, C., Polz, J., and Kunstmann, H.: Rainfall estimation from a German-wide commercial microwave link network: optimized processing and validation for 1 year of data, *Hydrol. Earth Syst. Sci.*, 24, 2931–2950, https://doi.org/10.5194/hess-24-2931-2020, 2020.

Abstract

Rainfall is one of the most important environmental variables. However, it is a challenge to measure it accurately over space and time. During the last decade commercial microwave links (CMLs), operated by mobile network providers, have proven to be an additional source of rainfall information to complement traditional rainfall measurements. In this study we present the processing and evaluation of a German-wide data set of CMLs. This data set was acquired from around 4000 CMLs distributed across Germany with a temporal resolution of one minute. The analyzed period of one year spans from September 2017 to August 2018. We compare and adjust existing processing schemes on this large CML data set. For the crucial step of detecting rain events in the raw attenuation time series, we are able to reduce the amount of miss-classification. This was achieved by a new approach to determine the threshold, which separates a rolling window standard deviation of the CMLs signal into wet and dry periods. For the compensation of wet antenna attenuation, we compare a time-dependent model with a rain-rate-dependent model and show that the rain-rate-dependent model performs better for our data set. As precipitation reference, we use RADOLAN-RW, a gridded gauge-adjusted hourly radar product of the German Meteorological Service (DWD), from which we derive the path-

averaged rain rates along each CML path. Our data processing is able to handle CML data across different landscapes and seasons very well. For hourly, monthly and seasonal rainfall sums we found a good agreement between CML-derived rainfall and the reference, except for the winter season with non-liquid precipitation. We discuss performance measures for different subset criteria and show, that CML derived rainfall maps are comparable to the reference. This analysis shows that opportunistic sensing with CMLs yields rainfall information with a good agreement to gauge-adjusted radar data during periods without non-liquid precipitation.

2.1 Introduction

Measuring precipitation accurately over space and time is challenging due to its high spatiotemporal variability. It is a crucial component of the water cycle and knowledge of the spatiotemporal distribution of precipitation is an important quantity in many applications across meteorology, hydrology, agriculture, and climate research.
Typically, precipitation is measured by rain gauges, ground-based weather radars or spaceborne microwave sensors. Rain gauges measure precipitation at the point scale. Errors can be caused for example by wind, solid precipitation or evaporation losses (Sevruk, 2006). The main disadvantage of rain gauges is their lack of spatial representativeness.
Weather radars overcome this spatial constraint, but are affected by other error sources. They do not directly measure rainfall, but estimate it from related observed quantities, typically via the Z-R relation, which links the radar reflectivity "Z" to the rain rate "R". This relation, however, depends on the rain drop size distribution (DSD), resulting in significant uncertainties. Dual-polarization weather radars reduce these uncertainties, but still struggle with the DSD-dependence of the rain rate estimation (Berne and Krajewski, 2013). Additional error sources can stem from the measurement high above ground, from beam blockage or ground clutter effects.
Satellites can observe large parts of the earth, but their spatial and temporal coverage also has limits. Geostationary satellites can provide a high temporal sampling rate of a specific part of the earth. However rain rate estimates show large uncertainties because they have to be derived from measurements of visible and infra red channels, which were not meant for this purpose. Satellites in Low Earth orbits typically use dedicated sensors for rainfall estimation (microwave radiometers and radars), but their revisiting times are constraint by their orbits. Typical revisit times are in the order of hours to days. As a result, even merged multi-satellite products have a latency of several hours, e.g. the Integrated Multi-satellite Retrievals (IMERG) early run of the Global Precipitation Measurement Mission (GPM) has a latency of 6 hours, while it is limited to a spatial resolution of 0.1 degrees. The employed retrieval algorithms are highly sophisticated and several calibration and correction stages are potential error sources (Maggioni et al., 2016).
Additional rainfall information, for example derived from commercial microwave links (CMLs) maintained by cellular network providers, can be used to compare and comple-

ment existing rainfall data sets (Messer et al., 2006). In regions with sparse observation networks, they might even provide unique rainfall information.

The idea to derive rainfall estimates via the opportunistic usage of attenuation data from CML networks emerged over a decade ago independently in Israel (Messer et al., 2006) and the Netherlands (Leijnse et al., 2007). The main research foci in the first decade of dedicated CML research were the development of processing schemes for the rainfall retrieval and the reconstruction of rainfall fields. The first challenge for rainfall estimation from CML data is to distinguish between fluctuations of the raw attenuation data during rainy and dry periods. This was addressed by different approaches which either compared neighbouring CMLs using the spatial correlation of rainfall (Overeem et al., 2016a) or which focused on analyzing the time series of individual CMLs (Chwala et al., 2012; Polz et al., 2020; Schleiss and Berne, 2010; Wang et al., 2012). Another challenge is to estimate and correct the effect of wet antenna attenuation. This effect stems from the attenuation caused by water droplets on the covers of CML antennas, which leads to rainfall overestimation (Fencl et al., 2019; Leijnse et al., 2008; Schleiss et al., 2013).

Since many hydrological applications require spatial rainfall information, several approaches have been developed for the generation of rainfall maps from the path-integrated CML measurements. Kriging was successfully applied to produce countrywide rainfall maps for the Netherlands (Overeem et al., 2016b), representing CML rainfall estimates as synthetic point observation at the center of each CML path. More sophisticated methods can account for the path-integrated nature of the CML observations, using an iterative inverse distance weighting approach (Goldshtein et al., 2009), stochastic reconstruction (Haese et al., 2017) or tomographic algorithms (D'Amico et al., 2016; Zinevich et al., 2010).

CML-derived rainfall products were also used to derive combined rainfall products from various sources (Fencl et al., 2017; Liberman et al., 2014; Trömel et al., 2014). In parallel, first hydrological applications were tested. CML-derived rainfall was used as model input for hydrologic modelling studies for urban drainage modeling with synthetic (Fencl et al., 2013) and real world data (Stransky et al., 2018) or on run-off modeling in natural catchments (Brauer et al., 2016; Smiatek et al., 2017).

With the exception of the research carried out in the Netherlands, where more than two years of data from a country-wide CML network were analyzed (Overeem et al., 2016b), CML processing methods have only been tested on small data sets. We advance the state of the art by performing an analysis of rainfall estimates derived from a German-wide network of close to 4000 CMLs. In this study one CML is counted as the link along one path with typically two sub-links, for the communication in both directions. The temporal resolution of the data set is one minute and the analyzed period is one year from September 2017 until August 2018. The network covers various landscapes from the North German Plain to the Alps in the south, which feature individual precipitation regimes.

The objectives of this study are (1) to compare and adjust selected existing CML data processing schemes for the classification of wet and dry periods and for the compensation of wet antenna attenuation and (2) to validate the derived rain rates with an established

rainfall product, namely RADOLAN-RW, both on the country-wide scale of Germany.

2.2 Data

2.2.1 Reference data set

The *Radar-Online-Aneichung* data set (RADOLAN-RW) of the German Weather Service (DWD) is a radar-based and gauge adjusted precipitation data set. We use data from the archived real-time product RADOLAN-RW as reference data set throughout this work (DWD). It is a compiled radar composite from 17 dual-polarization weather radars operated by DWD and adjusted by more than 1000 rain gauges in Germany and 200 rain gauges from surrounding countries. RADOLAN-RW does not use dual-pol information, though. It is based on the reflectivity observations in horizontal polarization from each radar site, which are available in real-time every five minutes. This data is then used to compile a national composite of reflectivities, from which rain rates are derived. For the hourly rainfall information of the RADOLAN-RW product, the national composite of 5-minute radar rain rates is then aggregated and adjusted with the hourly rain gauge observations. A weighted mixture of additive and multiplicative corrections is applied. The rain gauges used for the adjustment have a spatial density of approximately one gauge per 300 km^2.
The gridded data set RADOLAN-RW has a spatial resolution of 1 km, covering Germany with 900 by 900 grid cells. The temporal resolution is one hour and the rainfall values are given with a quantization of 0.1 mm. RADOLAN-RW is available with a lag time of around 15 minutes. Detailed information on the RADOLAN processing and products is availabel from DWD (Bartels et al., 2004; Winterrath et al., 2012).
Kneis and Heistermann (2009) and Meissner et al. (2012) compared RADOLAN-RW products to gauge-based data sets for small catchments and found differences in daily, area averaged precipitation sums of up to 50 percent, especially for the winter season. Nevertheless, no data set with comparable temporal and spatial resolution, as well as extensive quality control is available.
In order to compare the path integrated rainfall estimates from CMLs and the gridded RADOLAN-RW product, RADOLAN-RW rain rates are resampled along the individual CML paths. For each CML, the weighted average of all intersecting RADOLAN-RW grid cells is calculated, with the weights being the lengths of the intersecting CML path in each cell. As result, one time series of the hourly rain rate is generated from RADOLAN-RW for each CML. The temporal availability of this reference is 100 percent but we excluded the CML and RADOLAN-RW pairs in the evaluation, when CML data is not available. We chose the RADOLAN-RW product, because it provides both a high temporal and spatial resolution for entire Germany. This resolution is the basis for a good evaluation of the path-averaged rain rates derived from CMLs. The rain gauge adjustments, while not perfect, assures that the RADOLAN-RW rainfall estimates have an increased accuracy compared to a radar-only data set.

2.2.2 Commercial microwave link data

We present data of 3904 CMLs operated by Ericsson in Germany. Their distribution over Germany is shown in Fig. 2.1. The CMLs are distributed country-wide over all landscapes in Germany, ranging from the North German Plain to the Alps in the south. The uneven distribution, with large gaps in the north east can be explained by the fact that we only access one subset of all installed CMLs, the Ericsson MINI-LINK Traffic Node systems operated for one cell phone provider.

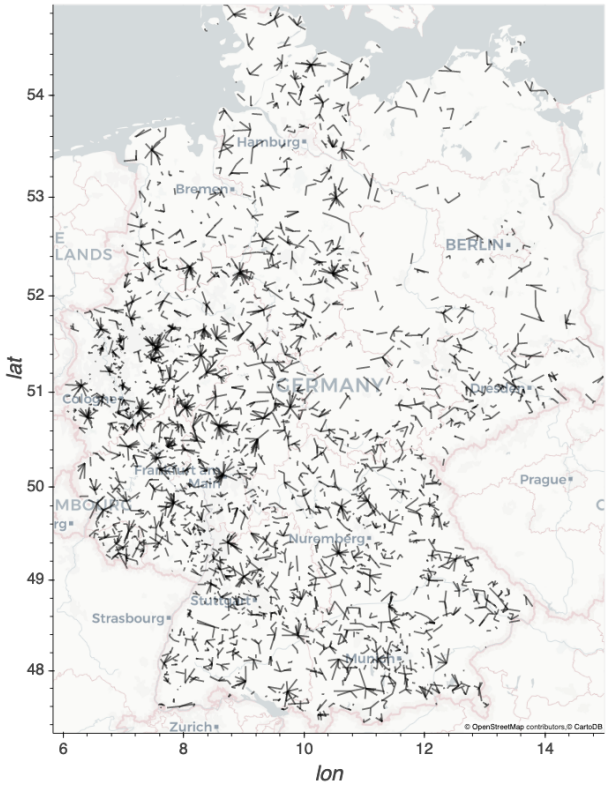

Figure 2.1: Map of the distribution of 3904 CMLs over Germany. © OpenStreetMap contributors 2019. Distributed under a Creative Commons BY-SA License.

CML data is retrieved with a real-time data acquisition system which we operate in cooperation with Ericsson (Chwala et al., 2016). Every minute, the current transmitted signal level (TSL) and received signal level (RSL) are requested from more than 4000 CMLs for both ends of each CML. The data is then immediately sent to and stored at our server. For the complete processing chain presented in this work, we used this 1-minute instantaneous data of TSL and RSL for the period from September 2017 to August 2018 for 3904 CMLs to derive rain rates with a temporal resolution of 1 minute. For comparison with the reference data, the 1-minute data is then aggregated. Due to missing, unclear or corrupted metadata we could not use all CML data. Furthermore, we only used data of one sub-link per CML. There was no specific criterion for selecting the sub-link. We simply used the pair of TSL and RSL that came first in our listing.
The available power resolution is 1 dB for TSL and 0.3 (with occasional jumps of 0.4

dB) for RSL. The TSL is constant for 25 percent of the CMLs. An Automatic Transmit Power Control (ATPC), which is able to increase TSL by several dB to prevent blackouts due to heavy attenuation, is active at 75 percent of the CMLs. While the length of the CMLs ranges between a few hundred meters to almost 30 km, most CMLs have a length of 5 to 10 km. They are operated with frequencies ranging from 10 to 40 GHz, depending on their length. Figure 2.2 shows the distributions of path lengths and frequencies. For shorter CMLs higher frequencies are used.

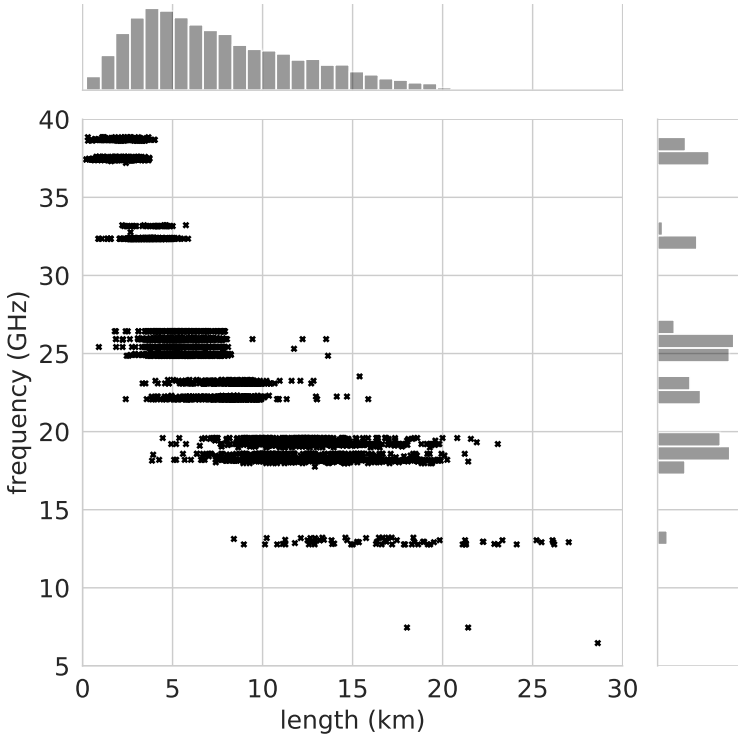

Figure 2.2: Scatterplot of the length against the microwave frequency of 3904 CMLs including the distribution of length and frequency.

To derive rainfall from CMLs, we used the difference between TSL and RSL, the transmitted minus received signal level (TRSL). An example of a TRSL time series is shown in Fig. 2.3a). To compare the rain rate derived from CMLs with the reference rain rate, we resampled it from a minutely to an hourly resolution after the processing.
In our CML data set 2.2 percent are missing time steps due to outages of the data acquisition systems. Additionally 1.2 percent of the raw data show missing values (Nan) and 0.1 percent show default fill values (e.g. -99.9 or 255.0) of the CML hardware, which we excluded from the analysis. In order to increase the data availability, we linearly interpolated gaps in raw TRSL time series which were up to five minutes long. This increased the data availability by 0.5 percent. On the one hand, these gaps can be the result of missing time steps and missing values but we also found cases where we suspect very high rainfall to be the reason for short blackouts of a CML.
The size of the complete CML data set is approximately 100 GB in memory. The data set

Table 2.1: Adopted confusion matrix

		reference	
		wet	*dry*
CML	*wet*	true wet (TP)	false wet (FP)
	dry	missed wet (FN)	true dry (TN)

is continuously extended by the operational data acquisition, allowing also the possibility of near-realtime rainfall estimation.

2.3 Methods

2.3.1 Performance measures

To evaluate the performance of the CML-derived rain rates against the reference data set, we used several measures which we calculated on an hourly basis. We defined a confusion matrix according to Tab. 2.1 where *wet* and *dry* refer to hours with and without rain, respectively. The Matthew's correlation coefficient (MCC) summarizes the four values of the confusion matrix in a single measure (2.1) and is typically used as measure of binary classification in machine learning. This measure is accounting for the skewed ratio of wet and dry events. It is high only if the classifier is performing well on both classes.

$$\mathrm{MCC} = \frac{\mathrm{TP} * \mathrm{TN} - \mathrm{FP} * \mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FP})(\mathrm{TP}+\mathrm{FN})(\mathrm{TN}+\mathrm{FP})(\mathrm{TN}+\mathrm{FN})}} \tag{2.1}$$

The mean detection error (MDE) (2.2) is introduced as a further binary measure focusing on the miss-classification of rain events.

$$\mathrm{MDE} = \frac{\frac{\mathrm{FN}}{\mathrm{n(wet)}} + \frac{\mathrm{FP}}{\mathrm{n(dry)}}}{2} \tag{2.2}$$

It is calculated as the average of missed wet and false wet rates of the contingency table from Tab. 2.1.

The linear correlation between CML-derived rainfall and the reference is expressed by the Pearson correlation coefficient (PCC). The coefficient of variation (CV) in (2.3) gives the distribution of CML rainfall around the reference expressed by the ratio of residual standard deviation and mean reference rainfall,

$$\mathrm{CV} = \frac{\mathrm{std}\sum(\mathrm{R_{CML}} - \mathrm{R_{reference}})}{\overline{\mathrm{R_{reference}}}} \tag{2.3}$$

where R_{CML} and $R_{reference}$ are hourly rain rates of the respective data set. Furthermore, we computed the mean absolute error (MAE) and the root mean squared error (RMSE) to measure the accuracy of the CML rainfall estimates. The relative bias is given as

$$\mathrm{bias} = \frac{\overline{(\mathrm{R_{CML}} - \mathrm{R_{reference}})}}{\overline{\mathrm{R_{reference}}}} \tag{2.4}$$

Often, in studies comparing CML derived rainfall and radar data, a threshold is used as a lower boundary for rainfall. The performance measures, summarized in Tab. 2.2, were calculated with different subset criteria or thresholds. This gives insight on how CML derived rainfall compares to the reference for different rain rates and on how the large number of data points without rain influence the performance measures. Another reason for listing the performance measures with several thresholds is the increased comparability with other studies on CML rainfall estimation, which do not uniformly use the same threshold, see e.g. Table A1 in de Vos et al. (2019). Therefore, we defined a selection of subset criteria and thresholds and show performance measures for data without any thresholds (*none*), for the data set with R_{CML} and $R_{reference} < 0.1$ mm/h set to 0 mm/h, for two thresholds where at least R_{CML} or $R_{reference}$ must be > 0 and $>= 0.1$ mm/h and two thresholds where $R_{reference}$ must be $>= 0.1$ and $>= 1$ mm.

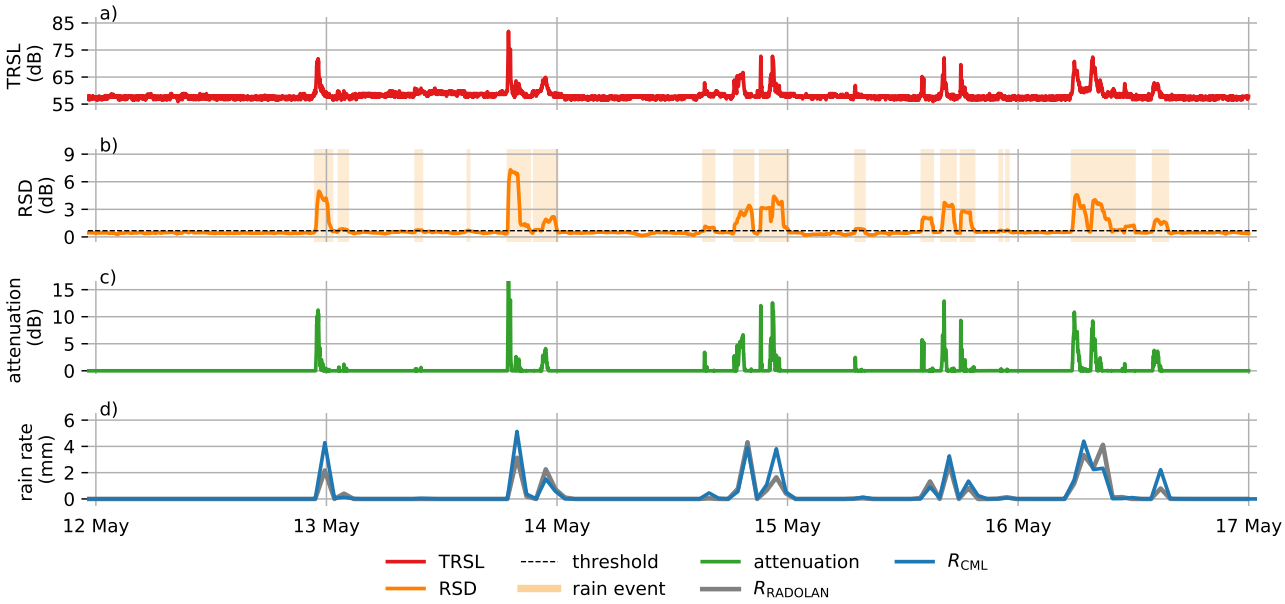

Figure 2.3: Processing steps from the TRSL to rain rate. a) The TRSL is the difference of TSL - RSL, the raw transmitted and received signal level of a CML. b) The RSD (rolling standard deviation) of the TRSL with an exemplary threshold shows the resulting wet and dry periods. c) The Attenuation is the difference between the baseline and the TRSL during wet periods. d) The derived rain rate is resampled to an hourly scale in order to compare it to the reference RADOLAN-RW.

2.3.2 From raw signal to rain rate

As CMLs are an opportunistic sensing system rather than part of a dedicated measurement system, data processing has to be done with care. Most of the CML research groups developed their own methods tailored to their needs and data sets. Overviews of these

methods are summarized by Chwala and Kunstmann (2019), Messer and Sendik (2015) and Uijlenhoet et al. (2018). The size of our data set is a challenge itself. As TRSL can be attenuated by rain or other sources, described in Sect. 2.3.2.1 and only raw TSL and RSL data is provided, the large size of the data set is of advantage but also a challenge. Developing and evaluating methods was significantly sped up by the use of an automated processing workflow, which we implemented as a parallelized workflow on a HPC system using the Python packages *xarray* and *dask* for data processing and visual exploration. The major challenges which arose from the processing of raw TRSL data into rain rates and the selected methods from literature are described in the following sections. We use parameters in this processing which are either based on literature, modified from the literature or which we developed in this study. An overview of all used parameters is given in Appendix 2.4.

2.3.2.1 Erratic behavior

Rainfall is not the only source of attenuation of microwave radio along a CML path. Additional attenuation can be caused by atmospheric constituents like water vapor or oxygen, but also by refraction, reflection or multi-path propagation of the beam (Upton et al., 2005). In particular, refraction, reflection and multi-path propagation can lead to strong attenuation in the same magnitude as from rain. CMLs that exhibit such behavior have to be omitted due to their noisiness.

We excluded erratic CML data which was extremely noisy or which showed drifts and jumps from our analysis on a monthly basis. To deal with this erratic data, we applied the following sanity checks: We exclude individual CMLs if 1) the five hour moving window standard deviation exceeds the threshold 2.0 for more then ten percent of a month, which typically is the case for CMLs with either a strong diurnal cycle or very noisy periods during a month, or if 2) a one hour moving window standard deviation exceeds the threshold 0.8 more than 33 percent of the time in a month. This filter is based on the approach for detecting rain events in TRSL time series from Schleiss and Berne (2010), which we also use later on in our processing. For the filter, a fairly high threshold was used, which should only be exceeded for fluctuations stemming from real rain events. The reasoning of our filter is, that if the threshold is exceeded too often, here 33 percent of the time per month, the CML data shows an unreasonably high amount of strong fluctuation. In total, the two sanity checks removed 1.1 percent from our CML data set. Together with the missing values that remain after interpolating data gaps of maximum five minutes in the TRSL time series, 4.2 percent of our data set are not available or not used for processing.

Jumps in data are mainly caused by single default values in the TSL which are described in Sect. 2.2.2. When we removed these default values, we are able to remove the jumps. TRSL can drift and fluctuate on daily and yearly scale (Chwala and Kunstmann, 2019). We could neglect the influence of these drifts in our analysis, because we dynamically derived a baseline for each rain event, as explained in Sect. 2.3.2.2. We also excluded CMLs having a constant TRSL over a whole month.

2.3.2.2 Rain event detection and baseline estimation

The TRSL during dry periods can fluctuate over time due to ambient conditions as mentioned in the previous section. Rainfall produces additional attenuation on top of the dry fluctuation. In order to calculate the attenuation from rainfall, a baseline level of TRSL during each rain event has to be determined. We derived the baseline from the precedent dry period. During the rain event, this baseline was held constant, as no additional information on the evolution of the baseline level is available. The crucial step for deriving the baseline is to separate the TRSL time series into wet and dry periods, because only then the correct reference level before a rain event is used. By subtracting the baseline from TRSL, we derived the attenuation caused by rainfall which is shown in Fig. 2.3c).
The separation of wet and dry periods is essential, because the errors made in this step will impact the performance of rainfall estimation. Missing rain events will result in rainfall underestimation. False detection of rain events will lead to overestimation. The task of detecting rain events in the TRSL time series is simple for strong rain events, but challenging when the attenuation from rain is approaching the same order of magnitude as the fluctuation of TRSL data during dry conditions.
There are two essential concepts to detect rain events. One compares the TRSL of a certain CML to neighbouring CMLs (Overeem et al., 2016a) and the other investigates the time series of each CML separately (Chwala et al., 2012; Schleiss and Berne, 2010; Wang et al., 2012). We choose the latter one and used a rolling standard deviation (RSD) with a centered moving window of 60 minutes length as a measure for the fluctuation of TRSL as proposed by Schleiss and Berne (2010).
It is assumed that RSD is high during wet periods and low during dry periods. Therefore, an adequate threshold can be defined, which differentiates the RSD time series in wet and dry periods. An example of an RSD time series and a threshold is shown in Fig. 2.3b) where all data points with RSD values above the threshold are considered as wet.
Schleiss and Berne (2010) proposed the use of a RSD threshold derived from rainfall climatology e.g. from nearby rain gauges. For our data set we assumed that it is raining 5 percent of all minutes in Germany, as proposed by Schleiss and Berne (2010) for their CMLs in France. Therefore, we used the 95 percent quantile of RSD as a threshold, assuming that the 5 percent of highest fluctuation of the TRSL time series refer to the 5 percent of rainy periods. We refer to this threshold as the climatologic threshold. We compared it to two new definitions of thresholds. We are aware that this threshold does not reflect the real climatology at each CML location, nevertheless this method is a rather robust and a simple approach which provides a first rain event detection.
For the first new definition, we derived the optimal threshold for each CML based on our reference data for the month of May 2018. We used the same approach as for the climatologic threshold, but for each CML we tested a range of possible thresholds and calculated the binary measure MCC for each. For each CML we picked the threshold which produced the highest MCC in May 2018 and used it over the whole analysis period.

The second new definition to derive a threshold is based on the quantiles of the RSD, similarly to the climatologic threshold describe above. However, we propose to not focus on the fraction of rainy periods for finding the optimal threshold, since a rainfall climatology is likely not valid for individual years and not easily transferable to different locations. We took the 80th quantile of the RSD of each CML, which can be interpreted as a measure of the strength of the TRSL fluctuation during dry periods, and multiplied it by a constant factor to derive the individual threshold. The 80th quantile can be assumed to be more robust against missclassification than the climatologic threshold, because this quantile represents the general notion of each TRSL time series to fluctuate, rather than the percentage of time in which it is raining. We chose the 80th quantile, since it is very unlikely that it is raining more than 20 percent of the time in a month in Germany.
To find the right factor, we selected the month of May 2018 and fitted a linear regression between the optimal threshold for each CML and the 80th quantile. The optimal threshold was derived beforehand with a MCC optimization from the reference. We then used this factor for all other months in our analysis. Additional, we found it to be similar for all months of the analyzed period.

2.3.2.3 Wet antenna attenuation

Wet antenna attenuation is the attenuation caused by water on the cover of a CML antenna. With this additional attenuation, the derived rain rate overestimates the true rain rate (Schleiss et al., 2013; Zinevich et al., 2010). The estimation of WAA is complex, as it is influenced by partially unknown factors, e.g. the material of the antenna cover. van Leth et al. (2018) found differences in WAA magnitude and temporal dynamics due to different sizes and shapes of the water droplets on hydrophobic and normal antenna cover materials. Another unknown factor for the determination of WAA is the information whether both, one or none of the antennas of a CML is wetted during a rain event. We selected and compared two parametric WAA correction schemes which do not rely on the use of auxiliary data like near-by rain gauges.
Schleiss et al. (2013) measured the magnitude and dynamics of WAA with one CML in Switzerland and derived a time-dependent WAA model. In this model, WAA increases at the beginning of a rain event to a defined maximum in a defined amount of time. From the end of the rain event on, WAA decreases again, as the wetted antenna is drying off. We ran this scheme with the proposed 2.3 dB of maximal WAA for both antennas together. This is also similar to the WAA correction value of 2.15 dB, which Overeem et al. (2016b) derived over a 12-day period in their data set. For τ, which determines the increase rate with time we chose 15 minutes. The decrease of WAA after a rain event is not explicitly modelled, because this WAA scheme is only applied for time steps, which are considered wet from the rain event detection, which has to be carried out in a previous step.
Leijnse et al. (2008) proposed a physical approach where the WAA depends on the microwave frequency, the antenna cover properties (thickness and refractive index) and the rain rate. A homogeneous water film is assumed on the antenna, with its thickness having

a power law dependence on the rain rate. Higher rain rates cause a thicker water film and hence higher WAA. A factor γ scales the thickness of the water film on the cover and a factor δ determines the non-linearity of the relation between rain rate and water film thickness. We adjusted the thickness of the antenna cover to 4.1 mm which we measured from one antenna provided by Ericsson. We are aware of the fact, that antenna covers have different thicknesses. But since we do not have this information for the actual antennas that are used by the CMLs of our data, we use this value, as the best one available. We further adjusted γ to 1.47E-5 and δ to 0.36 in such a way, that the increase of WAA with rain rates is less steep for small rain rates compared to the originally proposed parameters. The original set of parameters suppressed small rain events too much because the WAA compensation attributed all attenuation in the TRSL to WAA. For strong rain events (>10 mm/h), the maximum WAA that is reached with our set of parameters is in the same range as the 2.3 dB used as maximum in the approach of Schleiss et al. (2013). We want to note that several recent methods quantifying the WAA were developed using auxiliary information such as rain gauge data. This is the reason we did not consider these approaches, as we wanted our CML data processing to be as applicable to new regions as possible. The transferability of WAA estimation methods remains an open scientific question, though. Fencl et al. (2019) quantified the influence of WAA for eight very short (length < 500 m) CMLs using cumulative distribution functions from attenuation and rain gauge data. Their approach is not applicable to new CMLs as it requires calibration for each individual CML based on the local rainfall and attenuation statistics. Ostrometzky et al. (2018) used a rain gauge to estimate the WAA of an E-band CML. They calculated both the (dry, constant during rain events) baseline and the theoretical attenuation using rain gauge data and attributed the residual attenuation as WAA. Moroder et al. (2020) developed a model based on the dynamic antenna parameters reflectivity, efficiency and directivity based on a full-wave simulation and applied it on a dedicated experimental setup with CML antennas (Moroder et al., 2019). To apply this method it is required to continuously collect the individual properties of the CML antennas, which might only be possible in future CML hardware generations.

2.3.2.4 Derivation of rain rates

The estimation technique of rainfall from the WAA-corrected attenuation is based on the well known relation between specific path attenuation k in dB/km and rain rate R in mm/h

$$k = aR^b \tag{2.5}$$

with a and b being constants depending on the frequency and polarization of the microwave radiation (Atlas and Ulbrich, 1977). In the currently most commonly used CML frequency range between 15 GHz and 40 GHz, the constants only show a low dependence on the rain drop size distribution. Using the k-R relation, rain rates can be derived from the path integrated attenuation measurements that CMLs provide as shown in Fig. 2.3 d). We used

values for a and b according to ITU-R (2005) which show good agreement with calculations from disdrometer data in southern Germany (Chwala and Kunstmann, 2019, Fig. 3).

2.4 Results and Discussion

2.4.1 Comparison of rain event detection schemes

The separation of wet and dry periods has a crucial impact on the accuracy of the rainfall estimation. We compared an approach from Schleiss and Berne (2010) to three modifications on their success in classifying wet and dry events as explained in Sect. 2.3.2.2.
The climatologic approach by Schleiss and Berne (2010) worked well for CMLs with moderate noise and when the fraction of times with rainfall over the analyzed periods did correspond to the climatological value. The median MDE was 0.33 and the median MCC of 0.43. The distribution of MDE and MCC values from all CMLs of this climatologic threshold were compared with the performance of the two extensions, displayed in Fig. 2.4.

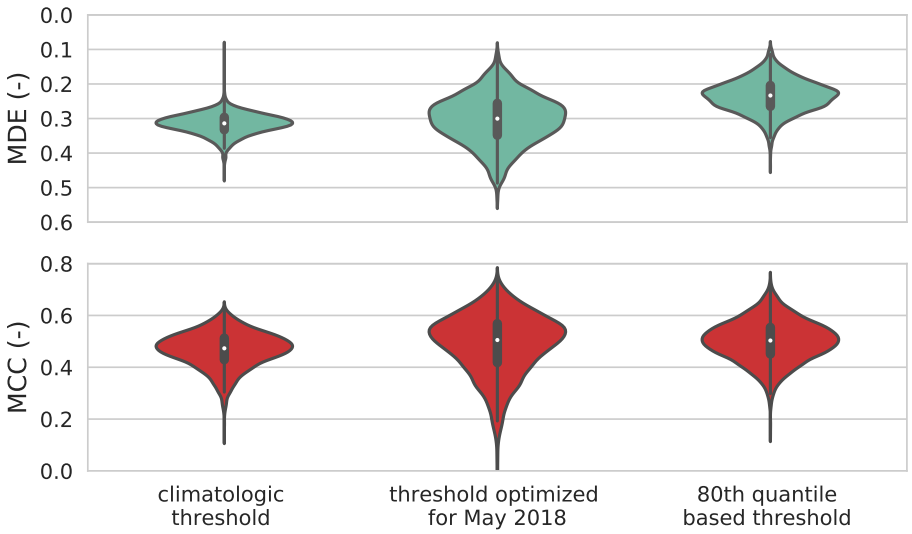

Figure 2.4: Mean detection error (MDE) and Matthews correlation coefficient (MCC) for three rain event detection schemes for the whole analysis period.

When we optimized the threshold for each CML for May 2018 and then applied these thresholds for the whole period, the performance increased with a median MDE of 0.32 and median MCC of 0.46. The better performance of MDE and MCC values highlights the importance of a specific threshold for each individual CML, accounting for their individual notion to fluctuate. The range of MDE and MCC values is wider than with the climatologic threshold, though. The wider range of MDE and MCC values, however, indicates that there is also a need for adjusting the individual thresholds over the course of the year.
The 80th quantile-based method had the lowest median MDE with 0.27 and highest median MCC with 0.47. Therefore it miss-classified the least wet and dry periods compared to

the other methods.
The threshold, which is based on the 80th quantile, is independent from climatology and depends on the individual notion of a CML to fluctuate. Although the factor used to scale the threshold was derived from comparison to the reference data set as described in Sect. 2.3.2.2, it was stable over all seasons and for CMLs in different regions of Germany. Validating the scaling factor with other CML data sets could be a promising method for data scarce regions, as no external information is needed.
For single months, the MDE was below 0.20 as shown in Tab. 2.2, which still leaves room for an improvement of this rain event detection method. Enhancements could be achieved by adding information of nearby CMLs, if available. Also data from geostationary satellite could be used. Schip et al. (2017) found improvements of the rain event detection when using rainfall information from Meteosat Second Generation (MSG) satellite, which carries the Spinning Enhanced Visible and InfraRed Imager (SEVIRI) instrument.
All further processing, presented in the next sections, uses the method based on the 80th quantile.

2.4.2 Performance of wet antenna attenuation schemes

Two WAA schemes are tested and adopted for the present CML data set. Both are compared to uncorrected CML data and the reference in Fig. 2.5. Without a correction scheme, the CML-derived rainfall overestimated the reference rainfall by a factor of two when considering mean hourly rain rates, as displayed in Fig. 2.5a).
The correction by Schleiss et al. (2013) produced comparable mean hourly rain rates with regard to the reference data set. Despite its apparent usefulness to compensate for WAA, this scheme worked well only for stronger rain events. The mean detection error is higher than for the uncorrected data set, because small rain events are suppressed completely throughout the year. The discrepancy can also be a result of the path length of 7.6 km in our data set which is four times the length of the CML Schleiss et al. (2013) used. This might have an impact, since shorter CMLs have a higher likeliness that both antennas get wet. Furthermore, the type of antenna and antenna cover impacts the wetting during rain, as discussed in section Sect. 2.3.2.3.
With the method of Leijnse et al. (2008) the overestimation of the rain rate was also compensated well. It incorporates physical antenna characteristics and, what is more important, depends on the rain rate. The higher the rain rate, the higher the WAA compensation. This leads to less suppression of small events. The MDE is close to the uncorrected data sets and the PCC is higher, as displayed in 2.5b) and c). Recent results from Fencl et al. (2019) also favor a dynamic, rain intensity depended WAA model, instead of a constant value for WAA compensation. Therefore, the scheme from Leijnse et al. (2008) is used for the evaluation of the CML-derived rain rates in the following sections.
Both methods are parameterized, neglecting known and unknown interactions between WAA and external factors like temperature, humidity, radiation and wind. Current re-

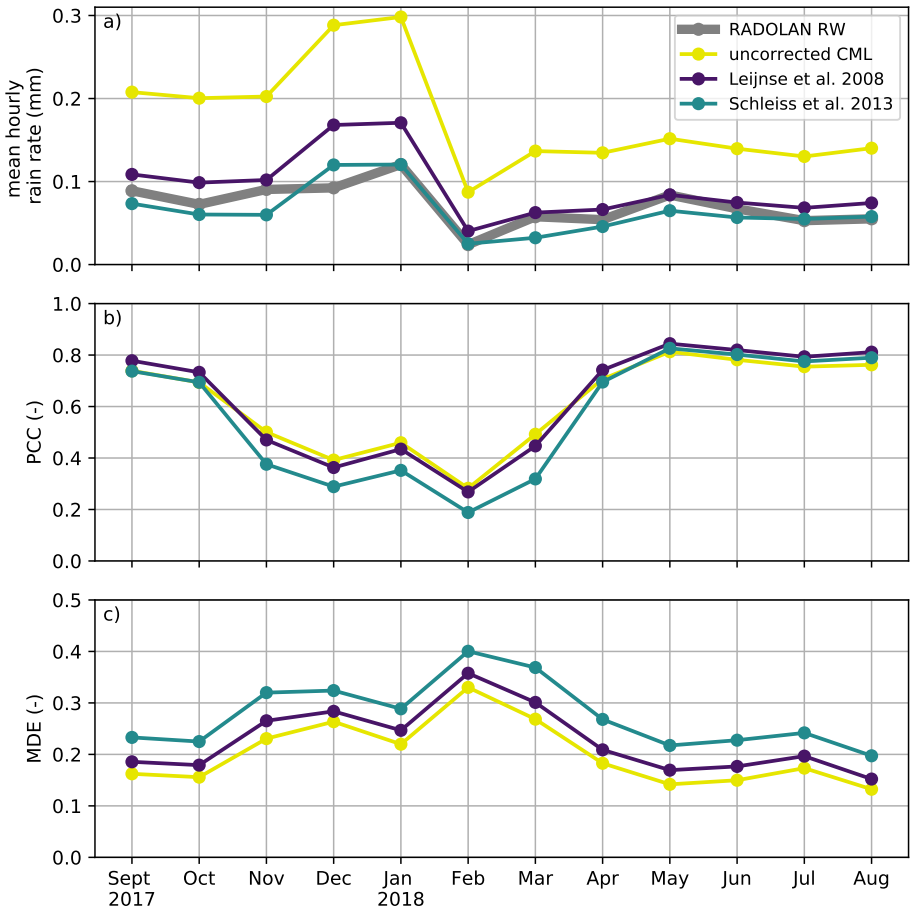

Figure 2.5: WAA compensation schemes compared on their influence on the a) mean hourly rain rate, b) the correlation between the derived rain rates and the reference and c) the mean detection error between the derived rain rates and the reference.

search aims to close this knowledge gap, but the feasibility for large scale networks like the one presented in this study is going to be a challenge as only TSL and RSL are available. A possible solution is the WAA model based on the reflectivityy, efficiency and directivity of the antenna proposed by Moroder et al. (2020), which would have to be measured by future CML hardware. Another approach could be to extend the analysis with meteorological model reanalysis products to be able to better understand WAA behavior in relation to meteorologic parameters like wind, air temperature, humidity and solar radiation.

2.4.3 Evaluation of CML derived rainfall

Path-averaged rainfall information obtained from almost 4000 CMLs is evaluated against a reference data set, RADOLAN-RW. In Fig. 2.6 we show scatter density plots of path averaged hourly rain rates, daily rainfall sums and seasonal sums of each CML with the respective performance measures. Furthermore, scatter density plots of hourly, path-averaged rain rates and rain rates from interpolated rainfall maps are compared for each month in Fig. 2.8 and Fig. 2.9. Looking at the differences between the seasons in 2.6, it is evident, that CMLs are prone to produce significant rainfall overestimation during the

cold season (DJF). This can be attributed to precipitation events with melting snow,

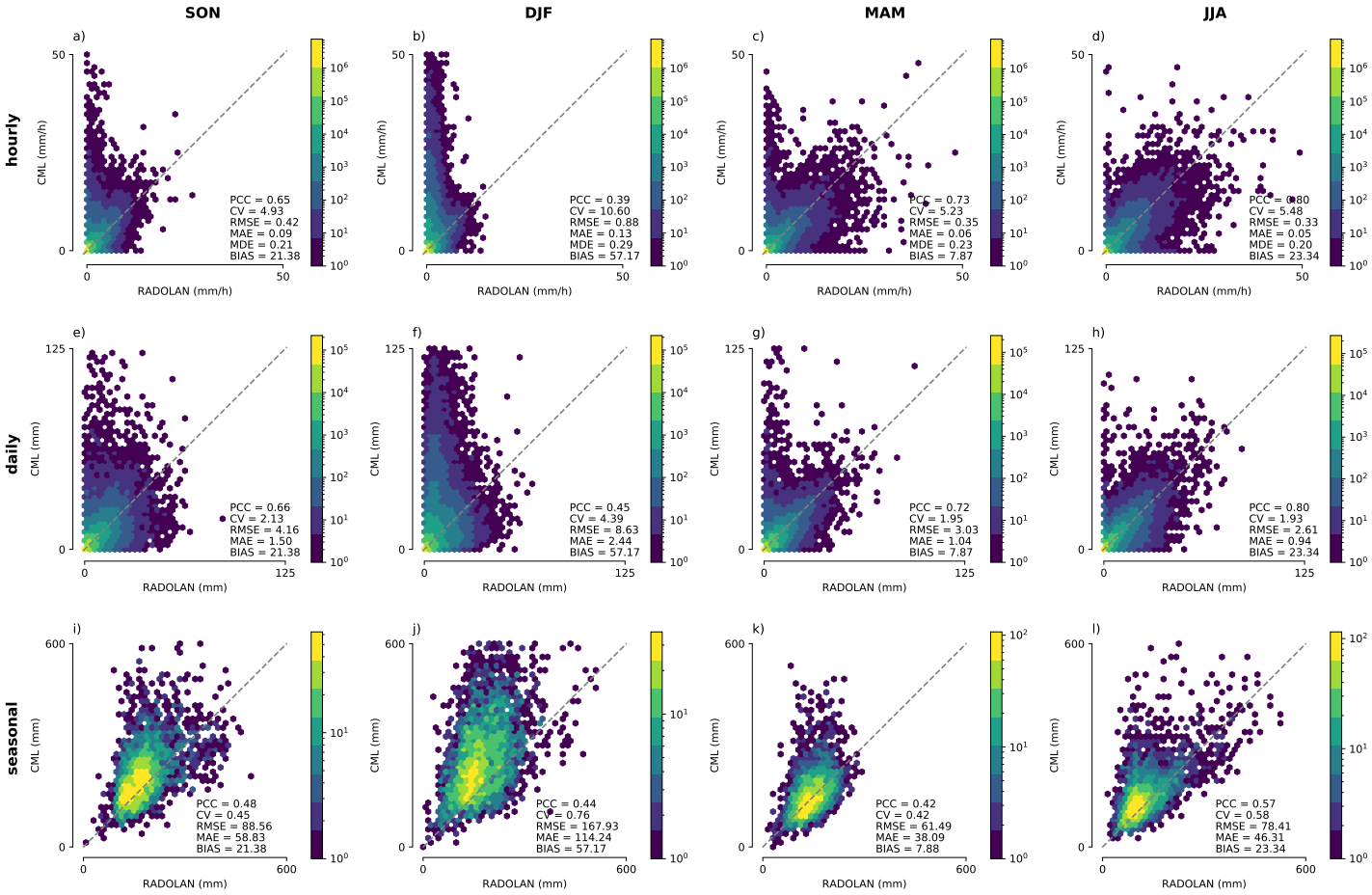

Figure 2.6: Seasonal scatter density plots of CML-derived rainfall and path-averaged RADOLAN-RW data for hourly, a) - d), daily, e) - h) and seasonal, i) - l) aggregations with respective performance metrics calculated from all available data pairs.

occurring mainly from November to March. Melting snow can potentially cause as much as four times higher attenuation than a comparable amount of liquid precipitation (Paulson and Al-Mreri, 2011). Snow, ice and their melt water on the covers of the antennas can also cause additional attenuation. A decrease of the seasonal performance measures also reflects this effect, as the lowest values for PCC and highest for CV, MAE, RMSE, BIAS and MDE are found for DJF. The largest overestimation occurs at low rain rates of the reference. At higher reference rain rates, which most likely are those stemming from liquid precipitation, there is far less overestimation. In spring (MAM) and fall (SON), overestimation by CML rainfall is still visible, but less frequent. This can be explained by the fact, that in the Central German Upland and the Alps, snowfall can occur from October to April. Best agreement between CML-derived rainfall and RADOLAN-RW is found for summer (JJA) months.

The temporal aggregation to daily rainfall sums and the respective performance measures are shown in 2.6e)-h). The general relation between CML derived rainfall and the refer-

ence is similar on both the hourly and daily scale. The BIAS is identical for the daily aggregation. The RMSE and MAE are higher due to the higher rain sums. The overestimation during the winter month is unchanged.
The accumulated rainfall sums of individual CMLs are compared against the reference rainfall accumulation for each season in Fig 2.6i) - l). The overestimation of the CML derived rainfall sums in DJF, and partly SON and MAM, can again be attributed to the presence of non-liquid precipitation. This overestimation is larger for higher rainfall sums. This could be the result of more extensive snowfall in the mountainous parts of Germany, which are also the areas with highest precipitation year round. Rainfall sums close to zero can be the result from the quality control that we have applied. Periods with missing data in CML time series are consequently not counted in the reference rainfall data set. Therefore, the rainfall sums in Fig. 2.6 are not representative for the rainfall sum over Germany for the shown period. The PCC for the four seasons shown in Fig. 2.6i)-l) range from 0.42 in MAM to 0.57 in JJA.

2.4.4 Performance measures for different subset criteria

Tab. 2.2 gives an overview of monthly performance measures for different subsets of CML-derived and path-averaged reference rainfall. In the following, we will discuss the effects of the different subset criteria and then compare our results to previous CML rainfall estimation studies.
For all subset criteria, best performance measures are found during late spring, summer and early fall. Highest PCC values are reached when all data pairs, including true dry events, are used to calculate the measures. When very light rain (< 0.1 mm/h) is set to zero on an hourly basis, the performance measures stay very similar, with the exception of CV and BIAS, which show a slight increase in performance. This means that, even when very small rain rates < 0.1 mm are produced, they do not change rainfall sums too much. When either R_{CML} or $R_{reference}$ have to exceed 0 mm/h, the performance measures are worse than with all data, because all 0 mm/h pairs are removed. When the same subset criterion is set to 0.1 mm/h, a good agreement in the range of very small rain rates below 0.1 mm/h between both data becomes apparent, because the performance measures get worse without them.
To examine the performance of the CML derived rainfall during rain events detected by the reference, two thresholds are selected, where the reference must be above 0.1 and 1 mm/h, respectively, for the period to be considered rainy. With these thresholds, all false wet classifications are removed before the calculation of the performance measures. The PCC with this thresholds is still high for the non-winter months. The CV is reduced, while MAE and RMSE are higher due to higher mean rain rates. The biggest differences can be observed in the bias, where the influence of false wet detection and the overestimation of CMLs over 0.1 and 1 mm/h reduce the bias.

Table 2.2: Monthly performance measures between path averaged, hourly CML-derived rainfall and RADOLAN-RW as reference for subset criteria and thresholds.

	subset criteria		**2017**				**2018**							
	(mm)	**mean**	**Sept**	**Oct**	**Nov**	**Dec**	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**
PCC (-)	none	**0.62**	0.78	0.73	0.46	0.36	0.43	0.27	0.45	0.74	0.85	0.81	0.79	0.81
	light rain to 0	**0.62**	0.78	0.73	0.46	0.36	0.43	0.27	0.45	0.74	0.85	0.81	0.79	0.81
	cml or ref > 0	**0.58**	0.74	0.68	0.38	0.28	0.35	0.20	0.37	0.71	0.83	0.80	0.78	0.79
	cml or ref >= 0.1	**0.54**	0.70	0.64	0.34	0.23	0.31	0.13	0.32	0.68	0.81	0.78	0.76	0.77
	ref >= 0.1	**0.58**	0.73	0.71	0.38	0.28	0.35	0.22	0.39	0.73	0.82	0.79	0.80	0.80
	ref >= 1	**0.51**	0.65	0.64	0.32	0.17	0.27	0.12	027	0.67	0.75	0.73	0.73	0.74
CV (-)	none	**7.01**	3.80	4.40	6.09	11.4	7.62	18.5	6.82	5.20	3.98	5.17	5.88	5.33
	light rain to 0	**7.19**	3.88	4.51	6.23	11.64	7.75	18.28	7.06	5.33	4.03	5.23	5.96	5.40
	cml or ref > 0	**3.03**	1.73	2.00	2.96	5.59	3.85	6.82	3.09	2.19	1.60	2.04	2.36	2.10
	cml or ref >= 0.1	**2.42**	1.40	1.64	2.51	4.78	3.35	5.19	2.53	1.67	1.18	1.50	1.71	1.54
	ref >= 0.1	**1.69**	1.05	1.06	1.92	3.61	2.67	3.25	1.90	1.11	0.88	1.01	0.96	0.92
	ref >= 1	**1.11**	0.73	0.69	1.24	2.27	1.73	2.18	1.14	0.70	0.63	0.72	0.67	0.65
MAE (mm/h)	none	**0.08**	0.08	0.08	0.11	0.17	0.17	0.05	0.07	0.05	0.06	0.06	0.05	0.05
	light rain to 0	**0.08**	0.08	0.07	0.11	0.17	0.16	0.05	0.07	0.05	0.05	0.05	0.05	0.05
	cml or ref > 0	**0.41**	0.38	0.36	0.46	0.71	0.64	0.37	0.35	0.30	0.34	0.36	0.33	0.33
	cml or ref >= 0.1	**0.64**	0.58	0.53	0.64	0.97	0.86	0.66	0.53	0.49	0.61	0.64	0.60	0.58
	ref >= 0.1	**0.72**	0.64	0.57	0.70	1.02	0.91	0.68	0.55	0.54	0.73	0.83	0.74	0.69
	ref >= 1	**1.40**	1.16	1.05	1.40	2.02	1.73	1.73	1.25	1.09	1.30	1.51	1.39	1.22
RMSE (mm/h)	none	**0.48**	0.34	0.33	0.56	1.08	0.94	0.46	0.41	0.29	0.36	0.35	0.32	0.30
	light rain to 0	**0.48**	0.35	0.33	0.56	1.08	0.94	0.46	0.41	0.29	0.34	0.35	0.32	0.30
	cml or ref > 0	**1.06**	0.75	0.71	1.16	2.18	1.84	1.25	0.90	0.68	0.84	0.89	0.78	0.75
	cml or ref >= 0.1	**1.34**	0.94	0.87	1.38	2.58	2.14	1.70	1.12	0.90	1.14	1.22	1.08	1.02
	ref >= 0.1	**1.45**	1.01	0.90	1.47	2.66	2.22	1.68	1.15	0.96	1.33	1.52	1.31	1.18
	ref >= 1	**2.33**	1.59	1.43	2.36	4.02	3.33	3.48	1.97	1.61	1.99	2.32	2.04	1.78
BIAS (%)	none	**30**	20	34	11	79	39	67	7	21	0	10	30	35
	light rain to 0	**29**	20	34	11	80	40	67	7	20	-2	8	27	32
	cml or ref > 0	**30**	20	34	11	79	39	67	7	21	0	10	30	35
	cml or ref >= 0.1	**29**	20	33	11	80	40	67	7	20	-2	8	27	32
	ref >= 0.1	**-4**	-1	-1	-15	36	14	-6	-20	-10	-16	-15	-13	-3
	ref >= 1	**-9**	-4	-9	-24	22	2	-16	-21	-12	-15	-17	-13	-5
MDE	none	**0.23**	0.20	0.19	0.24	0.27	0.23	0.35	0.29	0.22	0.19	0.19	0.22	0.17

Therefore, when discussing these performance measures in relation to previous studies on CML rainfall estimation, the selection of the threshold is of great importance. de Vos et al. (2019) showed a collection of dutch CML-studies in Table A1. In Tab. 2.3 we compare our performance measures to those of studies from de Vos et al. (2019) table which are similar to our study. 'Similar' in this context means considering the size and temporal aggregation of the CML data set as well as the use of radar data as a reference for path-averaged (link-based) rain rates from CMLs. The performance measures from our results with the respective thresholds are in the same range as the performance measures from de Vos et al. (2019) and Rios Gaona et al. (2015). The results thus should not be compared in a purely quantitative way, because both use different sampling strategies and span different time periods.

2.4.5 Rainfall maps

Interpolated rainfall maps of CML-derived rainfall compared to RADOLAN-RW are shown in Fig. 2.7, Fig. 2.8 and Fig. 2.9. The respective CML maps have been derived using inverse distance weighting (IDW) with the RADOLAN-RW grid as target grid and on an hourly basis. Each CML rainfall value is represented as one synthetic point observation at the center of the CML path. For each pixel of the interpolated rainfall field the nearest 12

Table 2.3: Comparison of the performance measures to similar CML validation studies (only link-based comparisons) with respective thresholds.

Study and Dataset	Comparison	Threshold	Bias (%)	CV (-)	PCC (-)
de Vos et al. (2019)					
Average of 1451 CMLs over 7 months (18 Feb–16 Oct 2016), 15 min instantaneously sampled	Link-based comparison with gauge-adjusted radar, 15 min	CML or ref > 0 mm	23	3.43	0.52
Rios Gaona et al. (2015)					
Average of 1514 CMLs over 12 rainy days (June to September 2011), min-max sampled	CML-based comparison with gauge-adjusted ref, 15 min	CML or ref > 0.1 mm	-13	1.44	0.66
This study					
Average of 3904 CMLs over one year (September 2017 - August 2018), one min instantaneously sampled	CML-based comparison with gauge-adjusted radar, hourly	CML or ref > 0 mm	30	3.03	0.58
		CML or ref >= 0.1 mm	19	2.42	0.54

synthetic CML observation points are taken into account. Weights decrease with the distance d in km, according to d^{-2}. After the interpolation, we masked out grid cells further away than 30 km from a CML path, for each individual time step. Hence, hourly rainfall maps derived from CMLs are only produced for areas with data coverage. We applied the same mask to the reference data set on an hourly basis to increase the comparability between both data sets. For the aggregated rainfall maps, we summed up the interpolated, individually masked, hourly rainfall fields. As an example, Fig. 2.7 shows 48 hours of accumulated rainfall in May 2018. The general distribution of CML-derived rainfall reproduces the pattern of the reference very well and the rainfall sums of both data sets are similar. Individual features of the RADOLAN-RW rainfall field are, however, missed due to the limited coverage by CMLs in certain regions. A video of this 48 hour showcase with hourly time steps is published alongside this study (Graf et al., 2020b).

A qualitative comparison of monthly aggregation of the hourly rainfall maps is shown in Fig. 2.8 and Fig. 2.9. The CML-derived rainfall fields resemble the general patterns of the RADOLAN-RW rainfall fields. Summer months show a better agreement than winter months. This is a direct result of the decreased performance of CML-derived rain rates during the winter season, explained in Sect. 2.4.3. Strong overestimation is also visible year round for a few individual CMLs, for which the filtering of erratic behavior was not successful.

A quantitative comparison of the CML-derived rainfall maps to the reference is shown in the third column of Fig. 2.8 and Fig. 2.9. For these scatter density plots we used all hourly pixel values of the respective month within the 30 km coverage mask. During the winter month, CMLs show strong overestimation. This is a direct result of non-liquid precipitation as described in Sect. 2.4.3. From May to August 2018 the reference shows very high rain intensities between 50 and 100 mm/h, which are not produced by the CML rainfall maps. This can be attributed to several reasons. First, CML-derived rainfall, which serves as basis for the interpolation, is path-averaged, with a typical path length in the range of 3-15 km. This means, that the rainfall estimation of a single CML represents an average of several RADOLAN-RW grid cells which smoothes out the extremes. Second, due to the interpolation, rainfall maxima in the CML rainfall maps can only occur at the synthetic

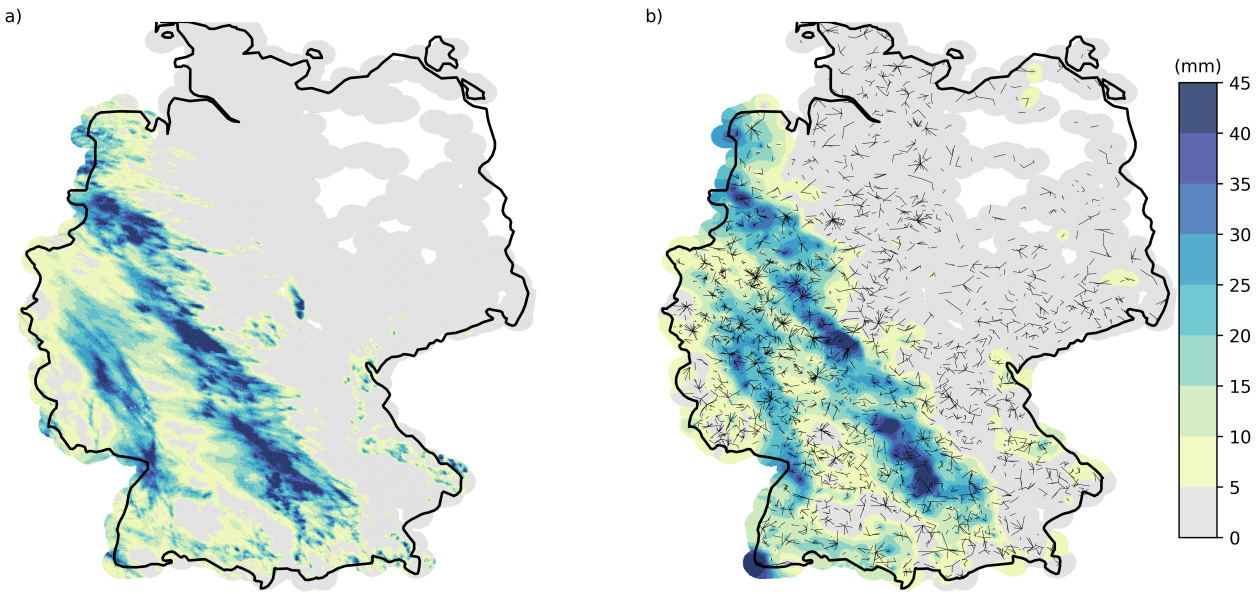

Figure 2.7: Accumulated rainfall for a 48 hour showcase from 12.05.2018 until 14.05.2018 for a) RADOLAN-RW and b) CML-derived rainfall. CML-derived rainfall is interpolated using a simple inverse distance weighting interpolation. A coverage mask of 30 km around CMLs is used.

observation points at the center of each CML. Third, rainfall is only observed along the path of CMLs and even with almost 4000 CMLs across Germany, the spatial variation of rainfall cannot be fully resolved. In particular in summer, small convective rainfall events might not intersect with CML paths and hence cannot appear in the CML-derived IDW interpolated rainfall fields.

Considering this, the effect of different coverage ranges around the CMLs has to be taken into account. For the map based comparison in Fig. 2.8 and Fig. 2.9 we tested several distances from 10 to 50 km. For the presented results we choose 30 km as a trade off between minimizing the uncertainty of the spatial interpolation and the goal to reach country wide coverage with the produced rainfall maps. van de Beek et al. (2012) found an averaged range of around 30 km for summer semi-variograms of 30 years of hourly rain gauge data in the Netherlands, which can be used to justify/enforce our choice. With a 10 km coverage range, the performance measures are better than the ones for 30 km, which are shown in Fig. 2.8 and Fig. 2.9. Monthly PCC values show an increase of around 0.05 and the bias is reduced by 3 to 5 percent. Nevertheless, with a coverage of 10 km around the CMLs, coverage gaps emerge not only in the north-eastern part of Germany, but also in the south eastern part. Vice versa, with a 50 km coverage range, the country wide coverage is almost given, while the performance measures are worse compared to 30 km (PCC shows a decrease between 0.03 and 0.05). Overall, the difference of the performance measures of the 10 and 50 km coverage mask is limited in most parts of Germany by the high density of CMLs, which already lead to an almost full coverage with the 10 km mask.

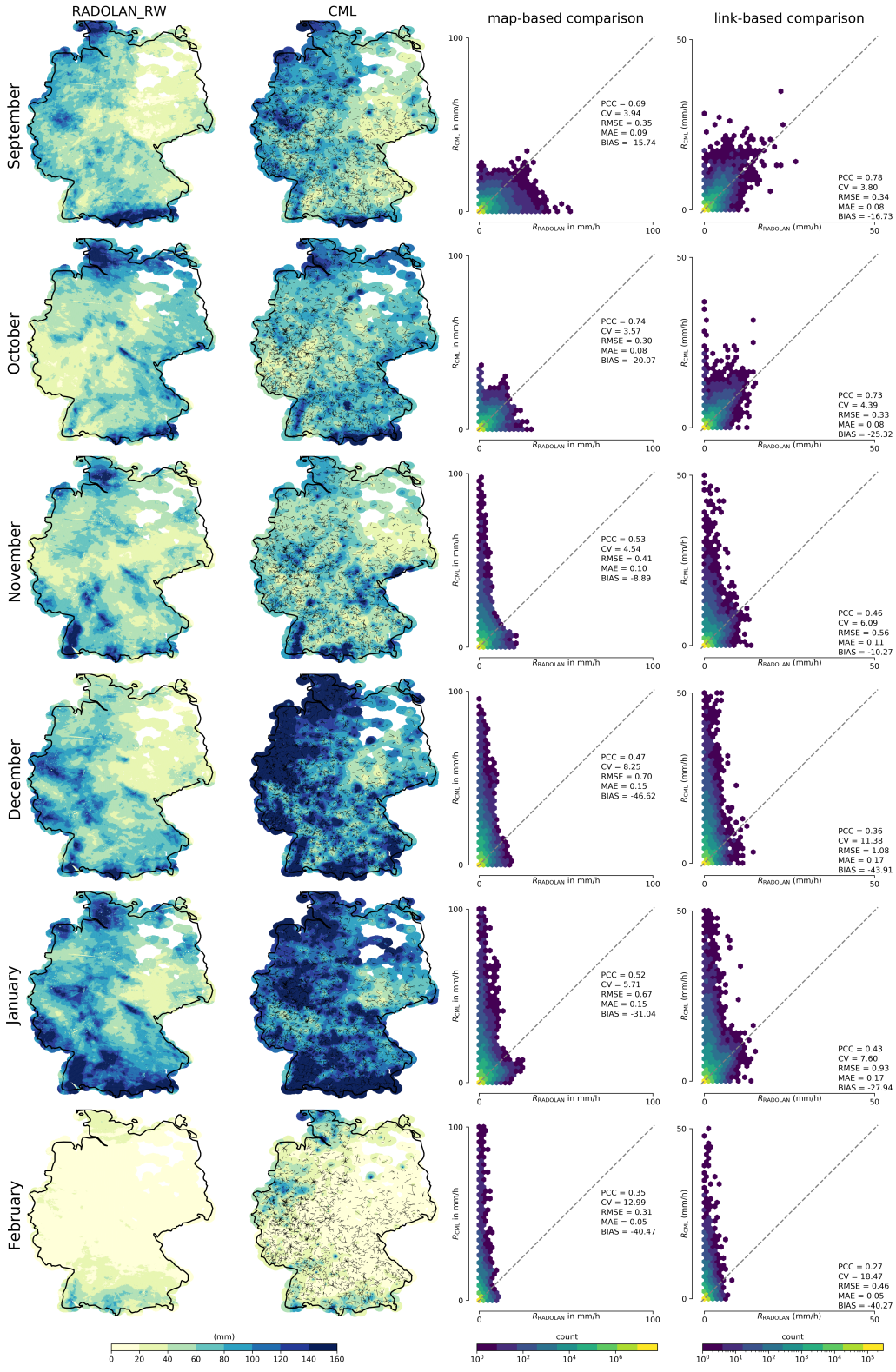

Figure 2.8: Monthly aggregations of hourly rainfall maps from CMLs compared to RADOLAN-RW from September 2017 until February 2018. For each month two scatter density plots are shown, one for pixel-by-pixel comparison of the hourly maps (map-based comparison), and one for the comparison of the hourly path-averaged rainfall along the individual CMLs (link-based comparison).

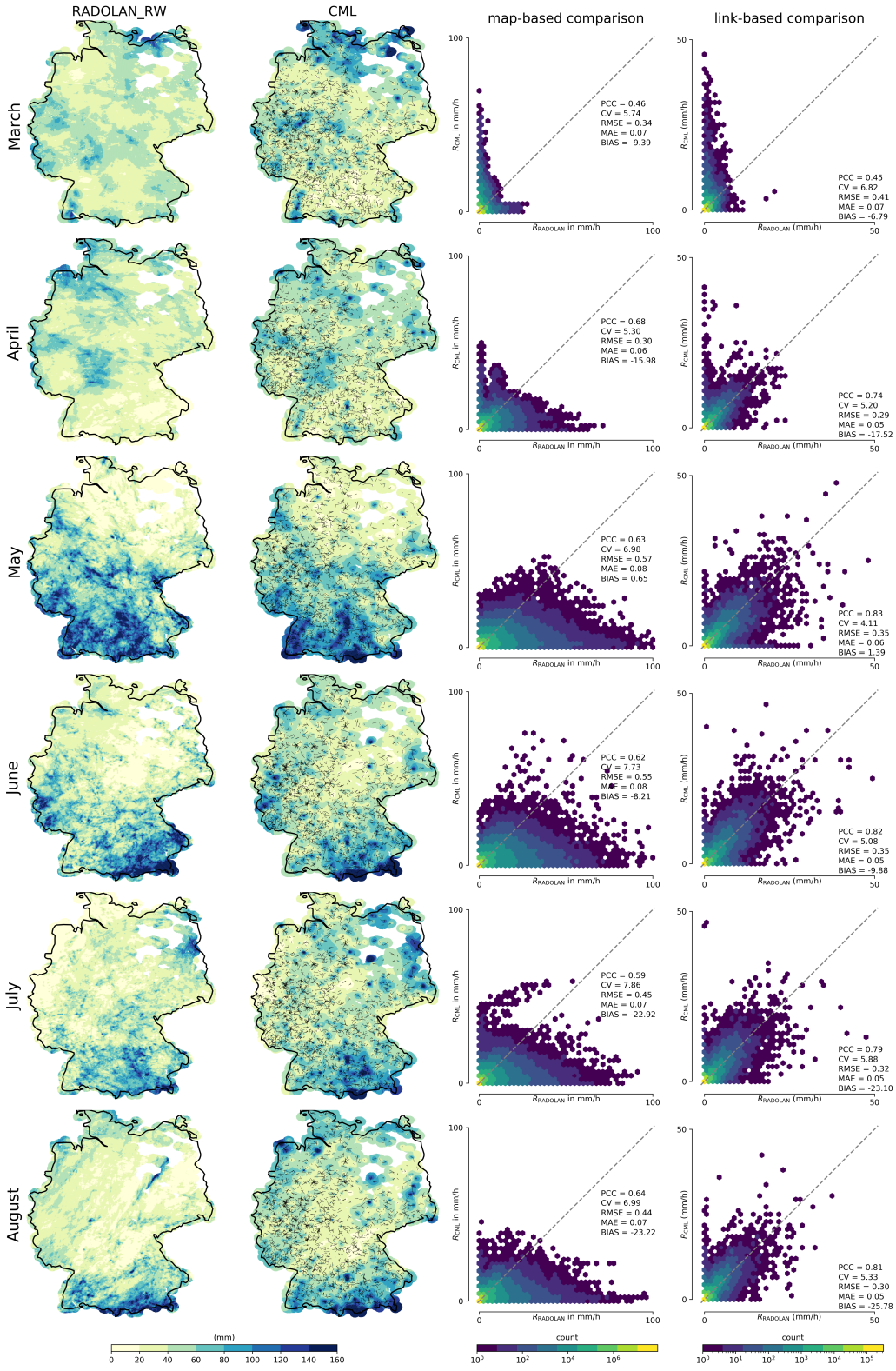

Figure 2.9: Monthly aggregations of hourly rainfall maps from CMLs compared to RADOLAN-RW from March until August 2018. For each month two scatter density plots are shown, one for pixel-by-pixel comparison of the hourly maps (map-based comparison), and one for the comparison of the hourly path-averaged rainfall along the individual CMLs (link-based comparison).

In order to highlight the differences between a map-based and link-based comparison Fig. 2.8 and Fig. 2.9 also show hourly link-based scatter density plots for each month. The differences in the performances measures for the warm months support the qualitative impression, that the map-based comparison perform worse. The interpolation is prone to introduce an underestimation for areas which are more distant to the CML observations. During the winter months, this underestimation compensates the overestimation of the individual CMLs which is due to wet snow and ice covered antennas. Hence, because the two errors compensate each other by chance, this results in slightly better map-based performance measures compared to the link-based measures for the winter months. Nevertheless, rainfall estimation with CMLs for months with non-liquid precipitation is considerably worse than for summer months in all spatial and temporal aggregations.
The derivation of spatial information from the estimated path-averaged rain rates could be improved by applying more sophisticated techniques as described in Sect. 2.1. We have already carried out several experiments using Kriging, to test one of these potential improvements over IDW. We followed the approach of Overeem et al. (2016b) and adjusted the semivariogram parameters on a monthly basis based on the values from van de Beek et al. (2012). We also tried fixed semivariogram parameters and parameters estimated from the individual CML rainfall estimates for each hour. In conclusion, we, however, only found marginal or no improvements of the performance metrics of the CML rainfall maps. Combined with the drawback of Kriging that the required computation time is significantly increased (approximately 10 to 100 times slower than IDW, depending e.g. on the number of neighboring points used by a moving kriging window), we thus decided to keep using the simple, yet robust and fast IDW interpolation. Furthermore, it is important to note that the errors in rain rate estimation for each CML contribute most to the uncertainty of CML-derived rainfall maps (Rios Gaona et al., 2015). Hence, within the scope of this work, we focused on improving the rainfall estimation at the individual CMLs. Taking into account that we compare to a reference data set derived from 17 C-band weather radars combined with more than 1000 rain gauges, the similarity with the CML-derived maps, which solely stem from the opportunistic usage of attenuation data, is remarkable.

2.5 Conclusion

German wide rainfall estimates derived from CML data compared well with RADOLAN-RW, a hourly gridded gauge-adjusted radar product of the DWD. The methods used to process the CML data showed promising results over one year and several thousand CMLs across all landscapes in Germany, except for the winter season.
We presented the data processing of almost 4000 CMLs with a temporal resolution of one minute from September 2017 until August 2018. We developed a parallelized processing work flow, which could handle the size of this large data set. This workflow enabled us to test and compare different processing methods over a large spatiotemporal scale.

A crucial processing step is the rain event detection from the TRSL, the raw attenuation data recorded for each CML. We used a scheme from (Schleiss and Berne, 2010) which uses the 60 minute rolling standard deviation RSD and a threshold. We derived this threshold from a fixed multiple of the 80th quantile of the RSD distribution of each TRSL. Compared to the original threshold using the 95th quantile, which is based on rainfall climatology, the 80th quantile reflects the general notion of each CML's TRSL to fluctuate. We were able to reduce the amount of miss-classification of wet and dry events, reaching a yearly mean MDE of 0.27, with an average of the MDE for summer months below 0.20. Potential approaches for further decreasing the amount of miss-classifications could be the use of additional data sets. For example, cloud cover information from geostationary satellites could be employed to reduce false wet classification, by, as a first simple approach, defining periods without clouds as dry. Another opportunity would be, to additionally implement algorithms exploiting information of neighboring CMLs.

For the compensation of WAA, the attenuation caused by water droplets on the cover of CML antennas, we compared and adjusted two approaches from literature. In order to evaluate WAA compensation approaches, we used the reference data set. We were able to reduce the overestimation caused by WAA, while maintaining the detection of small rain events, using an adjustment of the approach introduced by Leijnse et al. (2008). The compensation of WAA without an evaluation with a reference data set is not feasible with the CML data set we use.

Compared to the reference data set RADOLAN-RW, the CML-derived rainfall performs well for periods with only liquid precipitation. For winter months, the performance of CML-derived rainfall is limited. Melting snow and snowy or icy antenna covers can cause additional attenuation resulting in overestimation of precipitation, while dry snow cannot be measured at the frequencies and the TRSL quantizations the CMLs in our data set use. We found high correlations for hourly, monthly and seasonal rainfall sums between CML-derived rainfall and the reference. To increase the comparability of our analysis with existing and future studies on CML rainfall estimation we calculated all performance metrics for different subset criteria, e.g. requiring that either CML or reference rainfall is larger than 0 mm.

We found the performance measures of this study to be in accordance with similar CML studies, although the comparability is limited due to differences of the CML and reference data sets. CML-derived rainfall maps calculated with a simple, yet robust inverse distance weighting interpolation showed the plausibility of CMLs as an stand-alone rainfall measurement system.

With the analysis presented in this study, the need for reference data sets in the processing routine of CML data is reduced, so that the opportunistic sensing of country-wide rainfall with CMLs is at a point, where it should be transferable to (reference) data scarce regions. Especially in Africa, where water availability and management are critical, this task should be challenged as Doumounia et al. (2014) did already. The high temporal resolution of the presented data set can be used in future studies, e.g. for urban water

management. In addition, CML derived rainfall can also complement other rainfall data sets, e.g. to improve the radar data adjustment in RADOLAN in regions with high CML density and regions, like mountain ranges, where radar data is often compromised. Thus, CMLs can contribute substantially to improve the spatiotemporal estimations of rainfall.

2.6 Code availability

Code used for the processing of CML data can be found in the Python package pycomlink (pycomlink, 2021).

2.7 Data availability

CML data was provided by Ericsson Germany and is not publicly available. RADOLAN-RW is publicly available through the Climate Data Center of the German Weather Service

2.8 Acknowledgements

We thank Ericsson for the support and cooperation in the acquisition of the CML data. This work was funded by the Helmholtz Association of German Research Centres within the Project *Digital Earth*. We also like to thank the German Research Foundation for funding the projects *IMAP* and *RealPEP* and the Bundesministerium für Bildung und Forschung for funding the project *HoWa-innovativ*. We further thank the three anonymous reviewers for their valuable comments improving this paper.

2.9 Appendix to Chapter 2

Table 2.4: Comprehensive overview of used parameters, a short description and their reference from literature if applicable. Parameters with enumeration in parentheses are not used in the final processing.

	description	**parameter value**	**source**
parameters used in final processing routine			
1. Erratic behavior of CMLs (section 3.2.1)			
1.1	sanity check to remove CMLs with strong duirnal cycle or which have noisy periods	5 hour RSD >2 for at lest 10% per month	this study
1.2	sanity check to remove CMLs with high fluctuation over large parts of or the complete month	1 hour RSD >0.8 for at least 33% per month	this study
2. Rain event detection (section 3.2.2)			
2.1	RSD window length	60 min	Schleiss and Berne (2010)
2.2	scaled q80 threshold	1.12 * 80% quantile of RSD	this study
3. WAA compensation (section 3.2.3)			
3.1	thickness of antenna cover	4.1 cm	measured from one antenna cover
3.2	scale for water film thickness γ	1.47E-5	modified after Leijnse et al. (2008)
3.3	factor for the relation between	0.36	modified after Leijnse et al. (2008)
parameters used in alternative processing steps			
(2.3)	climatologic threshold	95% quantile of RSD	Schleiss and Berne (2010)
(3.4)	time for WAA to reach maximum τ	15 min	Schleiss et al. (2013)
(3.5)	maximal value of WAA	2.3 dB	Schleiss et al. (2013)

Chapter 3

Rain event detection in commercial microwave link attenuation data using convolutional neural networks (Polz et al., 2020)

Polz, J., Chwala, C., Graf, M., and Kunstmann, H.: Rain event detection in commercial microwave link attenuation data using convolutional neural networks, *Atmos. Meas. Tech.*, 13, 3835–3853, https://doi.org/10.5194/amt-13-3835-2020, 2020

Abstract

Quantitative precipitation estimation with commercial microwave links (CMLs) is a technique developed to supplement weather radar and rain gauge observations. It is exploiting the relation between the attenuation of CML signal levels and the integrated rain rate along a CML path. The opportunistic nature of this method requires a sophisticated data processing using robust methods. In this study we focus on the processing step of rain event detection in the signal level time series of the CMLs, which we treat as a binary classification problem. This processing step is particularly challenging, because even when there is no rain the signal level can show large fluctuations similar to that during rainy periods. False classifications can have a high impact on falsely estimated rainfall amounts. We analyze the performance of a convolutional neural network (CNN), which is trained to detect rainfall specific attenuation patterns in CML signal levels, using data from 3904 CMLs in Germany. The CNN consists of a feature extraction and a classification part with, in total, 20 layers of neurons and 1.4×10^5 trainable parameters. With a structure, inspired by the visual cortex of mammals, CNNs use local connections of neurons to recognize patterns independent of their location in the time-series (Fukushima, 1980). We test the CNNs ability to generalize to CMLs and time periods outside the training data. Our CNN is trained on four months of data from 800 randomly selected CMLs

and validated on two different months of data, once for all CMLs and once for the 3104 CMLs not included in the training. No CMLs are excluded from the analysis. As a reference data set we use the gauge adjusted radar product RADOLAN-RW provided by the German meteorological service (DWD). The model predictions and the reference data are compared on an hourly basis. Model performance is compared to a state of the art reference method, which uses the rolling standard deviation of the CML signal level time series as a detection criteria. Our results show that within the analyzed period of April to September 2018, the CNN generalizes well to the validation CMLs and time periods. A receiver operating characteristic (ROC) analysis shows that the CNN is outperforming the reference method, detecting on average 76% of all rainy and 97% of all non-rainy periods. From all periods with a reference rain rate larger than 0.6 mmh^{-1}, more than 90% were detected. We also show that the improved event detection leads to a significant reduction of falsely estimated rainfall by up to 51%. At the same time, the quality of the correctly estimated rainfall is kept at the same level in regard to the Pearson correlation with the radar rainfall. In conclusion, we find that CNNs are a robust and promising tool to detect rainfall induced attenuation patterns in CML signal levels from a large CML data set covering entire Germany.

3.1 Introduction

Rainfall is the major driver of the hydrologic cycle. Accurate rainfall observations are fundamental for understanding, modeling and predicting relevant hydrological phenomena, e.g. flooding. Data from commercial microwave link (CML) networks have proven to provide valuable rainfall information. Given the high spatio-temporal variability of rainfall, they are a welcome complement to support traditional observations with rain gauges and weather radars; particularly in regions where radar is hampered by beam blockage or ground clutter. In regions with sparse rainfall observation networks, like in developing countries, CMLs might even be the only source of small scale rainfall information.
Since the work of Messer et al. (2006) and Leijnse et al. (2007) more than a decade ago, several research groups have shown the potential of CML data for hydrometeorological usage. Prominent examples are the countrywide evaluations in the Netherlands (Overeem et al., 2016b) and Germany (Graf et al., 2020a), which demonstrated that CML-derived rainfall information corresponds well with gauge-adjusted radar rainfall products, except for the cold season with solid precipitation. CML-derived rainfall information was also successfully used for river runoff simulations in a pre-alpine catchment in Germany (Smiatek et al., 2017) and for pipe flow simulation in a small urban catchment in Czech Republic (Pastorek et al., 2019). A further important step was the first analysis of CML-derived rain rates in a developing country, carried out by Doumounia et al. (2014), with data from Burkina Faso.
In general, the number of CMLs available for research has increased significantly over the last years and researchers from several countries have gained access to CML attenuation

data. Currently, data from 4000 CMLs over Germany is recorded continuously with a temporal resolution of one minute via a real-time data acquisition system (Chwala et al., 2016). The number of existing CMLs over Germany is 30 times higher (Bundesnetzagentur, 2017), amounting to 130.000 registered CMLs. Consequently, it is envisaged to increase the number of CMLs included in the data acquisition.
With this large number of CMLs available in Germany and with new data being retrieved continuously, there is a need for optimized and robust processing of these big data sets. Several studies address the details of the processing steps which are required for deriving rainfall information from CMLs. These steps involve, e.g. the detection of rain events in noisy raw data, the filtering of artifacts, correcting for bias due to wet antenna attenuation (WAA) and the spatial reconstruction of rainfall fields. Uijlenhoet et al. (2018) give a general overview of the required processing steps and the existing methods and Chwala and Kunstmann (2019) discuss and summarize the related current challenges.

3.1.1 On the importance of rain event detection

The first of these processing steps, called rain event detection, is the separation of rainy (wet) and non-rainy (dry) periods. A static signal level baseline to derive attenuation that can be attributed to rainfall has proven to be ineffective due to e.g. daily or annual cycles and unexpected jumps in the time series like for CML B in Fig. 3.1. Therefore, after the rain events are localized correctly, an event specific attenuation baseline can be determined and actual rain rates can be derived via the k-R power law which relates specific attenuation k in dB km^{-1} to rain rate R in mm h^{-1}.
Detecting rain events is challenging, because CML signal levels can show high fluctuations, even when there is no rain, e.g. due to multi-path propagation (e.g. Chwala and Kunstmann, 2019, Fig. 6). Therefore, the main difficulty is to distinguish between noise and signal fluctuations caused by rain along the CML path. As seen in Fig. 3.1, the differences in noise levels can vary significantly, depending on the CML that is used. When looking at the magnitude of these fluctuations, we can see that a misclassification of wet and dry periods can easily lead to a large over- or underestimation of rainfall. These missed or falsely estimated quantities are often overlooked in scatter density comparisons of rainfall products like Figure 3.9 a) and b) below, which shows our own results. But when absolute amounts are compared, they represent an obvious issue with up to 30% of the total CML rainfall that can be attributed to false positives. As these misclassifications generate a bias different from the bias corrected in later processing steps like the WAA correction it is important to optimize the rain event detection as an isolated processing step first and to optimize subsequent processing steps afterwards.

3.1.2 State of the art

So far, several methods for rain event detection with CMLs have been proposed. The main difference that divides these methods into two groups, is the type of CML data that can

be used to estimate rainfall.

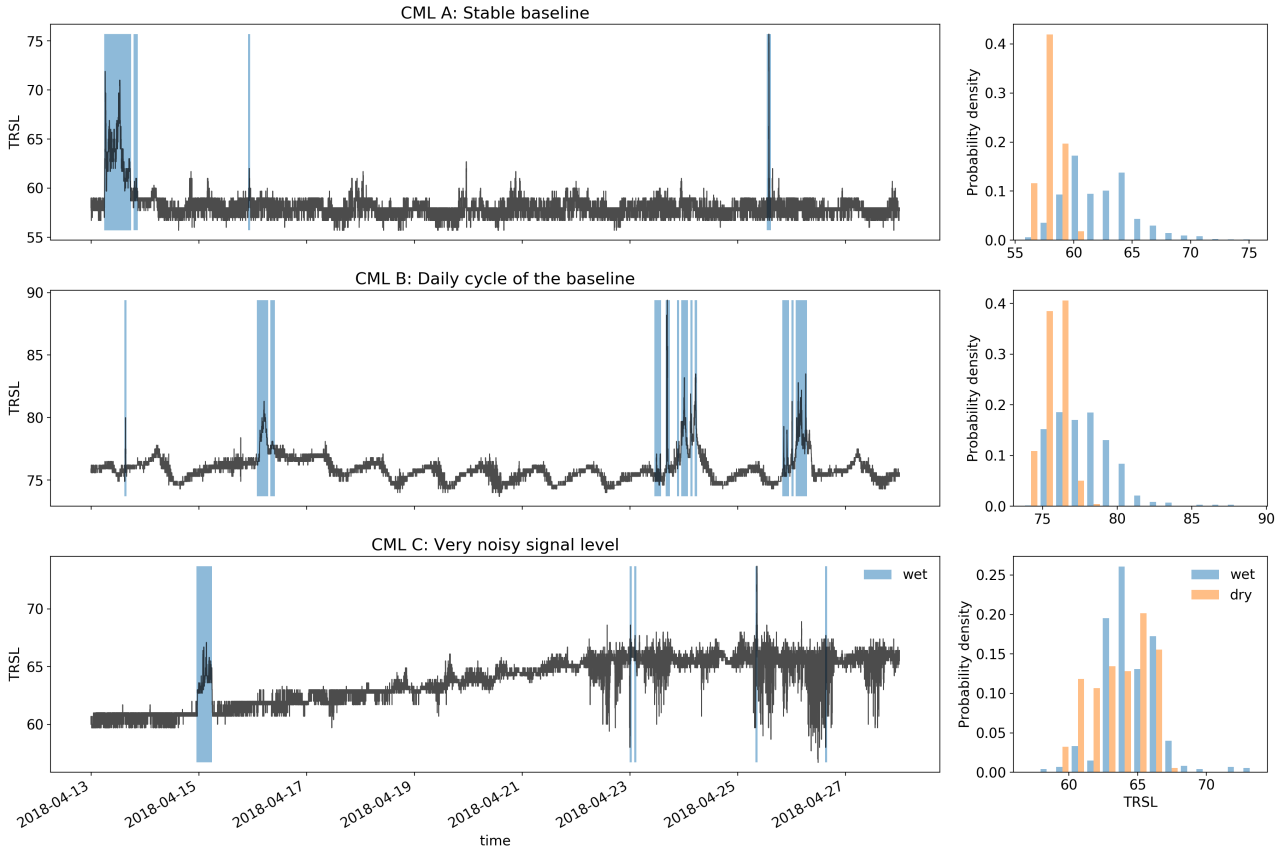

Figure 3.1: Three example signal level (TRSL) time series that illustrate the high variability in data quality when comparing different CMLs. The blue shaded periods indicate where the radar reference show rainfall along the CML paths. The challenge is to identify these periods by analysing the time series. Note that each attenuation event that is falsely classified as wet, will produce false rain rate estimates, which will lead to overestimation. The histograms show that for some CMLs the wet periods can be easily separated from the dry periods and for others the distribution of TRSL values is nearly identical for both classes. Fig. 3.2 below will show an example of how different detection methods deal with the challenging time series of CML C.

Depending on the available data acquisition, CML signal levels are either instantaneously sampled at a rate ranging from a few seconds up to 15 minutes or they are stored as 15-minute minimum and maximum values derived from a high instantaneous sampling rate in the background. In almost all cases only one of the two sampling strategies is available due to the type of data management through the network provider. The resulting rain event detection methods are highly optimized for one kind of sampling strategy and therefore in general incompatible with the other kind.

The following methods were developed for instantaneous measurements: Schleiss and Berne (2010) introduced a threshold for the rolling standard deviation (RSTD) of the

attenuation time-series as a criteria to detect rain events. Despite being one of the first methods that were developed, a large part the method is still the most commonly used within the CML research community, as it was used in very recent studies from different working groups such as Kim and Kwon (2018), Graf et al. (2020a) or Fencl et al. (2020). Chwala et al. (2012) introduced Fourier transformations on a rolling window of CML signal levels to detect the pattern of rain events in the frequency domain. Wang et al. (2012) used a Markov switching model, which was calibrated and validated for a single CML test site. Kaufmann and Rieckermann (2011) have shown the applicability of random forest classifiers and Gaussian factor graphs and validated their approach using 14 CMLs. Đorđević et al. (2014) used a simple Multilayer Perceptron (MLP) which was trained and validated on a single CML. Ostrometzky and Messer (2018) proposed a simple rolling mean approach to determine a dynamic baseline, also validated on a single CML. Most of these studies are based on a comparably low and sometimes pre-selected amount of CMLs ranging from one to a maximum of 50 devices, a number that is likely much larger in a possible operational setting.

As a detection scheme for 15 minute min/max sampled data with a 10 Hz background sampling rate Overeem et al. (2011) introduced the 'nearby link approach'. A period is considered wet if the increase of CML specific attenuation correlates with the attenuation pattern of nearby CMLs. They concluded that this is only applicable for dense CML networks with a high data availability. Later, they conducted the first evaluation of a rain event detection method on data from 2044 CMLs on a country scale Overeem et al. (2016b). Very recently the same approach was used in de Vos et al. (2019), showing that this approach works better in combination with min/max sampling than with 15 minute instantaneous sampling. Habi and Messer (2018) tested the performance of Long Short-Term Memory (LSTM) networks to classify rainy periods from 15 minute min/max values of CML signal levels for 34 CMLs.

All rain event detection methods have to make a similar trade-off: A liberal detection of wet periods is more likely to recognize even small rain rates, while it will produce more false alarms during dry periods. On the other hand, a conservative detection will accurately classify dry periods, but is more likely to miss small rain events. One can address this by two means. First, by increasing detection rates on both wet and dry periods as much as possible and therefore decreasing the impact of the trade-off. Second, by allowing the flexibility to easily adjust the model towards liberal or conservative detection, e.g. by only changing a single parameter.

In conclusion, until now, there have been few studies analyzing the performance of rain event detection methods on large data sets. Overeem et al. (2016b) tested the nearby link approach using 2044 CMLs distributed over the Netherlands with a temporal coverage of 2.5 years of data. Graf et al. (2020a) extended the RSTD method and applied it to one year of data from 3904 CMLs to set a benchmark performance on the same data set used in this study. By optimizing thresholds for individual CMLs the full potential of the RSTD method for one year of data was explored, yielding good results for the warm

season with liquid precipitation. While the RSTD method is simple to implement and has only two parameters (window length and threshold) to optimize, it is limited to measuring the amount of fluctuations, rather than the specific pattern. More room for optimization is expected using a data driven approach, such as machine learning techniques for pattern recognition.

3.1.3 Data driven optimization through deep learning

Deep learning is a rapidly evolving field that is becoming increasingly popular in the earth system sciences. A large field of application is remote sensing using artificial neural networks for image recognition (Zhu et al., 2017). Deep learning is also an established method in time-series classification (Ismail Fawaz et al., 2019). In both studies, convolutional neural networks (CNNs) are considered one of the leading neural network architectures for image and time-series classification. CNNs are inspired by the visual cortex of mammals and they are designed to recognize objects or patterns, regardless of their location in images or time-series. They are characterized by local connections of neurons, shared weights and a large number of layers of neurons, involving pooling layers (LeCun et al., 2015). CNNs with one dimensional input data (1D-CNNs) have already been used for time-series classification, e.g. for classifying environmental sounds (Piczak, 2015). This makes 1D-CNNs a promising candidate for the task of rain event detection in CML signal levels.

3.1.4 Research gap and objectives

Due to the opportunistic use of CMLs, the variety of signal fluctuations and possible occurrences of errors naturally increase in a CML data set with its size. Separating rainy from non-rainy periods is therefore a crucial step for rainfall estimation from CMLs. Although applicable on a large scale, recently applied methods still struggle with falsely estimated rainfall as can be seen in the evaluations from Graf et al. (2020a) and de Vos et al. (2019). Despite the amount of proposed methods, this processing step has not yet been investigated in detail using a large and diverse CML data set, especially for data driven approaches. Given their promising results in other applications, the usage of artificial neural networks (ANNs) for rain event detection in the CML attenuation time-series on a large scale provides a promising opportunity. It has been proven that in many cases ANNs allow for high-performance, fast and robust processing of a variety of suitable data sets. What is missing is a proof that they are applicable to a large and diverse CML data set. The question is, does a high variability of frequency, length and spatial distribution of the analyzed CMLs or a high variability of rain rates and event duration for a large amount of analyzed periods affect the performance of ANNs in this specific case or not? Additionally, the effect of rain event detection performance on the estimated rain rates has yet to be investigated.
The objective of this study is to evaluate the performance of 1D-CNNs to detect rainfall

induced attenuation patterns in instantaneously measured CML signal levels and to investigate the effect of an improved temporal event localization on the CML-derived rainfall amounts. Furthermore, we test the CNNs ability to generalize to new CMLs and future time periods in order to provide a validated open source model, that can be used on other data sets. To provide the CML community with comprehensible results, we compare the CNN to the method of Schleiss and Berne (2010), which we consider state-of-the-art due to the amount of recent applications. We aim to provide a high statistical robustness of the derived performance measures by using the, to date, largest available CML data set consisting of data from 3904 CMLs distributed over entire Germany.

3.2 Methods

The following definition of rain event detection with CMLs is the basis of our methodology: Rain event detection is a binary classification problem. Given a time window $X_{t,w,i}$ of CML signal data, where t is the starting time, w is the window length and i is the index specifying a unique CML path, we have to decide if there is attenuation caused by rain (wet) or not (dry). A time window is assigned the label 1 if it is wet or 0 if it is dry. The available information to do this classification depends on the used data acquisition and on which information is provided by the CML network operator. In the following, we describe how a CNN can be used as a binary classifier to succeed in this task.

3.2.1 Data set

We use a CML data set that has been collected in cooperation with Ericsson Germany through our custom CML data acquisition system (Chwala et al., 2016). It covers 3904 CMLs across entire Germany. The CML path length ranges from 0.1 km to more than 30 km, with an average of around 7 km. CML frequencies range from 10 to 40 GHz. The acquired data consists of two sub-links per CML, transmitting their signal in opposite directions along the CML path. For each sub-link a received signal level (RSL) and a transmitted signal level (TSL) is recorded at a temporal resolution of 1 minute and a power resolution of 0.3 dB for RSL and 1.0 dB for TSL. The recorded period used in this study starts in April 2018 and ends in September 2018, to focus on the periods which are dominated by liquid precipitation, where CMLs perform better than during the cold season (Graf et al., 2020a). The data is available at 97.1% of all time steps and gaps are mainly due to outages of the data acquisition system.
As reference data we use the gauge adjusted radar product RADOLAN-RW provided by the German meteorological service (DWD). It has a spatial resolution of 1x1 km, covering entire Germany on 900x900 grid cells. The temporal resolution is 60 minutes and the resolution for the rain amount is 0.1 mm (Winterrath et al., 2012). To compare to this reference, the window length w is set to 60 minutes and therefore w is omitted in the notation below.

Along each CML i, the path-averaged mean hourly rain rate $R_{t,i}$ is generated from the reference, using the weighted sum

$$R_{t,i} = \frac{\sum_k l_{k,i} r_{k,t}}{l_i}, \tag{3.1}$$

where k is indexing the RADOLAN grid cells intersected by the path of i. The rain rate of each grid cell is $r_{k,t}$. Furthermore, $l_{k,i}$ is the length of the intersect of k and i and l_i is the total length of i. A time window $X_{t,i}$ is considered wet if $R_{t,i} \geq 0.1$ mm h^{-1} and dry otherwise.

3.2.2 Pre-processing

Before training and testing an artificial neural network, the raw time-series data has to be pre-processed. We do this to sample time windows of a fixed size, which are normalized and labelled according to the reference.

First, the full data set, consisting of all available CMLs, is split into three subsets. One subset is used for training the CNN (TRG), one is used for validation and to optimize model hyper-parameters (VALAPR) and one is used for testing only (VALSEP). The data set TRG consists of data from 800 randomly chosen CMLs in the period from May to August 2018. VALAPR covers the remaining 3104 CMLs during April 2018 and VALSEP consists of data from all 3904 CMLs during September 2018. We used this splitting routine to avoid information leakage from the training to the validation data. There can be a high correlation of signal levels between CMLs that are situated close to each other (Overeem et al., 2011). Therefore, the measurements contained in VALAPR or VALSEP can not be taken from the same time range as for TRG. Using only 20% of all available CMLs for training allows us to analyze the CNNs generalization to the remaining CMLs in the validation data set. No CMLs were excluded from this analysis.

For each of the two sub-links of a CML, we compute a transmitted minus received signal level (TRSL). Within one TRSL time-series, randomly occurring gaps of up to five minutes of missing data are linearly interpolated to be consistent with with the preprocessing used in Graf et al. (2020a). We assume that the temporal variability of rainfall is not high enough such that entire rain events can be hidden in such short gaps. The next step is to normalize the data. Normalization of training and validation data is a commonly used procedure in deep learning to enhance the model performance. We perform the normalization as a pre-processing step and outside the CNN. After testing various normalization techniques it turned out that the best performance of the CNN can be achieved by subtracting the median of all available data from the preceding 72 hours from each time step. In rare cases of larger gaps in the data acquisition, we set a lower limit for the data availability to 120 minutes.

The set of starting time-stamps of the hourly reference data set is denoted T_{rad}. For each CML i and each starting time $t \in T_{rad}$ a sample of data $\bar{X}_{t,i}$ is composed from 60+k minutes of TRSL from the two sub-links starting at $t - k$. The first k minutes serve as a

reference to previous behaviour of the same CML and the last 60 minutes are the period $X_{t,i}$ that has to be classified. To investigate the impact of adding this additional information, we compare multiple setups with k ranging from 0 to 240 minutes. The results are given in Sect. 3. An example TRSL over a period of two weeks is shown in Fig. 3.2 (a).

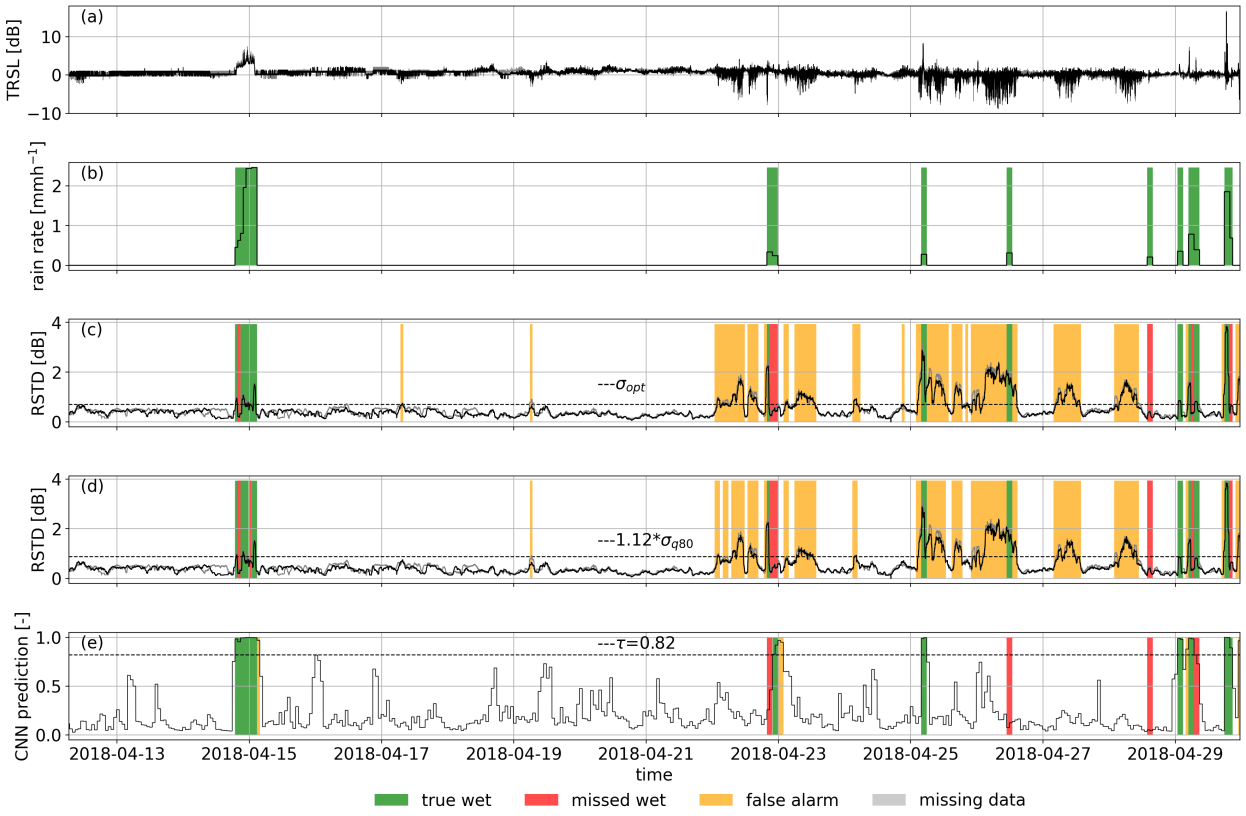

Figure 3.2: Performance of the CNN and the reference methods for the noisy example CML time-series from Fig. 3.1. a) shows the normalized TRSL time-series and b) is the radar reference. Predictions from the CNN (e) yield an MCC of 0.74. Predictions through σ_{opt} (c) and σ_{q80} (d), which are very similar in this case, both yield MCCs of 0.28. Note that the TRSL and RSTD time series of sub-link 2 are almost identical to those of sub-link 1 and are shown in light grey.

After interpolating short gaps, as described above, we exclude all samples with missing values from the analysis. Since we loose up to five hours of data whenever there is a gap, the interpolation routine increases the number of available samples from 75% to 94%.
To train the CNN we have to balance the wet and dry classes in the data set (Hoens and Chawla, 2013). The under-sampling approach to achieve an equalized (50:50) class ratio is to randomly discard samples of the majority class, i.e. dry samples. This approach is chosen since we assume that dry periods mostly consist of redundant samples with only small fluctuations. Later, we check that there is no loss in performance by evaluating

the unbalanced data. The initial percentage of wet samples is between 5-10%. We perform the balancing on TRG and VALAPR. The balanced version of VALAPR is denoted VALAPRB. VALAPR and VALSEP are kept as unbalanced data sets for validation. TRG already denotes the balanced data, since the original unbalanced training data set is not used in the analysis. In total, the number of samples is 2.3×10^5 for TRG, 3.9×10^5 for VALAPRB, 2.2×10^6 for VALAPR and 2.8×10^6 for VALSEP.

3.2.3 Neural Network

CNNs especially apply to time-series classification when patterns have to be recognized in longer sequences of data but the location of the occurring patterns is variable. They are therefore suitable classifiers for sensor data like the TRSL from CMLs. The expected advantage of the CNN over the reference method is that it is able to recognize the rainfall specific patterns, rather than just the amount of fluctuations. Like other neural network architectures they consist of a series of layers of neurons (Fig. 3.3). The first layer receives the input data and the last layer serves as an output for a prediction. The hidden layers in between are organized in two functional parts. The first part consists of a series of convolution and pooling layers and is used to extract features from the raw model input. Earlier convolution layers identify simple patterns in the data, which are used to identify more complex patterns in subsequent layers. The second part consists of fully connected layers of neurons and is used to classify the input based on the features extracted by the convolutional part.
Before a CNN can be used as a classifier, it has to be trained on data in a supervised learning process. All layers have a set of trainable parameters, so called weights, which are optimized during the training process according to a learning rule. To be able to monitor the model performance, a test data set is evaluated regularly during the training process. Training is stopped before the model starts to over-fit, i.e. the performance on the test data set either stagnates or drops, while it still rises for the training data.

3.2.3.1 Network architecture

We use a 1D-CNN, which has the same structure as the basic 2D-CNN, with alternating convolutional and pooling layers followed by fully connected layers. The only difference is that the input data of the convolutional layers is one dimensional. The specific architecture and parameterization was optimized experimentally. To give an intuitive description of our CNN, we follow the approach provided in (LeCun et al., 2015, p. 439):
The convolutional part of the CNN consists of four blocks of two convolutional layers followed by a max pooling layer and one block of one convolutional and one average pooling layer (see Fig. 3.3). Convolutional layers extract feature maps by passing local patches (3x1) of input from the preceding layer through a set of filters followed by a rectified linear unit. Each filter creates a different feature map.

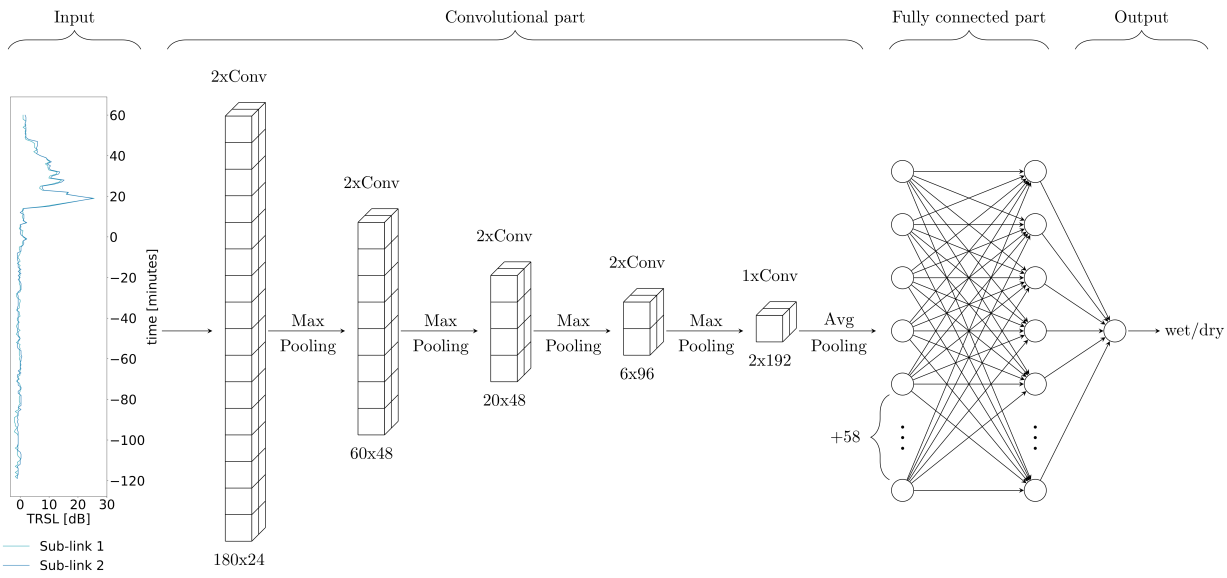

Figure 3.3: Graphical illustration of the CNNs architecture for $k = 120$. The Input shows one sample $\bar{X}_{t,i}$ of data consisting of 180 minutes of TRSL from the two sub-links of one CML. Convolutional and pooling layers reduce the input dimension from 180 to 2, while a total of 192 features are extracted. Numbers below convolutional layers are the layer output dimensions, i.e. input dimension times the number of filters. The size of the local patch in a convolutional layer is 3. Based on the extracted features, the fully connected layers predict a class, which is stored in the output layer.

The pooling layer then combines semantically similar features by taking the maximum (resp. average) within one local patch. This way, the dimension of the input is gradually reduced while, at the same time, the number of extracted features increases.
The fully connected part of the CNN consists of two layers with 64 neurons each and an output layer with one neuron. Its output is a prediction between zero and one, that can be interpreted as the likeliness for the input sample to be wet or dry. To avoid over-fitting to the training data two dropout layers are added, one after each fully connected layer, with a dropout ratio of 0.4 (Srivastava et al., 2014).
We implement the CNN in a Python framework using the Keras (version 2.3.1) backend for Tensorflow (version 2.1.0) (Chollet, 2022; Developers, 2022). For the model architecture, type, number and order of layers has to be chosen. There are several hyper-parameters that can be specified in the model setup. Each layer has a number of hyper-parameters that can be adjusted, e.g. the size of the local patch or the number of filters in a convolutional layer. We optimized all hyper-parameters iteratively by evaluating the performance of several reasonable configurations on the test data set VALAPRB, and by choosing the model with the best performance metrics (see 3.2.4). Depending on the length of the input time-series, which varies with k, the number of convolutional layers is different, i.e. $k < 60$ we omit the last two convolution layers. We trained one model for each value of k and one extra model,

that additionally receives the CML meta-data consisting of the length and the frequency of both channels through parallel fully connected layers and an add-layer before the fully connected part. For k set to 120 minutes the final CNN consists of 20 functional layers with a total of 140,033 trainable parameters. The organization of those layers is shown in the network graph in Fig. 3.3. For all model versions, the detailed model and training specifications, all hyper parameters and the weights of the trained CNN can be retrieved from the code example at `https://github.com/jpolz/cnn_cml_wet-dry_example`.

3.2.3.2 Training setup

CNNs are feed-forward neural networks, which are trained by a supervised learning algorithm (Goodfellow et al., 2016). Batches of samples are passed through the network and the outputs are compared to the reference labels. After each batch a loss function is computed and the weights are updated according to a learning rule. Here, the learning rule is stochastic gradient descent with binary cross-entropy as a loss function and an initial learning rate of 0.008 (Bottou et al., 2018). The training data set TRG consists of 7 batches with 10^4 samples each and the validation data set is VALAPRB. One training epoch is finished when the whole data set is used once. After each epoch the training and validation data sets are evaluated to compute the training and validation loss and the learning rate is decreased slightly.

The training is stopped if the validation loss does not equal or surpass an earlier minimal value for 50 epochs (stopping criterion). Afterwards the model which achieves the best validation Matthews correlation (see MCC below) is selected from all versions, that existed after the individual training epochs (model selection criterion). This model is then used for classification on the validation data sets.

3.2.4 Validation

Our CNN is a probabilistic classifier. The raw model output $\bar{Y}_{t,i}$ is on a continuous scale from 0 to 1 (see Fig. 3.5), representing the estimated likeliness that a sample $\bar{X}_{t,i}$ is wet. A threshold $\tau \in [0, 1]$ is then set to decide whether a sample is wet or not, leading to the prediction rule

$$\tilde{Y}_{t,i} = \begin{cases} 1, & \text{if } \bar{Y}_{t,i} > \tau \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

Classification results, in the form of true positives (TP), false positives (FP), false negatives (FN) and true negatives (TN) are compared to the reference in a confusion matrix, shown in Table 3.1, which is the basis for computing further metrics.

Table 3.1: Confusion matrix

		reference *wet*	**reference** *dry*
prediction	*wet*	True wet (TP): #{ detected wet| reference wet}	False wet (FP): #{ detected wet| reference dry}
	dry	Missed wet (FN): #{ detected dry| reference wet}	True dry (TN): #{ detected dry| reference dry}

The normalized version of the confusion matrix consists of the occurrence rates of TP, FP, FN and TN samples, defined as

$$TPR = \frac{TP}{TP + FN}, \tag{3.3}$$

$$FPR = \frac{FP}{FP + TN}, \tag{3.4}$$

$$FNR = \frac{FN}{TP + FN}, \tag{3.5}$$

and

$$TNR = \frac{TN}{FP + TN}. \tag{3.6}$$

As a first metric for validation we use the accuracy score, defined as

$$ACC = \frac{TP + TN}{\text{total population}} \in [0, 1]. \tag{3.7}$$

It is a measure for the percentage of correct classifications being made. It is dependent on the class balance of the data set. The balance of wet and dry samples in the data set is directly related to the regional and seasonal climatology. Therefore, this metric is not climatologically independent.

The second metric we use is the Matthews correlation coefficient (MCC), also known as ϕ-coefficient, which is a commonly used metric for binary classification (Baldi et al., 2000). It is acknowledging the possibly skewed ratio of the wet and dry periods and is high only if the classifier is performing good on both of those classes. It is defined as

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \in [-1, 1], \tag{3.8}$$

where an MCC of 0 represents random guessing and an MCC of 1 represents a perfect classification. A strong correlation is given at values above 0.25 (Akoglu, 2018). The advantage of the MCC is, that it is a single number which we use to optimize the threshold for the CNN.

The third metric we use is the receiver operating characteristic (ROC), defined by the pair $(FPR, TPR) \in [0, 1] \times [0, 1]$ (Fawcett, 2006). The domain of the ROC is called ROC space. The point (0,1) represents a perfect classifier, while the [(0,0),(1,1)] diagonal represents

random guessing. The ROC is independent of the ratio of wet and dry periods and therefore a climatologically independent measure for the classifier's performance on rain event detection. Each $\tau \in [0,1]$ leads to a ROC resulting in a ROC curve $\gamma \subset [0,1] \times [0,1]$ (e.g. Fig. 3.7). The performance of a classifier for different values of τ is measured by the area

$$AUC = \int_0^1 \gamma d\tau \in [0,1] \tag{3.9}$$

under the ROC curve. Since changing τ directly influences the prediction rule (Eq. 3.2), it can be adjusted causing the model to classify in a conservative (below [(0,1),(1,0)] diagonal in ROC space) or liberal (above diagonal) manner. We can therefore address the trade-off between true wet and true dry predictions as mentioned in the introduction. This way, the AUC becomes a measure of the flexibility of a classifier, i.e. the ability to show good performance with a more conservative or liberal threshold τ. The main purpose of the ROC is that we use it to compare different methods, e.g. different values of k, independent from a fixed threshold, by considering the ROC curve and the AUC.

3.2.5 Reference method

The reference method is a modification of Schleiss and Berne (2010) which is to date the most commonly used method to separate wet and dry periods as reviewed in the introduction. It is based on the following assumption: The standard deviation values of fixed-size windows of TRSL is bounded during dry periods, whereas it exceeds this boundary during wet periods and therefore allows for distinguishing the two classes. This assumption has proven to give good results on our data set, however there are known drawbacks. The method is limited to measuring the amount of signal fluctuations and there are multiple effects that can cause high signal fluctuations during dry periods, e.g. like for CML C in Fig. 3.1. Some of the factors are known, like multi-path propagation, but others are unknown and still need to be investigated.

The method is applied by computing a rolling standard deviation of the TRSL time-series. The normalization step is not necessary for this method. The window length is 60 minutes and the standard deviation value is written to the timestamp in the center of this window. A period $X_{t,i}$ is considered wet if at least one standard deviation value on one or both sub-links exceeds a threshold σ.

We compare two different thresholds σ, which are computed individually for each CML. The first one, denoted σ_{80}, is the 80th percentile of the 60-minute rolling standard deviation of one month for a certain CML multiplied by a scaling factor which is constant for all CMLs. In our case, the threshold is computed for VALAPR in April and VALSEP in September. The scaling factor of 1.12 is adopted from Graf et al. (2020a). The second one, denoted σ_{opt}, is optimized against the reference by maximizing the MCC. We computed it for April 2018 and then reapplied it to September 2018 to test its transferability to future time periods. To derive ROC curves, we applied a scaling factor τ_σ to each of the standard

deviation thresholds. In the following we will refer to σ_{80} and σ_{opt} as both the resulting detection method and the threshold.

3.2.6 Rain rate estimation

In the same way as the rolling standard deviation, the CNN can be used in a rolling window approach, classifying the timestamp t as wet or dry by using the sample with starting timestamp $t - 30$ as model input. With the resulting rain event detection information from either the CNN or the two reference methods, rain rates are estimated in several steps. We use the exact same processing scheme as described in Graf et al. (2020a), which we refer the reader to for all the technical details. This processing includes erratic treatment of CMLs and WAA compensation to derive rain rates with a temporal resolution of one minute. For each detected rain event a constant baseline of the TRSL is calculated from the preceding dry period. The attenuation above this baseline level is attributed to rain but also to WAA. The WAA is compensated depending on the rain rate using a method modified after Leijnse et al. (2008). The remaining specific attenuation k is used to derive the path averaged rain rate R using the $k - R$ relation from Eq. 3.10. The constants a and b are taken from ITU-R (2005).

$$k = aR^b \tag{3.10}$$

For the CMLs used in this study this relation is close to linear, i.e. b is close to one. For a comparison to RADOLAN-RW the one minute rain rates are then aggregated by taking the hourly average.

Only from this analysis data from 45 CMLs (1.1 %) is discarded due to substantially erratic signal levels to be able to follow the same procedure as in Graf et al. (2020a). Additionally, we justify this procedure with the following observation: For the rain event detection we want periods of erratic behavior to be included in both training and validation data, since also CMLs that are not discarded by the erratic treatment can show periods of erratic behavior, such as CML C from Fig. 3.1. Each erratic training and validation sample contributes to the final statistics as one sample and the erratic CMLs do not distort the analysis. This is very different for the rainfall amount, since erratic links are prone to a very high overestimation of the final rain rates even when a low amount of time periods is detected wet. Since erratic CMLs are a small fraction of the available CMLs and they can be detected automatically, we decided to exclude their bias when analyzing the contribution of false positives to absolute rainfall amounts. An example of such a time series can be found in Fig. 3.11.

3.3 Results

During training on TRG, the performance of the CNN was evaluated on VALAPRB after each epoch. The resulting graphs of loss, ACC, TPR and TNR during the training process are shown in Fig. 3.4.

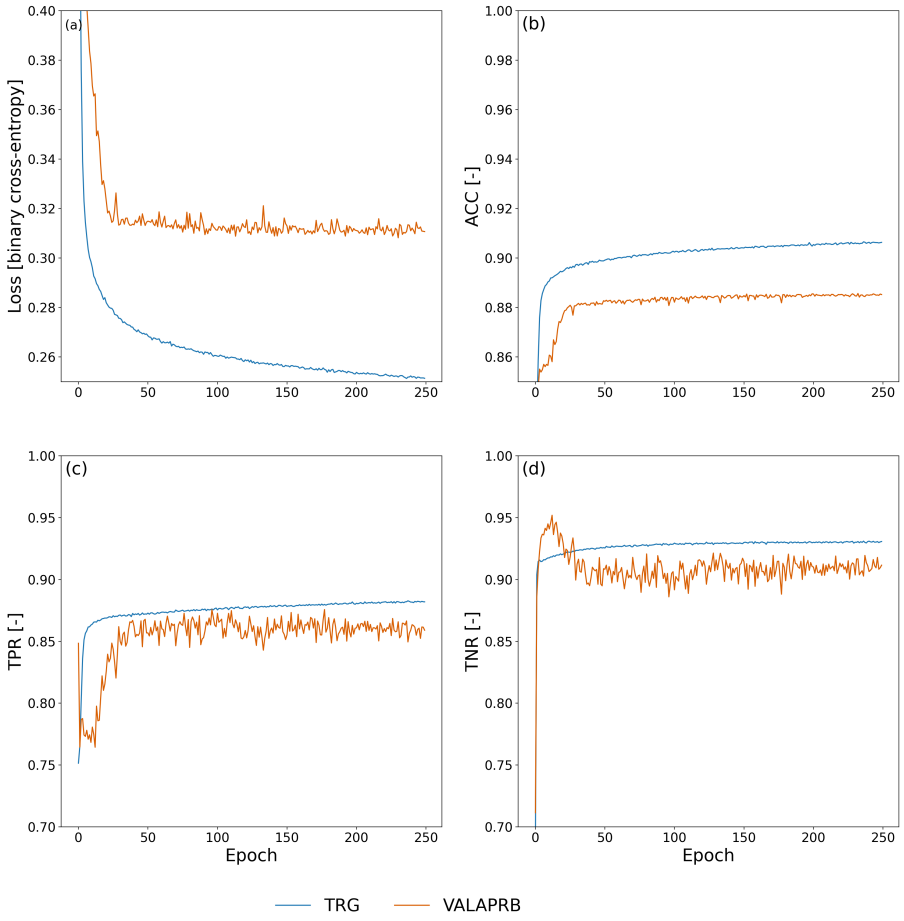

Figure 3.4: Statistics of variables that were monitored during the training process.

For all three variables the performance on TRG and VALAPRB were similar across all epochs with slightly higher performance on TRG. The threshold τ was optimized using VALAPR, by maximizing the MCC, with resulting values of τ shown in Tab. 3.3. The results from that table and the ROC curves in Fig. 3.7 c) show that in general the performance of the CNN is increasing with higher values of k, but the performance gain was insignificant for raising the value higher than 120 minutes or adding meta data as model input. We therefore decided to set $k = 120$ and not to use added meta data for evaluating further results and comparing them to the reference methods.

Table 3.2: Performance metrics of rain event detection methods on VALAPR and VALSEP

	Method	TPR	TNR	ACC	MCC	AUC
VALAPR	CNN	0.74	**0.97**	**0.95**	**0.69**	**0.94**
	σ_{q80}	**0.79**	0.79	0.79	0.38	0.85
	σ_{opt}	0.61	0.95	0.91	0.52	0.83
VALSEP	CNN	0.77	**0.97**	**0.96**	**0.69**	**0.96**
	σ_{q80}	**0.82**	0.78	0.78	0.35	0.87
	σ_{opt}	0.63	0.92	0.90	0.44	0.84

Fig. 3.5 shows the distribution of the CNNs predictions on VALAPRB. The threshold τ is set to 0.82. The final number of training epochs was 248 and the model from epoch 212 was selected (see Fig. 3.4 (a)). On one Nvidia Titan Xp GPU the training time was 30 minutes. Classifying 3904 samples, i.e. a one minute time-step for all CMLs, took 20ms which can be considered extremely fast allowing for a real-time application of the method. For further verification, we repeated the training multiple times with a different randomization (selection of CMLs and balancing) of TRG and VALAPRB but no significant changes in performance could be observed.

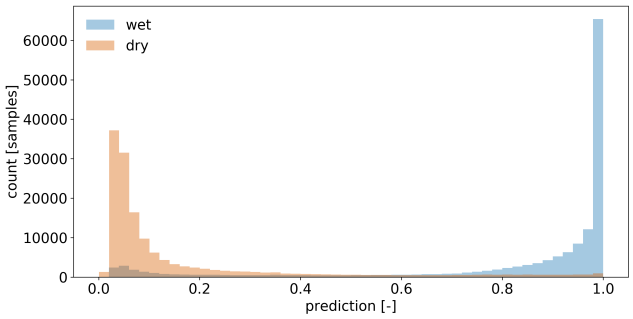

Figure 3.5: Raw CNN predictions on VALAPRB, coloured according to the reference.

We evaluated the performance of the CNN and both reference methods using the unbalanced data sets VALAPR and VALSEP. The complete list of the achieved performance metrics is presented in Table 3.2. Applying the threshold τ to the CNN predictions yielded TPRs of 0.74 (VALAPR) and 0.77 (VALSEP) and TNRs of 0.97 (VALAPR and VALSEP) (see also Fig. 3.10). On average, only 3% of the dry periods were falsely classified as wet and 24% of the wet periods were missed. With a scaling factor $\tau_{\sigma_{q80}}$ of 1.12, σ_{q80} achieved a balanced TPR and TNR with a value of around 0.79 for both rates in April and September. σ_{opt} on the other hand achieved similar TNRs than the CNN but at the cost of lower TPRs.

For both data sets, the CNN's ROC showed a higher TPR for any fixed FPR than the reference methods (see Fig. 3.7). As a consequence, the AUC was largest for the CNN. On VALAPR, σ_{opt} yielded a better ROC than σ_{q80}, but only for low FPR values. On VALSEP σ_{q80} achieved a better ROC than σ_{opt}. The ROC curves of the CNN and σ_{q80} had a very similar convex shape. Compared to the other two curves the ROC curve of σ_{opt} showed a higher asymmetry. The CNN achieved the highest ACC and MCC scores with an average of 0.95 and 0.69 on both data sets. While σ_{opt} has the second highest ACC and MCC scores, the area below the ROC curve is lowest for both data sets.

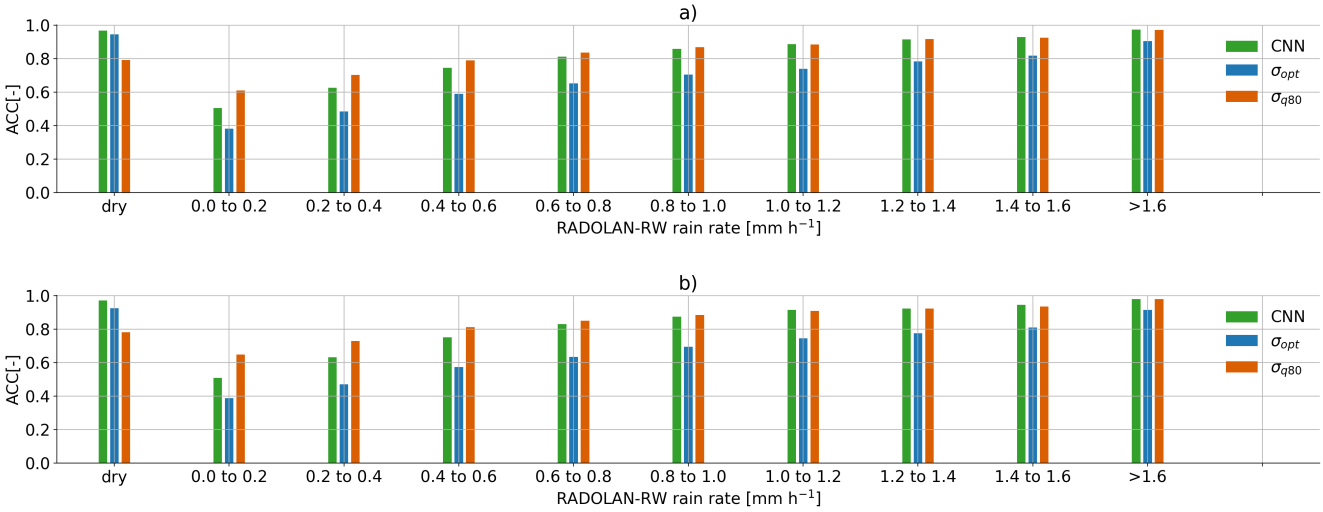

Figure 3.6: Each bar shows the ACC score on samples from a) VALAPR and b) VALSEP, grouped by the reference rain rate. An ACC of 0.5 represents random guessing.

We compare the ACC on detecting samples with a specific RADOLAN-RW rain rate of $x < R_{t,i} < x + 0.1$ in Fig. 3.6. From all rain events where $R_{t,i} \geq 0.6$ mm 90.4% were correctly detected by the CNN. On the other hand around 38.9% of all rain events with $R_{t,i} < 0.6$ mm were missed. All three methods have a lower ACC, the lower the rain rate is. While σ_{q80} shows an ACC for wet periods of different rain intensities, that is very similar to that of the CNN, σ_{opt} misses more small events. On the other hand σ_{q80} is producing more false wet classifications than the CNN or σ_{opt}.
The MCC was computed individually for each CML and each validation data set. Figure 3.8 shows scatter density plots comparing the individual MCC scores of the CNN and σ_{opt}. The CNN's MCC on VALAPR is higher for 95.9% of all CMLs and on VALSEP it is higher for 96.7% of all CMLs. We focus our analysis on hourly rainfall rates from all non-erratic CMLs in September 2018. The resulting rain rates using either the CNN or the σ_{q80} detection scheme are shown in Fig. 3.9. For both methods the distribution of false positive and false negative samples is centered around 0.1 mmh^{-1} and the distribution of true positives is centered around 1 mmh^{-1}.

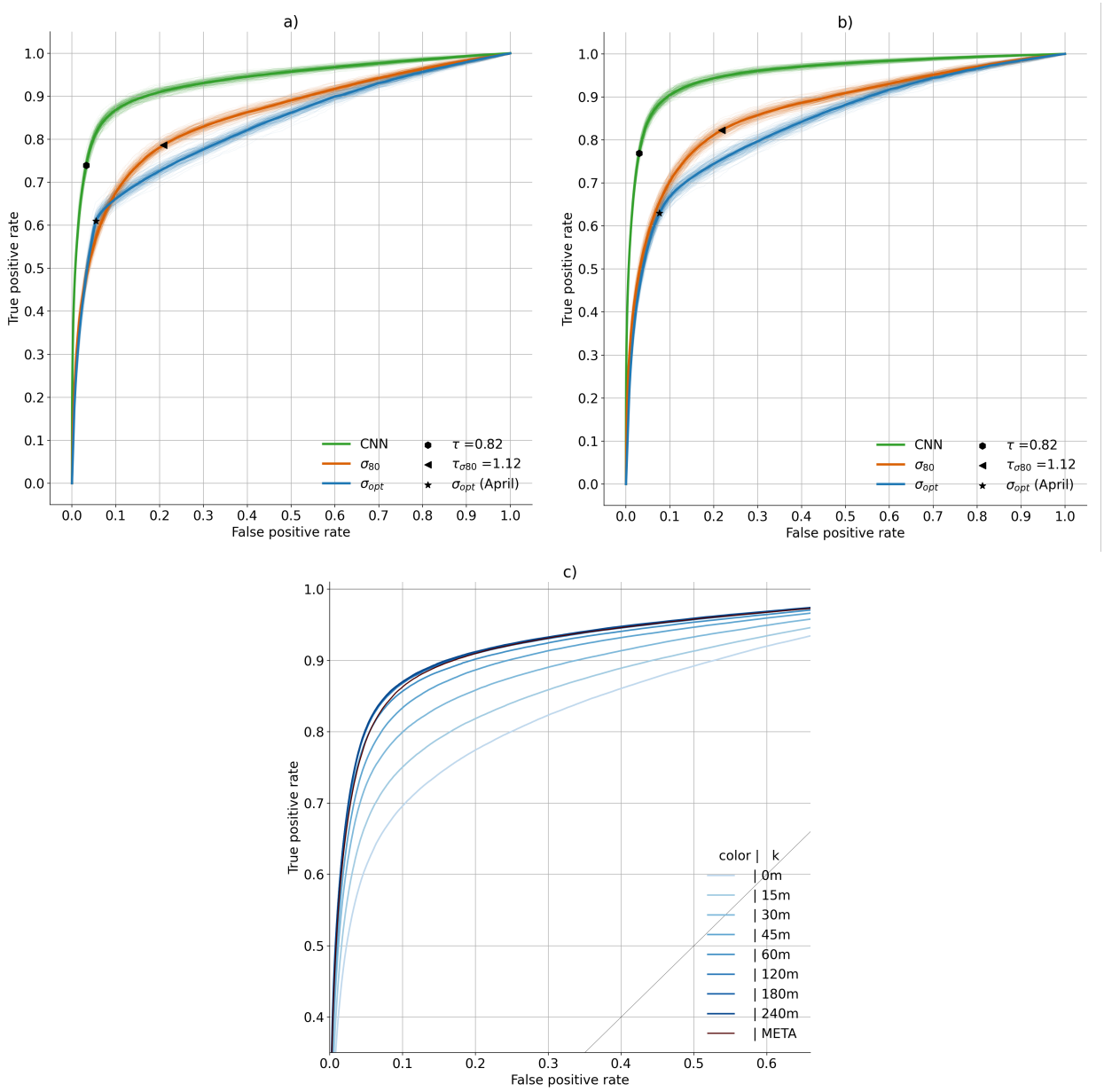

Figure 3.7: Receiver Operating Characteristic curves on VALAPR a) and VALSEP b). Fine lines are generated by 200 random selections (bootstrapping) of 1% of the samples and account for the variability of the model performance during a random short period ($\sim$ eight hours) of data. The performances of the CNN for different values of k and the added meta data are shown in c) and the AUC values are given in Table 3.3

While the percentage of CML derived rainfall estimated during false positive events is 29.9% for σ_{q80}, it is significantly less for the CNN (see Fig. 3.9 d) and f). This constitutes a reduction of 51% of falsely estimated rainfall for the month of September 2018. At the same time the amount of missed rainfall is reduced by 27.5%. The amount of rainfall in the true positive category could therefore be raised by 4.7%. The Pearson correlation for the hourly rainfall estimates between radar and CMLs is 0.83 using σ_{q80} and 0.84 using the CNN.

3.4 Discussion

3.4.1 Performance

We evaluate the performance of the CNN to detect rain events by two means. First, we compare it to the performance of a reference method. Second, we estimate if the model is performing in a near optimal state or if we expect that a higher performance could be achieved. The comparison to the results of previous studies, e.g. Overeem et al. (2016a), is difficult since the overall performance is depending on the distribution of the intensity of rain events (see Fig. 3.6) and since there is a large variability of performance between the CMLs (see Fig. 3.8).

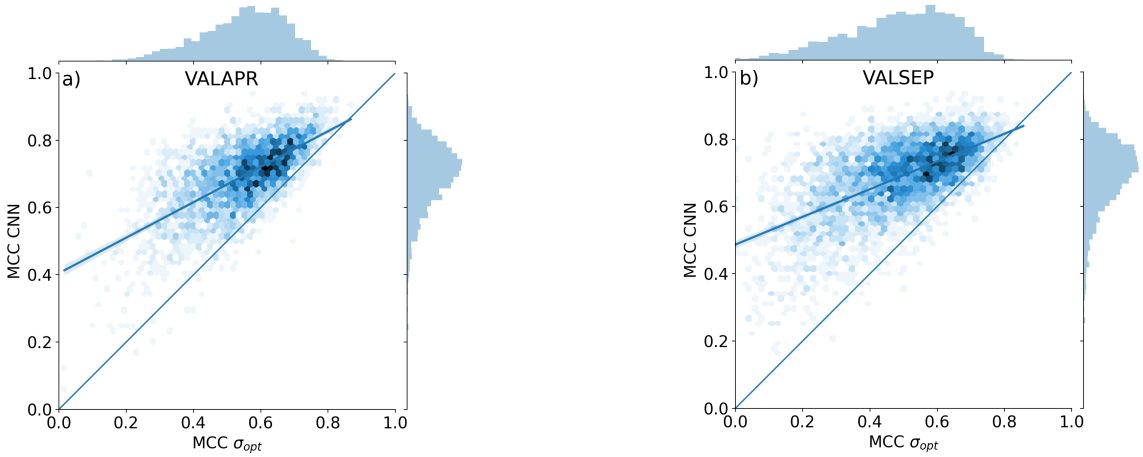

Figure 3.8: Scatter density plots of the MCC achieved by the CNN and σ_{opt} on data from individual CMLs. Both methods are MCC optimized for the unbalanced data from VALAPR, while the CNN keeps the optimized performance in September, the performance of σ_{opt} drops.

Since the results on both validation data sets are very similar (see Table 3.2) we further focus on VALSEP, which was not used to optimize the model hyper-parameters. With an ACC of 0.95 and an MCC of 0.69 the correlation between the CNN predictions and the reference data set RADOLAN-RW can be considered as very high. A TPR of 0.74 might not appear very good at first sight, but considering that the detection accuracy for samples with a rain rate of smaller than 0.6 mmh^{-1} is only 0.61, we actually achieve an accuracy of over 0.9 for all rain rates higher than 0.6 mmh^{-1}.
The CNN and the reference method σ_{opt} have a similar ACC value. At the same time the CNN's MCC is higher, despite the fact that σ_{opt} is MCC optimized for each CML. The high ACC of σ_{opt} is due to the high TNR and the fact that 95% of all samples are negative (dry). At a similar ACC and TNR we could increase the TPR, or rain event detection rate, by 0.13. This constitutes a major improvement by the CNN.

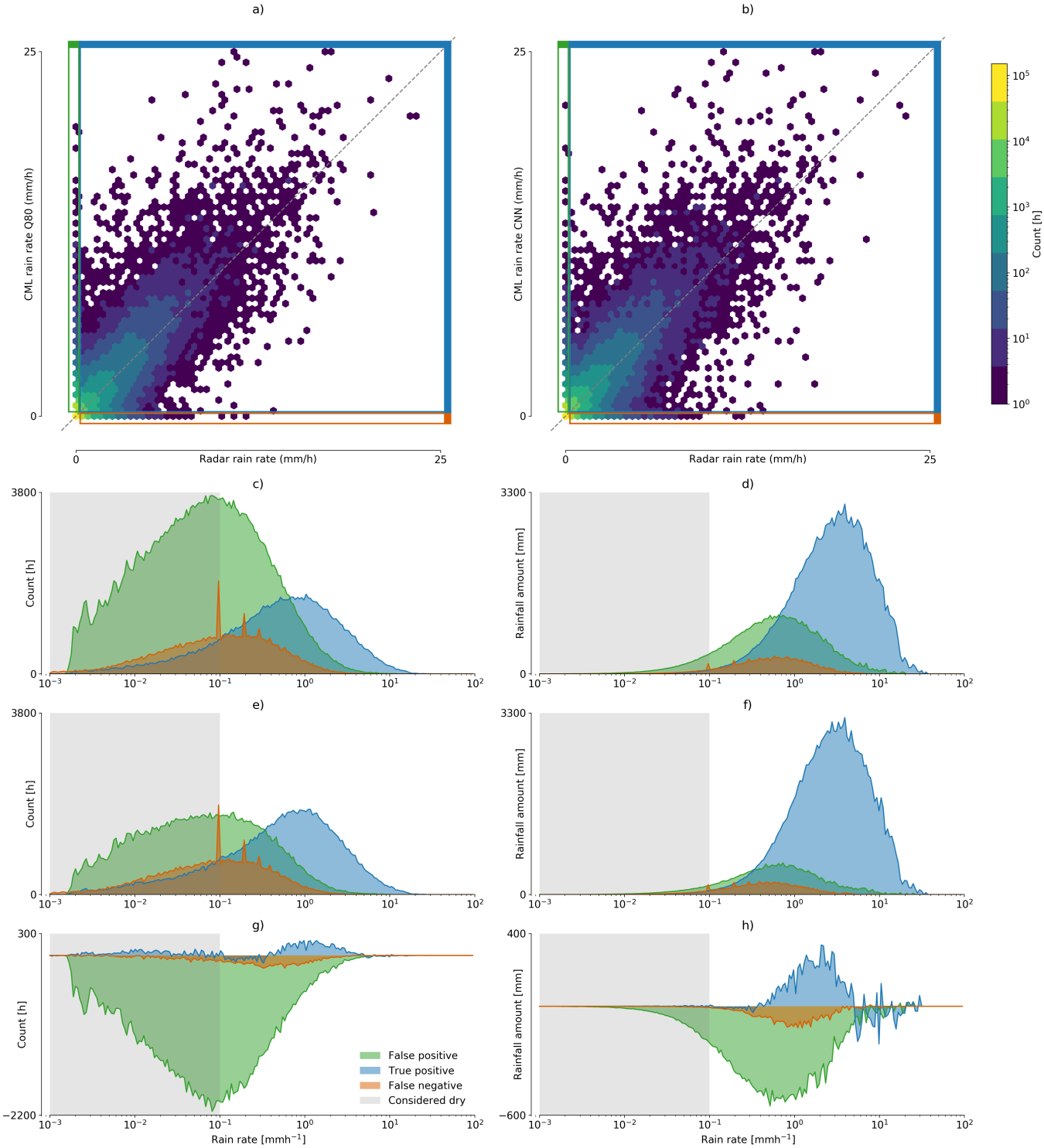

Figure 3.9: Scatter density comparison between hourly CML and radar rain rate estimates derived from a) σ_{q80} and b) the CNN. On the left hand side the amount of FP, TP and FN hours with a specific rain rate are compared for c) σ_{q80}, e) the CNN and g) their difference). On the right hand side the amount of rainfall these hours contribute are shown for d) σ_{q80}, f) the CNN and h) their difference. The rain rates for false positives and true positives are estimated by the CML, while the rain rates for false negatives are taken from the reference.

As shown in Fig. 3.8 the improvement is higher for CMLs with lower MCC, making the whole CML data set more balanced in performance and therefore more trustworthy for quantitative precipitation estimation. The CNNs distribution of MCC values of individual CMLs is the same in April and September, while performance drops for σ_{opt}. The CNN's improvement in ACC and MCC over σ_{q80} was even higher with 0.17 and 0.32. While the TPR of σ_{q80} is slightly higher than the TPR of the CNN, the TNR is much lower for σ_{q80}. Thus the CNN shows substantial improvement in correctly classifying dry periods.
While the RSTD method can be set up to either have a high TPR (σ_{q80}) or a high TNR (σ_{opt}), the ROC curves show that CNN achieves both rates at the same time. Thus, the CNN shows a better overall performance than the reference methods and therefore improves on the trade-off as mentioned above. This observation is illustrated by the example in Fig. 3.2, which shows a very noisy CML time-series that produces a high amount of false positives for the reference method, while the CNN does not attribute these fluctuations to rainfall. All three methods have limitations to detect events with rain rates smaller than 0.3mm. This is likely due to the detection limit of CMLs in our data set which is in the same range. The detection limit depends on frequency, length and signal quantization of a CML. For example, at a frequency of <20 GHz and at a length of <10 km a path averaged rain rate of 1 mm h^{-1} creates a maximum of 1 dB of attenuation (Chwala and Kunstmann, 2019, Fig. 7). In some cases the quantization (0.3dB for RSL and 1dB for TSL) might therefore not allow for a detectable signal.
Differences in the performance on VALAPR and VALSEP can be traced back to a different distribution of occurring rain rates. While in April 35.5% of all events are in the critical range from 0.1mm to 0.3mm, there are only 32% in September. In both data sets the performance on higher rain rates (> 1.6 mm) and dry periods is almost identical. Therefore the loss of performance in April is due to the slightly worse performance of the CNN on smaller rain rates which occur more often in VALAPR than in VALSEP.
It should not be expected that the rain events detected through CMLs and the events detected by the radar coincide completely. Both methods produce artifacts that are mistaken as rainfall, or they miss events due to their detection limits. From all false classifications that the CNN makes on VALSEP there are 50% with a raw model output between 0.2 and 0.8. Here the CNN does not give a certain prediction. This is due to very similar signal patterns in noisy dry periods and small rain rates. The other 50% of those samples are, according to the CNN, very likely to belong to the falsely predicted class. Despite this being an issue for many CMLs about 10% have a ROC of $(> 0.97, < 0.1)$ and correlate very well with the RADOLAN reference. Therefore, we expect that less errors could be made when training with a perfect reference data set, but there would still be errors due to artifacts or insensitivity in CML measurements.
Despite those errors, which occur mostly for small rain rates, the correlation of wet and dry periods between RADOLAN-RW and our CML data set is very high. The performance boost in rain event detection gained through the CNN is very promising for future applications in quantitative precipitation estimation with CMLs.

3.4.2 Robustness

The CNNs ability to generalize to previously unknown CMLs is very high. As seen in the training results the learning curves for both training and validation show a similar dynamic (see Fig. 3.4). As expected the training data showed better performance, but the validation was close at all epochs.
Only 20% of all available CMLs were used for training. The remaining 80% were only used to prevent the model from over-fitting to the training data, to choose the model architecture and to optimize the single parameter τ. Thus no information about the validation data was given directly to the model. The resulting model architecture and hyper-parameters are not specific enough to store this information. The high performance in ACC, MCC and ROC on data set VALAPR, together with the learning curves in Fig. 3.4), therefore prove that the CNN was able to recognize the attenuation pattern in the signal levels of a large number of previously unknown CMLs.
The stability of the CNNs performance for future time periods is analyzed using the results on VALSEP. While the training was done with TRG including the period of May to August 2018, the performance in September was similar. Compared to the results on VALAPR the CNN shows even higher performance on VALSEP, which can be explained by the lower percentage of samples with small rain rates in September, which are challenging to classify (see Fig. 3.6 a)). When we compare the CNNs accuracy per rain rate between VALAPR and VALSEP, we see that there are no major differences in the individual scores. Therefore the method can be considered as very stable throughout the analyzed time period, while differences in overall performance mostly stem from different distributions of the occurring rain rates. The reference method σ_{opt}, which was optimized in April, loses performance in September, where it is outperformed by the adaptive method σ_{q80}. The bootstrapping in Fig. 3.7 shows that all three methods perform almost equally well on small random subsets of the validation data. The CNN shows the lowest variability.
As a measure for the flexibility of a classifier we adopted the ROC analysis in Sect. 3.2.4. A model is called flexible if it has a high area below its ROC curve and if the curve is axis-symmetric with respect to the [(0,1),(1,0)] diagonal of the ROC space. As observed both the CNN and σ_{q80} show a symmetrical ROC curve. Therefore they perform almost equally well with a liberal or conservative threshold with a slight tendency to the conservative side. On the other hand σ_{opt} shows a skewed performance, with a strong tendency to the conservative side. The area AUC below the ROC curve was highest for the CNN, making it the most flexible classifier. We can adjust τ for a ROC of either $(0.03, 0.7)$ or $(0.3, 0.94)$ and a smooth, concave transition in between (see Fig. 3.7).
We conclude that within the analyzed period the CNN shows a temporally stable performance, with a good generalization to previously unknown CMLs. The σ_{opt} method performs well only if it is re-calibrated for different months and to individual CMLs, while σ_{q80} is by definition an adaptive method. Even with re-calibration or adaption, the reference methods are outperformed by the CNN.

3.4.3 Impact of the detection scheme on the derived rainfall amounts

The difference between the scatter density plots in Fig. 3.9 a) and b) seems to be quite low at first sight. What this representation of the data is not stressing enough is the amount of rainfall generated by false positives. But they are an issue that is clearly visible from Fig. 3.9 c)-h). Considering that the amount of rainfall estimated during time periods falsely classified as wet can be reduced by 51.0% and that the amount of rainfall from missed events can be reduced by 27.4%, the CNN shows a major improvement over the reference method. The 4.1% of additional rainfall in the correctly classified wet periods stem from time periods that were originally harder to classify, i.e. from small rain events, and it should be expected, that the correlation between CML and radar rainfall drops. Instead, the Pearson correlation coefficient increased slightly showing that the quality of the estimated hourly rainfall could be improved. We omitted the same analysis for a comparison of the CNN and σ_{opt} for which, based on the ROC values in Fig 3.7, we anticipate a similar result, but with a higher pronunciation of missed rain events instead of the strong impact of false positives.
Overall, we could observe that the improvement in rain event detection has a considerable effect on the amount of over- or under estimation through falsely detected or missed rain events. The improvement on the trade-off between false positives and false negatives directly translates to the impact of their respective rainfall amounts. This is shown by the false positive and false negative distributions in Fig. 3.9 c)-f) which are centered around the same value, but are different in their amount depending on the used detection method.

3.5 Conclusion

In this study, we explore the performance and robustness of 1D-CNNs for rain event detection in CML attenuation time-series using a large and diverse data set, acquired from 3904 CMLs distributed over entire Germany. We prove that, compared to a reference method, we can minimize the trade-off between false wet and missed wet predictions. While the reference method needs to be adjusted for different months of the analyzed period to provide optimal results, the trained CNN generalizes very well to CMLs and time periods not included in the training data. On average, 76% of all wet and 97% of all dry periods were detected by the CNN. For rain rates higher than 0.6 mmh^{-1} more than 90% were correctly detected. This underlines the strong agreement between rain events that can be detected in the CML time-series and rain events in the RADOLAN-RW data set.
In future work, we plan to investigate the potential of using reference data with higher temporal resolution to improve the temporal localization of the rain events. Data with higher temporal resolution will, however, magnify the uncertainties that arise due to the different spatial and temporal coverage of the different rainfall observation techniques. In order to address these uncertainties, it will be important to further explore the relationship

between weather radar and CML derived rainfall products. In the study presented here, we focused on the optimization of rain event detection as an isolated processing step, which provides the basis for a successful rain rate estimation. All subsequent processing steps, including WAA correction, k-R relation and spatial interpolation, have an effect on the CML derived rain rate, that can also lead to over or under-estimation. While 29.9% of the estimated rainfall through the reference method can be attributed to false positive classifications, the CNN reduces this amount by up to 51% and, at the same time, improves on true positive and false negatives. We anticipate, that this improvement will lead to new insights into other effects that may disturb the quality of this opportunistic sensing approach.

Our study shows that using data driven methods like CNNs in combination with the good coverage of the highly developed weather radar network in Germany can lead to robust CML data processing. We anticipate that this robustness enhances the chance that we can transfer processing methods to data from other CML networks, particularly in developing countries like Burkina Faso, where rainfall information is still scarce despite its high importance to the local population (Gosset et al., 2016).

3.6 Code availability

Interactive code to build the CNN and an example evaluation using the trained CNN are available at `https://github.com/jpolz/cnn_cml_wet-dry_example`. CML data was provided by Ericsson Germany and is not publicly available in its full extent. RADOLAN-RW is publicly available through the Climate Data Center of the German Weather Service (DWD) `https://opendata.dwd.de/climate_environment/CDC/grids_germany/hourly/radolan/`. We include a small example data set with modified CML locations, the trained model weights and the pre-processed RADOLAN-RW reference data together with the interactive code at `https://github.com/jpolz/cnn_cml_wet-dry_example`.

3.7 Acknowledgements

We thank Ericsson, especially Reinhard Gerigk, Michael Wahl and Declan Forde for their support and cooperation in the acquisition of the CML data. This work was funded by the German research foundation within the RealPEP research group. Furthermore, we like to thank the German Research Foundation for funding the project IMAP, the Helmholtz Association of German Research Centres for funding the project Digital Earth and the Bundesministerium für Bildung und Forschung for funding the project HoWa-innovativ. Special thanks are given to Bumsuk Seo for his valuable advice and for providing the Titan Xp GPU used for this research, which was donated by the NVIDIA Corporation.

3.8 Appendix to Chapter 3

Table 3.3: Number of training epochs, MCC optimized threshold and resulting metrics for different values of k, evaluated on VALAPR.

Method	k	Training epochs	Threshold τ	TPR	TNR	ACC	MCC	AUC
CNN	0	269	0.77	0.53	0.97	0.93	0.55	0.86
	15	158	0.78	0.59	0.97	0.94	0.60	0.88
	30	274	0.79	0.64	0.97	0.94	0.64	0.91
	45	271	0.79	0.67	0.97	0.94	0.66	0.92
	60	128	0.84	0.71	0.97	0.95	0.68	0.93
	120	212	0.85	0.72	0.97	0.95	0.69	0.94
	180	211	0.86	0.72	0.97	0.95	0.69	0.94
	240	170	0.84	0.73	0.97	0.95	0.69	0.94
CNN+Meta	180	321	0.79	0.70	0.97	0.95	0.68	0.93
σ_{q80}	-	-	-	0.79	0.79	0.79	0.38	0.85
σ_{opt}	-	-	-	0.61	0.95	0.91	0.51	0.83

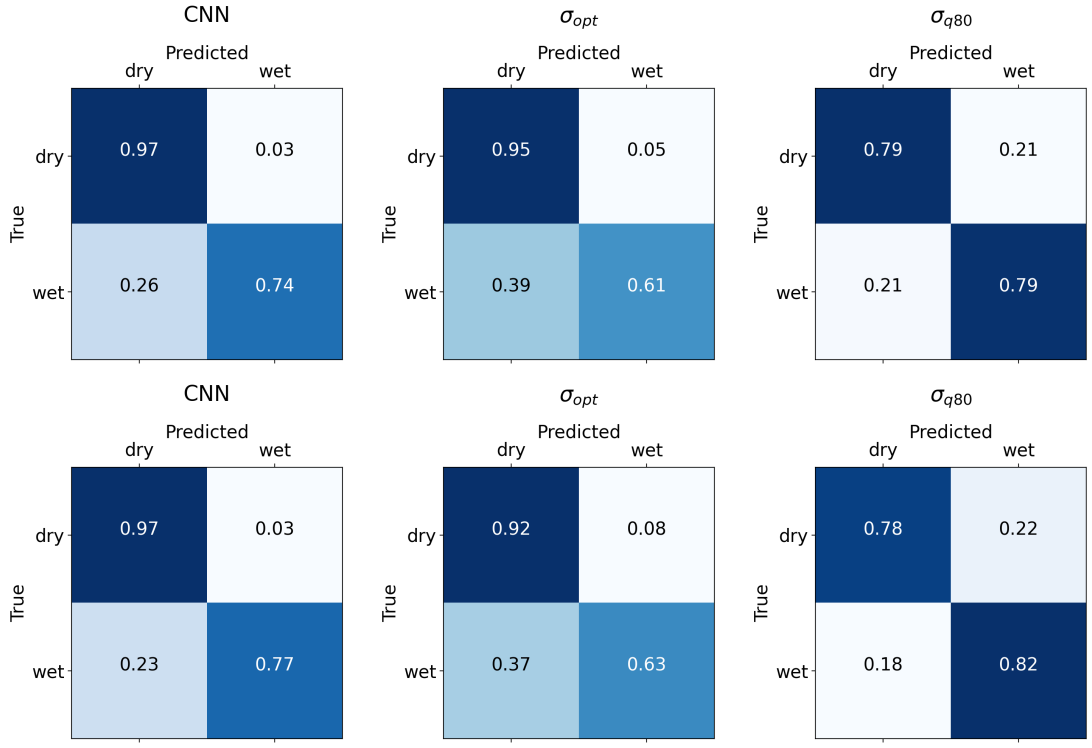

Figure 3.10: Normalized confusion matrices of VALAPR (top) and VALSEP (bottom).

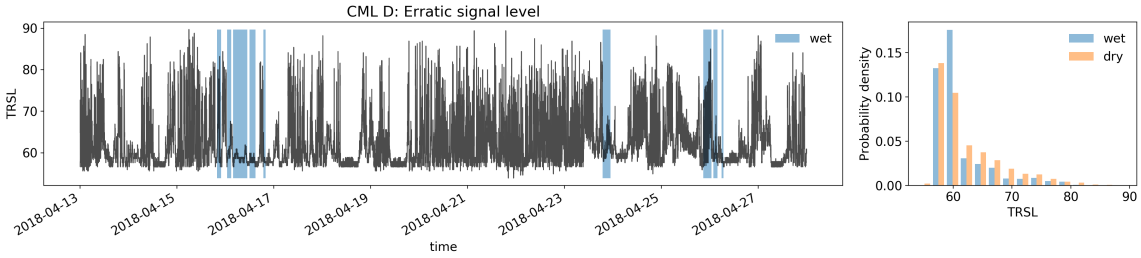

Figure 3.11: Time series of a CML that is considered as erratic and is removed by the simple filter for erratic CML data introduced in Graf et al. (2020a). There are no time periods, where a reasonable rainfall estimation would be possible.

Chapter 4

Missing rainfall extremes in commercial microwave link data due to total loss of signal (Polz et al., 2023b)

Polz, J., Graf, M., Chwala, C.: Missing rainfall extremes in commercial microwave link data due to total loss of signal, Earth and Space Science, 10, e2022EA002456. https://doi.org/10.1029/2022EA002456 , 2023

Abstract

An important aspect of rainfall estimation is to accurately capture extreme events. Commercial microwave links (CMLs) can complement weather radar and rain gauge data by estimating path-averaged rainfall intensities near ground. Our aim with this paper was to investigate attenuation induced total loss of signal (blackout) in the CML data. This effect can occur during heavy rain events and leads to missing extreme values. We analyzed three years of attenuation data from 4000 CMLs in Germany and compared it to a weather radar derived attenuation climatology covering 20 years. We observed that the average CML experiences 8.5 times more blackouts than we would have expected from the radar derived climatology. Blackouts did occur more often for longer CMLs (e.g. > 10 km) despite their increased dynamic range. Therefore, both the hydrometeorological community and network providers can consider our analysis to develop mitigation measures.

Plain Language Summary

Commercial microwave links (CMLs) are used to transmit information between towers of cellphone networks. If there is rainfall along the transmission path, the signal level is attenuated. By comparing the transmitted and received signal levels, the average rainfall intensity along the path can be estimated. If the attenuation is too strong, no signal is received, no information can be transmitted and no rainfall estimate is available. This

is unfavorable both for network stability and rainfall estimation. In this study, we investigated the frequency of such blackouts in Germany. How many blackouts per year are observed in a three year CML dataset covering around 4000 link paths and how many are expected from 20 years of weather radar data? We observed that the average CML experiences 8.5 times more blackouts than we would have expected from the radar derived climatology. Blackouts did occur more often for long CMLs, which was an unexpected finding. While only one percent of the annual rainfall amount is missed during blackouts, the probability that a blackout occurs was very high for high rain rates. Both, the hydrometeorological community and network providers can consider our analysis to develop mitigation measures.

4.1 Introduction

Microwave radiation is attenuated by hydrometeors through scattering and absorption processes. For raindrops an advantageous relationship between specific attenuation k in dB/km and rainfall rate R in mm h^{-1} exists. This power law known as the k-R relation is close to linear at frequencies between 20 and 35 GHz (Chwala and Kunstmann, 2019). Commercial microwave links (CMLs) use frequencies from 7 to 80 GHz and thus can be used to derive path averaged rainfall intensities by comparing transmitted and received signal levels (TSL and RSL) (Uijlenhoet et al., 2018). In theory, the k-R relation is valid for arbitrary rainfall intensities occurring in the underlying drop size distribution simulations. In practice, the measurement of high attenuation values at a given transmitted signal level has an upper bound when the signal cannot be distinguished from the receiver's background noise.

CML rainfall estimates were derived for many countries around the globe, e.g. the Netherlands (Overeem et al., 2016b), Sri Lanka (Overeem et al., 2021), Burkina Faso (Doumounia et al., 2014) and Germany (Graf et al., 2020a). CML-derived rainfall information can be used for applications like streamflow prediction, urban drainage modeling, agricultural purposes and rainfall nowcasting (Fencl et al., 2013; Brauer et al., 2016; Stransky et al., 2018; Imhoff et al., 2020). Especially for flash flood prediction, precise precipitation maxima are of great importance (Cristiano et al., 2017). While rainfall estimates from weather radars are known to underestimate high intensities (Schleiss et al., 2020), rain gauges lack spatial representativeness (Sevruk, 2006). CMLs can fill this information gap by estimating path averaged intensities at path lengths of a few kilometres.

Recent studies on the quality of CML rainfall estimates suggest a good agreement with radar and rain gauge estimates (Graf et al., 2021a; Overeem et al., 2021). However, missing periods in the signal level time series might be excluded e.g. when comparing CML time series against a path-averaged radar reference or rain gauges. Such periods can occur due to hardware failure, maintenance or outages in the data acquisition. Additionally, network providers usually design the hardware in such a way that transmission outages due to high attenuation (blackouts) are allowed to occur for a certain amount of time

per year. The International Telecommunication Union (ITU) recommends a minimum availability of 99.99% which would allow up to 52 minutes of total loss of signal per year (ITUßR, 2017).
Rainfall is the prevalent reason for CML signal attenuation. Hence, the amount of missing data is in a close relationship with the local rainfall climatology. Because of blackouts rainfall estimates from CMLs miss peak intensities, an error which propagates to further applications. Figure 4.1 shows examples of such blackouts in CML attenuation time series and the rainfall intensity according to a weather radar reference. To date, it is unclear to what extent rain events are missed due to blackouts.
Our aim is to answer two questions related to CML blackouts using a country-wide CML network in Germany. The first question is how many blackouts each CML is experiencing in practice and how this affects rainfall estimates. The second question is how much blackout time is expected considering 20 years of high-resolution weather radar rainfall climatology and how this expectation compares to the results derived from the CML data.

4.2 Data and Methods

Our analysis was based on observed blackouts within CML data collected in Germany and a comparison to the expected frequency derived from weather radar climatology (Sec. 4.2.1). We detected gaps in CML data that are assumed to be caused by attenuation (Sec. 4.2.2) and derived path integrated attenuation values from path averaged weather radar rain rates (Sec. 4.2.3). Note that all calculations were repeated for each CML individually.

4.2.1 Data

CML data has been collected in cooperation with Ericsson Germany. The data acquisition system described by (Chwala et al., 2016) has been used to record three years of instantaneously measured RSL and TSL of 3904 CMLs distributed over Germany (2018 to 2020). The temporal resolution is one minute and the power resolution is 0.3 or 0.4 dBm for RSL and 1 dBm for TSL. 25% of the CMLs have a constant TSL value (e.g. Figure 4.1b). The other 75% use an automatic transmit power control (ATPC), which can increase TSL if RSL decreases due to attenuation (e.g. Figure 4.1a,c,d). The CML path lengths range from 0.1 to 30 kilometers with frequencies from 7 to 40 GHz as shown in Figure 4.2d). In the context of rainfall estimation, CMLs are characterized by two main features. First, the signal level sensitivity to rainfall, see e.g. Fig. 7 in Chwala and Kunstmann (2019), which depends on the frequency, polarization and path length. Second, the dynamic range of the signal level D_{range}, i.e. the difference between clear sky attenuation and maximum measurable attenuation. The communication along a CML requires (de-)modulation of information onto the carrier frequency. Different modulation schemes are used. In case of increased attenuation along their path, the CMLs switch to more simple modulation.

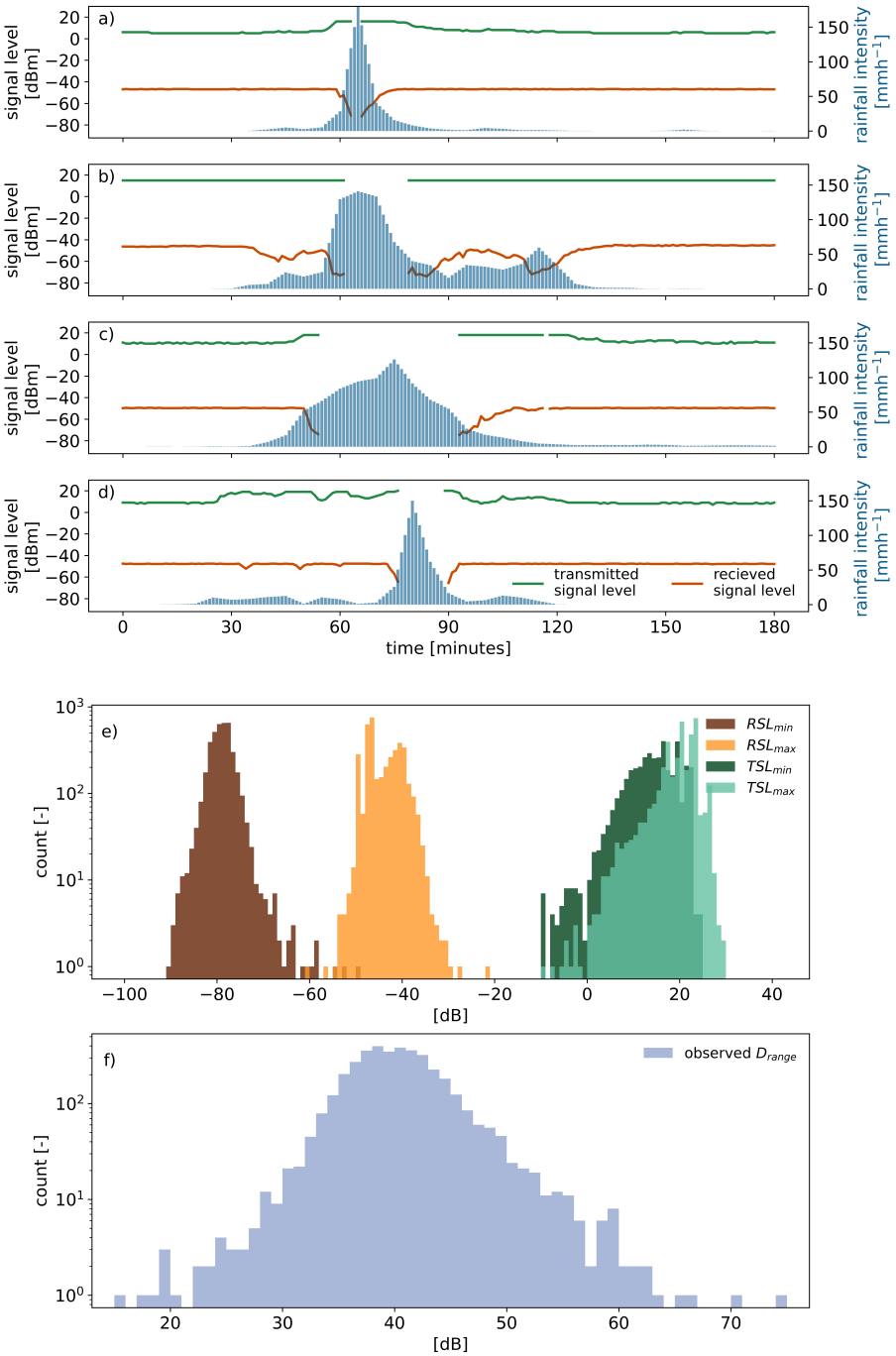

Figure 4.1: a)-d) show TSL and RSL time series during blackout gaps from four CMLs. Rainfall intensities are derived from RADKLIM-YW along the CML's paths. e) gives the minimal and maximal TSL and RSL values of all 3904 CMLs for the analysed period of three years. f) shows the distribution of the dynamic range directly calculated from CML signal levels with Equation 4.1.

This decreases the usable bandwidth but increases the robustness against transmission errors stemming from the lower signal-to-noise ratio during attenuation events (Bao et al., 2015). If the RSL is too low, i.e. close to the noise floor of the receiver, the error rate

for demodulation becomes too large and communication is cut off. Datasheets of CML hardware (e.g. from Ericsson (2012)) guarantee a certain error rate at defined low RSL values rather than a fixed lower RSL limit where this cutoff happens. Therefore, we need to estimate the empirical D_{range} of each CML as

$$D_{range} = TSL_{max} - RSL_{min} - TSL_{min} + RSL_{max}. \tag{4.1}$$

We removed TSL and RSL outliers outside the intervals [-20 dBm, 50 dBm] and [-99 dBm, 0 dBm] respectively. TSL_{max} and RSL_{min} were the highest (lowest) values which occurred during heavy attenuation events representing the strongest observed attenuation event for each single CML (see Fig 4.1a)- d) as examples). Contrarily, we can assume that TSL_{min} and RSL_{max} are occurring frequently during clear sky conditions. To account for individual outliers we removed values for TSL_{min} and RSL_{max} when they occurred less often than approximately one hour in total during the three years, i.e. using the 99.995% quantile. Without this filter D_{range} would be overestimated for about half of all CMLs because there are individual rarely occurring high RSL_{max} or low TSL_{min} values. With the potentially abrupt onset of heavy rainfall causing a complete loss of signal, RSL_{min} may have been undersampled by the 1-minute instantaneous data sampling. Therefore, the derived D_{range} can be assumed to be the minimal dynamic range a CML has.
As reference we used RADKLIM-YW (Winterrath et al., 2018) from the German Meteorological Service (DWD) which we linearly interpolated from a 5- to a 1-minute resolution to match the CML resolution. RADKLIM-YW is a gauge-adjusted, climatologically corrected radar product with a temporal resolution of five minutes and a spatial resolution of 1 km. The underlying radar precipitation scans have been carried out every five minutes. Therefore, the radar rainfall intensities can be considered to be instantaneous measurements without temporal averaging. The product is composed of 17 weather radars and adjusted by more than 1000 rain gauges with additive and multiplicative corrections. The climatological correction accounts for range-dependent underestimation and radar spokes caused by beam blockage, among others. RADKLIM-YW was considered the best and highest resolved rainfall reference for this analysis and was available from 2001 to 2020. Following Graf et al. (2020a) we derived the path averaged rain rate R for each CML as the sum of radar grid cell rainfall intensities r_i weighted by their lengths of intersection l_i with a given CML path of total length L as described by Eq. 4.2.

$$R = \frac{1}{L} \sum_i r_i l_i \tag{4.2}$$

To be able to investigate a potential temperature dependence of observed blackouts we used the 2 meter temperature from the ERA5 analysis dataset provided by the European Centre for Medium-Range Weather Forecasts Muñoz-Sabater et al. (2021). The temporal resolution is instantaneous at a one hour frequency and the spatial resolution is 9 kilometers. Similar to Eq. 4.2 an average along the CML path was computed by a weighted sum

of ERA5-land grid cells intersected by the CML path.

4.2.2 Detecting blackouts in CML data

Gaps in CML signal level time series can have various causes. In this analysis we were interested in gaps caused by strong attenuation during heavy rainfall and therefore excluded periods which could be attributed to one of the following causes. Gaps longer than 24 hours were assumed not to be caused by heavy rain events. When more than 400 CMLs exhibited a gap at the same time, we excluded this time step. The reasoning behind this value is that we assumed a partial or complete outage of our data acquisition system which polls the data in several batches of 400 to 500 CMLs every minute. Gaps occurring during a period where a seven-day rolling mean of the RSL was below -60 dBm were removed. This was done, because we can assume that there is a long-term transmission disturbance, i.e. partial beam blockage due to a growing tree or due to ice cover on the antenna during consecutive winter days with temperatures below freezing point, since none of the CMLs in our dataset has a 3-year median RSL below -60 dBm. That is, all our CMLs have their long-term baseline RSL level during clear sky conditions above -60 dBm. Around 0.2% of all RSL values are removed from the analysis by filtering data acquisition gaps and long term transmission disturbances.

The actual detection of blackout gaps is done with the remaining CML data based on the following rule. A gap is defined as a blackout gap if either the last valid RSL before, the first valid RSL after this gap, or both values were below -65 dBm. Examples of such automatically detected gaps are shown in Figure 4.1a-d). The median RSL levels within our dataset are well above -65 dBm. Therefore, we chose this threshold to select only events with attenuation typical of heavy rain events. The thresholds we chose for filtering the data and detecting the blackout gaps proved to be robust when applied to our dataset where the CML hardware and data acquisition are homogeneous (Chwala et al., 2016). However, they might need adjustment if our method is applied for CML datasets with different characteristics.

We grouped observed blackouts into reference rainfall intensity bins and computed the average amount of observed blackout minutes n_{obs} per year for each CML. In addition, n_{obs} was normalized by applying the factor

$$f_{avail} = \frac{\#\{\ minutes\ in\ observation\ period\}}{\#\{\ minutes\ with\ valid\ observations\}} \tag{4.3}$$

for each CML to account for missing time steps in the CML data.

4.2.3 Deriving a blackout climatology from radar data

In theory, a blackout due to heavy rainfall should be expected whenever the path integrated attenuation (PIA) exceeds the CML's dynamic range D_{range}. We estimated a blackout climatology using 20 years of instantaneous radar measurements. A radar derived PIA was calculated by individually applying the k-R relation to the rain rate r_i of the i-th radar

grid cell intersected by a CML path. This procedure was chosen over applying the k-R relation to the path averaged rain rate to minimize errors due to the spatial variability of rainfall along the path as explored by Berne and Uijlenhoet (2007). Hence, we calculated

$$PIA = \frac{1}{L} \sum_i a r_i^b l_i + w_{aa} \tag{4.4}$$

using coefficients a and b, derived from the ITU recommendation ITU-R (2005), which depend on the CMLs frequency and polarization. The intersection length of CML path and radar grid cell i is denoted l_i. Additionally, a constant w_{aa} =3 dB accounting for the wet antenna attenuation (WAA) caused by rain drops on the cover of the CML antennas was added (van Leth et al., 2018). We chose a value similar to Leijnse et al. (2008); Schleiss et al. (2013). We assumed a high constant value which is reasonable for peak rainfall intensities. Whenever PIA was larger than D_{range}, the CML was expected to show a blackout gap. Thus, we derived the cumulative number of expected blackout minutes $n_{exp}(D_{range})$ as the average number of timestamps per year where PIA> D_{range} multiplied by five due to the radar's instantaneous sampling rate of five minutes. We applied Eq. 4.3 to n_{exp} according to the radar availability along CML paths. Due to RSL_{min} undersampling, D_{range} might be higher in reality than estimated. In turn, n_{exp} should be lower than estimated, i.e. we would expect n_{obs} to be smaller than n_{exp}.

4.3 Results

4.3.1 CML signal levels and dynamic ranges

The distribution of TSL_{min} and TSL_{max} is defined by hardware configuration. The distribution of RSL_{min} and RSL_{max} depends on TSL, path length and path loss. The spread of observed RSL_{max} is lower than the spread of observed RSL_{min}. The distribution of the dynamic range estimate is shown in Figure 4.1f). The observed D_{range} was on average 40.5 dB with a minimum of 15.2 dB and a maximum of 74.3 dB.

4.3.2 Observed CML blackout gaps

Figure 4.2a) shows a histogram of path-averaged radar rainfall intensities. The higher the path-averaged rainfall intensity the less frequently it occurred. For each bin the fraction of CML data gaps which were detected as blackout gaps are shown (dark blue). In addition, the fraction of all gaps that have not been detected as blackout are shown (light blue). Note that gaps that were attributed to, e.g. failure of the data acquisition, have been removed as described in Sec. 4.2.1. The fraction of gaps is increasing quickly until 50 mm h^{-1} and then less steep up to 125 mm h^{-1}. For very high intensities above 125 mm h^{-1} the sample size was less than 50 minutes per bin. Therefore, the fraction of all gaps, including detected blackout gaps, was becoming sensitive to the occurrence of individual events and hence the statistics were less robust. Overall, around 95% of the gaps during rainfall in

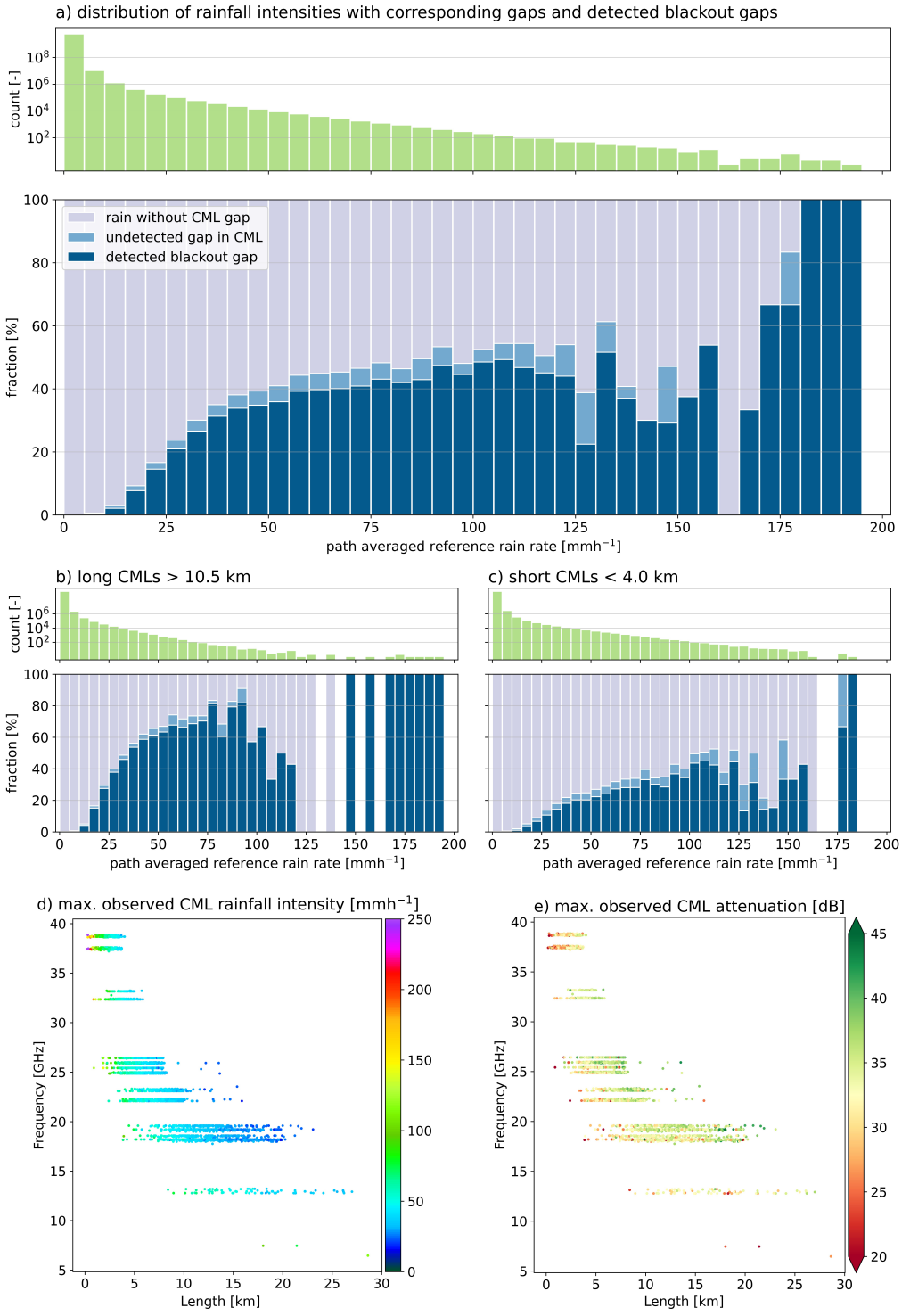

Figure 4.2: a) shows the distribution of the reference rainfall intensities in green. For each bin the fraction of gaps in the CMLs *RSL* time series and the fraction of the detected blackout gaps are shown in light and dark blue. b) and c) show the same for the longest (> 10.5 km) and shortest (< 4.0 km) quartile of all CMLs, respectively. Note that gaps that were attributed to, e.g. failure of the data acquisition, have been removed as described in Sec. 2.1 for a), b) and c). d) shows the maximal rainfall intensity derived from the CMLs estimated with the rainfall retrieval methodology from Graf et al. (2020a) and Polz et al. (2020). e) shows the respective maximal attenuation observed at each CML during the analysed three years.

the radar reference were detected as blackout gaps. This fraction varied for the highest observed rainfall intensities due to the small sample size. Based on the statistics from Figure 4.2a), CMLs missed on average 1% of the yearly rainfall sum during blackout gaps. The quartile of long CMLs, i.e. longer than 10.5 km, in 4.2b) showed a higher fraction of (blackout) gaps. Additionally, path-averaged rainfall intensities are lower on average as longer paths average out peak intensities. The quartile of short CMLs , i.e. shorter than 4.0 km, shows fewer (blackout) gaps and higher rainfall intensities. This pattern is also visible in 4.2d) and e) where the maximum instantaneous rainfall intensity and attenuation from each CMLs observations are shown. While the maximum attenuation increased with length, the maximum observed path-averaged rainfall intensity decreased. The maximum observed rainfall intensity from CMLs with 600 $\mathrm{mm\,h^{-1}}$ (and several events above 250 $\mathrm{mm\,h^{-1}}$ all beyond the figures colorscale) is well above the maximum intensity of the path averaged reference product. Overall, shorter CMLs show fewer blackouts during heavy rainfall.

4.3.3 Expected blackout gaps derived from radar based attenuation climatology

Expected PIA values along each CML path were derived using Equation 4.4 and 20 years of RADKLIM-YW data. Figure 4.3 shows path-averaged rain rate and PIA percentiles of the full 20-year dataset corresponding to the highest 60, five or one minutes per year and the 20-year maximum for individual CMLs (i.e. 60 minutes per year corresponds to the highest 0.011415526 percent in the data). The expected PIA was increasing with CML length, while the path averaged rain rate was decreasing. The five-minute PIA exceedance level (see Figure 4.3 second column) was between 10 dB (1st percentile), occurring mostly for shorter CMLs, and 53 dB (99th percentile), occurring mostly for longer CMLs. On average, a path-average rain rate of 42.8 $\mathrm{mm\,h^{-1}}$ and a PIA of 32.7 dB were exceeded for five minutes per year and a path-averaged rain rate of 17.9 $\mathrm{mm\,h^{-1}}$ and a PIA of 13.5 dB were exceeded for 60 minutes per year.

The decrease of path-averaged rain rate exceedance levels with CML length was similar for all frequency bins, while the increase of PIA was higher for higher frequencies. For all frequencies, the respective de- or increase was higher for more extreme values, i.e. comparing the 20 year maximum to 60 minutes per year, while the correlation was decreasing. The p-value was close to zero in all cases showing very high statistical significance.

Using the expected PIA values and our estimates of D_{range} we calculated n_{exp} which is shown in the second column of Figure 4.4). The majority of D_{range} was between 30dB and 50dB with higher values for longer CMLs (Figure 4.4 first column). Even though D_{range} was increasing with length, n_{exp} was also increasing with length.

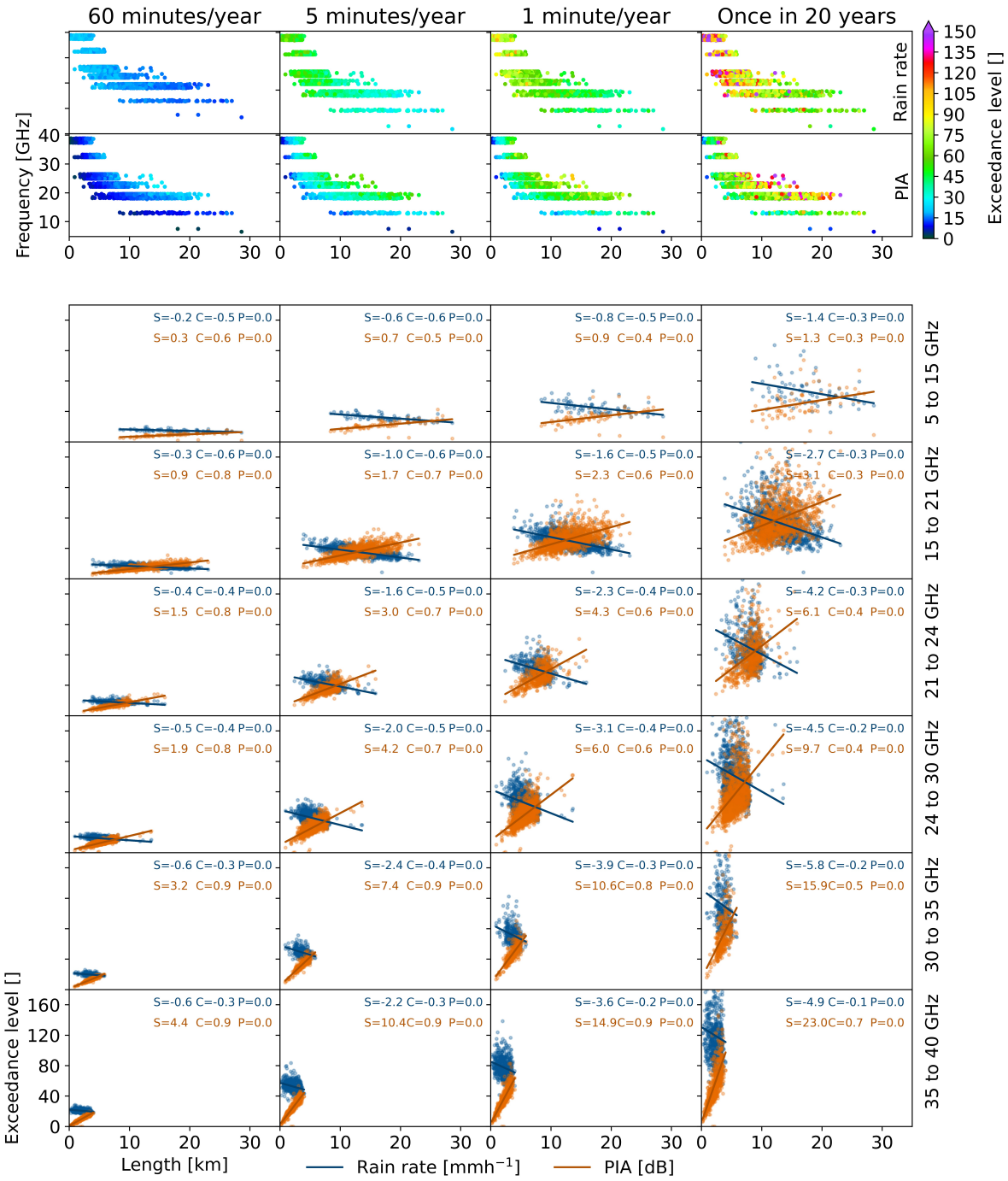

Figure 4.3: Rainfall and attenuation climatology for individual CMLs based on 20 years of RADKLIM-YW. The exceeded path-averaged rain rate and PIA along each CML path of a given length and frequency for at least 60, five and one minutes per year and the maximum rain rate occurring once in 20 years are shown in the four columns. The top two rows show rain rate and PIA (same color scale) as length against frequency scatterplots. Below, the same rain rate and PIA exceedance levels are compared to the CML length data points are shown for six frequency bins. The respective linear regression lines are shown with values for slope (S), correlation (C) and p-value (P).

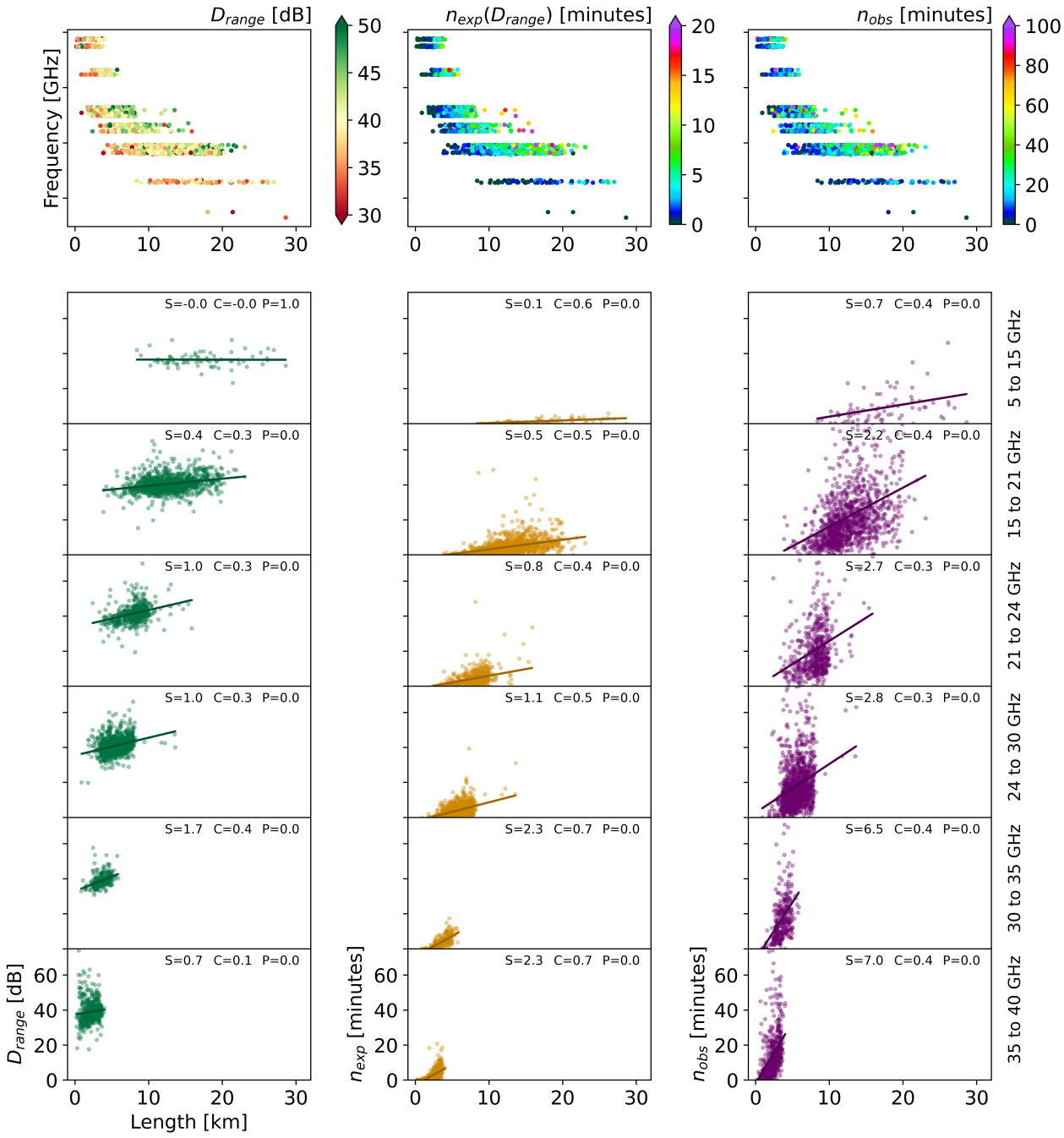

Figure 4.4: In the three columns D_{range} (left), n_{obs} (middle) and n_{exp} (right) are compared for each CML. The top row shows the respective variable on the color scale of a length against frequency scatterplot. Below, the three variables are shown against the CMLs length for six different frequency bins. A linear regression line and its values for slope (S), correlation (C) and p-value (P) is given for each of the scatter plots.

4.3.4 Comparison of observed and expected blackouts

D_{range} and the number of observed (n_{obs}) and expected n_{exp} blackout minutes per year are shown in Figure 4.4 for the individual CMLs length and frequencies and shown for each length in six frequency bins. D_{range} was increasing with CML length for all frequency bins except for 5 GHz to 10 GHz. We observed that n_{exp} and n_{obs} increased with CML length. However, n_{exp} showed a smaller slope than n_{obs}. The correlation between length and n_{obs} was low, but significant. The slope was strictly positive and increasing with higher frequencies, though.
Longer CMLs missed a higher percentage of of high rainfall intensities than shorter CMLs (see Figure 4.2b) and c). According to n_{exp} a 99.99% availability margin (as recommended by the ITU which is less than 60 minutes of blackouts per year) should have been observed for all CMLs. In practice, i.e. for n_{obs}, the 99.99% margin (60 minutes) was exceeded for the longest CMLs in each frequency band. We found this to be true throughout all frequency bins except 5 GHz to 15 GHz.

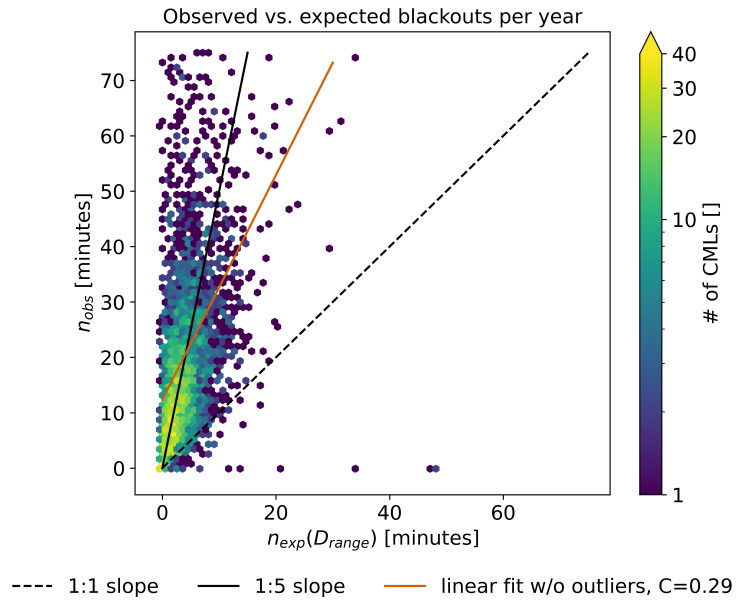

Figure 4.5: The observed number of blackout minutes per CML shown is compared to the number of expected blackout minutes in the scatter density plot, where the dashed black line corresponds to a 1:1 relation and the solid black line corresponds to a 5:1 relation. The orange line shows the best linear regression fit with a slope of 2.0 and a correlation (C) of 0.29. Outliers above the 99th percentile, i.e. $n_{obs} > 207.2$ or $n_{exp} > 17.5$, are excluded from the linear regression.

In Figure 4.5, n_{exp} is directly compared to n_{obs}. The mean of n_{obs} was 6 times higher than the mean of n_{exp}. The 99th percentiles of n_{obs} and n_{exp} were 207.2 and 17.5 minutes. Higher values are considered as outliers. On average n_{obs} was twelve times higher than

n_{exp} for all CMLs where $n_{exp} > 0$ and 8.5 times higher if outliers were excluded. Taking the median instead of the mean, the value is 4.6 independent of outliers included or not. The average n_{obs} for CMLs where $n_{exp} = 0$ was 19.4 minutes and the median was 3.2 minutes. 95.0% of all CMLs showed more observed blackout minutes than expected, i.e. $\frac{n_{obs}}{n_{exp}} > 1$. For 47.6% of the CMLs we observed more than five times more blackouts than expected and for 22.8% we observed more than 10 times more blackouts. A linear regression excluding outliers shows an additive increase of 11 minutes and a multiplicative increase of a factor 2. However, the correlation of 0.29 was low.

4.4 Discussion

4.4.1 Effects of CML length on blackout gaps and network design

The result that short CMLs have a lower likeliness to experience a blackout gap than longer CMLs was unexpected, because we expected the dynamic range to increase with CML length to account for the increasing PIA. Our empirical dynamic range estimates indeed show an increase with length, but it is not sufficient to compensate the even larger increase of PIA. Also, the path-averaging effect results in lower peak intensities of the path-averaged rain rates which decreases the attenuation per kilometer of CML length.
We found this difference between short and long CMLs in both our CML dataset and our radar-based attenuation climatology. Since observed and expected blackouts are based on independent methodological assumptions, we are confident that the effect is real. One potential explanation is that the path-averaging effect of peak intensities is overestimated during planning of the CMLs availability, so that longer CMLs experience more PIA than expected.
Our findings show potential to improve planning for future CML installations. Most prominently, our results suggest to increase the dynamic range of long CMLs. ITU recommends that the actual path length is multiplied by a so-called *distance factor* when calculating long-term statistics of rain attenuation (ITUßR, 2021). This factor significantly reduces the effective length (which is used for the calculation of path attenuation exceedance levels from rain rate exceedance statistic) of longer CMLs, e.g. the factor is approximately 0.5 for a 10 km CML with 20 GHz. Our findings, that longer CMLs experience more blackouts then shorter ones, suggest that this reduction of effective length of a CML for the calculation of path attenuation statistics is too strong, resulting in longer CMLs being planned with a too low D_{range}. Our radar-based exceedance probability can be used to estimate the potential increase of blackouts with CML length on the one hand. The total number of blackouts should be expected to be much higher on the other hand, which requires an additional increase of the dynamic range for all CMLs. As the ITU-recommended 99.99% availability was satisfied in most cases, this recommendation may be more urgent for hydrometeorological applications than network stability.

4.4.2 Implications of blackouts on CML rainfall estimation

Previous studies which compared CML rainfall information against reference data, naturally considered blackouts as missing values and little attention was payed to their implication on CML rainfall estimation. Our results confirmed that their impact on annual precipitation sums is in fact low with around 1% .
However, blackout gaps do impact CML-derived rainfall maps on shorter time scales and extreme value statistics in general, because extreme values are lost. The importance of this effect is illustrated by Figure 4.2 which shows the occurrence of blackouts during certain radar rainfall rates. The probability of a blackout at path-averaged rainfall intensities beyond 100 $\mathrm{mm\,h^{-1}}$ is higher than 40%. To interpret such maximum observable path-averaged rainfall rates the path-averaging effect of the CML observation needs to be taken into account, which is different from point-like observations.
Since we observed that shorter CMLs have a much lower probability of blackout gaps, there cannot be a general conclusion about the capability of a CML network to capture rainfall extremes. We suggest several possibilities to deal with blackouts associated with higher rainfall estimates. For applications requiring estimates of rainfall maxima with high temporal resolution, only short CMLs could be used. Another solution could be to fill RSL during detected blackout gaps with the minimal observable RSL value. Although the true maxima cannot be recovered, this could be a reasonable first step to reduce the considerable underestimation of high rain rates in CML-derived rainfall maps.

4.4.3 Underestimation of blackouts through radar-based attenuation climatology

Our results also have potential implications for radar rainfall estimates. We observed that the average CML experienced 8.5 times more blackouts than expected from the radar-based climatology. The underestimation occurs even though our dynamic range estimate is lower than in reality due to undersampling of RSL_{min} and the consideration of 3dB WAA. Although false positive blackout detection can not be excluded with certainty, manual checks of the blackout gap detection (see Graf et al. (2022a)) confirmed the correct magnitude of observed blackouts for the vast majority of CMLs.
Therefore, there is evidence that radar-derived path-averaged rain rates and the related PIA could underestimate extreme values. This is supported by studies reporting that gauge-adjusted radar products often underestimate heavy rainfall (e.g. Schleiss et al. (2020)). This underestimation could be explained by the different spatial integration characteristic of CML and radar. Another reason for the underestimation are effects that occur in combination with rainfall, e.g. hail, that may lead to unexpected high attenuation values, but they may not lead to high weather radar rainfall estimates due to quality control and attenuation of the radar signal.

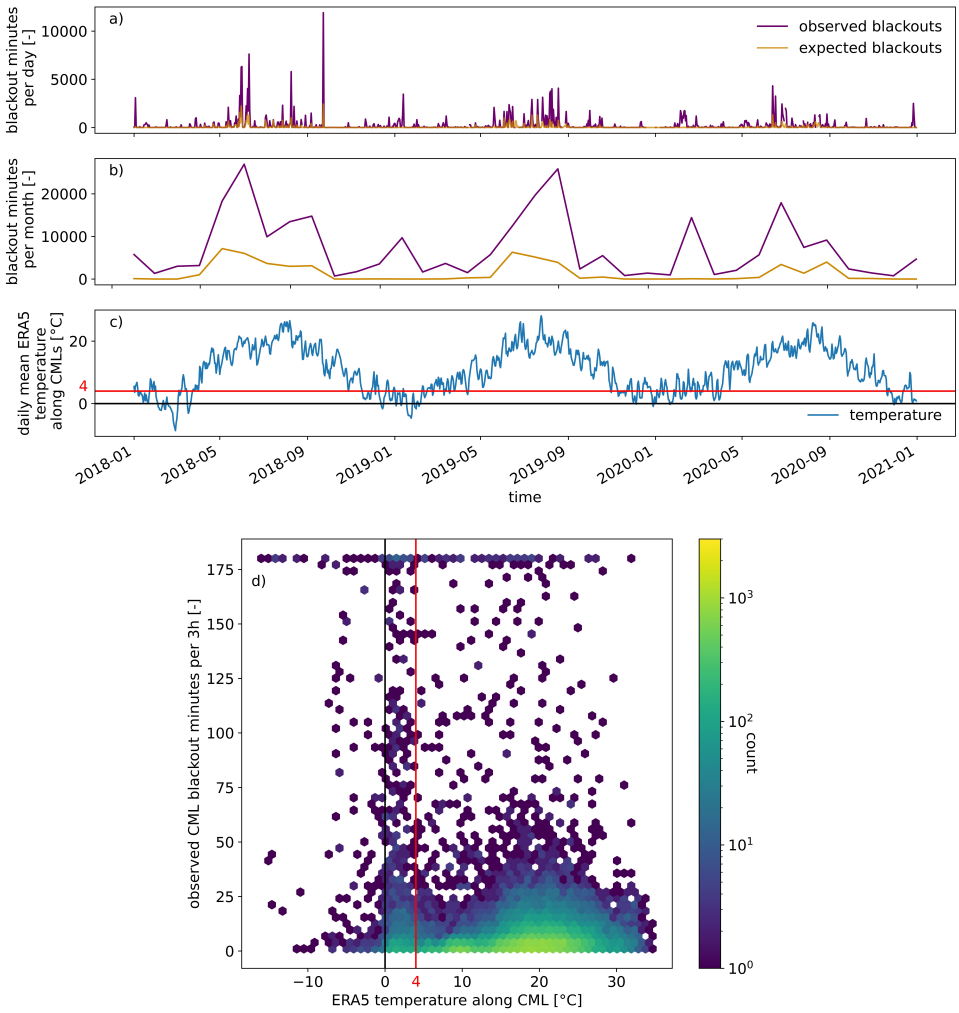

Figure 4.6: a) and b) show the observed and expected number of blackouts per day and month between 2018 and 2020. c) shows the mean 2 meter temperature along all CMLs derived from ERA-5-land. d) shows the observed number of blackout minutes per CML per 3 hours compared to the average ERA5-land 2 meter temperature along the link path during the same period. The red line in c) and d) indicates the 4°C threshold below which mixed type precipitation is more likely. 17.7% of all observed blackouts occurred below this threshold.

Melting hydrometeors like wet snow or sleet cause attenuation of CML signal larger than their rainfall equivalent (e.g.Tjelta and Bacon (2010)). We tested whether this effect influenced the number of observed blackouts by comparing blackout occurrences to the temperature along the CML paths derived from ERA5-land. Figure 4.6 shows the three-year time series of daily and monthly observed and expected blackouts with the daily mean temperature from ERA5 as well as a scatter density comparison of temperature and observed blackouts occurring in 3-hour periods. Similar to van Leth et al. (2018) we assumed that almost all precipitation above 4°C is liquid. We found that 17.7% of all observed blackouts occurred below 4°C, while the majority was centered around 20 degrees. This shows that the contribution of melting hydrometeors at temperatures below

4°C is not high enough to fully explain the underestimation of blackouts through weather radar data. It could also be observed that, as expected, blackouts rarely occur at negative temperatures. We conclude that despite the high spatial and temporal resolution, the weather radar data is not sufficient to fully explain CML blackouts.

4.5 Conclusions

During extreme heavy rain events, CMLs may experience blackouts, i.e., complete loss of signal. Our objectives were to determine the impact on rainfall estimation, the occurrence of blackouts in a country-wide network of 3904 CMLs and to determine if these numbers were consistent with the theoretical number of occurrences of blackouts derived from a 20-year climatology of a high-resolution weather radar product. On average, CMLs experienced 20 minutes of blackout per year and the average CML experienced 8.5 times more blackouts than the radar climatology suggested. Shorter CMLs showed fewer blackouts in both the observed and theoretically derived data. Although the amount of rainfall missed was small compared to annual sums, the observed probability of blackouts during path-averaged radar rainfall intensities beyond 100 $\mathrm{mm\,h^{-1}}$ was more than 40%, which impacts CML rainfall estimates of individual heavy rainfall events on short timescales. Especially surprising was the increase of blackouts with CML length. Therefore, we suggest that the CML research community should be aware of this limitation and the proposed mitigation measures. Finally, this study fills a knowledge gap on the distribution of blackouts in CML data and weather radar derived attenuation climatology which can be considered in future CML infrastructure planning.

4.6 Acknowledgments

We thank Ericsson, especially Reinhard Gerigk, Michael Wahl, and Declan Forde for their support in the CML data acquisition. This research has been supported by the Helmholtz Association (grant ZT-0025), the German Research Foundation (grant CH-1785/1-2) and the Federal Ministry of Education and Research (grant 13N14826). We acknowledge support by the KIT-Publication Fund of the Karlsruhe Institute of Technology.

4.7 Open research

Software for the blackout gap detection routine (Graf et al., 2022a) is available within the CML rainfall retrieval Python-package pycomlink (Chwala et al., 2022) under BSD-3-Clause License. The CML data supporting this research was provided to the authors by Ericsson, restricting the distribution of this data due to their commercial interest. In order to obtain CML data for research purposes a separate and individual agreement with the network provider has to be established. To allow for an independent evaluation of our methodology we published data from 500 CMLs over ten days and two CMLs for the

full period of this study (Chwala et al., 2022; Graf et al., 2022b) under CC BY 4.0. The RADKLIM-YW dataset used in this research is publicly available and can be downloaded from Winterrath et al. (2018).

Chapter 5

Expert Flagging of Commercial Microwave Link Signal Anomalies: Effect on Rainfall Estimation and Ambiguity of Flagging (Polz et al., 2023a)

Polz, J., Glawion, L., Graf, M., Blettner, N., Lasota, E., Schmidt, L., Kunstmann, H., and Chwala, C.,: Expert flagging of commercial microwave link signal anomalies: Effect on rainfall estimation and ambiguity of flagging. In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), pages 1–5, https://doi.org/10.1109/icasspw59220.2023.10193654, 2023,

Abstract

Accurate detection of signal anomalies in the attenuation time-series from commercial microwave links (CMLs) is crucial for high quality rainfall estimates. Example causes of such anomalies include dew or ice on the antenna and multi-path propagation. In a first effort to catalog examples of CML signal anomalies, four experts flagged suspicious segments in the time-series of 20 CMLs in Germany. The results show that the agreement between experts depends on the definition of the anomaly class. Removing the flagged anomalies increased the Pearson correlation coefficient between CML and radar rainfall estimates from 0.61 to 0.70 and reduced the BIAS by 40%. An implication of our study is that expert uncertainty is an important factor for the quality control of environmental sensor data.

5.1 Introduction

Commercial microwave links (CMLs) have been established as an important tool for rainfall estimation in data scarce regions and to support traditional rainfall sensors like rain gauges or weather radars (Messer et al., 2006; Uijlenhoet et al., 2018). Processing CML data is challenging due to the opportunistic nature and the indirect measurement of a

rainfall intensity averaged along the link path based on the path integrated attenuation of microwave radiation. The separation of wet and dry periods in the attenuation time-series of CMLs is the most common strategy for quality control of CML data. The purpose of wet-dry detection is to avoid the false estimation of rainfall from signal fluctuations during dry periods (Chwala and Kunstmann, 2019; Polz et al., 2020). Signal anomalies may disturb a reasonable rainfall estimation even during rain events (Engström et al., 2020). Arguably, the definition of an anomaly is not straight forward as rainfall itself is the main factor causing signal fluctuations that disturb the transmission of information along the link path. However, it is equally important to learn more about the potential reasons for fluctuations in CML signal levels which are not related to rainfall. We will denote only fluctuations caused by such phenomena as anomalies in the following. A series of possible causes for anomalies have been proposed including dew or ice on the antenna of a CML, which creates high attenuation. Furthermore, the sensitivity to water vapor increases with higher transmission frequencies (Fencl et al., 2021). Multi-path propagation above water bodies, interference at the receiver by the reflection of the beam or the decrease of the signal level by beam blockage may also cause anomalies. Additionally, wind, temperature, or solar radiation can affect the hardware.

Many of the potential anomalies are assumed to be caused by phenomena for which suitable reference data is often missing to investigate their impact. For example, the wind that causes the tower to swing may be estimated using data from weather models, but other factors like the mechanical properties of the tower, that is, the height and construction material, are largely unknown. Progress in the investigation of CML signal anomalies has been steady due to the interest in CML rainfall estimates, but slow compared to the overall development in the field (Chwala and Kunstmann, 2019). This work tries to follow a different approach to shed light on CML anomalies. Experts in the field of CML based rainfall estimation usually develop a good understanding of the patterns in the time-series caused by rainfall. Despite being able to tell which segments of the time-series resemble rain events, the causes for anomalies remain less well understood.

The research questions for this work are: If individual experts are provided with the same CML data and a predefined set of anomaly classes, do they usually agree on the same classes? How can expert flagging improve CML rainfall estimates?

Figure 5.1: Anomaly flagging tool used for this study. The two drop down menus (upper left) were used to select a CML ID and a month. The green and white buttons allow for the selection, saving, and deleting of anomaly flags. The flagged periods were visualized in the 2nd panel. Besides the TL time-series of both channels (1st panel), the rainfall estimates derived from CML, RADKLIM-YW and RADOLAN-RW reference data were given (3rd panel). Additionally, the ERA5-Land variables $t2m$ and $d2m$ (4th panel), and horizontal wind speed and surface solar radiation (5th panel) were given.

5.2 Data and Methods

The CML data used in this study is a subset of a German CML data set, which in its entirety contains 3904 link paths. The data acquisition allowed for an instantaneous sampling of received (RSL) and transmitted (TSL) signal levels every minute with a signal quantization of 1 dB for TSL and 0.3 or 0.4 dB for RSL (Chwala et al., 2016). Data was collected from two sub-links with opposing transmission directions. For this study, we selected 20 CMLs in March, May, and July 2019. For an expert flagging of anomalies, this amount of data was expected to take about one day of work. Rainfall estimates were derived following the processing routine from Graf et al. (2020a) using the rain event detection from Polz et al. (2020). The wet antenna compensation we used was the rain rate dependent model from Leijnse et al. (2008).

As a reference data set, we used the gauge-adjusted and climatology corrected weather radar rainfall product RADKLIM-YW provided by the German Meteorological Service (DWD) (Winterrath et al., 2017). It has a temporal resolution of 5 minutes and a grid spacing of 1 km by 1 km. Its domain has an extent of 1100 km by 900 km and covers the whole of Germany. A path-averaged rain rate R along a CML of total length L was derived by a sum of radar rainfall intensities r_i weighted by the lengths of intersection l_i of the CML path and the radar grid cells.

$$R = \frac{1}{L} \sum_i r_i l_i \tag{5.1}$$

We compared CML rainfall estimate (1-minute resolution) to the preceding radar measurement (5-minute resolution). To quantify the agreement between CML and radar estimates we used the mean squared error (MSE) and Pearson correlation coefficient (PCC). The relative BIAS is defined as the average difference between radar and CML as a percentage of the average radar rainfall intensity. The Matthews correlation coefficient compares the quality of the rain event detection (radar wet vs. CML wet).

Additional reference data along link paths is derived from the enhanced land component of the fifth generation of European ReAnalysis (ERA5-Land) provided by the European Centre for Medium-Range Weather Forecasts (Muñoz-Sabater et al., 2021). Its spatial resolution is 9km and the temporal resolution is 1 hour. We derived path averaged quantities using the same procedure as in Eq. 5.1. We used the 2 metre temperature ($t2m$), the 2 metre dewpoint temperature ($d2m$), the surface net solar radiation and the 10 metre U wind component (wind speed).

All signal patterns that could not be attributed to rainfall were called anomalies. We categorized CML anomalies considering four different classes that we defined before flagging the data. Examples are given in Figure 5.2

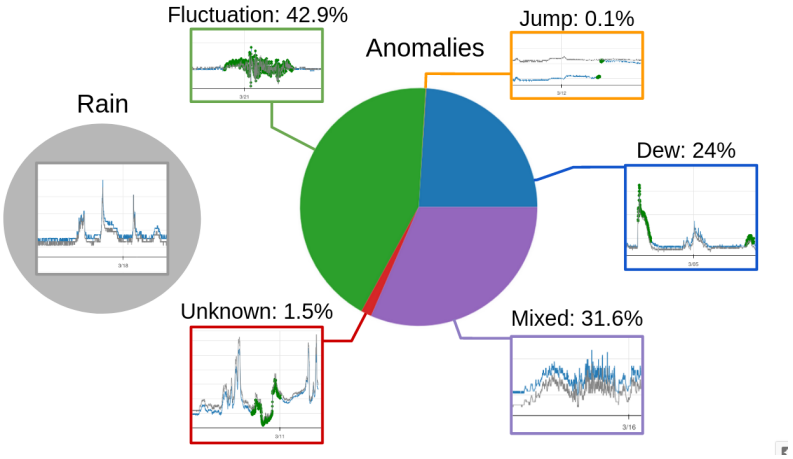

Figure 5.2: Share of classes flagged by the experts. The percentage of the main classes DEW (blue), JUMP (orange), FLUCTUATION (green), and UNKNOWN (red) among the time steps flagged as anomalous are shown. The MIXED class (purple) describes all timestamps, where more than one class was selected by the experts. Examples of TL time-series of the sub-links (grey and blue) are shown for all anomaly classes and additionally for typical patterns caused by rainfall. All time-series segments are 24 hours long and the green selection indicates the anomaly location. © 2023, IEEE.

The DEW class was described as a segment of the time-series which comprises a period of a constant increase in attenuation without heavy fluctuation. Such, typically bell-shaped, patterns can be caused by dew or snow on the antenna.
The JUMP class describes all segments that had a large change in signal level within a single time step. It includes spikes and jumps larger than 5 dB or long-term baseline changes, i.e. multiple hours, larger than 2 dB.
The FLUCTUATION class describes all random or noise-like signal fluctuations which are larger than 2 dB and the expert did not associate with rainfall.
The UNKNOWN class was meant to be an open category to include all signal patterns which can not be attributed to rainfall and which did not fit into any of the classes above.
Data was flagged using a custom built tool for visualizing the time-series. The code is available online[1]. The user interface and provided information are shown in Figure 5.1. It was allowed to select multiple classes for any given time step and CML. The flagging was conducted by the first four co-authors listed for this paper who all have worked with CML data for multiple years on a daily basis. After an initial round of flagging 10 CMLs different from the ones used in this study, the tool, the available classes, and their interpretation were discussed. In the second round of flagging the data presented in this study was conducted independently by each expert. Therefore, up to four votes per class and time-step were possible.

[1] https://github.com/LGlawion/anomaly_selection_tool/tree/main/notebooks/cml

5.3 Results

Figure 5.1 illustrates the expert flagging by showing a four-day segment from one of the CML time-series. We shortly describe the reasoning behind the shown classification. In the night between the 2nd to 3rd of March, after a small rain event detected by radar and CML, the baseline attenuation slowly starts to rise until it stabilizes in the morning hours. During the rise two distinct peaks are visible which might be due to semi frozen particles which were not detected by the radar when they were completely frozen at higher altitudes. At the same time, $t2m$ is declining and $d2m$ is very close to $t2m$. Therefore, the expert selected the dew class since they assumed that dew, ice, or wet snow was covering the CML antenna. This assumption was backed up by the sudden drop of TL 36 hours later after surface solar radiation and temperature rose, likely causing the frozen ice cover of the antenna to fall off. The sudden drop was classified as a jump. It could be observed that the rain event detection of our CML data processing classified the beginning and end of the "dew" event as wet which led to a large estimated rainfall intensity while the radar did not detect rainfall.

From all time steps in the data set, 10.84% were classified as anomalous by at least one expert. Figure 5.2 shows the share of the flagged anomaly in each class. If multiple classes were selected for a time step, it was referred to as a MIXED class. The most frequent anomalies were FLUCTUATIONS with a share of 42.9%. The second largest category was the MIXED class with 31.6% followed by DEW with 24%. The UNKNOWN and JUMP classes were very seldom selected with 1.5% and 0.1%.

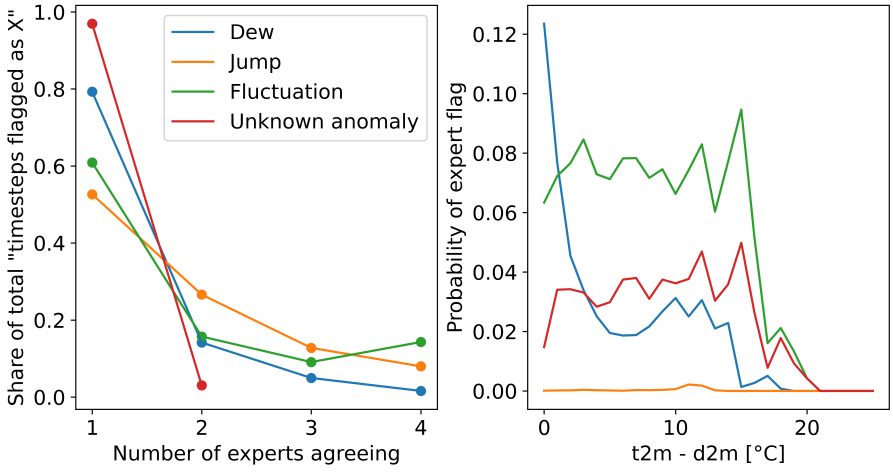

Figure 5.3: Left side: Frequency of anomaly classes (including MIXED cases) compared to the difference of ERA5-Land 2m surface temperature ($t2m$) and dew point temperature ($d2m$). Right side: Agreement of experts for pure classes (excluding MIXED cases). To be read like: for 80% of all time steps where DEW and only DEW were flagged, there was only one vote from one expert. © 2023, IEEE.

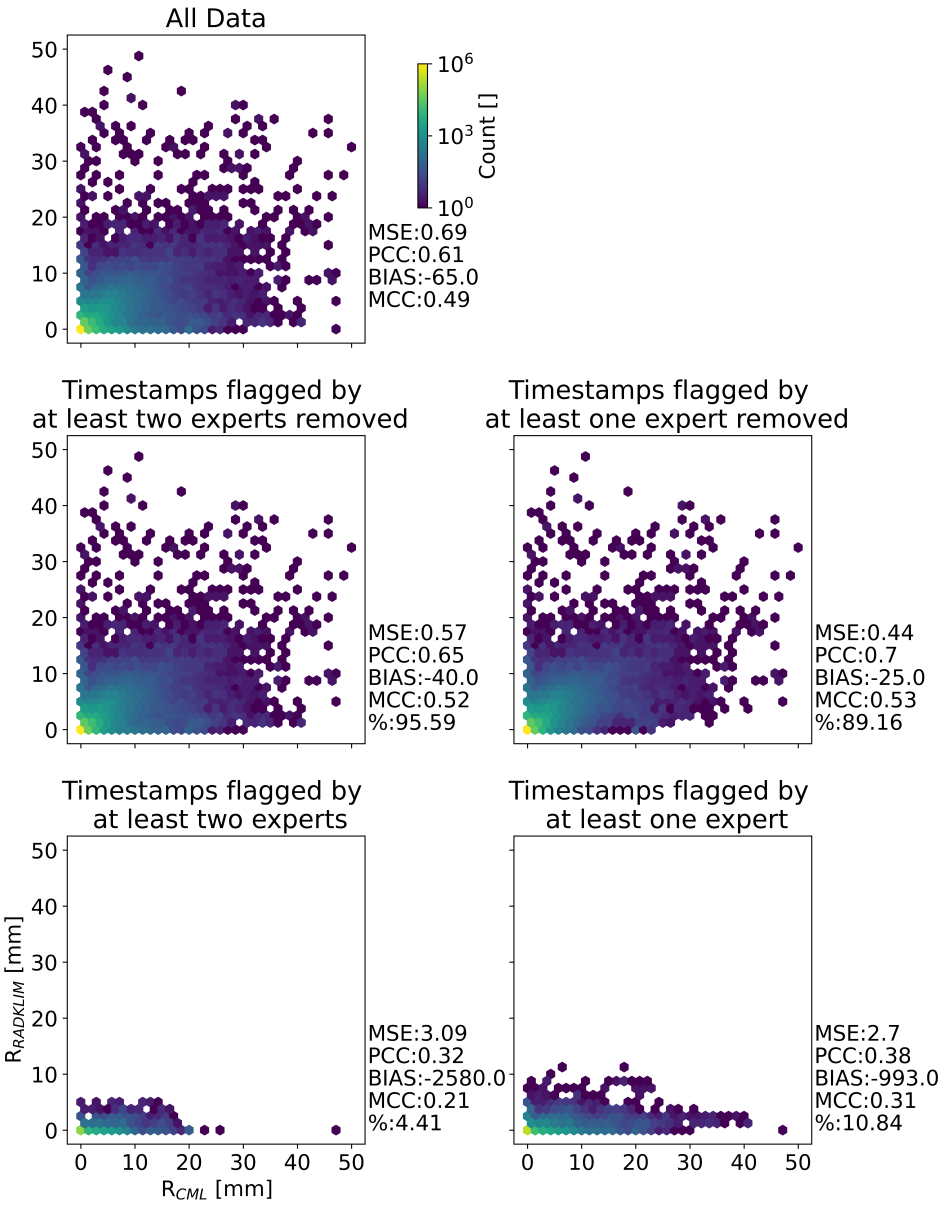

Figure 5.4: Scatter density comparison of CML and RADKLIM-YW rainfall intensities. The top panel shows all time steps of the data set. Below, all remaining time steps after removing anomalies flagged by at least one (right) or at least two (left) experts are compared. The bottom row shows the time steps that have been removed above. © 2023, IEEE.

The right hand side of Figure 5.3 shows how many experts agreed on the classification of each time step. While no more than two experts classified the same time step as UNKNOWN, showing the lowest agreement, there was a high agreement for the FLUCTUATION class. If more than one expert agreed on FLUCTUATION, then it was most likely that all four agreed. For the two remaining classes (DEW and JUMP), the likelihood of agreement declined with the number of experts.

The left panel of Figure 5.3 compares the frequency of the anomaly classes to the difference

between $t2m$ and $d2m$, indicating a possible formation of dew on the antenna. While JUMP, FLUCTUATION, and UNKNOWN are independent of this quantity, the DEW class was selected more frequently when $d2m$ and $t2m$ were close.
The impact of the expert flagging on rainfall estimation is illustrated in Figure 5.4. All panels compare the 1-minute rainfall intensities estimated from the CMLs (R_{CML}) to the 5-minute rainfall intensity of the RADKLIM-YW $R_{RADKLIM}$. Removing anomalies flagged by at least one expert leads to the largest improvement increasing the PCC from 0.61 to 0.70 and reducing the BIAS by 40%. The elimination of time steps flagged by at least two experts seems to show less improvement. However, when comparing the removed rainfall estimates (bottom row), it becomes obvious that all metrics worsen the more experts agree on a flag. Overall, almost all removed data is close to the x-axis where $R_{CML} \gg R_{RADKLIM}$.

5.4 Discussion

Our results showed that the agreement between the experts was low since more than half of all flags had only one vote for the same class from one expert. However, in the majority of cases, the elimination of flagged time steps improved the PCC, MSE, BIAS and MCC. A reason for the low agreement could be an imprecise flagging of the experts, that is, in general, the same events were flagged, but the start and end points were chosen differently. An event-specific analysis of this issue would require a larger and more complex evaluation of the data set than we were able to present here.
In general, the benefit of the flagging for CML rainfall estimation can be considered high. While wet-dry detection only removes false positives and false negatives (on the x and y-axis of Figure 5.4), the expert flagging also discards correctly classified rain events that produce a very large overestimation due to anomalous data.

5.5 Conclusion

The objective of this study was to investigate the impact of CML signal anomalies on rainfall estimation as they were manually flagged by four experts in the field. Our results show that while there is a low agreement among experts, the potential benefit for rainfall estimation is very high considering an improvement of the PCC from 0.61 to 0.70 and a reduction of the BIAS by 40%. Further implications of our results are that potential algorithmic automation of the flagging procedure, e.g. by a machine learning approach, could benefit from multiple experts flagging the same data. Previous approaches for anomaly detection mainly focused on false classifications of the rain event detection, which seems to be an incomplete view of the problem. Additionally, we showed that in case of an automated detection, not only model uncertainty, but also expert uncertainty has to be considered. Our results provide a baseline for future machine learning applications and for the CML community to judge how ambiguous manual flags can be.

5.6 Acknowledgments

This research has been supported by the German Research Foundation within the projects RealPEP (Grant CH-1785/1-2) and SpraiLINK (Grant CH-1785/2-1), by the Helmholtz Association within the project RESEAD (Grant ZT-I-PF-5-148-RESEAD) and by the Federal Ministry of Education and Research within the project HoWa-PRO (Grant 13N16432).

Chapter 6

Temporal Super-Resolution, Ground Adjustment and Advection Correction of Weather Radar Observations using 3D-Convolutional Neural Networks (Polz et al., 2024)

Polz, J., Glawion, L., Gebisso, H., Altenstrasser, L., Graf, M., Kunstmann, H., Vogl, S., and Chwala, C. Temporal Super-Resolution, Ground Adjustment, and Advection Correction of Radar Rainfall Using 3-D-Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 62:1–10, https://doi.org/10.1109/TGRS.2024.3371577, 2024.

Abstract

Weather radars are highly sophisticated tools for quantitative precipitation estimation and provide observations with unmatched spatial representativeness. However, their indirect measurement of precipitation high above ground leads to strong systematic errors compared to direct rain gauge measurements. Additionally, the temporal undersampling from 5-minute instantaneous radar measurements requires advection correction. We present ResRadNet, a 3D-convolutional residual neural network approach, to reduce these errors and, at the same time, increase the temporal resolution of the radar rainfall fields by a 5-minute short-range prediction of 1-minute time-steps. The network is trained to process spatiotemporal sequences of radar rainfall estimates from a composite product derived from 17 C-band weather radars in Germany. In contrast to previous approaches, we present a method that emphasizes the generation of spatiotemporally consistent and advection-corrected country-wide rainfall maps. Our approach significantly increased the Pearson correlation coefficient of the radar product (from 0.63 to 0.74) and decreased the root mean squared error by 22 percent when compared to 247 rain gauges at a 5-minute

resolution. An additional large-scale comparison to 8 years of data from 1138 independent manual daily gauges confirmed that the improvement is robust and transferable to new locations. Overall, our study shows the benefits of using 3D convolutional neural networks for weather radar rainfall estimation to provide 1-minute, ground-adjusted, that is, bias-corrected with respect to on-ground sensors, and advection-corrected radar rainfall estimates.

6.1 Introduction

Weather radars are one of the most advanced tools for quantitative precipitation estimation (QPE) at high spatial and temporal resolutions. Their observations are not only essential for state agencies providing information for disaster or water management. They also became one of the most popular rainfall observation tools for citizens. The key to their success is the high spatial representativeness of the derived measurements. Even though they integrate over large volumes, depending on the distance from the radar (Pejcic et al., 2020), they provide several observations per km^2 which is significantly higher than rain gauge networks. For example, the average density of rain gauges is $\frac{1}{330}\ km^{-2}$ in Germany, which is high in a global comparison. However, the raw measurements of weather radars are subject to systematic and random errors due to their indirect measurement high above ground (Wilson and Brandes, 1979; Pulkkinen et al., 2016). Additionally, weather radar measurements are instantaneous, typically sampled every 5 minutes, which leads to a temporal undersampling. This becomes obvious when aggregating rain fields containing small fast-moving cells which leads to gaps in the rain field. As a solution, advection correction methods have been suggested (Anagnostou and Krajewski, 1999; Seo and Krajewski, 2015; Pulkkinen et al., 2019).

A common approach to correct radar QPE biases is to merge radar and rain gauge data to provide optimal rainfall estimates near the ground. Most of them consider both radar and rain gauge data as sources of information to achieve a combined product. Others correct radar data to remove biases compared to rain gauge observations using a statistical approach (Vogl et al., 2012; McKee and Binns, 2016; Pulkkinen et al., 2016). However, scenario-dependent and advection-driven biases between radar and rain gauge data ask for a more dynamic approach like Wang et al. (2015) propose. It is reasonable to assume that the spatiotemporal dynamics contained in the rainfall fields measured by weather radars can provide valuable information to correctly map radar rainfall estimates to the ground. Deep neural networks have been used effectively for modeling dynamic physical systems. They provide a flexible and computationally efficient modeling framework that can outperform state-of-the-art physical models with comparatively small computational effort and at low latency (Pathak et al., 2022). The concept of using deep neural networks to derive improved reflectivity-rainfall (Z-R) retrieval for radar-derived QPE has been used recently by Yo et al. (2021) who targeted hourly observations of 45 rain gauges from radar reflectivity at S-band using one radar. Chen and Chandrasekar (2021) applied a 2D-convolutional

neural network (CNN) to derive point scale rainfall from reflectivity and differential reflectivity in the vicinity of one radar, also at S-band. Vogl et al. (2022) applied a neural network approach deriving rainfall estimates from reflectivity measured at X-band using one reference rain gauge and investigated the influence of different temporal aggregations on the neural network performance. Hassan et al. (2022) applied classical machine learning algorithms to determine suitable physical retrieval algorithms in different scenarios. Moraux et al. (2021) present a multi-modal approach for merging rain-gauge, satellite, and radar rainfall estimates using all three data sources as model input. In summary, these studies directly used radar reflectivity as input for their retrievals without increasing the temporal resolution. They provide a proof of concept that a learned Z-R retrieval can improve radar rainfall estimates at S- or X-band. The respective case studies involved either single radar stations, a low amount of reference stations, or a short time period. It remains unclear if an improved Z-R retrieval or an estimate of the spatiotemporal dynamics and a reduced spatial or temporal mismatch between radar and on-ground reference is responsible for the improvement. It is also unclear if radar rainfall estimates from a radar composite, which can include additional errors or discontinuities, can be used to consistently improve surface rainfall at high spatial and temporal resolutions (e.g. 1 km and 1 minute) at a country-wide scale.
The research questions addressed by this study target these knowledge gaps:

1. Are 3D-convolutional residual neural networks a suitable tool to increase the temporal radar resolution of 5 minutes by a short-term prediction of five 1-minute timesteps?

2. Can biases between high-resolution C-band weather radar rainfall estimates and rain gauges on the ground be effectively reduced?

3. Does the method provide spatiotemporally consistent country-wide rainfall maps over Germany which can be used to correct advection-driven undersampling?

In summary, we aim to show that 3D-convolutional residual neural networks are capable of simultaneous temporal super-resolution, ground adjustment, and advection correction of radar rainfall. Our evaluation is based on 8 years of country-wide radar and rain gauge observations. For the sake of brevity, we denote our proposed method by ResRadNet.

6.2 Data and Methods

6.2.1 Data

All data used in this study was provided by the German Meteorological Service (DWD). We used two rain gauge and two radar data sets. The 1-minute rain gauge data is freely available on DWD's opendata online archive (`https://opendata.dwd.de/`). The network operated Ott Pluvio^2 sensors until 2018. From 2018 to 2020 these rain gauges were replaced by rain[e] sensors from Lambrecht meteo. Both sensors are weighing gauges that provide the accumulated rainfall amount on a one-minute basis. While they have a different resolution (Pluvio^2 0.01 mm and rain[e] 0.001 mm) the quantization of the data set remains the same with 0.01 mm.

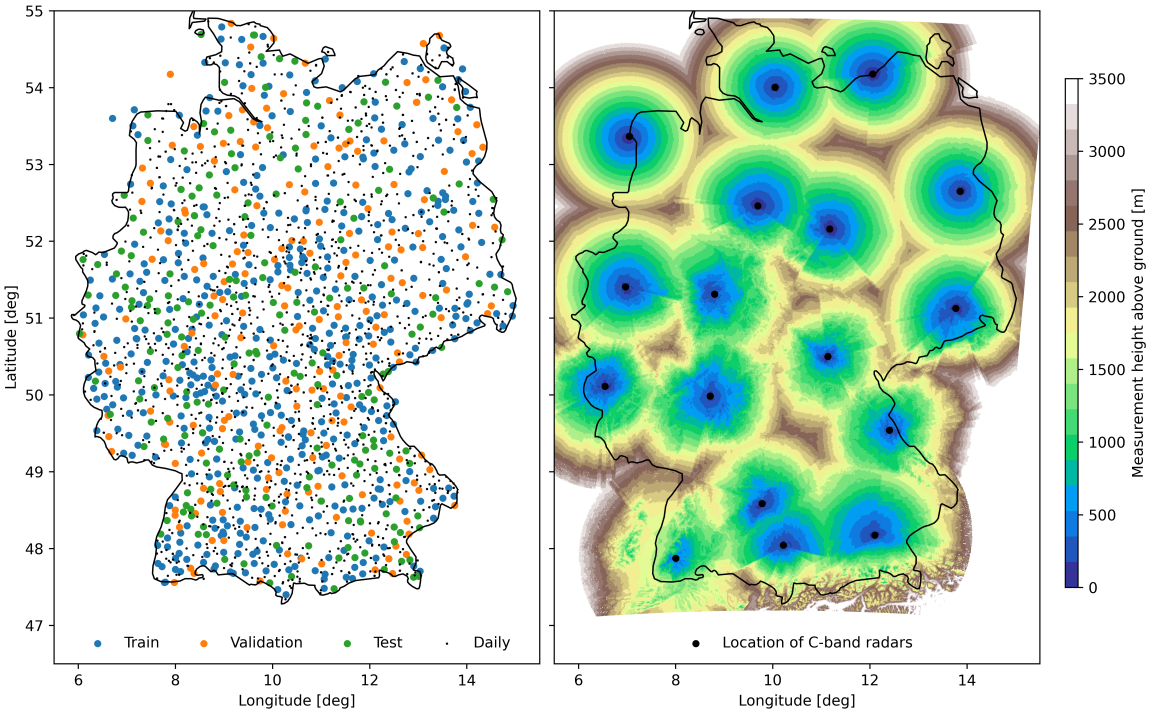

Figure 6.1: Map of Germany showing the locations of the rain gauges used for the train, validation and test split (left) and the minimal radar measurement height above ground of the composite derived from the 17 C-band radars (right). Training data is taken from 2020 and validation and test data are from 2021.

The average station density was one rain gauge per $330km^2$.
The daily rain gauge data is also freely available on DWD's opendata online archive (DWD, 2021). The network consisted of Hellmann rain gauges with a manual readout of the accumulated rainfall amount of the last 24 hours at 5:50 UTC. Therefore, the aggregation

time was from 5:50 UTC to 5:50 UTC on the following day. The quantization of the data set was 0.1 mm.

For the period from 2013 to 2021, there were 1066 1-minute and 2150 daily rain gauge stations available. Around 50% of the daily rain gauges were located at the same site as the 1-minute rain gauges. In order to obtain an independent network of daily rain gauges we removed these stations from the analysis. The remaining 1138 locations are shown as black dots in Fig. 6.1. Note that not all stations were available for the full period due to maintenance, redistribution, and new installation. The number of missing gauges per year was less than 5%.

The radar products from DWD used in this study are RADOLAN-RY, RADOLAN-RW, and RADKLIM-YW. A detailed explanation of the operational routines of these three radar products from DWD, e.g. on the used three-part Z-R relation, the rain gauge adjustment or climatological corrections can be found in Bartels et al. (2004) for RADOLAN-RY and RADOLAN-RW and in Winterrath et al. (2017) for RADKLIM-YW.

RADOLAN-RY is a quality-checked and attenuation-corrected composite of 17 weather radars in Germany, with a temporal resolution of 5 minutes and a spatial resolution of 1 km by 1 km on a 900 km by 900 km equally spaced grid. DWD derives the rainfall rate from a Z-R relationship based on measured reflectivity at C-band.

RADOLAN-RW is an hourly aggregated and gauge-adjusted version of RADOLAN-RY. The adjustment consists of a weighted combination of multiplicative and additive factors derived from the comparison of the 1-minute rain gauges described above with the radar grid at an hourly aggregation.

For RADKLIM-YW, the DWD uses the same radar and gauge data, as well as the daily rain gauges described above to derive a gauge-adjusted and climatologically corrected radar product at a temporal resolution of 5 minutes and a spatial resolution of 1 km by 1 km on an equally spaced grid of 1100 km by 900 km extending 100 km to the east and west compared to RADOLAN-RY. It should be noted that the adjustment weights are calculated at a 1-hour resolution and remain constant for the 12 5-minute timesteps within this hour. Thus, the 5-minute observations still deviate from the rain gauges used for adjustment. The climatological corrections aim to reduce errors for example radar spokes from beam blockage and range-dependent underestimation. We used RADKLIM-YW as a visual reference for high-resolution maps.

The data set covers the years 2001 to 2021, but we omitted the time period between 2001 and 2012 since the number of available rain gauges was much lower due to the ongoing set-up of the gauge network. To compare our results to an advection-corrected version of RADOLAN-RY we used the Lucas-Kanade algorithm (Lucas and Kanade, 1981) implemented in PYSTEPS version 1.7.1 (Pulkkinen et al., 2019). We used a base resolution of 5 minutes and 5 intermediate timesteps. To compare 1-minute neural network estimates to RADOLAN-RY we interpreted the 5-minute resolution as a constant average during the 5 minutes (see $R_{ry}/5$ in Fig. 6.2e).

The minimal measurement height above ground (see Fig. 6.1) is derived from the terrain

following the beam angle of the precipitation scan and is provided on the same grid as RADOLAN-RY. Where no terrain is blocking the beam an angle of 0.8 degrees is used. In the radar composite, the minimum of the measurement heights of two overlapping radar scans is used. This usually gives the measurement height of the closest radar which also has the highest weight in the composite of measurements. In some parts of Germany, the measurement height exceeds 2.5 kilometers.

6.2.2 Data Pre-processing

We train our model on sequences of RADOLAN-RY radar images with the objective to predict an assigned sequence of rain gauge measurements. The selection of radar sequences and the assigned rain gauge were defined as shown in Fig. 6.1: If a rain gauge Y was contained in the radar pixel $X_{t,i,j}$ at time t in minutes, then $X_{i,j}$ was associated with rain gauge measurements $(Y_{t'})_{t' \in \{t,\dots,t+4\}}$. For grid calculations we used ωradlib (Mühlbauer et al., 2022). As model input, the neighboring radar pixels $\{X_{k,l}\}_{k \in \{i-5,\dots,i+5\}, l \in \{j-5,\dots,j+5\}}$ at times $t-20, t-15, \dots, t$ were concatenated to form a sample of shape (5,11,11) (see Fig. 6.2b). Additional model input was given by the minimal measurement height above ground from the 11 by 11 radar pixels (see Fig. 6.2a).
We excluded all samples where the radar input contained missing values or where all values were zero. We also excluded all samples where the associated rain gauge had a missing value, but no additional requirements were made for the rain gauge reference. Data was split into spatially and temporally separated train, validation, or test sets by a random selection of 581 stations for training, 238 for validation, and 247 for testing of the model. The locations of the selected rain gauges are shown in 6.1. For training only the year 2020 was considered and, if not indicated otherwise, the full year 2021 was used for both validation and testing. The training data was used to train the neural network and the validation data was used for model selection and bias correction as described below. The test dataset was an independent dataset only used for validation. The training data set contained 8,773,000 samples in randomized order.

6.2.3 Model Architecture

The model architecture of ResRadNet consisted of a two-branch convolutional neural network with two 3D-residual blocks in the radar input branch and a 2D convolution block in the measurement height branch. A schematic overview is given in Fig. 6.2c, a more precise description of the layers and connections is given in Tab. 6.2, and code to re-build the model or retrieve the trained model is given in Polz (2023b). The model architecture was designed using the Keras API in Tensorflow version 2.7.0 (Abadi et al., 2016) and optimized with the AMSgrad version of the Adam algorithm and a learning rate of 0.0001 (Kingma and Ba, 2017; Reddi et al., 2018) using a mean squared error (MSE) loss (see Eq. 6.3). The residual blocks consisted of two convolutional branches receiving the same input (see Fig. 6.2d). The first uses two 3D convolutional layers with a kernel

size of 3 and ReLU activation. The second uses one 3D convolutional layer with a kernel size of 1. They are followed by an add layer that combines the two branches. In general, such a design allows for the use of deeper networks to learn more complex features without running into gradient vanishing problems (He et al., 2015). More specifically, as Bronstein et al. (2021) describe this network design encourages the representation of features as a perturbation of the input. Therefore, in our case, the model is encouraged to learn to represent the optical flow of a field rather than to produce a precise representation of the field itself in every layer.

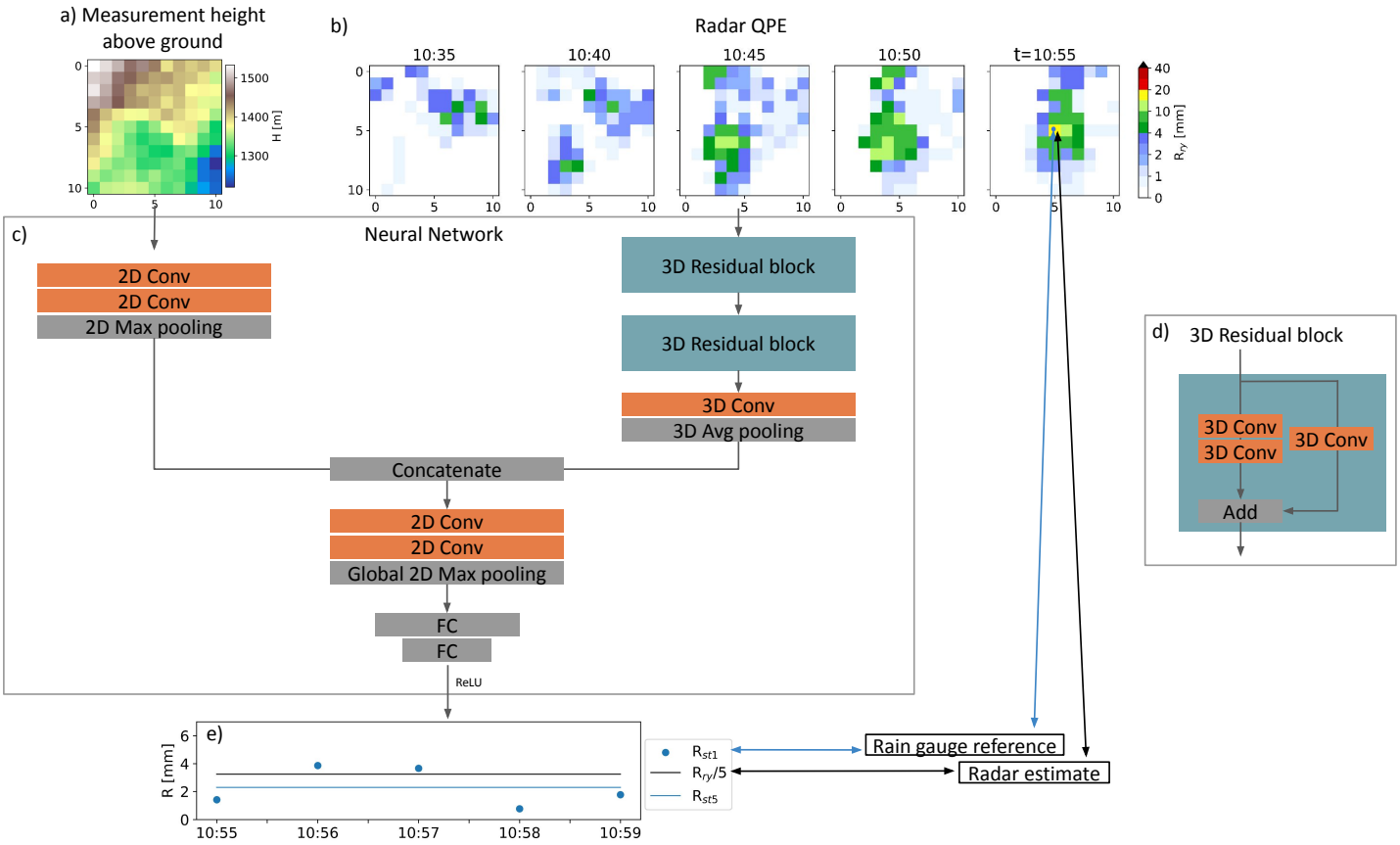

Figure 6.2: Schematic overview of the ResRadNet architecture, the input and the target data. To predict one set of five 1-minute rain gauge observations (e) the model uses the minimal measurement height above ground (a) and five 5-minute radar time-steps (b) with a spatial extent of 11 by 11 pixels ($\equiv$ 11 km), centered at the rain gauge location. The neural network (c) is using convolutional layers (orange) and skip connections with add layers organized as residual blocks (d). The height and rainfall information are processed by separate input branches and later concatenated before being processed by another convolutional block and a final fully connected part. Beneath the 1-minute rain gauge observations R_{st1}, e) also shows the 5-minute average rainfall from the center radar pixel $R_{ry/5}$ and the rain gauge reference R_{st5}. ResRadNet was trained to predict R_{st1} from a) and b) and thus improve $R_{ry/5}$.

We trained the model with a batch size of 1000. After each epoch, the same 100,000 random validation samples were evaluated and the model was saved. We did not use the full validation set to speed up training time. After training, the model with the best validation loss was selected. Due to the skewed distribution of rainfall estimates, there was a multiplicative bias of the form $\overline{R_{st}}/\overline{R_{nn}} = \alpha \neq 1$, where R_{st} is the rain gauge reference, R_{nn} is the neural network prediction and where the $\overline{R}$ indicates the mean of the quantity. We decided to remove this bias by using a bias correction factor $R_{nn} \mapsto \alpha R_{nn}$

which is equivalent to a common mean field bias reduction (McKee and Binns, 2016). We computed this factor using the validation dataset.
In inference mode, the final model was used like a spatio-temporal filter kernel that was applied to the neighborhood of every radar pixel. The computation time for a full radar image was 5 seconds per time step using an NVIDIA Tesla V100 32GB GPU. The performance loss on smaller GPUs is negligible due to the small size of the input data. An approximate 20-second overhead for serialization and de-serialization of the samples could be omitted by parallel computation.

6.2.4 Evaluation

The derived QPE is evaluated for different temporal aggregations (1 minute, 5 minute, and daily) using a set of pixel-wise error metrics which are commonly used in the field (Chen et al., 2021). The Pearson correlation coefficient (PCC) measuring the linear correlation between predicted values and ground truth is defined as

$$PCC = \frac{\sum_n (R_{st} - \overline{R_{st}})(R - \overline{R})}{\sqrt{\sum_n (R_{st} - \overline{R_{st}})^2 \sum_n (R - \overline{R})^2}}, \tag{6.1}$$

where R_{st} is the rain gauge reference and R is the radar (R_{ry}) or neural network (R_{nn}) prediction. Accordingly, the mean squared error (MSE), (normalized) root mean squared error (NRMSE), and normalized bias (NBIAS) are given by

$$MSE = \overline{(R_{st} - R)^2}, \tag{6.2}$$

$$NRMSE = \frac{1}{\overline{R_{st}}} \sqrt{\overline{(R_{st} - R)^2}}, \tag{6.3}$$

and

$$NBIAS = \frac{\overline{R_{st}} - \overline{R}}{\overline{R_{st}}} * 100. \tag{6.4}$$

The $NBIAS$ is independent of the temporal resolution and, therefore, also reflects the relative bias for the whole evaluation period. Let R^+ and R_{st}^+ indicate if radar or rain gauges detect rainfall, i.e. let them be equal to 1 if the value of R (resp. R_{st}) is larger than zero and zero otherwise. Then the mean detection error (MDE) is defined as the frequency of cases, where $R_{st}^+ \neq R^+$. Additionally, we investigate the similarity, or rather maximal dissimilarity, between the two distributions using the Kolmogorov-Smirnov (KS) test. It is defined as the maximum deviation between the cumulative distribution of reference and predicted values.

6.3 Results

6.3.1 Model Selection and Comparison at 1-Minute Resolution

ResRadNet achieved the best MSE on the 100,000 sample validation set after 15 epochs. For this validation data, the bias correction factor α was close to 1 with a value of 1.176 and we applied it to all predicted rainfall estimates presented hereafter. Therefore, the model achieved zero $NBIAS$, a PCC of 0.68, and an MDE of 0.16 when compared to the full 1-minute validation set. A constant prediction of five 1-minute values using the average RADOLAN-RY value achieved worse results with an $NBIAS$ of -24.3%, a PCC of 0.57, and an MDE of 0.26. For the test data results at a 1-minute resolution were similar with a slightly increased $NBIAS$ (see Table 6.1). The 2D histogram comparing R_{nn1} and R_{st1} shown in Fig. 6.3f showed that missed extremes ($R_{nn1} \ll R_{st1}$) were more frequent than false extremes ($R_{nn1} \gg R_{st1}$) considering the higher values of bins close to the x-axis.

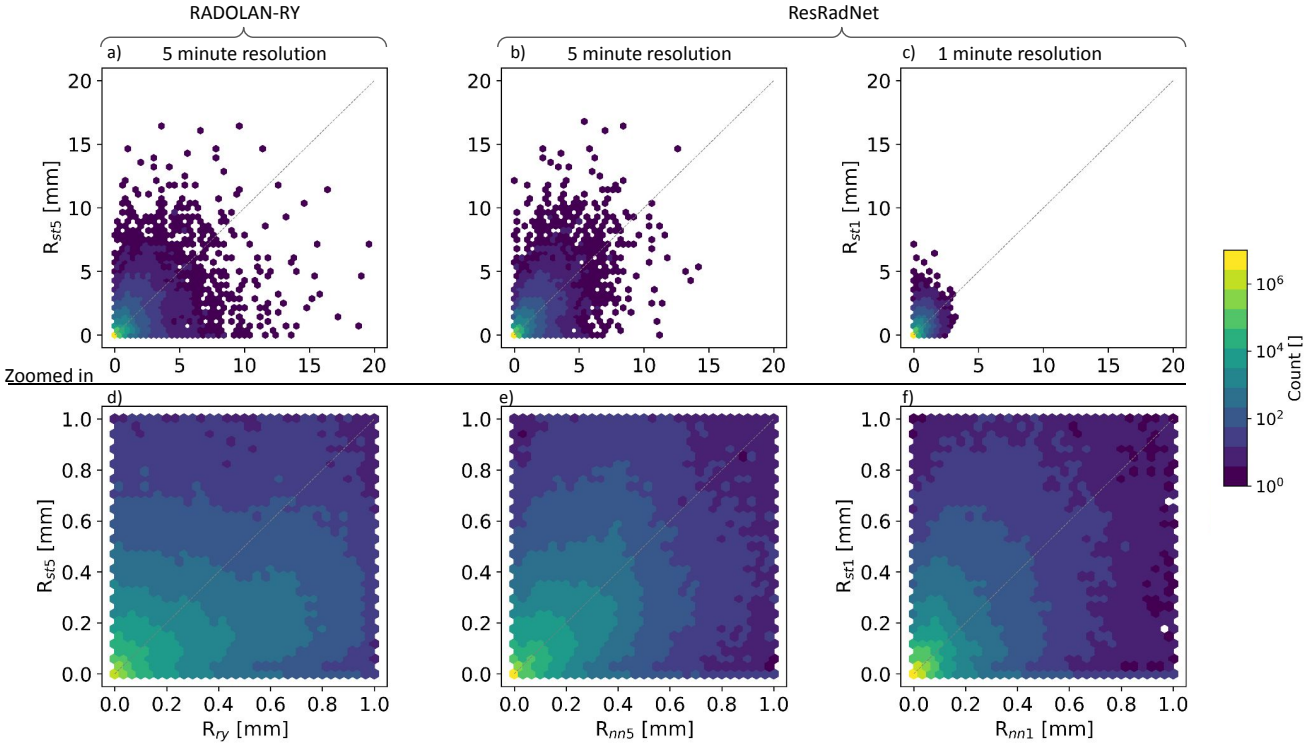

Figure 6.3: 2D-histograms showing the relationship between RADOLAN-RY (R_{ry} in a and d) or ResRadNet (R_{nn} in b,c,e and f) and rain gauge (R_{st}) observations. For a and b the temporal resolution is 5 minutes. For c it is 1-minute. d, e and f are a zoomed in version of a, b and c with smaller bin widths to visualize what is happening for smaller rainfall amounts below 1 mm.

Table 6.1: Results for the 1- and 5-minute resolution of the test data and the independent daily gauges.

Resolution	1-Minute		5-Minute		Daily	
Model	RADOLAN-RY	ResRadNet	RADOLAN-RY	ResRadNet	RADOLAN-RY	ResRadNet
MDE	0.26	0.16	0.23	0.19	0.15	0.15
$NBIAS$	-18.0	4.4	-18.0	4.4	-15.0	7.5
PCC	0.57	0.67	0.63	0.74	0.84	0.86
$RMSE$	0.033	0.027	0.141	0.110	3.275	2.650
KS-Test	0.177	0.067	0.051	0.073	0.044	0.048

6.3.2 Comparison at 5-Minute Resolution

The results of the test data set at a 5-minute resolution are shown in Table 6.1. Both RADOLAN-RY's and ResRadNet's PCC improved compared to the 1-minute results. The KS-test and MDE improved for the radar and worsened for the neural network. The $RMSE$ increased in both cases and the $NBIAS$ was independent of the resolution. Except for the KS-test the ResRadNet achieved better results than RADOLAN-RY.
Fig. 6.3 shows a scatter density comparison of the rain gauge and the radar or neural network predictions. The most obvious improvement of ResRadNet could be observed near the x- and y-axis of the plots (comparing panels a) and b) and for the smaller values comparing panels d) and e). The panels showing ResRadNet data show a higher concentration of points along the diagonal and a lower concentration closer to the axes. While there is still a large uncertainty for extreme values, outliers close to both axes could be reduced. Fig. 6.4 shows a long-term analysis using the test stations and the years 2013 to 2021. It confirms that for the 5-minute resolution, ResRadNet performed better than the radar for all years and all scores. The average improvement was 0.2 for the PCC, 2 for the $NRMSE$, 20% for the $NBIAS$ and 7% for the MDE. The year 2016 showed exceptionally poor PCC and $NRMSE$ for RADOLAN-RY while ResRadNet seemed to be able to correct this issue.

6.3.3 Comparison to Daily Rain Gauge Data

The comparison to the independent set of daily rain gauge measurements confirmed the 5-minute test results. ResRadNet improved the PCC with a value of 0.86 compared to 0.84 for RADOLAN-RY and decreased the $RMSE$ from 3.28 to 2.65. The MDE and KS-Test were very similar for both products. The resolution-independent $NBIAS$ was similar to the 1-minute gauges despite the different station locations (see Tab. 6.1). The long-term evaluation from 2013 to 2021 (excluding the training year 2020) showed that the improvement was very consistent (see Fig. 6.4). The 2021 scores were, again, the highest for both RADOLAN and ResRadNet and the improvement was smaller than in previous years. The $NBIAS$ shows the same temporal dynamics as for the 5-minute data but is slightly higher for both products.

As an additional reference, we compared the operational gauge adjustment routine of the DWD (RADOLAN-RW) and the daily rain gauge data, which was not used for this adjustment (see Fig. 6.4). It could be observed that, at this temporal resolution, the MDE is only marginally better than for RADOLAN-RY or ResRadNet. The $NBIAS$ and $NRMSE$ are lower than for ResRadNet and the PCC is higher with values around 0.9. An exception is the year 2016 when RADOLAN-RW suffers from a similar performance decrease as RADOLAN-RY and does not outperform ResRadNet.

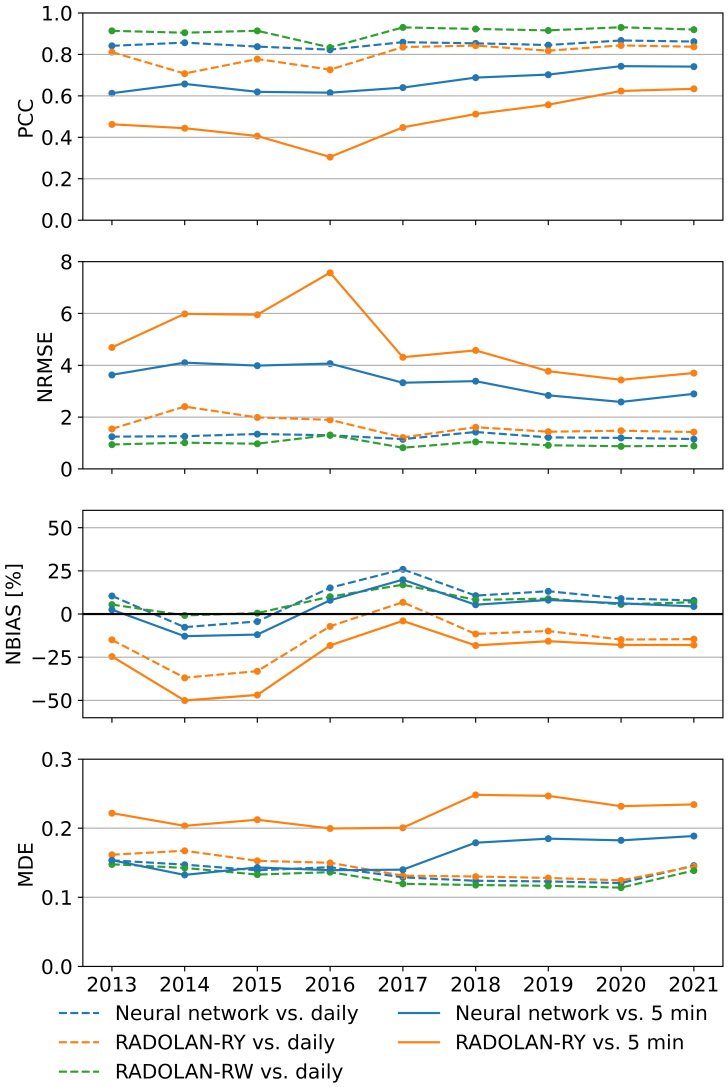

Figure 6.4: Long-term performance comparison of RADOLAN-RY (orange), RADOLAN-RW (green), and ResRadNet (blue). The PCC, $NRMSE$, $NBIAS$ and MDE metrics are shown using the 247 5-minute test (solid lines) and 1138 daily (dashed lines) rain gauges as a reference. RADOLAN-RW is not available at a 5-minute resolution.

6.3.4 Spatial and Temporal Coherence

The spatial and temporal coherence of the rainfall fields can be seen in the maps shown in Fig. 6.5 and the animation provided on Zenodo (Polz, 2023a). Since no radar reference at a 1-minute resolution was available, the analysis of the 1-minute rainfall maps from the neural network was done by visual inspection.

The maps showed that, compared to RADOLAN-RY and RADKLIM-YW, ResRadNet produced smoother structures with more gentle gradients. However, no additional structures with an artificial character like previously discovered in Glawion et al. (2023) were produced and the spatial distribution and connection of rain cells looked reasonable. Fig. 6.5 shows that the neural network decreased the overestimation (less red colors) in the southwest of Germany and around the upper two rain cells in the southeast without an increased underestimation (more blue colors). It even improved the severe misplacement of the small rain cell with the highest intensity close to the border of the study area in the southeast. The temporal coherence of the neural network predictions was judged by visual inspection of the provided animation. It was similar to both radar products, that is, between 5-minute time steps similar discontinuities could be observed. However, the 1-minute neural network resolution removed these discontinuities at the presented scale creating a fluid motion.

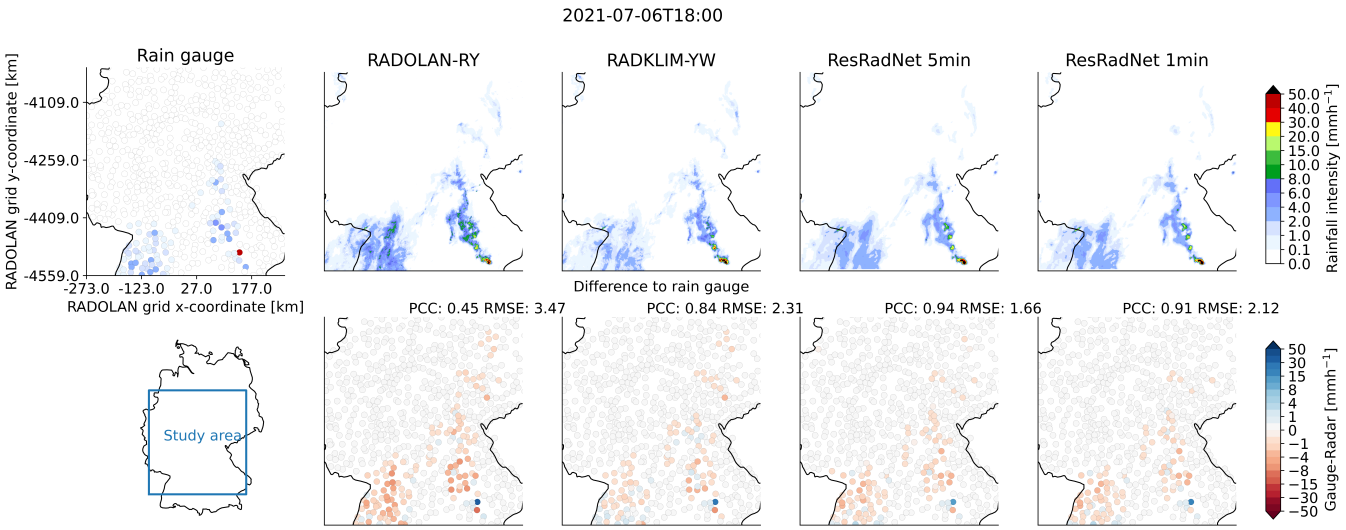

Figure 6.5: Maps of rainfall intensity for 18:00 on 6 July 2021. The upper row shows the 5-minute rain gauges, RADOLAN-RY, RADKLIM-YW, an aggregation of the neural network predictions from 18:00 to 18:05 (ResRadNet 5min), and the single neural network prediction from 18:00 (ResRadNet 1min). The bottom row shows the selected study area in Germany and the difference between the rain gauge value and the grid cell it is contained in using the product on top of the respective map. Red colors indicate an overestimation compared to the rain gauges and blue colors an underestimation.

To investigate if this fluid motion presented a plausible advection scheme we compared a 180-minute aggregation of rainfall fields containing small, fast-moving cells. The products we used were RADOLAN-RY, RADKLIM-YW, an aggregation of every fifth neural network prediction (ResRadNet every 5 min) to simulate a temporal undersampling, an aggregation of all ResRadNet predictions, and the advection corrected RADOLAN-RY. The maps are shown in Fig. 6.6. The two most prominent observations were that the neural network was moving the center of mass of the rainfall field on the bottom towards the center of mass of RADKLIM-YW, which was gauge-adjusted. And, the discontinuities that are visible for RY, YW and "ResRadNet every 5 min" are gone for the "ResRadNet every 1 min" and the advection corrected RADOLAN-RY and RADKLIM-YW predictions. The smoothing of the rain field and the attenuation of extreme values is less pronounced for the neural network than for the Lucas-Kanade advection correction.

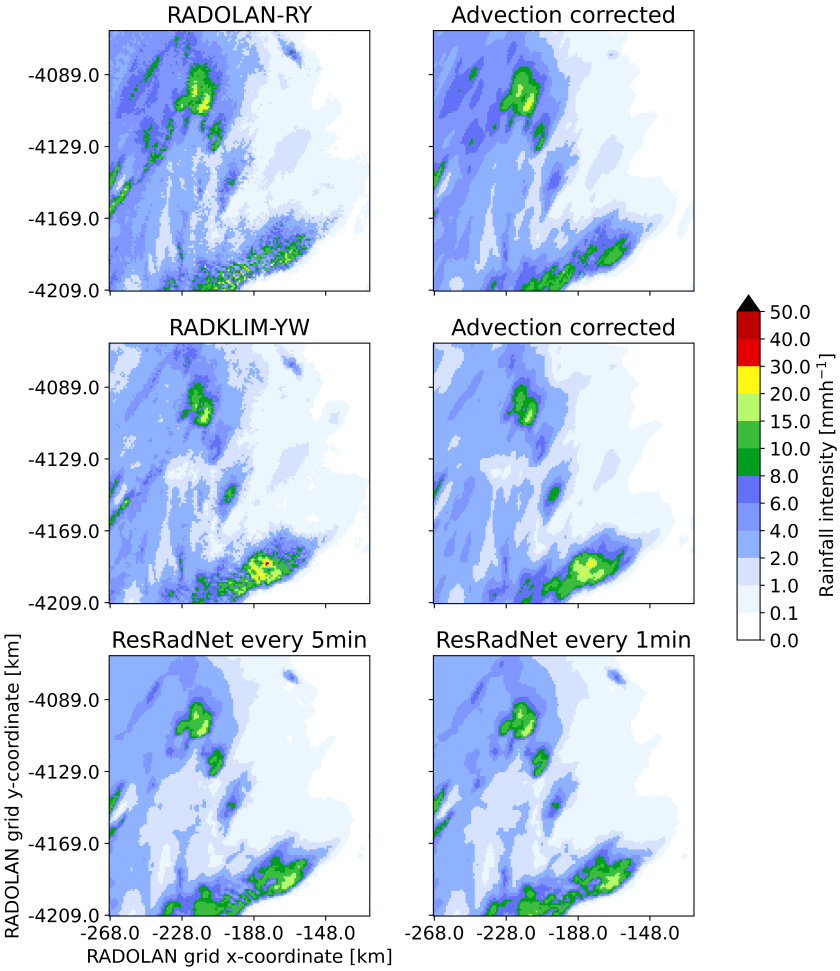

Figure 6.6: Maps of rainfall sum between 16:00 and 19:00 on 6 July 2021. The products shown are from top to bottom: RADOLAN-RY, RADKLIM-YW and the neural network predictions. The left column shows an aggregation of 5 minute instantaneous measurements and the right column shows the advection corrected radar products and all 1-minute time steps.

6.3.5 Influence of Measurement Height Above Ground

To investigate the influence of the measurement height above ground we analyzed all time steps in the test dataset with a reference rainfall amount of at least 0.1mm. The considered variables in those time steps were the minimal measurement height above ground and the absolute error between R_{nn1} and R_{st1}. Fig. 6.7 shows a linear fit to the two variables that minimizes the squared error. The slope of the linear model can be interpreted as a 3.7% increase in the mean absolute error per kilometer.
In addition to the version of ResRadNet that was presented above, we trained a second model in exactly the same way but excluded the height information by omitting the concatenate layer. The results show that the slope and intercept of the linear fit are higher for the model that does not use the height information. Here, the slope represents an increase of 4.6% of the MAE per kilometer. This model also showed a decreased PCC of 0.60 which was computed using the 1-minute test data analogous to the first column of Tab. 6.1 where the model with the height information achieved a PCC of 0.67 and RADOLAN-RY achieved a PCC of 0.57.

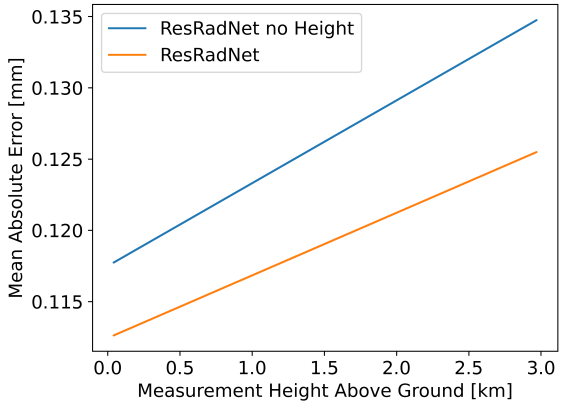

Figure 6.7: Impact of measurement height above ground on rainfall estimates at a 5-minute resolution. The lines show a linear fit to the absolute error between ResRadNet and the rain gauge reference in the test dataset, where the rainfall amount exceeded 0.1mm. The orange line shows ResRadNet using the minimal measurement height above ground as described in the method section and the blue line shows a ResRadNet version where the height information is not given to the model (by omitting the concatenate layer).

6.4 Discussion

We evaluated the ability of a 3D-convolutional neural network to produce spatially and temporally coherent rainfall fields at an increased temporal resolution while reducing biases to rain gauge measurements on the ground.
The spatial structure of the neural network images is smoother than for the unadjusted instantaneous radar rainfall measurements, but less smooth than the Lucas-Kanade advection corrected RADOLAN-RY product. The comparison to the rain gauges shows that the

point-wise accuracy measured by PCC, $RMSE$, and MDE is much better for ResRadNet. We see two possible factors explaining this. On the one hand, a diffusive process due to the advection, collision, and coalescence of raindrops may lead to an actual diffusion of the distribution of rainfall on the ground. On the other hand, the location of the rain gauge in the radar pixel is not considered in this study, which leads to the neural network prediction as a maximum likely rain gauge measurement anywhere within the pixel. This may of course lead to a smoother distribution than the actual radar measurement.
Despite this slight imprecision of the rain gauge location, we could show that the added information about the minimal measurement height above ground significantly increases the model performance, thus increasing the confidence in a meaningful ground adjustment. The temporal coherence of the neural network at a 5-minute resolution is similar to RADOLAN-RY. However, the fluid motion created by the 1-minute predictions indicates that the neural network acts as a reasonable optical flow estimator, even though it was trained to provide pixel-wise predictions and not spatiotemporally consistent fields. RADOLAN-RY achieved a better MDE at a 5-minute resolution than at a 1-minute resolution, while the opposite holds for the neural network. This also shows that the neural network is able to perform useful temporal interpolation.
The scores show in general that the neural network is able to significantly improve the rainfall estimates of RADOLAN-RY. Compared to Seo and Krajewski (2015) we found that our combined advection correction and ground adjustment method does impact and improve light rain estimation. The comparison at the daily resolution showed that the performance gain is a significant step towards the performance of the operational hourly gauge-adjusted product RADOLAN-RW while operating at a much higher temporal resolution. While RADKLIM-YW is gauge-adjusted and climatology-corrected and therefore assumed to be much better than the un-adjusted product, the improvement of the neural network towards RADKLIM-YW is larger than we would have anticipated. Our evaluation strategy using a set of independent rain gauges and a separate time period shows that ResRadNet is transferable to new locations with similar rainfall climatology.
A clear limitation of the neural network is that the uncertainty for extreme values could not be reduced and that they are underestimated. However, as explained in the results section outliers close to the axis (see Fig. 6.4) could be reduced effectively. Another drawback is that the multiplicative correction factor was necessary to correct the $NBIAS$. We assume that both the extreme value underestimation and the correction factor are due to the heavily skewed distribution of precipitation. In the progress of this work, we experimented with log transformations to reduce this problem, but without reasonable success. These drawbacks clearly need to be considered when using ResRadNet. However, the considerably improved scores that proved to be consistent over a period of 8 years and the temporal super-resolution we achieved make our approach very valuable. Additionally, we were able to show that the model can be used to correct for advection-based temporal undersampling when aggregating multiple time steps with less smoothing of gradients than common advection correction techniques.

Finally, we emphasize the real-time applicability of our method: With the low latency of 5 seconds for the production of five 1-minute time-steps of rainfall maps covering all of Germany, an application to a real radar system with measurements performed every 5 minutes is realistic. Additionally, the required data consisting of 25 minutes of available radar data and an estimate for the measurement height above ground should be available to potential users.

6.5 Conclusion

In this study, we evaluated the performance of a 3D-convolutional residual neural network for simultaneous ground adjustment, advection correction, and temporal super-resolution of weather radar images. In an attempt to solve all three issues by training a single neural network, we were able to significantly increase the quality of the gridded country-wide 5-minute radar product RADOLAN-RY.

We were able to show that 3D-convolution in a residual network architecture is a suitable tool to increase the temporal radar resolution of 5 minutes by a short-term prediction of five 1-minute time-steps. While our neural network ResRadNet is only trained to predict one pixel value at a time, it generates continuous predictions for neighboring pixels, resulting in spatially and temporally consistent rainfall fields. Using the raw RADOLAN-RY radar product as a baseline, the model was able to effectively reduce biases between country-wide C-band weather radar rainfall estimates and 247 1-minute and 1138 daily rain gauges on the ground. By using a separate set of rain gauges for training and evaluation we demonstrated the transferability of the network to new locations. With plausible advection schemes and a 1-minute resolution produced by the neural network, an exemplary case study showed that the model acts as a suitable optical flow estimator that can be used for advection correction. Despite the significant improvements that ResRadNet provides, we experienced a common issue with using neural networks for modeling precipitation. The heavily skewed distribution of rainfall leads to an underestimation of extremes and makes an additional mean field bias correction necessary. Logarithmic transformations of the input data did not yield the desired improvements. Higher potential for an accurate representation of extreme values may be given by probabilistic (ensemble) neural networks like generative adversarial networks. A potential point of view is that the current deterministic approach naturally favors the maximum likely predictions over extremes. As for this study, we point out that the produced rainfall fields may be understood as a well-performing maximum likely estimate. Future studies should aim to provide stochastic approaches to extremes.

6.6 Acknowledgments

We want to thank Tanja Winterrath and Elmar Weigel from DWD for the helpful discussions about the DWD radar and rain gauge data. We also want to thank Kai Mühlbauer

and everyone else involved in the ωradlib project for providing the great radar processing software. This study was supported by the German Research Foundation within the DFG research unit RealPEP (Grant CH 1785/1-2), the HGF-Innopool, and the Federal Ministry of Education and Research (Grant 13N14826).

6.7 Appendix of Chapter 6

6.7.1 Model Architecture

Table 6.2: Architecture of ResRadNet. The table resembles the model summary given by the Keras API of Tensorflow version 2.7.0. The total number of trainable parameters was 177,985.

Layer	Type	Output Shape	Kernel size	Activation	nParam	Connected to	Note
input-1 (QPE)	(InputLayer)	(5, 11, 11,1)			0		
conv3D-1	(Conv3D)	(5, 11, 11, 32)	3x3x3	-	896	'input-1'	
relu1	(ReLU)	(5, 11, 11, 32)			0	'conv3D-1'	
conv3D-2	(Conv3D)	(5, 11, 11, 32)	3x3x3	-	27680	'relu-1'	
relu-2	(ReLU)	(5, 11, 11, 32)			0	'conv3D-2'	
conv3D-3	(Conv3D)	(5, 11, 11, 32)	1x1x1	-	64	'input-3'	
add-1	(Add)	(5, 11, 11, 32)			0	'relu-2','conv3D-3'	
conv3D-4	(Conv3D)	(5, 11, 11, 32)	3x3x3	-	27680	'add-1'	
relu-3	(ReLU)	(5, 11, 11, 32)			0	'conv3D-4'	
conv3D-5	(Conv3D)	(5, 11, 11, 32)	3x3x3	-	27680	'relu-3'	
relu-4	(ReLU)	(5, 11, 11, 32)			0	'conv3D-5'	
conv3D-6	(Conv3D)	(5, 11, 11, 32)	1x1x1	-	1056	'add-1'	
add-2	(Add)	(5, 11, 11, 32)			0	'relu-4', 'conv3D-6'	
conv3D-7	(Conv3D)	(3, 9, 9, 64)	3x3x3	ReLU	55360	'add-2'	valid pad
averagepool3D-1	(AveragePooling3D)	(1, 4, 4, 64)	2x2x2		0	'conv3D-7'	
reshape-1	(Reshape)	(4, 4, 64)			0	'averagepool3D-1'	
input-2 (Height)	(InputLayer)	(11, 11, 1)			0		
conv2D-1	(Conv2D)	(10, 10, 4)	2x2	-	20	'input-2'	valid pad
conv2D-2	(Conv2D)	(9, 9, 8)	2x2	ReLU	136	'conv2D-1'	valid pad
maxpool2D-1	(MaxPooling2D)	(4, 4, 8)	2x2		0	'conv2D-2'	
concatenate-1	(Concatenate)	(4, 4, 72)			0	'reshape-1','maxpool2D-2'	
conv2D-3	(Conv2D)	(4, 4, 64)	2x2	-	18496	'concatenate-1'	
conv2D-4	(Conv2D)	(4, 4, 32)	2x2	ReLU	8224	'conv2D-3'	
maxpool2D-3	(MaxPooling2D)	(2, 2, 32)	2x2		0	'conv2D-4'	
conv2D-5	(Conv2D)	(2, 2, 32)	2x2	-	4128	'maxpool2D-3'	
conv2D-6	(Conv2D)	(2, 2, 32)	2x2	ReLU	4128	'conv2D-5'	
globalmaxpool2D-1	(Global MaxPooling2D)	(32)			0	'conv2D-6'	
dense-1	(Dense)	(64)		ReLU	2112	'globalmaxpool2D-1'	
dense-2	(Dense)	(5)		ReLU	325	'dense-1'	

6.7.2 Animation

The animation is available at Zenodo (Polz, 2023c).

Chapter 7

Synthesis

7.1 Discussion of key findings and answers to specific research questions

The main objective of this thesis is the improvement of commercial microwave link (CML) and weather radar-derived precipitation estimation using deep learning approaches. Section 1.2.2 reviewed sources of systematic and random errors in CML and weather radar QPE. To achieve the main goal several studies that target either the quantification of errors (see Chapters 2, 4, and 5) or their correction using deep learning approaches (see Chapters 3 and 6) were designed. While the improvement of weather radar QPE is posed as a post-processing problem (see Chapter 6), the improvement of CML QPE targets the optimization of individual processing steps in the QPE retrieval (see Chapters 2 and 3). In the following discussion the research questions are stated (bold) and a short summarizing answer (italic) is given before going into more details.

SQ1: How do established signal processing techniques compare in a large-scale evaluation of one year of CML data in Germany?

Rain event detection and wet antenna compensation were identified to have the highest impact on the quality of rainfall estimates. It was shown that a self-adapting version of the standard deviation method by Schleiss and Berne (2010) and the rain-rate dependent model of Leijnse et al. (2008) were superior. The resulting CML rainfall estimates showed a very high quality when compared to gauge adjusted weather radar QPE in the warm season.

Any optimization process needs to start with an accurate estimate of the current state and appropriate means to measure progress. To improve rainfall estimates from Commercial Microwave Links (CMLs) the first step was the comparison of state-of-the-art processing methods using a large and representative data set and a first evaluation of country-wide rainfall observations. Designing a suitable evaluation framework to quantify the perfor-

mance of individual processing steps was an additional necessity.
Chapter 2 showed that for liquid precipitation opportunistic rainfall estimates from CMLs are in very good agreement with established rainfall estimates from gauge-adjusted and quality-checked weather radar information which was later backed up by the findings of Chapter 3. The validation in Chapter 2 uses one full year of data from country-wide CML observations to optimize and benchmark existing methods used to process the challenging raw CML data.
The detection of rain events and the compensation of wet antenna attenuation were identified to have the highest impact on the quality of the resulting rainfall estimates. The focus was on methods that do not rely on auxiliary data like weather radar, satellite images, or neighboring CMLs because of potential data scarcity in regions where CMLs are situated in future applications. This turned out to be the right path for an additional study conducted in Burkina Faso, where the first high-resolution rainfall maps in West Africa based on CML data were presented (see further related articles: Djibo et al. (2023)).
A comparison of different versions of the rain event detection algorithm of Schleiss and Berne (2010) showed that the best performance could be achieved by using the self-adapting 80th quantile of the 60-minute rolling standard deviation of the total loss (TL) and an optimized scaling factor. It was preferred over the climatologic threshold of Schleiss and Berne (2010) because of better performance and over a completely optimized threshold because of the need for re-calibration and the resulting dependency on the weather radar reference. The introduction of the MCC as an objective function for optimizing and evaluating rain event detection was crucial, since it is more robust against an imbalance of wet and dry periods, only yielding good values if the detection performs well on both classes. Chapter 3 showed that seasonal differences in performance measures like accuracy are mostly due to a change in climatology, which is a bias that could be avoided using climatology-robust measures like MCC or AUC. By using the MCC, the scaling factor for the standard deviation threshold turned out to be similar for all months.
As possible candidates for WAA compensation methods, the time-dependent model of Schleiss et al. (2013) and the rain-rate dependent model of Leijnse et al. (2008) were considered. Both schemes performed well for higher rainfall intensities, but only the rain rate dependent model could be calibrated to not suppress small rainfall rates completely, which is why it was selected for processing the CML data in Germany. Note that the transferability of these WAA estimation methods remains an open question. In general, WAA compensation is harder to optimize than rain event detection which can be treated as an isolated problem. Due to missing direct measurements of WAA, the effectiveness of the compensation can only be estimated indirectly using its impact on the resulting rainfall estimate. However, the quality of the rainfall estimate depends on multiple factors ranging from miscalibration of other processing steps to DSD uncertainties and systematic errors like wind, humidity, or radiation that can interact with WAA.
Using these optimized processing schemes the performance of one year of hourly, daily, and seasonal rainfall sums from country-wide CML observations was evaluated against

the gauge-adjusted weather radar product RADOLAN-RW. Taking into account that the reference data set is derived from a dense network of C-band weather radars combined with more than 1000 rain gauges in Germany and 200 from neighboring countries, the opportunistic CML-derived rainfall estimates turned out to be of remarkable quality. The Pearson Correlation Coefficient for hourly data was best during the warm period with a value of 0.80 and worst during the cold season with a value of 0.39. This difference in performance was confirmed by an evaluation of additional performance measures. The worse performance in the cold season is due to a large overestimation of small reference rain rates which is likely due to melting snow that creates high specific attenuation which is in accordance with the findings of Overeem et al. (2016b). The computation of performance metrics for different subsets of the rainfall data increased the comparability to previous studies by de Vos et al. (2019) and Rios Gaona et al. (2015) conducted in the Netherlands. The remaining differences in the performance metrics are due to the sampling strategy and local rainfall regime. However, the comparison showed that a similar performance of CML QPE in different countries should be expected.

The following limitations to the study conducted in Chapter 2 led to further research questions SQ2-SQ4 answered by the studies conducted in Chapters 3, 4, and 5:

First, all considered methods for detecting rain events rely on the rolling standard deviation approach. Consequently, they are based on the assumption that rain events can be separated from dry periods solely based on the amount of fluctuation within the considered time window.

Second, missing periods in the CML data were not considered for the evaluation using the assumption that they are caused by failures of the data acquisition system and, therefore, appear randomly. Not considering these periods would, then, not create a bias in the data set. It turns out that this assumption is wrong since the total loss of signal during heavy rain events leads to a higher probability of missing extreme values (see Chapter 4 and SQ3 below).

Third, data quality was controlled by a static filtering approach that excludes single sensors entirely from the data set if the amount of signal fluctuation significantly exceeds an amount that can be justified by the local rainfall climatology. However, when analyzing interpolated rainfall maps over different seasons the same study confirmed that data quality is heterogeneous in space and time with a strong overestimation of some CMLs that was visible throughout the year. Therefore, a dynamic filtering approach is necessary.

SQ2: Are convolutional neural networks able to distinguish between CML signal level patterns caused by rainfall and fluctuations during dry periods?

The hypothesis that rain event detection can be improved significantly by recognizing the signal pattern of rain events rather than measuring the amount of fluctuation proved to be true. The CNN-based deep learning approach performed well in recognizing these patterns, significantly reduced the detection error, and improved the resulting rainfall estimates.

Chapter 3 was concerned with dynamic quality control of CML data by separating wet and dry periods in the CML attenuation time series and the resulting improvement of CML QPE. As identified by Chapter 2 fluctuations during dry periods present a major challenge for CML rainfall estimation. Previously, the assumption made to separate wet and dry periods is that fluctuations of a significant magnitude are only due to attenuation caused by rainfall (Schleiss and Berne, 2010). The aim was to challenge this assumption and replace it with the hypothesis that rainfall-induced attenuation creates particular patterns in the CML attenuation time series that can be distinguished from fluctuations caused by phenomena not related to rainfall. While it was known that there were significant fluctuations during dry periods it remained unclear if these patterns can be distinguished from rainfall-induced attenuation (van Leth et al., 2018).

A 1D-Convolutional neural network designed for pattern recognition was applied to test the hypothesis. The method was chosen, because the feature extraction part of the neural network is a location equivariant function. That is, rain events in a time sequence could be recognized independent of their precise location and only depending on the pattern. The time-series pattern recognition task is a high-dimensional learning problem where each time step of the fixed-length sequence presented to the neural network acts as a feature dimension. Utilizing this high-dimensional feature space was assumed to be superior to, for example, reducing the complexity of the sequence to a single standard deviation value. The CNN efficiently reduces the dimensionality of the problem which is essentially achieved by the convolution algorithm which acts locally on the time series and is able to discriminate redundant information in the sequence.

The results showed that the 1D-CNN approach was able to significantly outperform the reference method by Schleiss and Berne (2010) which was optimized in Chapter 2. The trained CNN was able to generalize well to sensors and time periods not included in the training dataset correctly classifying, on average, 76% of all wet and 97% of all dry periods. For rain rates higher than 0.6 mmh^{-1} more than 90% were correctly detected. An increasingly worse performance for low rain rates was likely caused by a detection limit of the CMLs and not by a failure of the CNN-based rain event detection.

The reduction in false wet and missed wet events also led to a significant reduction of false (51%) and missed (27%) accumulated rainfall over the analyzed period. The correctly classified rainfall was increased by 4%, while, at the same time, increasing the Pearson Correlation to the reference compared to the standard deviation method. Considering that the added rainfall stems from "hard to detect" rain events with reduced quality, this shows that the temporal localization of rain events was improved.

In summary, CNNs proved to be a robust tool for classifying rain events in CML attenuation time series and improved the quality of CML rainfall estimates. The deep learning approach was used to replace one processing step in the retrieval algorithm in that case. To test the potential of CNNs for improving reflectivity-based radar rainfall estimates a post-processing approach was chosen (see Chapter 6 and the answer to SQ5 below).

SQ3: How do blackouts impact CML rainfall estimates and how frequent are they?

It was found that only a small fraction of the annual rainfall is missed due to blackouts, but that the chances of missing rainfall extremes are very high introducing a strong bias to the dataset. The occurence of blackouts also largely exceeded the climatologically expected value.

In previous studies comparing CML rainfall estimates at the link level, periods with missing data for either CML or reference have mostly been excluded from the analyses. However, when extreme events are analyzed at a high temporal resolution, it becomes apparent that an increased amount of CMLs drop out when peak intensities should be reached. This is unfortunate, because capturing extremes is one of the most important aspects of rainfall estimation.

In Chapter 4 the occurrence of blackouts in CML data and their impact on CML rainfall estimates were analyzed as well as their expected frequency of occurrence according to a radar-derived attenuation climatology. While only 1% of the annual rainfall sum is missed during blackouts, the likelihood of missing rainfall estimates rises with the rainfall intensity measured by the radar reference. It reached up to 40% for 5-minute, path averaged radar rainfall rates of 100mmh^{-1} and higher which can occur multiple times per year.

Overall, the average CML in the dataset missed 20 minutes of extreme precipitation each year. While this was found to be in accordance with the planned margin for mobile network stability, which allows for up to 60 minutes of blackout per year, the missing extreme values are clearly a more serious problem for hydrometeorological applications. The likelihood of missing rainfall extremes is increased for longer CMLs, an unexpected finding that suggests that in extreme precipitation the reliability of shorter CMLs is higher. A too-small increase of the dynamic range of the signal power with CML length was identified as a potential reason.

The comparison to the radar-derived attenuation climatology showed that the observed occurrence of blackouts is much higher than expected. This coincides with the findings of Schleiss et al. (2020) that weather radar QPE can underestimate extreme values. Therefore, radar data has proven to be ineffective for planning a CML network that is robust against blackouts. However, the effect that longer CMLs experience more blackouts was also visible in this comparison indicating that the effect is not due to an incorrect detection of blackouts.

A comparison to the two-meter surface temperature along the CML path derived from the ERA5 reanalysis dataset showed that less than 20% of all blackouts occur below 4°C where mixed-type precipitation can be expected. This excludes the possibility that non-liquid precipitation that is not detected by the radar largely contributes to the number of observed blackouts when melting while falling closer to the ground. This increases the

confidence that the attribution of blackouts to heavy precipitation is correct.
In a later study, it was shown that even simple interpolation of blackout gaps improves the rainfall estimates from CMLs considerably (see further related articles: Graf et al. (2023)). An interpolation using fill values that assume the minimum observable received signal level performed best. This observation agrees with the assumption that the true total loss during blackouts surpasses the detection limit of the CMLs receiver.

SQ4: How robust is manual quality control of CML data and how does it affect rainfall estimation?

While the robustness of manual quality control proved to be low due to ambiguity in quality flags from different experts the possible improvement when excluding flagged data from the analysis was significant.

The separation of the attenuation time series into wet and dry periods is the state-of-the-art quality control mechanism for CMLs. An aspect that is missed in rain event detection methods is the possibility of additional disturbances during rain events and a resulting error in the rainfall estimate. A number of potential causes for signal anomalies not related to rain have been proposed in the introduction to this thesis. It would be desirable to have compensation schemes for these anomalies like there are for wet antenna attenuation. However, suitable reference data is often missing and the sporadic occurrence of such anomalies does not allow for global application of a correction like it is performed for wet antenna attenuation.
In Chapter 5 CML signal anomalies were detected using expert knowledge and their impact on rainfall estimates was analyzed. CML data was quality-checked in a manual flagging process using a custom-built data explorer with pre-defined anomaly classes. To see if such an expert classification is robust, the flags of four individuals were compared and the degree of uncertainty between experts was investigated. The dataset was required to be relatively small to be able to be processed within one day of work. It covered data from 20 CMLs and the months of March, May, and July 2020. Only CMLs with a proportionally high amount of fluctuation were considered. The TL time series was used as the primary indicator for anomalies. The results showed that 11% of all time steps of the considered CMLs were flagged by at least one expert. The most frequently selected class contained fluctuations, that is, random fluctuations of the signal (42.9%). The second largest class contained patterns created by the formation of dew, or ice, on the antenna. This class showed a higher agreement with the provided dew point indicator derived from the ERA5-Land reanalysis. Sudden jumps in the time series or phenomena that did not fit into the before-mentioned classes were found seldom. The agreement between experts was highest for the fluctuation class.
To analyze the impact of anomalies on the quality of the CML QPE a comparison to gauge-adjusted and climatology-corrected radar rainfall at a 5-minute resolution was con-

ducted. To process the data, the routines developed in Chapters 2 and 3 were used. The improvement when anomalies flagged by at least one expert were discarded was very high, raising the Pearson Correlation Coefficient from 0.61 to 0.70 and reducing the relative bias by 40%. In almost all cases, the removed anomalies produced an overestimation of rainfall and a low correlation with the reference.

SQ5: Are convolutional neural networks capable of simultaneous temporal super-resolution, ground-adjustment, and advection-correction of radar rainfall estimates?

The residual neural network architecture used to post-process the radar QPE showed a high-performance and significantly reduced several error estimates. It was shown that the spatial and temporal consistency of produced rainfall maps is very high and that the method acts as a suitable optical flow estimator. A limitation of the model was an underestimation of extreme values.

This question was answered in Chapter 6, where ResRadNet, a residual 3-D convolutional neural network, was applied to post-process radar rainfall estimates. This post-processing routine targeted a minimization of the mean squared error compared to rain gauges with a 1-minute temporal resolution. However, the method itself was designed to address specific sources of systematic errors such as vertical and horizontal variability of rainfall when the vertical distance to ground sensors is large and temporal sampling errors. The reduction of the overall error was very successful increasing the Pearson Correlation Coefficient from 0.63 to 0.74 and reducing the root mean squared error by 22% and the normalized mean bias by 20%. A long-term comparison of 8 years of data from 247 1-minute and 1138 daily rain gauges showed that the performance of the method is stable in space and time and that a transfer to new locations is possible within the study area.
A reason for the good performance of the method was the use of residual blocks in the neural network architecture that emphasizes learning to represent the optical flow of a field rather than the production of a precise representation of the field itself in every layer. A comparison to the Lucas-Kanade algorithm showed that this prior assumption was correct, as the neural network proved to be a good optical flow estimator. It is therefore a suitable method for advection correction, that is, the reduction of temporal sampling errors by an imputation of intermediate timesteps.
The largest limitation of ResRadNet is the underestimation of extreme precipitation. There are two potential reasons that may be addressed in future studies: The first is the imbalance of frequent zeros and small rain rates in the training dataset compared to the sparse occurrence of extremes. The second is the deterministic nature of the method that targets a maximum likelihood for predicted rainfall and, therefore, suppresses extremes. A probabilistic approach that produces a well-calibrated ensemble similar to the one used in the related study Glawion et al. (2023) is the logical next step. Transferability

of the method to different climate zones may not be given and is yet to be tested. This also holds for a different radar system and the application of a different Z-R relation that may introduce a different bias to the ResRadNet model input. In general, users of the method are required to test for a covariate shift (input distribution changes while target distribution stays) or concept drift (input distribution stays while target distribution changes) when applied to a different dataset.

7.2 Answers to Overarching Research Questions

OQ1: Can systematic errors of CML and weather radar QPE be quantified, attributed to causes, and corrected?

The results showed that there is a high potential to reduce systematic errors with respect to a reference, and that the error can be quantified in a suitable evaluation framework, but that the attribution to causes is very difficult in most cases.

The long answer starts with calling to mind the fact that both types of rainfall sensors that were considered in this thesis have a complex error structure and it is difficult to disentangle the individual causes of errors. The opportunistic nature of CMLs increases the difficulty because only the measured signal levels are available to researchers in order to control the quality of rainfall estimates. One particular error in CML rainfall estimates is due to an incorrect attribution of attenuation to rainfall. Chapter 2 identified erroneous rain event detection and wet antenna attenuation to have the highest impact on rainfall estimates and the largest potential for correction. The study focused on time-series-based methods for rain event detection and wet antenna attenuation correction without auxiliary data to improve the applicability in data-scarce regions. The country-wide evaluation quantified the overall error compared to a gauge-adjusted weather radar product with remarkable similarity in the warm season. A large overestimation of annual precipitation sums could be attributed to the cold season. Highly erroneous CMLs could be filtered statically to exclude their error in produced rainfall maps.
Chapter 3 improved the detection of rain events in CML attenuation data using deep learning-based pattern recognition to correct errors caused by large fluctuations during dry periods. An attribution of the dry fluctuations to other causes was not possible since the weather radar-based wet-dry reference did not allow for this kind of classification. However, the large improvement in the detection of rain events compared to previously used methods shows that it is possible to distinguish between rain events and anomalies during dry periods that disturb the measurement. The CNN approach reduced the rainfall estimated falsely due to such anomalies by more than 50% showing that their error can be corrected to a large part.
Further quantification of different types of signal patterns and an effort to relate them to atmospheric phenomena are presented in Chapter 5. The quantification of anomalies

relied on the judgment of experts. However, there is a high ambiguity between experts which hinders quantification, attribution, and correction of anomalies. An attribution to causes such as dew was only partly successful. The example time series provided in Fig. 1.4 shows how different phenomena contribute to one pattern classified as anomalous. The automated detection of manually flagged anomalies using ANNs is the subject of ongoing research.
Chapter 4 treated "blackouts", another systematic error of CMLs that is easier to quantify and where the attribution to a cause is less ambiguous. Yet, it is even more surprising that previous studies completely neglected this error. The developed detection algorithm showed a high agreement of blackouts with high rainfall intensities measured by weather radars which shows that the cause of the missing data was indeed strong rainfall-induced attenuation. However, it is unclear if the error can be corrected by reconstructing the missing extremes. Graf et al. (2023) shows that a large improvement can be made by filling blackout gaps with the highest measurable attenuation values. Further research is needed to investigate if any skillful reconstruction of extreme rain events beyond the highest measurable attenuation is possible.
The post-processing approach to correct systematic errors of weather radar rainfall estimates described in Chapter 6 is in contrast to the CML approach, where individual phenomena and processing steps are treated separately. The focus was on simultaneous advection correction and ground adjustment which is not a cause, but a consequence of a multitude of potential errors. The residual neural network approach was successful in reducing the overall error (e.g. RMSE) and in increasing the temporal resolution to effectively reduce advection-driven sampling errors. However, not only the strength of the deep learning method in increasing the predictive skill could be observed. A common weakness of such models to not allow for the discovery of causal relations between model input and output was visible.

OQ2: Are artificial neural networks a robust and transferable tool to provide quantitative rainfall estimates?

Both applications showed that the methods can be transferred to new locations and future time periods within the same climatic region without losing skill. While the CML rain event detection performed better for higher rainfall intensities, the opposite was true in the radar application. Despite these limitations, the methods performed very well in the respective tasks.

Transferability was defined as model predictions that do not lose skill in future time periods and new locations. Both applications of neural networks in Chapters 3 and 6 have shown a very high transferability with almost no loss of skill outside the temporal and spatial training domain. Chapter 3 showed that despite training on a small subset

of the available CMLs, a good generalization to new sensors could be achieved. This is especially remarkable because individual CMLs can exhibit very different characteristic behavior over longer time periods.

Chapter 6 showed that the correction of radar estimates at a subset of all pixels that contain rain gauges leads to an equal gain in skill at all pixels in the domain covered by the radar network. It was shown that, despite individual predictions being made for a single pixel, the transition to neighboring pixels is free of discontinuities and shows a high spatial autocorrelation. The improvement was also visible when comparing the results to a long time series (8 years) of daily rain gauges which were a completely independent set of observations.

The robustness of the deep learning approach is judged in terms of correctly classifying and predicting extremes that occur with a low frequency. While the detection of extreme rain events in CML data was achieved with a very high probability the prediction of extreme radar rainfall was less skillful. A potential reason for the missing extremes in the ResRadNet predictions was that the regression problem using the mean squared error as an objective function was not optimal. In both cases, the skewed distribution of the rainfall data was problematic for model training. The solution was straightforward for the classification task where a 1:1 ratio of rainy and non-rainy periods in the training data was achieved. For ResRadNet such a sampling approach did not improve the skill of the model. However, it was also observed that established advection correction methods based on the Lucas-Kanade algorithm lead to an even higher suppression of extreme values. When comparing the two deep learning applications, where rain event detection is a classification task and the correction of radar rainfall is a regression task, it is not surprising that in the first task, more extreme signals lead to a higher model skill, whereas in the second case, more extreme signals lead to more cautious predictions to minimize the average error.

However, a generative approach that produces ensembles of solutions shows much more skill in correctly predicting extreme values as shown in Glawion et al. (2023).

In summary, the two conducted studies showed that neural networks trained on past data and a separate set of sensors can be transferred to future times and new sensor locations without a significant loss of skill. The handling of extreme values was very good in the CML case and, in the radar case, not worse than previously established methods. The results indicate that suitable objective functions for rainfall intensity distributions are yet to be established.

7.3 Conclusion and Outlook

The studies presented in this thesis show that state-of-the-art rainfall products based on CML or weather radar measurements of attenuation and reflectivity can be outperformed using deep learning approaches. In addition to the performance gain best practices and important limitations for using deep learning for rainfall estimation have been identified. For CMLs, the detection of rain events and the compensation of attenuation caused by other effects such as antenna wetting have been identified and successfully mitigated. The application of a convolutional neural network for deep learning-based pattern recognition outperformed the reference method and reduced missed or false rain events greatly enhancing the trust in CML QPE. Progress was also made in the detection and compensation of blackouts showing that while only a small fraction of the annual rainfall is missed, they significantly hinder the measurement of the most extreme precipitation events. However, they can be localized and a lower bound for the missed rain rate can be computed. A reconstruction of missing data in the measurement using a generative deep learning approach is the subject of current research.
The detection of signal anomalies caused by other atmospheric phenomena has proven to be difficult due to a lack of suitable reference data. This is even more so the case in parts of the world where climate observations are much more sparse compared to Germany. A proposed solution is manual quality control of the dataset, but the presented study in Germany showed that there is ambiguity in the expert flagging of the data. Despite this ambiguity, the improvement that can be made when excluding flagged data from the set of valid observations is large.
Overall, the performance of the improved CML QPE shows a remarkable quality compared to gauge-adjusted weather radar QPE which has a much longer history of research and incremental improvements.
The improvement of weather radar QPE using a deep learning approach was also very successful. Instead of improving a particular processing step in the QPE retrieval a post-processing approach using a residual 3D-CNN was chosen. The prediction of five future 1-minute time steps of a rain gauge on the ground could be achieved with a high performance which showed that the neural network worked as a suitable optical flow estimator. A temporal sampling error could be corrected achieving a higher quality product compared to traditional methods.
The transferability of the deep learning approaches proved to be high when they were applied to new locations or sensors in the same climatic region. It is an open question how the trained neural networks would respond to a covariate shift in a different climate. For rain event detection (see SQ2) the detection of more extreme rainfall seems unproblematic. Larger issues could be caused by new sources of signal anomalies that are caused by phenomena not present in Germany, for example, dust storms that occur during the West African dry season. The correction of radar estimates is very much bound to the radar input used for training which is biased by the processing of the raw reflectivity data and

the radar calibration (e.g. the radar constant used to derive reflectivity).
Despite these problems that may occur when applying the trained networks to new climatic regions this thesis provides a clear proof of concept for the presented deep learning applications. There is no reason to assume that re-training the methods in new regions should fail other than due to missing training data.
The real-time applicability of the presented methods is very high since the computations are efficient and fast enough to not introduce additional latency in the provision of rainfall products. A 1-minute ground adjustment of weather radar data was achieved which is a significant improvement over the hourly gauge adjustment performed by the German meteorological service.
Future work will focus on enabling the full potential of the improvements presented in this thesis by investigating the automated detection of CML signal anomalies and the imputation of missing data using deep learning techniques. Additionally, the transferability to new climate zones has to be ensured by collecting more CML data around the globe and by using new verification techniques such as satellite observations. A first effort was made in Djibo et al. (2023) and Graf et al. (2024).
The next step in the improvement of radar QPE is the application of a generative deep learning approach to polarimetric radar QPE. The suitability of a probabilistic approach for rainfall extremes has been demonstrated in Glawion et al. (2023) and the superiority of polarimetric QPE is shown by Chen et al. (2021). The potential to overcome effects like partial beam blockage and to reduce DSD uncertainty is promising in combination with the superresolution ground adjustment presented in Chapter 6 and should be pursued once a large base of polarimetric radar data is available.
The expected impact of the published studies presented in this dissertation targets closing the observational gap for rainfall estimation as demonstrated by the study conducted in Burkina Faso which used processing routines developed for the German dataset (Djibo et al., 2023). The study presents data from a dense network of CMLs located of the city area of Ouagadougou which were used to achieve high-resolution rainfall maps where previously only one daily rain gauge was available. The study highly benefits from the experience and developed processing techniques that are presented in this work. The processing of the data to achieve the presented quality standard would not have been possible five years ago.
The availability of CML data for operational rainfall estimation is an urgent issue. While a large part of the global population lives in areas with CML coverage, only a tiny fraction of the data is being recorded. One problem is that network providers do not benefit directly from providing the CML data which is due to the lack of a business model. Given the public interest in building capabilities to adapt to climate change this must not remain a limiting factor. Hopefully, the demonstration of the potential of CMLs and weather radars will inspire researchers, stakeholders, and policymakers to acquire the necessary CML data that is currently lost and to build new infrastructure for conventional sensors like rain gauges and weather radars.

List of Symbols

Symbol	Description	Unit
$\mathbb{N}$	Natural numbers	
$\mathbb{R}$	Real numbers	
∇	Gradient	
sup	Supremum	
$\gg$	Much greater	
$\ll$	Much less	
A	Total attenuation	dB
C	Radar constant	-
f	Frequency	GHz
k	Specific attenuation	$\mathrm{dB\,km^{-1}}$
K_W	Dielectric factor	-
P	Power	W
PIA	Path integrated attenuation	dB
R	Rain rate	$\mathrm{mm\,h^{-1}}$
T	Temperature	°C
Z	Reflectivity	$\mathrm{mm^6\ m^{-3}}$
ν	Fall speed	$\mathrm{m\,s^{-1}}$
σ_{ext}	Extinction cross section	$\mathrm{cm^2}$
σ_{bsc}	Back-scattering cross section	$\mathrm{cm^2}$

List of Abbreviations

ACC	Accuracy
AI	Artificial intelligence
ANN	Artificial neural network
API	Application programming interface
ATPC	Automatic transmit power control
AUC	Area under curve
CNN	Convolutional neural network
CML	Commercial microwave link
CV	Coefficient of variation
DWD	German Meteorological Service
DSD	Drop-size distribution
ECMWF	European Center for Medium-range Weather Forecasting
ERA5	Fifth generation of European reanalysis
FN	False negative
FP	False positive
GPM	Global Precipitation Measurement Mission
GPU	Graphics processing unit
GSM	Groupe Speciale Mobile
IDW	Inverse distance weighting
IEEE	Institute of Electrical and Electronics Engineers
IFRC	International Federation of Red Cross and Red Crescent Societies
IFS	Integrated forecasting system
IMERG	Integrated multisatellite retrievals
ITU	International Telecommunication Union
KIT	Karlsruhe Institute of Technology
KS-test	Kolmogorov-Smirnov test
LSTM	Long short-term memory
MAE	Mean absolute error
MCC	Matthews Correlation Coefficient
MDE	Mean detection error

MLP	Multilayer perceptron
MSE	Mean squared error
MSG	Meteosat Second Generation
NBIAS	Normalized relative bias
nD	n-dimensional
NRMSE	Normalized root mean squared error
PARSIVEL	Particle size velocity
PCC	Pearson Correlation Coefficient
QPE	Quantitative precipitation estimation
QPF	Quantitative precipitation forecasting
QPN	Quantitative precipitation nowcasting
Radar	Radio detection and ranging
RADKLIM	Radarklimatologie
RADOLAN	Radar Online Aneichung
RealPEP	Near realtime quantitative precipitation estimation and prediction
ReLU	Rectified linear unit
ResNet	Residual neural network
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
TN	True negative
TL	Total path loss
TP	True positive
TRSL	Transmitted minus received signal level
TSL	Transmitted signal level
RMSE	Root mean squared error
ROC	Receiver operating characteristic
RSD, RSTD	Rolling standard deviation
RSL	Received signal level
UNFCCC	United Nations Framework Convention on Climate Change
UTC	Coordinated Universal Time
WAA	Wet antenna attenuation
WMO	World Meteorological Organization
WSR-88D	Weather Surveillance Radar – 1988 Doppler

List of Figures

List of Tables

Bibliography

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, 2016. URL https://doi.org/10.48550/arXiv.1603.04467.

Adler, R. F., Kidd, C., Petty, G., Morissey, M., and Goodman, H. M. Intercomparison of Global Precipitation Products: The Third Precipitation Intercomparison Project (PIP-3). Bulletin of the American Meteorological Society, 82(7):1377–1396, July 2001. doi: 10.1175/1520-0477(2001)082⟨1377:iogppt⟩2.3.co;2.

Agapiou, A. Remote sensing heritage in a petabyte-scale: satellite data and heritage Earth Engine© applications. International Journal of Digital Earth, 10(1):85–102, Jan. 2017. doi: 10.1080/17538947.2016.1250829.

Aggarwal, C. C., Hinneburg, A., and Keim, D. A. On the Surprising Behavior of Distance Metrics in High Dimensional Space. In Van den Bussche, J. and Vianu, V., editors, Database Theory — ICDT 2001, Lecture Notes in Computer Science, pages 420–434, Berlin, Heidelberg, 2001. Springer. doi: 10.1007/3-540-44503-x_27.

Akoglu, H. User's guide to correlation coefficients. Turkish Journal of Emergency Medicine, 18(3):91–93, Sept. 2018. doi: 10.1016/j.tjem.2018.08.001.

Allan, R. P. and Soden, B. J. Atmospheric Warming and the Amplification of Precipitation Extremes. Science, 321(5895):1481–1484, Sept. 2008. doi: 10.1126/science.1160787.

Ams. Glossary of Meteorology, 2021a. URL https://glossary.ametsoc.org/wiki/Rain.

Ams. Precipitation - Glossary of Meteorology, 2021b. URL https://glossary.ametsoc.org/wiki/Precipitation.

Anagnostou, E. N. and Krajewski, W. F. Real-Time Radar Rainfall Estimation. Part I: Algorithm Formulation. Journal of Atmospheric and Oceanic Technology, 16(2):189–197, Feb. 1999. doi: 10.1175/1520-0426(1999)016⟨0189:rtrrep⟩2.0.co;2.

Andersson, J. C. M., Olsson, J., van de Beek, R. C. Z. ., and Hansryd, J. OpenMRG: Open data from Microwave links, Radar, and Gauges for rainfall quantification in

Gothenburg, Sweden. Earth System Science Data, 14(12):5411–5426, Dec. 2022. doi: 10.5194/essd-14-5411-2022.

Atlas, D. and Ulbrich, C. W. Path- and Area-Integrated Rainfall Measurement by Microwave Attenuation in the 1–3 cm Band. Journal of Applied Meteorology and Climatology, 16(12):1322–1331, Dec. 1977. doi: 10.1175/1520-0450(1977)016⟨1322:paairm⟩2.0.co;2.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A. F., and Nielsen, H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics, 16(5):412–424, May 2000. doi: 10.1093/bioinformatics/16.5.412.

Bao, L., Hansryd, J., Danielson, T., Sandin, G., and Noser, U. Field trial on adaptive modulation of microwave communication link at 6.8GHz. In 2015 9th European Conference on Antennas and Propagation (EuCAP), pages 1–5, Apr. 2015.

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N. Routineverfahren zur online-aneichung der radarniederschlagsdaten mit hilfe von automatischen bodenniederschlagsstationen (ombrometer). Abschlussbericht, Selbstverlag, Deutscher Wetterdienst Offenbach, 2004. URL `https://www.dwd.de/DE/leistungen/radolan/radolan_info/abschlussbericht_pdf.pdf?__blob=publicationFile&v=2`.

Bellman, R. Dynamic programming. Princeton Univ. Pr, Princeton, NJ, 1984. ISBN 978-0-691-07951-6.

Berg, P., Moseley, C., and Haerter, J. O. Strong increase in convective precipitation in response to higher temperatures. Nature Geoscience, 6(3):181–185, Mar. 2013. doi: 10.1038/ngeo1731.

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

Berne, A. and Uijlenhoet, R. Path-averaged rainfall estimation using microwave links: Uncertainty due to spatial rainfall variability. Geophysical Research Letters, 34(7), Apr. 2007. doi: 10.1029/2007gl029409.

Blettner, N., Chwala, C., Haese, B., Hörning, S., and Kunstmann, H. Combining Commercial Microwave Link and Rain Gauge Observations to Estimate Countrywide Precipitation: A Stochastic Reconstruction and Pattern Analysis Approach. Water Resources Research, 58(10):e2022WR032563, 2022. doi: 10.1029/2022wr032563.

Bottou, L., Curtis, F. E., and Nocedal, J. Optimization Methods for Large-Scale Machine Learning. SIAM Review, 60(2):223–311, Jan. 2018. doi: 10.1137/16m1080173.

Brauer, C. C., Overeem, A., Leijnse, H., and Uijlenhoet, R. The effect of differences between rainfall measurement techniques on groundwater and discharge simulations in a lowland catchment. Hydrological Processes, 30(21):3885–3900, 2016. doi: 10.1002/hyp.10898.

Bringi, V. N., Chandrasekar, V., Balakrishnan, N., and Zrnić, D. S. An Examination of Propagation Effects in Rainfall on Radar Measurements at Microwave Frequencies. Journal of Atmospheric and Oceanic Technology, 7(6):829–840, Dec. 1990. doi: 10.1175/1520-0426(1990)007⟨0829:aeopei⟩2.0.co;2.

Bronstein, M. M., Bruna, J., Cohen, T., and Veličković, P. Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges, May 2021.

Bundesnetzagentur. Bundesnetzagentur: Tätigkeitsbericht Telekommunikation 2016/2017, Tech. rep., Report 2016/2017, Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen, Bonn, 2017. URL https://www.bundesnetzagentur.de/SharedDocs/Downloads/DE/Allgemeines/Bundesnetzagentur/Publikationen/Berichte/2017/TB_Telekommunikation20162017.pdf?__blob=publicationFile&v=3.

Chen, H. and Chandrasekar, V. Deep learning for surface precipitation estimation using multidimensional polarimetric radar measurements. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pages 359–362, July 2021. doi: 10.1109/igarss47720.2021.9554847.

Chen, J.-Y., Trömel, S., Ryzhkov, A., and Simmer, C. Assessing the Benefits of Specific Attenuation for Quantitative Precipitation Estimation with a C-Band Radar Network. Journal of Hydrometeorology, 22(10):2617–2631, Oct. 2021. doi: 10.1175/jhm-d-20-0299.1.

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E. Assessing objective techniques for gauge-based analyses of global daily precipitation. Journal of Geophysical Research: Atmospheres, 113(D4), 2008. doi: 10.1029/2007jd009132.

Chollet, F. Keras: Deep Learning for humans, Jan. 2022. URL https://github.com/keras-team/keras.

Chwala, C. and Kunstmann, H. Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges. WIREs Water, 6(2):e1337, 2019. ISSN 2049-1948. doi: 10.1002/wat2.1337. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wat2.1337.

Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., et al. Precipitation observation using microwave backhaul links

in the alpine and pre-alpine region of Southern Germany. Hydrology and Earth System Sciences, 16(8):2647–2661, Aug. 2012. doi: https://doi.org/10.5194/hess-16-2647-2012.

Chwala, C., Kunstmann, H., Hipp, S., and Siart, U. A monostatic microwave transmission experiment for line integrated precipitation and humidity remote sensing. Atmospheric Research, 144:57–72, July 2014. doi: 10.1016/j.atmosres.2013.05.014.

Chwala, C., Keis, F., and Kunstmann, H. Real-time data acquisition of commercial microwave link networks for hydrometeorological applications. Atmospheric Measurement Techniques, 9(3):991–999, Mar. 2016. doi: 10.5194/amt-9-991-2016.

Chwala, C., Polz, J., Graf, M., Sereb, D., Blettner, N., Keis, F., and Boose, Y. pycomlink/pycomlink: v0.3.4, Jan. 2022.

Ciach, G. J. Local Random Errors in Tipping-Bucket Rain Gauge Measurements. Journal of Atmospheric and Oceanic Technology, 20(5):752–759, May 2003. doi: 10.1175/1520-0426(2003)20⟨752:lreitb⟩2.0.co;2.

Cireşan, D. C., Meier, U., Masci, J., Gambardella, L. M., and Schmidhuber, J. Flexible, high performance convolutional neural networks for image classification. In Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two, Ijcai'11, pages 1237–1242, Barcelona, Catalonia, Spain, July 2011. AAAI Press. ISBN 978-1-57735-514-4.

Cristiano, E., ten Veldhuis, M.-C., and van de Giesen, N. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review. Hydrology and Earth System Sciences, 21(7):3859–3878, July 2017. doi: 10.5194/hess-21-3859-2017.

Cybenko, G. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems, 2(4):303–314, Dec. 1989. doi: 10.1007/bf02551274.

David, N. and Gao, H. O. Using Cellular Communication Networks To Detect Air Pollution. Environmental Science & Technology, 50(17):9442–9451, Sept. 2016. doi: 10.1021/acs.est.6b00681.

de Vos, L. W., Overeem, A., Leijnse, H., Uijlenhoet, R., Overeem, A., Leijnse, H., and Uijlenhoet, R. Rainfall Estimation Accuracy of a Nationwide Instantaneously Sampling Commercial Microwave Link Network: Error Dependency on Known Characteristics. Journal of Atmospheric and Oceanic Technology, July 2019. doi: 10.1175/jtech-d-18-0197.1.

Developers, T. TensorFlow, Jan. 2022. URL `https://zenodo.org/record/5898685`.

Djibo, M., Chwala, C., Graf, M., Polz, J., Kunstmann, H., and Zougmoré, F. High-resolution rainfall maps from commercial microwave links for a data-scarce region in West Africa. Journal of Hydrometeorology, -1(aop), Aug. 2023. doi: 10.1175/

jhm-d-23-0015.1. URL https://journals.ametsoc.org/view/journals/hydr/aop/JHM-D-23-0015.1/JHM-D-23-0015.1.xml.

Doumounia, A., Gosset, M., Cazenave, F., Kacou, M., and Zougmore, F. Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed. Geophysical Research Letters, 41(16):6016–6022, 2014. doi: 10.1002/2014gl060724.

Doviak, R. J., Zrnic, D. S., and Schotland, R. M. Doppler radar and weather observations. Applied Optics, 33(21):4531, 1994.

DWD. DWD Climate Data Center (CDC): Historical daily precipitation observations for Germany v21.3, 2021. URL https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/more_precip/historical/DESCRIPTION_obsgermany_climate_daily_more_precip_historical_en.pdf.

DWD, C. D. C. Historische stündliche RADOLAN-Raster der Niederschlagshöhe (binär). URL https://opendata.dwd.de/climate_environment/CDC/grids_germany/hourly/radolan/historical/bin/.

D'Amico, M., Manzoni, A., and Solazzi, G. L. Use of Operational Microwave Link Measurements for the Tomographic Reconstruction of 2-D Maps of Accumulated Rainfall. IEEE Geoscience and Remote Sensing Letters, 13(12):1827–1831, Dec. 2016. doi: 10.1109/lgrs.2016.2614326.

Ekelund, R., Eriksson, P., and Kahnert, M. Microwave single-scattering properties of non-spheroidal raindrops. Atmospheric Measurement Techniques, 13(12):6933–6944, Dec. 2020. doi: 10.5194/amt-13-6933-2020.

Engström, O., Tahvili, S., Muhammad, A., Yaghoubi, F., and Pellaco, L. Performance Analysis of Deep Anomaly Detection Algorithms for Commercial Microwave Link Attenuation. In 2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS), pages 47–52, Oct. 2020. doi: 10.1109/icacsis51025.2020.9263209.

Ericsson. Receiver Performance; Receiver Thresholds Rau1 - Ericsson MINI-LINK E Technical Description [Page 136] | ManualsLib, 2012. URL https://www.manualslib.com/manual/1620197/Ericsson-Mini-Link-E.html?page=136#manual.

Fawcett, T. An introduction to ROC analysis. Pattern Recognition Letters, 27(8):861–874, June 2006. doi: 10.1016/j.patrec.2005.10.010.

Fencl, M., Rieckermann, J., Schleiss, M., Stránský, D., and Bareš, V. Assessing the potential of using telecommunication microwave links in urban drainage modelling. Water Science and Technology, 68(8):1810–1818, Oct. 2013. doi: 10.2166/wst.2013.429.

Fencl, M., Dohnal, M., Rieckermann, J., and Bareš, V. Gauge-adjusted rainfall estimates from commercial microwave links. Hydrology and Earth System Sciences, 21(1):617–634, Jan. 2017. doi: https://doi.org/10.5194/hess-21-617-2017.

Fencl, M., Valtr, P., Kvičera, M., and Bareš, V. Quantifying Wet Antenna Attenuation in 38-GHz Commercial Microwave Links of Cellular Backhaul. IEEE Geoscience and Remote Sensing Letters, 16(4):514–518, Apr. 2019. doi: 10.1109/lgrs.2018.2876696.

Fencl, M., Dohnal, M., Valtr, P., Grabner, M., and Bareš, V. Atmospheric observations with E-band microwave links – challenges and opportunities. Atmospheric Measurement Techniques, 13(12):6559–6578, Dec. 2020. doi: 10.5194/amt-13-6559-2020.

Fencl, M., Dohnal, M., and Bareš, V. Retrieving Water Vapor From an E-Band Microwave Link With an Empirical Model Not Requiring In Situ Calibration. Earth and Space Science, 8(11):e2021EA001911, 2021. ISSN 2333-5084. doi: 10.1029/2021ea001911. URL `https://onlinelibrary.wiley.com/doi/abs/10.1029/2021EA001911`.

Forster, O. Sätze über stetige Funktionen. In Forster, O., editor, Analysis 1: Differential- und Integralrechnung einer Veränderlichen, Grundkurs Mathematik, pages 114–124. Springer Fachmedien, Wiesbaden, 2016. doi: 10.1007/978-3-658-11545-6_11.

Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics, 36(4): 193–202, Apr. 1980. doi: 10.1007/bf00344251.

Glawion, L., Polz, J., Kunstmann, H., Fersch, B., and Chwala, C. spateGAN: Spatio-Temporal Downscaling of Rainfall Fields Using a cGAN Approach. Earth and Space Science, 10(10):e2023EA002906, 2023. doi: 10.1029/2023ea002906.

Goldshtein, O., Messer, H., and Zinevich, A. Rain Rate Estimation Using Measurements From Commercial Telecommunications Links. IEEE Transactions on Signal Processing, 57(4):1616–1625, Apr. 2009. doi: 10.1109/tsp.2009.2012554.

Goodfellow, I., Bengio, Y., and Courville, A. Deep Learning. Adaptive Computation and Machine Learning series. MIT Press, Cambridge, MA, USA, Nov. 2016. ISBN 978-0-262-03561-3.

Goodfellow, I. J., Bulatov, Y., Ibarz, J., Arnoud, S., and Shet, V. Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks, Apr. 2014.

Gosset, M., Kunstmann, H., Zougmore, F., Cazenave, F., Leijnse, H., Uijlenhoet, R., Chwala, C., Keis, F., Doumounia, A., Boubacar, B., et al. Improving Rainfall Measurement in Gauge Poor Regions Thanks to Mobile Telecommunication Networks. Bulletin of the American Meteorological Society, 97(3):Es49–es51, Mar. 2016. doi: 10.1175/bams-d-15-00164.1.

Graf, M., Chwala, C., Polz, J., and Kunstmann, H. Rainfall estimation from a German-wide commercial microwave link network: optimized processing and validation for 1 year of data. Hydrology and Earth System Sciences, 24(6):2931–2950, June 2020a. ISSN 1027-5606. doi: 10.5194/hess-24-2931-2020. URL https://hess.copernicus.org/articles/24/2931/2020/.

Graf, M., Chwala, C., Polz, J., and Kunstmann, H. Showcase video of hourly RADOLAN and CML rainfall maps, Apr. 2020b. URL https://zenodo.org/record/3759208.

Graf, M., El Hachem, A., Eisele, M., Seidel, J., Chwala, C., Kunstmann, H., and Bárdossy, A. Rainfall estimates from opportunistic sensors in Germany across spatio-temporal scales. Journal of Hydrology: Regional Studies, 37:100883, Oct. 2021a. ISSN 2214-5818. doi: 10.1016/j.ejrh.2021.100883. URL https://www.sciencedirect.com/science/article/pii/S2214581821001129.

Graf, M., Polz, J., and Chwala, C. Regenmessung im Mobilfunknetz. Physik in unserer Zeit, 52(2):88–93, 2021b. ISSN 1521-3943. doi: 10.1002/piuz.202001602. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/piuz.202001602.

Graf, M., Polz, J., and Chwala, C. Blackout gap detection example notebook [Software], Mar. 2022a. URL https://github.com/pycomlink/pycomlink/blob/12fc302539851b19f7656cf7e2438c0ddbaa48bf/notebooks/Blackout%20gap%20detection%20examples.ipynb.

Graf, M., Polz, J., and Chwala, C. Data for a CML blackout gap detection example [Dataset], Mar. 2022b. URL https://doi.org/10.5281/zenodo.6337557.

Graf, M., Blettner, N., Polz, J., and Chwala, C. Potential and Limitations of Filling Gaps in Commercial Microwave Link Data Stemming From Complete Loss of Signal During Heavy Rainfall. In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), pages 1–5, June 2023. doi: 10.1109/icasspw59220.2023.10193696.

Graf, M., Wagner, A., Polz, J., Lliso, L., Lahuerta, J. A., Kunstmann, H., and Chwala, C. Improved rain event detection in commercial microwave link time series via combination with MSG SEVIRI data. Atmospheric Measurement Techniques, 17(7):2165–2182, Apr. 2024. ISSN 1867-1381. doi: 10.5194/amt-17-2165-2024. URL https://amt.copernicus.org/articles/17/2165/2024/. Publisher: Copernicus GmbH.

Gsma. The Mobile Economy 2022, 2022. URL https://www.gsma.com/mobileeconomy/wp-content/uploads/2022/02/280222-The-Mobile-Economy-2022.pdf.

Habi, H. V. and Messer, H. Wet-Dry Classification Using LSTM and Commercial Microwave Links. In 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM), pages 149–153, July 2018. doi: 10.1109/sam.2018.8448679.

Haese, B., Hörning, S., Chwala, C., Bárdossy, A., Schalge, B., and Kunstmann, H. Stochastic Reconstruction and Interpolation of Precipitation Fields Using Combined Information of Commercial Microwave Links and Rain Gauges. Water Resources Research, 53 (12):10740–10756, 2017. doi: 10.1002/2017wr021015.

Hassan, D., Isaac, G. A., Taylor, P. A., and Michelson, D. Optimizing Radar-Based Rainfall Estimation Using Machine Learning Models. Remote Sensing, 14(20):5188, Jan. 2022. doi: 10.3390/rs14205188.

He, K., Zhang, X., Ren, S., and Sun, J. Deep Residual Learning for Image Recognition, Dec. 2015.

Hoens, T. R. and Chawla, N. V. Imbalanced Datasets: From Sampling to Classifiers. In Imbalanced Learning, pages 43–59. John Wiley & Sons, Ltd, 2013. doi: 10.1002/9781118646106.ch3.

Hogg, D. C. Millimeter-Wave Communication through the Atmosphere. Science, 159 (3810):39–46, Jan. 1968. doi: 10.1126/science.159.3810.39.

Hornik, K., Stinchcombe, M., and White, H. Multilayer feedforward networks are universal approximators. Neural Networks, 2(5):359–366, Jan. 1989. doi: 10.1016/0893-6080(89)90020-8.

IFRC. World Disasters Report 2020 | IFRC, May 2021. URL `https://www.ifrc.org/document/world-disasters-report-2020`.

Imhoff, R. O., Overeem, A., Brauer, C. C., Leijnse, H., Weerts, A. H., and Uijlenhoet, R. Rainfall Nowcasting Using Commercial Microwave Links. Geophysical Research Letters, 47(19):e2020GL089365, 2020. doi: 10.1029/2020gl089365.

IPCC. Summary for policymakers. In Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., editors, Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2021. doi: 10.1017/9781009157896.001.

IPCC. Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022a. doi: 10.1017/9781009325844.

IPCC. Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, USA, 2022b. doi: 10.1017/9781009157926.

Iqbal, H. HarisIqbal88/PlotNeuralNet v1.0.0, Dec. 2018.

Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., and Muller, P.-A. Deep learning for time series classification: a review. Data Mining and Knowledge Discovery, 33(4): 917–963, July 2019. doi: 10.1007/s10618-019-00619-1.

ITU-R. Specific attenuation model for rain for use in prediction methods (Recommendation P.838-3). Geneva, Switzerland: ITU-R. Retrieved from https://www.itu.int/rec/R-REC-P.838-3-200503-I/en, 2005. URL https://www.itu.int/rec/R-REC-P.838-3-200503-I/en.

ITUßR. Characteristics of precipitation for propagation modelling (Recommendation P.837-7). Geneva, Switzerland: ITU-R. Retrieved from https://www.itu.int/rec/R-REC-P.837/en, 2017. URL https://www.itu.int/rec/R-REC-P.837/en.

ITUßR. P.530 : Propagation data and prediction methods required for the design of terrestrial line-of-sight systems, Sept. 2021. URL https://www.itu.int/rec/R-REC-P.530-18-202109-I/en.

Joss, J. and Waldvogel, A. Ein Spektrograph für Niederschlagstropfen mit automatischer Auswertung. pure and applied geophysics, 68(1):240–246, Dec. 1967. doi: 10.1007/bf00874898.

Kaufmann, M. and Rieckermann, J. Identification of dry and rainy periods using telecommunication ... In 12nd International Conference on Urban Drainage, Porto Alegro, Brazil, Sept. 2011. International Water Association. URL https://www.yumpu.com/en/document/view/6346597/identification-of-dry-and-rainy-periods-using-telecommunication-.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G., and Kirschbaum, D. B. So, How Much of the Earth's Surface Is Covered by Rain Gauges? Bulletin of the American Meteorological Society, 98(1):69–78, Jan. 2017. doi: 10.1175/bams-d-14-00283.1.

Kiefer, J. and Wolfowitz, J. Stochastic estimation of the maximum of a regression function. The Annals of Mathematical Statistics, 23(3):462–466, 1952. ISSN 00034851.

Kim, M.-S. and Kwon, B. H. Rainfall Detection and Rainfall Rate Estimation Using Microwave Attenuation. Atmosphere, 9(8):287, Aug. 2018. doi: 10.3390/atmos9080287.

Kingma, D. P. and Ba, J. Adam: A Method for Stochastic Optimization, Jan. 2017. URL http://doi.org/10.48550/arXiv.1412.6980.

Kneis, D. and Heistermann, M. Bewertung der Güte einer Radar-basierten Niederschlagsschätzung am Beispiel eines kleinen Einzugsgebiets. Hydrologie und Wasserbewirtschaftung. Hydrologie und Wasserbewirtschaftung, 53(3):160–171, 2009.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M. Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrology and Earth System Sciences, 22(11):6005–6022, Nov. 2018. doi: 10.5194/hess-22-6005-2018.

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Pritzel, A., Ravuri, S., Ewalds, T., Alet, F., Eaton-Rosen, Z., et al. GraphCast: Learning skillful medium-range global weather forecasting, Dec. 2022. URL http://arxiv.org/abs/2212.12794.

Lecun, Y. Generalization and network design strategies. Elsevier, 1989.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278–2324, Nov. 1998. doi: 10.1109/5.726791.

LeCun, Y., Bengio, Y., and Hinton, G. Deep learning. Nature, 521(7553):436–444, May 2015. doi: 10.1038/nature14539.

Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M. Rainfall measurement using radio links from cellular communication networks. Water Resources Research, 43(3), 2007. doi: 10.1029/2006wr005631.

Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M. Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation. Advances in Water Resources, 31(11):1481–1493, Nov. 2008. doi: 10.1016/j.advwatres.2008.03.004.

Leinonen, J., Nerini, D., and Berne, A. Stochastic Super-Resolution for Downscaling Time-Evolving Atmospheric Fields With a Generative Adversarial Network. IEEE Transactions on Geoscience and Remote Sensing, 59(9):7211–7223, Sept. 2021.

Lenderink, G., Mok, H. Y., Lee, T. C., and van Oldenborgh, G. J. Scaling and trends of hourly precipitation extremes in two different climate zones – Hong Kong and the Netherlands. Hydrology and Earth System Sciences, 15(9):3033–3041, Sept. 2011. doi: 10.5194/hess-15-3033-2011.

Liberman, Y., Samuels, R., Alpert, P., and Messer, H. New algorithm for integration between wireless microwave sensor network and radar for improved rainfall measurement and mapping. Atmospheric Measurement Techniques, 7(10):3549–3563, Oct. 2014. doi: https://doi.org/10.5194/amt-7-3549-2014.

Lorenz, C. and Kunstmann, H. The Hydrological Cycle in Three State-of-the-Art Reanalyses: Intercomparison and Performance Analysis. Journal of Hydrometeorology, 13(5):1397–1420, Oct. 2012. ISSN 1525-7541, 1525-755x. doi: 10.1175/jhm-d-11-088.1. URL https://journals.ametsoc.org/view/journals/hydr/13/5/jhm-d-11-088_1.xml.

Lucas, B. D. and Kanade, T. An iterative image registration technique with an application to stereo vision. In Proceedings of the 7th international joint conference on Artificial intelligence - Volume 2, Ijcai'81, pages 674–679, San Francisco, CA, USA, Aug. 1981. Morgan Kaufmann Publishers Inc.

Löffler-Mang, M. and Joss, J. An Optical Disdrometer for Measuring Size and Velocity of Hydrometeors. Journal of Atmospheric and Oceanic Technology, 17(2):130–139, Feb. 2000. doi: 10.1175/1520-0426(2000)017⟨0130:AODFMS⟩2.0.CO;2.

Maggioni, V., Meyers, P. C., and Robinson, M. D. A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. Journal of Hydrometeorology, 17(4):1101–1117, Feb. 2016. doi: 10.1175/jhm-d-15-0190.1.

Marshall, J. S., Langille, R. C., and Palmer, W. M. K. Measurement Of Rainfall By Radar. Journal of the Atmospheric Sciences, 4(6):186–192, Dec. 1947. doi: 10.1175/1520-0469(1947)004⟨0186:morbr⟩2.0.co;2.

McKee, J. L. and Binns, A. D. A review of gauge–radar merging methods for quantitative precipitation estimation in hydrology. Canadian Water Resources Journal / Revue canadienne des ressources hydriques, 41(1-2):186–203, Apr. 2016. doi: 10.1080/07011784.2015.1064786.

Meissner, D., Gebauer, S., Schumann, A. H., and Rademacher, S. Analyse radarbasierter Niederschlagsprodukte als Eingangsdaten verkehrsbezogener Wasserstandsvorhersagen am Rhein. Hydrologie und Wasserbewirtschaftung, 1(02), 2012. doi: 10.5675/HyWa_2012,1_2.

Messer, H. and Sendik, O. A New Approach to Precipitation Monitoring: A critical survey of existing technologies and challenges. IEEE Signal Processing Magazine, 32 (3):110–122, May 2015. doi: 10.1109/msp.2014.2309705.

Messer, H., Zinevich, A., and Alpert, P. Environmental Monitoring by Wireless Communication Networks. Science, 312(5774):713–713, May 2006. doi: 10.1126/science.1120034.

Moraux, A., Dewitte, S., Cornelis, B., and Munteanu, A. A Deep Learning Multimodal Method for Precipitation Estimation. Remote Sensing, 13(16):3278, Jan. 2021. doi: 10.3390/rs13163278.

Moroder, C., Siart, U., Chwala, C., and Kunstmann, H. Microwave Instrument for Simultaneous Wet Antenna Attenuation and Precipitation Measurement. IEEE Transactions on Instrumentation and Measurement, pages 1–1, 2019. doi: 10.1109/tim.2019.2961498.

Moroder, C., Siart, U., Chwala, C., and Kunstmann, H. Modeling of Wet Antenna Attenuation for Precipitation Estimation From Microwave Links. IEEE Geoscience and Remote Sensing Letters, 17(3):386–390, Mar. 2020. doi: 10.1109/lgrs.2019.2922768.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9):4349–4383, Sept. 2021. doi: 10.5194/essd-13-4349-2021.

Mühlbauer, K., heisterm, thpfaff, Heistermann, M., Goudenhoofdt, E., Chwala, C., Velthuizen, F. v., Helmus, J. J., Sinclair, S., miub radar, et al. wradlib/wradlib: wradlib v1.16.2, Aug. 2022.

Nystuen, J. A., Proni, J. R., Black, P. G., and Wilkerson, J. C. A Comparison of Automatic Rain Gauges. Journal of Atmospheric and Oceanic Technology, 13(1):62–73, Feb. 1996. doi: 10.1175/1520-0426(1996)013⟨0062:acoarg⟩2.0.co;2.

Olsen, R., Rogers, D., and Hodge, D. The aRbrelation in the calculation of rain attenuation. IEEE Transactions on Antennas and Propagation, 26(2):318–329, Mar. 1978. doi: 10.1109/tap.1978.1141845.

Ostrometzky, J. and Messer, H. Dynamic Determination of the Baseline Level in Microwave Links for Rain Monitoring From Minimum Attenuation Values. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(1):24–33, Jan. 2018. doi: 10.1109/jstars.2017.2752902.

Ostrometzky, J., Raich, R., Bao, L., Hansryd, J., and Messer, H. The Wet-Antenna Effect—A Factor to be Considered in Future Communication Networks. IEEE Transactions on Antennas and Propagation, 66(1):315–322, Jan. 2018. doi: 10.1109/tap.2017.2767620.

Overeem, A., Leijnse, H., and Uijlenhoet, R. Measuring urban rainfall using microwave links from commercial cellular communication networks. Water Resources Research, 47 (12), 2011. doi: 10.1029/2010wr010350.

Overeem, A., Leijnse, H., and Uijlenhoet, R. Retrieval algorithm for rainfall mapping from microwave links in a cellular communication network. Atmospheric Measurement Techniques, 9(5):2425–2444, June 2016a. doi: 10.5194/amt-9-2425-2016.

Overeem, A., Leijnse, H., and Uijlenhoet, R. Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks. Water Resources Research, 52(10):8039–8065, 2016b. doi: 10.1002/2016wr019412.

Overeem, A., Leijnse, H., Leth, T. C. v., Bogerd, L., Priebe, J., Tricarico, D., Droste, A., and Uijlenhoet, R. Tropical rainfall monitoring with commercial microwave links in Sri Lanka. Environmental Research Letters, 16(7):074058, July 2021. doi: 10.1088/1748-9326/ac0fa6.

Pastorek, J., Fencl, M., Rieckermann, J., and Bareš, V. Commercial microwave links for urban drainage modelling: The effect of link characteristics and their position on

runoff simulations. Journal of Environmental Management, 251:109522, Dec. 2019. doi: 10.1016/j.jenvman.2019.109522.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., Kurth, T., Hall, D., Li, Z., Azizzadenesheli, K., et al. FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, Feb. 2022.

Paulson, K. and Al-Mreri, A. A rain height model to predict fading due to wet snow on terrestrial links. Radio Science, 46(4), 2011. doi: 10.1029/2010rs004555.

Pejcic, V., Saavedra Garfias, P., Mühlbauer, K., Trömel, S., and Simmer, C. Comparison between precipitation estimates of ground-based weather radar composites and GPM's DPR rainfall product over Germany. Meteorologische Zeitschrift, pages 451–466, Nov. 2020. doi: 10.1127/metz/2020/1039.

Peters, D. P. C., Pielke, R. A., Bestelmeyer, B. T., Allen, C. D., Munson-McGee, S., and Havstad, K. M. Cross-scale interactions, nonlinearities, and forecasting catastrophic events. Proceedings of the National Academy of Sciences, 101(42):15130–15135, Oct. 2004. doi: 10.1073/pnas.0403822101.

Piczak, K. J. Environmental sound classification with convolutional neural networks. In 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), pages 1–6, Sept. 2015. doi: 10.1109/mlsp.2015.7324337.

Polz, J. Supplementary animation for resradnet, Mar. 2023a. URL `https://doi.org/10.5281/zenodo.7723004`.

Polz, J. jpolz/ResRadNet: Example ResRadNet architecture and trained model, Mar. 2023b. URL `https://zenodo.org/record/7705457`.

Polz, J. Supplementary animation for ResRadNet, Mar. 2023c. URL `https://zenodo.org/record/7723005`.

Polz, J., Chwala, C., Graf, M., and Kunstmann, H. Rain event detection in commercial microwave link attenuation data using convolutional neural networks. Atmospheric Measurement Techniques, 13(7):3835–3853, July 2020. ISSN 1867-1381. doi: 10.5194/amt-13-3835-2020. URL `https://amt.copernicus.org/articles/13/3835/2020/`.

Polz, J., Glawion, L., Graf, M., Blettner, N., Lasota, E., Schmidt, L., Kunstmann, H., and Chwala, C. Expert flagging of commercial microwave link signal anomalies: Effect on rainfall estimation and ambiguity of flagging. In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW), pages 1–5, 2023a. doi: 10.1109/icasspw59220.2023.10193654.

Polz, J., Graf, M., and Chwala, C. Missing Rainfall Extremes in Commercial Microwave Link Data Due To Complete Loss of Signal. Earth and Space Science,

10(2):e2022EA002456, 2023b. ISSN 2333-5084. doi: 10.1029/2022ea002456. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2022EA002456.

Polz, J., Glawion, L., Gebisso, H., Altenstrasser, L., Graf, M., Kunstmann, H., Vogl, S., and Chwala, C. Temporal Super-Resolution, Ground Adjustment, and Advection Correction of Radar Rainfall Using 3-D-Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing, 62:1–10, 2024. doi: 10.1109/TGRS.2024.3371577. URL https://ieeexplore.ieee.org/document/10453974.

Potthast, R., Vobig, K., Blahak, U., and Simmer, C. Data Assimilation of Nowcasted Observations. Monthly Weather Review, 150(5):969–980, May 2022. ISSN 1520-0493, 0027-0644. doi: 10.1175/mwr-d-21-0017.1. URL https://journals.ametsoc.org/view/journals/mwre/150/5/MWR-D-21-0017.1.xml.

Pulkkinen, S., Koistinen, J., Kuitunen, T., and Harri, A.-M. Probabilistic radar-gauge merging by multivariate spatiotemporal techniques. Journal of Hydrology, 542:662–678, Nov. 2016. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2016.09.036. URL https://www.sciencedirect.com/science/article/pii/S002216941630590X.

Pulkkinen, S., Nerini, D., Pérez Hortal, A. A., Velasco-Forero, C., Seed, A., Germann, U., and Foresti, L. Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0). Geoscientific Model Development, 12(10):4185–4219, Oct. 2019. ISSN 1991-959x. doi: 10.5194/gmd-12-4185-2019. URL https://gmd.copernicus.org/articles/12/4185/2019/.

pycomlink, 2021. URL https://github.com/pycomlink/pycomlink.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., et al. Skilful precipitation nowcasting using deep generative models of radar. Nature, 597(7878):672–677, Sept. 2021. ISSN 1476-4687. doi: 10.1038/s41586-021-03854-z. URL https://www.nature.com/articles/s41586-021-03854-z.

Reddi, S. J., Kale, S., and Kumar, S. On The Convergence Of Adam And Beyond. In International Conference on Learning Representations, 2018.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat. Deep learning and process understanding for data-driven Earth system science. Nature, 566(7743):195–204, Feb. 2019. doi: 10.1038/s41586-019-0912-1.

Rios Gaona, M. F., Overeem, A., Leijnse, H., and Uijlenhoet, R. Measurement and interpolation uncertainties in rainfall maps from cellular communication networks. Hydrology and Earth System Sciences, 19(8):3571–3584, Aug. 2015. doi: 10.5194/hess-19-3571-2015.

Rubin, Y., Rostkier-Edelstein, D., Chwala, C., and Alpert, P. Challenges in Diurnal Humidity Analysis from Cellular Microwave Links (CML) over Germany. Remote Sensing, 14(10):2353, Jan. 2022. doi: 10.3390/rs14102353.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning representations by back-propagating errors. Nature, 323(6088):533–536, Oct. 1986. doi: 10.1038/323533a0.

Ryde, J. The attenuation of centimetre radio waves and the echo intensities resulting from atmospheric phenomena. Journal of the Institution of Electrical Engineers-Part IIIA: Radiolocation, 93(1):101–103, 1946.

Ryzhkov, A., Diederich, M., Zhang, P., and Simmer, C. Potential Utilization of Specific Attenuation for Rainfall Estimation, Mitigation of Partial Beam Blockage, and Radar Networking. Journal of Atmospheric and Oceanic Technology, 31(3):599–619, Mar. 2014. doi: 10.1175/jtech-d-13-00038.1.

Ryzhkov, A. V. and Zrnic, D. S. Radar Polarimetry for Weather Observations. Springer Atmospheric Sciences. Springer International Publishing, Cham, 2019. doi: 10.1007/978-3-030-05093-1.

Schip, T. I. v. h., Overeem, A., Leijnse, H., Uijlenhoet, R., Meirink, J. F., and Delden, A. J. v. Rainfall measurement using cell phone links: classification of wet and dry periods using geostationary satellites. Hydrological Sciences Journal, 62(9):1343–1353, July 2017. doi: 10.1080/02626667.2017.1329588.

Schleiss, M. and Berne, A. Identification of Dry and Rainy Periods Using Telecommunication Microwave Links. IEEE Geoscience and Remote Sensing Letters, 7(3):611–615, July 2010. doi: 10.1109/lgrs.2010.2043052.

Schleiss, M., Rieckermann, J., and Berne, A. Quantification and Modeling of Wet-Antenna Attenuation for Commercial Microwave Links. IEEE Geoscience and Remote Sensing Letters, 10(5):1195–1199, Sept. 2013. doi: 10.1109/lgrs.2012.2236074.

Schleiss, M., Olsson, J., Berg, P., Niemi, T., Kokkonen, T., Thorndahl, S., Nielsen, R., Ellerbæk Nielsen, J., Bozhinova, D., and Pulkkinen, S. The accuracy of weather radar in heavy rain: a comparative study for Denmark, the Netherlands, Finland and Sweden. Hydrology and Earth System Sciences, 24(6):3157–3188, June 2020. doi: https://doi.org/10.5194/hess-24-3157-2020.

Schneider, R., Bonavita, M., Geer, A., Arcucci, R., Dueben, P., Vitolo, C., Le Saux, B., Demir, B., and Mathieu, P.-P. ESA-ECMWF Report on recent progress and research directions in machine learning for Earth System observation and prediction. npj Climate and Atmospheric Science, 5(1):1–5, June 2022. ISSN 2397-3722. doi: 10.1038/s41612-022-00269-z. URL `https://www.nature.com/articles/s41612-022-00269-z`.

Seo, B.-C. and Krajewski, W. F. Correcting temporal sampling error in radar-rainfall: Effect of advection parameters and rain storm characteristics on the correction accuracy. Journal of Hydrology, 531:272–283, Dec. 2015. doi: 10.1016/j.jhydrol.2015.04.018.

Sevruk, B. Rainfall Measurement: Gauges. In Encyclopedia of Hydrological Sciences. John Wiley & Sons, Ltd, 2006. doi: 10.1002/0470848944.hsa038.

Smiatek, G., Keis, F., Chwala, C., Fersch, B., and Kunstmann, H. Potential of commercial microwave link network derived rainfall for river runoff simulations. Environmental Research Letters, 12(3):034026, Mar. 2017. doi: 10.1088/1748-9326/aa5f46.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. Journal of Machine Learning Research, 15(56):1929–1958, 2014. ISSN 1533-7928.

Strangeways, I. A history of rain gauges. Weather, 65(5):133–138, 2010. doi: 10.1002/wea.548.

Stransky, D., Fencl, M., and Bares, V. Runoff prediction using rainfall data from microwave links: Tabor case study. Water Science and Technology, 2017(2):351–359, Apr. 2018. doi: 10.2166/wst.2018.149.

Testud, J., Bouar, E. L., Obligis, E., and Ali-Mehenni, M. The Rain Profiling Algorithm Applied to Polarimetric Weather Radar. Journal of Atmospheric and Oceanic Technology, 17(3):332–356, Mar. 2000. doi: 10.1175/1520-0426(2000)017⟨0332:trpaat⟩2.0.co;2.

Tjelta, T. and Bacon, D. Predicting Combined Rain and Wet Snow Attenuation on Terrestrial Links. IEEE Transactions on Antennas and Propagation, 58(5):1677–1682, May 2010. doi: 10.1109/tap.2010.2044316.

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., and Sheffield, J. Global warming and changes in drought. Nature Climate Change, 4 (1):17–22, Jan. 2014. doi: 10.1038/nclimate2067.

Trömel, S., Ziegert, M., Ryzhkov, A. V., Chwala, C., and Simmer, C. Using Microwave Backhaul Links to Optimize the Performance of Algorithms for Rainfall Estimation and Attenuation Correction. Journal of Atmospheric and Oceanic Technology, 31(8): 1748–1760, June 2014. doi: 10.1175/jtech-d-14-00016.1.

Trömel, S., Chwala, C., Furusho-Percot, C., Henken, C. C., Polz, J., Potthast, R., Reinoso-Rondinel, R., and Simmer, C. Near-Realtime Quantitative Precipitation Estimation and Prediction (RealPEP). Bulletin of the American Meteorological Society, 102(8):E1591–e1596, Aug. 2021. doi: 10.1175/bams-d-21-0073.1.

Uijlenhoet, R., Overeem, A., and Leijnse, H. Opportunistic remote sensing of rainfall using microwave links from cellular communication networks. WIREs Water, 5(4):e1289, 2018. doi: 10.1002/wat2.1289.

UNFCCC. Sharm el-Sheikh Implementation Plan. Revised draft decision -/CMA.4 | UNFCCC, Nov. 2022. URL https://unfccc.int/documents/621908.

Upton, G., Holt, A., Cummings, R., Rahimi, A., and Goddard, J. Microwave links: The future for urban rainfall measurement? Atmospheric Research, 77(1-4):300–312, Sept. 2005. doi: 10.1016/j.atmosres.2004.10.009.

Valtr, P., Fencl, M., and Bareš, V. Excess Attenuation Caused by Antenna Wetting of Terrestrial Microwave Links at 32 GHz. IEEE Antennas and Wireless Propagation Letters, 18(8):1636–1640, Aug. 2019. doi: 10.1109/lawp.2019.2925455.

van de Beek, C. Z., Leijnse, H., Torfs, P. J. J. F., and Uijlenhoet, R. Seasonal semivariance of Dutch rainfall at hourly to daily scales. Advances in Water Resources, 45:76–85, Sept. 2012. ISSN 0309-1708. doi: 10.1016/j.advwatres.2012.03.023. URL http://www.sciencedirect.com/science/article/pii/S0309170812000784.

van Leth, T. C., Overeem, A., Leijnse, H., and Uijlenhoet, R. A measurement campaign to assess sources of error in microwave link rainfall estimation. Atmospheric Measurement Techniques, 11(8):4645–4669, Aug. 2018. ISSN 1867-1381. doi: 10.5194/amt-11-4645-2018. URL https://amt.copernicus.org/articles/11/4645/2018/.

Villarini, G. and Krajewski, W. F. Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall. Surveys in Geophysics, 31(1):107–129, Jan. 2010. doi: 10.1007/s10712-009-9079-x.

Villarini, G., Mandapaka, P. V., Krajewski, W. F., and Moore, R. J. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research: Atmospheres, 113(D11), 2008. doi: 10.1029/2007jd009214.

Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H. Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. Hydrology and Earth System Sciences, 16(7):2311–2328, July 2012. doi: 10.5194/hess-16-2311-2012.

Vogl, S., Laux, P., Bialas, J., and Reifenberger, C. Modelling Precipitation Intensities from X-Band Radar Measurements Using Artificial Neural Networks—A Feasibility Study for the Bavarian Oberland Region. Water, 14(3):276, Jan. 2022. doi: 10.3390/w14030276.

Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B. Global Water Resources: Vulnerability from Climate Change and Population Growth. Science, 289(5477):284–288, July 2000. doi: 10.1126/science.289.5477.284.

Wallace, J. M. and Hobbs, P. V. 6 - Cloud Microphysics. In Wallace, J. M. and Hobbs, P. V., editors, Atmospheric Science (Second Edition), pages 209–269. Academic Press, San Diego, Jan. 2006. doi: 10.1016/b978-0-12-732951-2.50011-9.

Wang, L.-P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R. D., Pessemier, M., Kroll, S., Willems, P., and Onof, C. Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment. Journal of Hydrology, 531:408–426, Dec. 2015. doi: 10.1016/j.jhydrol.2015.05.049.

Wang, Z., Schleiss, M., Jaffrain, J., Berne, A., and Rieckermann, J. Using Markov switching models to infer dry and rainy periods from telecommunication microwave link signals. Atmospheric Measurement Techniques, 5(7):1847–1859, July 2012. doi: 10.5194/amt-5-1847-2012.

Watson, R. T., Zinyowera, M. C., and Moss, R. H. Climate change 1995. Impacts, adaptations and mitigation of climate change: scientific-technical analyses. Cambridge University Press, 1996. ISBN 0521564379.

Wilson, J. W. and Brandes, E. A. Radar Measurement of Rainfall—A Summary. Bulletin of the American Meteorological Society, 60(9):1048–1060, Sept. 1979. doi: 10.1175/1520-0477(1979)060⟨1048:rmors⟩2.0.co;2.

Winterrath, T., Rosenow, W., and Weigl, E. On the DWD quantitative precipitation analysis and nowcasting system for real-time application in German flood risk management. IAHS Publ., 351:7, 2012.

Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Walawender, E., Weigl, E., and Becker, A. Erstellung einer radargestützten Niederschlagsklimatologie, 2017. URL `https://opendata.dwd.de/climate_environment/GPCC/radarklimatologie/Dokumente/Endbericht_Radarklimatologie_final.pdf`.

Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Lengfeld, K., Walawender, E., Weigl, E., and Becker, A. Radar climatology (RADKLIM) version 2017.002; gridded precipitation data for Germany: Radar-based quasi gauge-adjusted five-minute precipitation rate (YW), Sept. 2018.

WMO. WMO Radar Database. URL `https://wrd.mgm.gov.tr/Home/Wrd`.

Wolpert, D. and Macready, W. No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation, 1(1):67–82, Apr. 1997. doi: 10.1109/4235.585893.

Yo, T.-S., Su, S.-H., Chu, J.-L., Chang, C.-W., and Kuo, H.-C. A Deep Learning Approach to Radar-Based QPE. Earth and Space Science, 8(3):e2020EA001340, 2021. doi: 10.1029/2020ea001340.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., and Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. IEEE Geoscience and Remote Sensing Magazine, 5(4):8–36, Dec. 2017. doi: 10.1109/mgrs.2017.2762307.

Zimek, A., Schubert, E., and Kriegel, H.-P. A survey on unsupervised outlier detection in high-dimensional numerical data. Statistical Analysis and Data Mining: The ASA Data Science Journal, 5(5):363–387, 2012. doi: 10.1002/sam.11161.

Zinevich, A., Messer, H., and Alpert, P. Prediction of rainfall intensity measurement errors using commercial microwave communication links. Atmospheric Measurement Techniques, 3(5):1385–1402, Oct. 2010. doi: 10.5194/amt-3-1385-2010.

Đorđević, V., Pronic-Rancic, O., Marinković, Z., Milijić, M., Markovic, V., Siart, U., Chwala, C., and Kunstmann, H. New Method for Detection of Precipitation Based on Artificial Neural Networks, 2014. URL https://www.semanticscholar.org/paper/New-Method-for-Detection-of-Precipitation-Based-on-%C4%90or%C4%91evi%C4%87-Pronic-Rancic/f775687286ca47c551c3a080a5f249e969850ff8.

Špačková, A., Bareš, V., Fencl, M., Schleiss, M., Jaffrain, J., Berne, A., and Rieckermann, J. A year of attenuation data from a commercial dual-polarized duplex microwave link with concurrent disdrometer, rain gauge, and weather observations. Earth System Science Data, 13(8):4219–4240, Aug. 2021. ISSN 1866-3508. doi: 10.5194/essd-13-4219-2021. URL https://essd.copernicus.org/articles/13/4219/2021/.

Danksagung

Mein Dank geht an alle, die zur Entstehung dieser Arbeit beigetragen haben und alle, die mich während dieser Zeit begleitet haben.
Danke Harald, dass du mir diese Chance ermöglicht hast und mich jederzeit voll unterstützt hast. Dass ich mit viel Freiheit und trotzdem sehr guter Unterstützung wo ich sie gebraucht habe die Welt der Wissenschaft entdecken durfte empfinde ich als etwas Besonderes, das ich sehr schätze. Vielen Dank für die Betreuung und Begutachtung dieser Arbeit.
Mein Dank geht auch an Prof. Dr. Sabine Timpf und Prof. Dr. Remko Uijlenhuit für die wichtige und aufwendige Arbeit diese Dissertation zu begutachten.
Danke Christian, dafür, dass du dich als Mentor weit mehr reinhängst als man guten Gewissens erwarten kann. Danke, für die vielen Diskussionen auf Augenhöhe, dass ich so viel von dir lernen konnte und dass du mich in die Forschungscommunity eingebunden hast.
Danke Max, ich hatte großes Glück, dass ich in dir einen Freund gefunden habe der nur zu gut versteht wo man gerade hängt und der jederzeit ein offenes Ohr hat. Danke an Ingo, für die vielen Klettermeter, die du mich gesichert hast, und den ganzen Rest und an Luca, Karina, Johannes, Nico, Tanja, Verena, David und alle anderen, die mehr Freunde sind als Kollegen und die dieser Reise zum Erfolg verholfen haben.
Danke an das CML team für den Teamgeist und den regen Austausch. Danke, an alle Kollegen aus der AG Kunstmann, am IFU und in Augsburg. Danke an die RealPEP Forschergruppe für die Zusammenarbeit im Projekt und die Diskussionen bei den Projekttreffen. Danke an das KIT und die Helmholtz Gemeinschaft für die exzellente Forschungsumgebung.
Danke an meine Familie, meine Eltern, meine Geschwister, meine Tante Hucki und meine Schwiegereltern, für das Interesse und die Unterstützung. Danke, dass ihr nie an mir gezweifelt habt und mich schon immer ermutigt zu tun was ich für richtig halte.
Danke an die "Bois d'Amour III" oder wie wir gerade heißen, den Stammtisch, die Viertorspitz WG und alle anderen Freunde für die gute Zeit.
Mein größter Dank geht an meine Partnerin Anna, die mir stets den Rücken freigehalten hat. Wir sind ein gutes Team! Danke an meinen Sohn Linus, für die viele gute Laune und das Relativieren von unsinnigen Problemen.