Improved rain event detection in Commercial Microwave Link time series via combination with MSG SEVIRI data

Andreas Wagner[1,*], Maximilian Graf[1,*], Julius Polz[2,*], Llorenç Lliso[3], José Alberto Lahuerta[3], Harald Kunstmann[1,2], and Christian Chwala[2]

[1]University of Augsburg, Institute of Geography (IGUA), Alter Postweg 118, 86159 Augsburg, Germany
[2]Karlsruhe Institute of Technology, Campus Alpin (IMK-IFU), Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany
[3]Agencia Estatal De Meteorología (AEMET Spain)
[*]These authors contributed equally to this work.

Correspondence: Christian Chwala (christian.chwala@kit.edu)

Abstract. The most reliable areal precipitation estimation is usually generated via combinations of different measurements~~and devices by merging their individual advantages~~. Path-averaged rain ~~rate~~ rates can be derived from Commercial Microwave Links (CML), where attenuation of the emitted radiation is strongly related ~~with~~ to rainfall rate. CMLs can be combined with data from other rainfall measurements or used individually. They are available almost worldwide and often represent the only opportunity ~~of~~ for ground-based measurement in ~~data scarce regions. Deriving~~ data-scarce regions. However, deriving rainfall estimates from CML data requires extensive data processing~~, though~~. The separation of the attenuation time series ~~in~~ into rainy and dry periods (rain event detection) is the most important step in this processing and ~~largely determines the quality of the~~ 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. ~~The investigation uses data from 3901~~ We used data from 3748 CMLs in Germany for four months in ~~summer~~ the summer of 2021 and ~~is carried out for~~ data from the two SEVIRI-derived products ~~PC and PC-Ph. We analyse~~ *PC* and *PC-Ph*. We analyzed all rain event detection methods for different ~~precipitation intensity~~rainfall intensities, differences between day and night, as well as their influence on the performance of rainfall estimates from individual CMLs. The radar product RADKLIM-YW ~~is~~ was used for validation. The results ~~show~~ showed that both SEVIRI products are promising candidates for ADB rainfall detection~~methods and led to at least equivalent results as~~ , 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 ~~, which is caused by dew formation on CML antennas and~~ due to the reduced availability of SEVIRI channels at night. In general, the ADB methods ~~lead~~ 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 ~~were able to improve~~ improved the Matthews Correlation Coefficient of the rain event detection from ~~0.53 (0.57~~ 0.49 (0.51 resp.) to ~~0.62~~ 0.59 during the day and from ~~0.47 (0.55~~ 0.41 (0.50 resp.) to ~~0.6~~ 0.55 during the night. Our

results show that ~~utilising~~ 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.

1 Introduction

~~Precipitation~~ Rainfall is the most important variable for hydrology and water management. It is characterized by ~~a~~ high variability in space and time, especially in the case of convective rain events. The quality of hydrological modeling results depends heavily on high-resolution and reliable areal ~~precipitation (Fu et al., 2011; Bruni et al., 2015; Rafieeinasab et al., 2015; Cristiano et al., 201~~ rainfall data (Fu et al., 2011; Bruni et al., 2015; Rafieeinasab et al., 2015; Cristiano et al., 2017).

There are a variety of rainfall measurement methods ~~which~~ that serve as a basis for the derivation of rainfall fields, each with specific advantages but also drawbacks. ~~Tipping buckets for instance usually provide the best~~ Rain gauges can provide point measurements of precipitation with high accuracy, but they are prone to errors due to wind and evaporation (Sevruk, 2006) and primarily lack spatial representativeness (Pollock et al., 2018).

Satellite data provide areal precipitation ~~patterns~~ estimates almost worldwide with a spatial resolution in the order of several kilometers. But they either suffer from a poor temporal resolution ~~(~~like the GPM core satellite that has a revisit time of approximately 1 day in the tropics ~~)~~ or from heterogeneous data quality and delayed availability ~~(~~like merged satellite products like IMERG ~~)~~ (Hou et al., 2014). Additionally, complex retrieval and calibration algorithms have to be applied which cause additional uncertainties (Maggioni et al., 2016).

Weather radars derive areal precipitation ~~patterns~~ estimates with a high resolution of 5 minutes and 1 km (Atlas, 1990; Bartels et al., 2004; Winterrath et al., 2012). However, the calculation of rain ~~rate~~ rates from radar reflectivity is non-trivial (Uijlenhoet et al., 2003; Steiner et al., 2004) and false echoes, clutter, and other measurement effects cause further problems (Villarini and Krajewski, 2010; Wagner et al., 2012; Wagner, 2018). Schleiss et al. (2020) showed that radar data tends to underestimate particularly heavy rainfall in Scandinavian countries. Nevertheless, gauge-adjusted radar ~~data is~~ products are considered to be one of the best possible data basis ~~(Bartels et al., 2004; Winterrath et al., 2012). It is precisely the combination of different measurements that makes it possible to~~ for spatial rainfall estimates, because they leverage the advantages ~~and neglect disadvantages of individual systems.~~

from individual measurement devices (Bartels et al., 2004; Winterrath et al., 2012).

The opportunistic sensing of ~~precipitation is becoming available on country-wide scales (Overeem et al., 2016; Graf et al., 2021). One example is the use of attenuation data from~~ rainfall with Commercial Microwave Links (CML) ~~measurements to determine precipitation. It was first initiated~~ was first demonstrated in Israel (Messer et al., 2006) and the Netherlands (Leijnse et al., 2007). ~~Attenuation of CML data is a disturbing effect for the operating telecommunication providers but can be used to calculate precipitation intensity~~In recent years CML rainfall estimation has become available on country-wide scales (Overeem et al., 2016; Graf et al., 2020). Rainfall attenuates the microwave radiation between two antennas of a CML. The

relationship between attenuation and ~~precipitation is significantly more~~ rainfall is close to linear for signals between 10 and 40 GHz ~~than between radar reflectivity and precipitation (Atlas and Ulbrich, 1977). Correspondingly, the line integrals of precipitation~~ (Atlas and Ulbrich, 1977). CMLs have already proven their potential as stand-alone rainfall sensors in multiple regions of the world (Overeem et al., 2013; Rios Gaona et al., 2015; Overeem et al., 2016; D'Amico et al., 2016; Graf et al., 2020; Roversi

.

Additionally, the path averaged rainfall information from CMLs can complement conventional measurement methods (Liberman et al., 2014; Haese et al., 2017; Graf et al., 2021). Kumah et al. (2022) for instance derived rain intensities from MSG satellite data by a random forest algorithm trained with CML rainfall estimates. ~~However, a number of publications confirm that CML data is also applicable as the only source for areal precipitation (Overeem et al., 2013; Rios Gaona et al., 2015; Overeem et al., 2016; . In large areas of Africa, for instance, radar data~~ For regions with sparse observation networks like most parts of Africa, where radar data is missing and even station data is ~~missing or~~ only sparsely available, ~~so a mobile phone network might be the only addition to satellite data. The importance of CML can therefore be rated much higher there. Chwala and Kunstmann (2019) and Uijlenhoet et al. (2018) give an overview of challenges and background information on the usage~~ CMLs can deliver additional ground-based rainfall estimates (Djibo et al., 2023). Thus, the use of CML data ~~as a source of precipitation measurements.CML data is affected by reflection of the beam, dew formation on the antenna or fog which can be wrongly interpreted as precipitation echoes. Changes in propagationcharacteristics due to the influence of water vapour or temperature and drifts in signal characteristics due to strong solar radiation or air temperature cause further non-precipitation variations of the signals (Chwala and Kunstmann, 2019) .Therefore, the rain event detection is a major step in estimating rainfall rates from CML signals. The different available approaches are "individual methods" for~~ is a good opportunity to reduce the gap in the global availability of climate information as recently emphasized by UNFCCC (2022).

The detection of rain events in CML attenuation time series is an important processing step for several reasons. First, it defines the rainy periods for which a baseline has to be defined, typically from preceding dry time steps. Second, it filters fluctuations that are not caused by rain, but by other disturbances e.g. from refraction, multi-path propagation, or mast sway (c.f. Chwala and Kunstmann (2019) for a detailed list). For an overview of available rain event detection ~~, since they are based only on one method or data set. They can roughly be divided into~~ methods, we divide them into two categories, time series-based (TSB) methods and methods based on auxiliary data of ~~precipitation~~ rainfall patterns (ADB).

~~The TSB methods include~~ ADB methods that are based on globally available data like the satellite data presented in this study are especially promising for CML processing in regions with otherwise sparse rainfall information.

~

Examples of TSB methods are the simple threshold models (Leijnse et al., 2008), an approach using the rolling standard ~~approach (RS) (Schleiss and Berne, 2010)and more complex ones such as~~ deviation (Schleiss and Berne, 2010), Markov switching models (Wang et al., 2012) ~~and~~ or short-term Fourier Transform approaches (Chwala et al., 2012). Messer and Sendik (2015) provide a detailed description of these approaches. Machine Learning approaches to distinguish between wet and dry time steps emerged in recent years, usually outperforming the previous methods (Habi and Messer, 2018; Polz et al., 2020; Song et al., 2020). The "nearby-link" approach (Overeem et al., 2016) is ~~usually not viewed as a TSB, but nevertheless uses~~

a hybrid TSB and ADB method because it compares CML attenuation time series ~~by comparing neighboring CMLsregarding attenuation effects within a radius of 15 km.~~

~~Regarding ADB approaches, best results are usually achieved with radar data (Overeem et al., 2011) , but satellite data is also used (van het Schip et al., 2017; Kumah et al., 2021). Geostationary~~ of neighboring CMLs. Similar to CMLs, satellite microwave links (SMLs) can be exploited to derive rainfall estimates. SML processing also includes rain event detection and several methods are available (Giannetti et al., 2019; Giro et al., 2022).

~

Data sources in ADB methods can be weather radar (Overeem et al., 2011) or satellite data (van het Schip et al., 2017; Kumah et al., 2021) . Regarding satellite data, geostationary satellites such as MSG SEVIRI offer a temporal resolution of 15 minutes at 4x6 km spatial resolution in mid-latitudes. This data is also used as areal precipitation (Roebeling et al., 2008; Roebeling and Holleman, 2009), although the derivation of precipitation from VIS and IR channels is often uncertain. According to NWC SAF (Lahuerta García, 2021) even a distinction between light, moderate, and heavy precipitation is difficult. That is the reason why they determine the probability of precipitation in their post-processed SEVIRI products.~~However, such a product can be very useful as an indicator whether it is raining or not. van het Schip et al. (2017) has carried out analyzes~~

van het Schip et al. (2017) analyzed applying post-processed SEVIRI products as a wet-dry indicator in the Netherlands. ~~A short period of~~ They evaluated the resulting rainfall maps for 12 ~~days is analyzed with focus only on rainfall maps not allowing any conclusion on individual CMLs and their specific rain event detection performance. Their satellite approach performs better than doing the CML processing without a rain event detection, but worse than the~~ days and found improvements compared to not separating the time series into wet and dry periods but a decreased performance compared to radar-based rain event detection~~approach. So, it is not straightforward to evaluate the quality of their approach compared to standard rain event detection approaches. In addition, there is no differentiation according to precipitation intensity or between~~ . However, they did not compare the methods for individual CMLs, different rainfall intensities, or day and night ~~, although there are no visual channels at night and the satellite product methodologically varies greatly.~~ periods for which the satellite products use different channels and methodologies. They also did not combine their ADB method with a TSB method.

Kumah et al. (2021) obtained improved rain intensities for convective rain events when applying MSG SEVIRI data for rain event detection of CML data. However, their results are based only on one CML in Kenya for daytime and rain intensities above 0.5 ~~mm/h.~~

mmh^{-}1. The limitation of both studies regarding the analyzed period, the number of CMLs, intensity classes, and day and night differences limit the transferability of their results. None of the presented rain event detection methods can provide a high-quality classification for CML datasets with varying characteristics (e.g. sparse or dense network, various temporal resolutions, different frequency ranges, etc.). CMLs with frequencies below 10 GHz which are commonly deployed in sparse CML networks in rural Sub-Saharan Africa still pose great challenges for all available rain event detection methods. Yet, these regions are often associated with a high potential for CML as the sole source of rainfall estimates.

In this study, ~~two precipitation products (PC and PC-Ph) from NWCSAF derived from~~ the precipitation products *PC* and *PC-Ph*, which are computed by NWC SAF using data from the SEVIRI radiometer ~~on-board~~ onboard the geostationary satellite

Table 1. Overview of rainfall sensors and products.

Sensor	Product	Spatial resolution (longitude x latitude)	Temporal resolution	Data points (including missing values)
C-Band weather radar	RADKLIM-YW	1km x 1km	5 minutes	30,808,350,000
MSG SEVIRI	*PC* & *PC-Ph*	3.4-5.1km x 4.0-6.4km	15 minutes	1,330,560,000
CML	-	3748 CML paths	1 min	765,649,270

METEOSAT~~are applied for rain event detection~~, are used to classify CML attenuation time series in rainy and dry periods. In addition to comparing ADB (based on ~~the PC and PC-Ph products~~*PC* and *PC-Ph*) and TSB (CML time series processing) methods, this study presents a novel way of combining TSB and ADB rain event detection approaches to improve rain event detection. ~~In order to~~ To analyze the applicability of such new rain event detection methods, data sets of high data quality are necessary. The present analysis is based on country-wide CML data from 4 months in ~~summer~~ the summer of 2021 in Germany~~and~~ . As reference a high-resolution gauge-adjusted radar product is used~~as reference.~~

The research questions of this investigation are: (1) are ~~PC and PC-Ph~~ *PC* and *PC-Ph* products suitable as ~~a~~ wet-dry ~~indicator~~ indicators for CML data? (2) Do the results vary with rain intensity? (3) Are there noticeable differences between day and night? (4) Can a combination of TSB and ADB rain event detection methods outperform TSB-only and ADB-only methods?

2 Data

~~The evaluations in this study are~~ This study is based on CML~~data, radar data~~ , weather radar, and SEVIRI data in Germany ~~for a period of 19 weeks between 31st of May~~ covering the period from the 30th of April 2021 ~~and~~ to the 10th of ~~October~~ September 2021. Tab. 1 summarizes the properties of these products.

2.1 CML data

~~3901 CMLs in Germany are used,which are operated by Ericsson. This is only~~ The CML dataset consists of a subset of ~~all CMLs operated~~ CMLs operated by Ericsson in Germany. ~~The path lengths~~ For the analyzed period, 3748 CMLs were available for more than 30% of the time and thus considered in this study. The path length of the CMLs ~~vary~~ varies between 0.2 km and about 30 km ~~(see Fig. 1). 72 % are shorter than 10 km and over 99 % are shorter than 20 km. The frequencies typically range between 10~~ and the frequencies range from 10 to 40 GHz with shorter CMLs using higher frequencies ~~.~~ (see Fig.1). The transmitted signal level (TSL) with a power resolution of 1 dB and the received signal level (RSL) with a power resolution of 0.3 dB are ~~saved~~ instantaneously measured every minute using a custom real-time CML data acquisition system (Chwala et al., 2016). The total loss (TL) is the difference between TSL and RSL. Each CML consists of two sublinks for two-way data transmission. The processing of the data is described in ~~more detail in chapter 3.1~~Section 3.2.

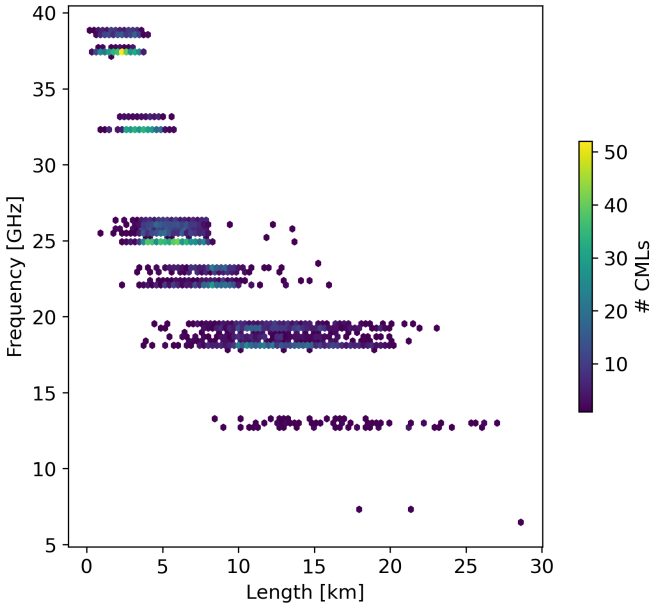

Figure 1. ~~Histogram revealing~~ Scatter density plot showing the distribution of length and frequency of the ~~3901~~ 3748 CMLs in Germany that are used in this study.

2.2 ~~Radar~~ Weather radar data

RADKLIM-YW is a gauge-adjusted, climatologically-corrected radar product of the German Meteorological Service (DWD)~~and it is used to validate SEVIRI and CML data. An adjustment to hourly data from precipitation stations similar to that of the more well-known~~ . The hourly adjustment with data from automatic rainfall stations is identical to the one of the *RADOLAN-RW* product ~~is used (Bartels et al., 2004; Winterrath et al., 2012), but the *RADKLIM-YW* data is available every 5 minutes. In addition, climatologically-based~~ (Bartels et al., 2004; Winterrath et al., 2012). Daily sums are adjusted with daily measurements of manual rainfall stations and climatology-based corrections of spokes and range-dependencies are carried out (Winterrath et al., 2018). *RADKLIM-YW* data has a temporal resolution of 5 minutes and a spatial resolution of 1 km covering Germany with 1100 times 900 grid cells. According to Kreklow et al. (2020), the range-dependent effects, which are particularly strong in winter, are reduced. This improves ~~precipitation~~ rainfall patterns and better represents orography compared to *RADOLAN-RW*. At the same time, however, a slight underestimation of the total ~~precipitation~~ rainfall was determined.

RADKLIM-YW is used to validate the binary classifications of rain event detection methods and the rainfall estimates derived from CMLs using these methods. To compare the gridded *RADKLIM-YW* with the CMLs we averaged the grid for each individual CML path weighted by the length of intersecting path segments in each pixel. *RADKLIM-YW* was aggregated to a resolution of 15 minutes.

To answer the research question (2), whether the performance of rain event detection methods shows a rain rate dependency,

we define intensity classes and group the 15-minute radar rainfall intensities. The classes light (0.1 to 2.5 mmh^{-}1), moderate (2.5 to 10.0 mmh^{-}1), and heavy (more than 10.0 mmh^{-}1) rainfall are based on the (DWD, 2023) classification. Light rainfall is further subdivided into the classes light1 (0.1 to 1.0 mmh^{-}1) and light2 (1.0 to 2.5 mmh^{-}1). Values below 0.1 mmh^{-}1 are considered dry. The resulting classes are shown in Fig.7.

2.3 SEVIRI data

The Spinning Enhanced Visible and Infrared Imager (SEVIRI) radiometer from the geostationary satellite METEOSAT provides image data in two ~~Visible~~ visible (VIS), one ~~high resolution~~ high-resolution channel (HRVIS), and nine Infrared (IR) channels including ~~also one Near Infrared~~ one near infrared channel (NIR). The channels range from 0.5 to 14.4 ~~micrometres~~ micrometers with a resolution of 3 km at the sub-satellite point. The high-resolution channel is not used for our purposes. Every 15 minutes ~~the~~ a calibrated image of the full earth disc (lon: -79° ~~–~~to 79°, lat: -81° ~~–~~to 81°) is available ~~including calibration~~ (Schmid, 2000). No ~~specific device~~ space-borne radar or radiometer for precipitation measurement is ~~on-board~~on board, such as for GPM (Hou et al., 2014). ~~The derivation of precipitation products is~~ Precipitation products are derived based on two approaches: either by regression of different channels and adjustment to precipitation measurements or, more ~~sophisticated, through the derivation of~~ sophisticatedly, by deriving microphysical parameters (Roebeling et al., 2008; Hernanz et al., 2019). For the calculation of microphysical parameters, the 0.6 um channel (VIS) and the 1.6 um channel (NIR) ~~is~~ are mandatory (Roebeling et al., 2008).

The first product ~~PC-Ph~~ *PC-Ph* is derived from the microphysical parameters Effective Radius (Reff) and Cloud Optical Thickness (COT) ~~at~~ during daytime. At night, when visible channels are missing, it is calculated by a regression of IR and Water Vapour channels (WV). The second product ~~PC~~ *PC* relies on different regression functions of available SEVIRI channels, ~~also~~ including VIS, IR, and WV channels at daytime and only IR and WV channels at night. ~~Both products~~ The definition of day and night time is derived from the variable *pccond* that is provided for both *PC* and *PC-Ph*. Both products are provided by NWC SAF. Recent data is freely available, but long-term records must be requested individually.

We chose both products for this study because they provide the probability of precipitation in percent: ~~PC~~ *PC* in increments of 10 and ~~PC-Ph~~ *PC-Ph* in increments of ~~1. A detailed description of the respective products and their validation can be found at Thoss (2014); Hernanz et al. (2019); Lahuerta García (2019)~~

3 ~~Methods~~

2.1 ~~CML data processing~~

~~The processing of CML data must be handled with care. Various groups developed CML processing methods that depend on e.g. the sampling strategy. In Germany, TSL and RSL of about 4000 CMLs are sampled instantaneously with a temporal resolution of~~ 1 ~~minute. Missing or corrupted data reduces the amount to 3901 CMLs used in this evaluation. Each CML has two sublinks which each have one TSL and RSL time series. From a purely practical point of view, we only processed data~~

from the first sublink. A more detailed description of the processing of the data can be found in Graf et al. (2020). The most important aspects of data processing are briefly outlined in the following.

The high temporal resolution of the raw data is preserved for the processing of the data. Gaps of up to 5 minutes due to missing time steps or measurements are interpolated to obtain continuous time series . The total path loss along the CML (TRSL) is then calculated as TSL minus RSL.

As a next step, wet and dry time steps have to be detected (rain event detection). This might be difficult, if attenuation signals from noise and precipitation are very similar. High similarities occur for very low rain intensities or noisy CML data. Overestimation of total precipitation and of the number of wet time steps result if noise is classified as precipitation. Underestimation occurs if time steps with precipitation are missed. Consequently, the reliable separation of precipitation-induced attenuation and noise from fluctuations in TRSL largely determines the quality of CML data. Being the central element of this manuscript, it will be discussed individually in Chapter 3.2.

Then, the baseline attenuation is dynamically identified for the preceding dry period of each rain event in order to derive the rainfall induced attenuation along the path. This baseline is the last dry time-step of the TRSL time-series and is set to be constant during the rain event.

The resulting attenuation may be overestimated due to wetting effects on the radome. To overcome this problem, a correction of this wet antenna attenuation (WAA) according to Leijnse et al. (2008) is used. In this physical approach, the dependence of WAA on antenna cover properties (refractive index and thickness), microwave frequency and rain intensity is defined. The original parameters from Leijnse et al. (2008) are adopted.

The corrected attenuation from CML was used to derive the precipitation rate using a k-R relationship in the form of percent. Compared to a power-law fit to drop-size-distribution simulations. The parameter settings were based on ITU recommendations (ITU-R, 2005) with a close-to-linear relationship between attenuation and rain rate. According to Chwala and Kunstmann (2019); Graf et al they show good agreements for Germany. pure precipitation product, the precipitation probability products enabled us to consistently alter the classification threshold similar to how it can be done for certain TSB methods. Example time series of the precipitation probability are shown in Fig 2 b) and c). More detailed descriptions and the evaluation of the two products are available from (Thoss, 2014; Hernanz et al., 2019) and (Lahuerta García, 2019). Similar to *RADKLIM-YW*, we derived *PC* and *PC-Ph* values along the CML paths. The validation of *PC* or *PC-Ph* wet and dry labels is based on the path averages of *RADKLIM-YW*.

3 Methods

3.1 Rain event detection

3.1.1 Individual methods for rain event detection

Two time series based We used two existing time series-based (TSB) methods as a baseline for rain event detectionare applied for comparison. The first one is the Rolling Standard Deviation method (*RS*)according to Schleiss and Berne (2010), modified

by Graf et al. (2020). Based on this approach , the threshold is determined based on the rolling standard deviation of total loss (TL). Time steps for which the standard deviation of a 60-minute rolling window exceeds a certain threshold are considered wet. This approach was originally suggested by Schleiss and Berne (2010) who used a fixed threshold. Later, Graf et al. (2020) determined the threshold based on the 80th percentile of a 60 min the 60-minute rolling standard deviation of TL multiplied by a scaling factor . The constant scaling factor of 1.12 is adopted from Graf et al. (2020) and was also used by Polz et al. (2020). The second one that adopted the threshold to the general amount of noisiness of each CML. This method will be called (*RS*). The second method is a machine learning approach based on Convolutional Neural Networks from Polz et al. (2020)which is called *CNN* in the following. The a convolutional neural network that was trained to classify TL time series into rainy and dry time steps from Polz et al. (2020). This model provides a continuous probability between 0 and 1 that describes the likeliness that a time step is rainy. Therefore, the choice of threshold that divides the probability values into rainy and dry time steps determines whether classification is more liberal or conservative. We adopted the classification threshold of 0.82 was adopted. Both which was found to be optimal by Polz et al. (2020). This method will be called *CNN*. Fig. 2a) shows an example of the *CNN* probability and the three thresholds used later in the combination of rain event detection methods. Both TSB approaches were compared in Polz et al. (2020) based on hourly data with significantly better performance of *CNN* compared to *RS*.

The use of We computed *RS* and *CNN* on a 1-minute basis.

We used *PC* and *PC-Ph* products from SEVIRI data as a wet/dry indicator can be applied based on different ADB rain event detection by applying a threshold on their precipitation probabilities. We use the thresholds 40 %, used the thresholds 30 %, 20 %, 10 % and 0.1 %. The last threshold represents probabilities greater than 0 %. The abbreviations for the thresholds are e.g., *P40P30* or *P01*, and the abbreviations for the specific data sets of PC and PC-Ph *PC* and *PC-Ph* are e.g., *PC10* or *PC-Ph20PC-Ph10*. Before SEVIRI data can be used as a wet/dry indicator, the SEVIRI pixels must be remapped to the CML paths. This is done by a weighted-averaging, where the weights are determined from the proportion of the CML length covered by the respective SEVIRIpixels. If several pixels are involved, the probability of precipitation often deviates from the original binning into 1 % and 10 % steps. Because of the We forward-filled the 15-minute resolution of SEVIRI data, the probability is kept constant for 15 1-minute time steps of CML data. classification to a 1-minute resolution in the CML processing described in Sec. 3.2.

3.1.2 Combinations of rain event detection methods

According to Polz et al. (2020) it is hardly possible to detect light rain events and accurately define dry periods in CML data at the same time . All existing rain event detection methods use a threshold that allows to choose between a more liberal detection (with better detection of small events but at the expense of more false positives) or a more conservative one (with better detection of dry periods, but at the expense of more false negatives). In The main goal of this study is to improve rain event detection compared to already available methods by combining TSB and ADB methods. Using TSB methods, detecting light rain events is more difficult than detecting strong rain events because it is harder to differentiate smaller attenuation from fluctuations induced by other factors than rain. ADB methods do not use TL for rain event detection, overcoming this issue, but in the case of SEVIRI, uncertainties are introduced by the indirect measurement principle and the difficulties of

separating light rain from no rain. We therefore propose to combine TSB and ADB methods to exploit their advantages. We use different probability thresholds for CNN and the two SEVIRI products to derive rain event detection variants with either high confidence in the correct classification of rainy time steps or high confidence regarding dry time steps. The used thresholds are shown in 2. *CNN10* and *PC01* are liberal variants in the ~~end, a compromise is unavoidable. This is not necessarily the case when combinations of different methods are applied. In order to exploit the advantages of different methods, a combination seems to be particularly promising if auxiliary-data-based (ADB) methods are used in addition to TSB methods. For the implementation, those time steps which can be classified dry or wet with a high degree of reliability have to be identified in both approaches. The challenge is defining the specific threshold and deciding when and where one method is more reliable than the other one.~~sense that they classify rainy time steps already for low probabilities, potentially introducing many FPs. Hence, time steps that are predicted to be dry have a lower chance of being FN. Vice versa, *CNN94* and *PC30* are considered conservative variants because they only classify time steps with a very high probability as rainy. This leads to a low number of FP while introducing more FN.

Our procedure of combining the rain event detection variants consists of five individual steps presented in Fig. 3 and algorithm 1 shown in Fig. A1. We either combined *CNN* and *PC* or *CNN* and *PC-Ph*. In the following the combination of *CNN* and *PC* is explained. In step 1, we choose a method with an intermediate threshold as a starting point which is either *CNN82*, *PC01*, or *PC10*. Step 2 uses the dry time steps from the liberal variant *PC10* (even if it was used as the starting point) that has a high confidence for dry time steps to replace rainy time steps from Step 1. Steps 3 and 4 use time steps with high confidence to be rainy from *CNN94* and *PC30* to replace time steps classified as dry after Step 2. In step 5, dry time steps from *CNN10* are used to replace previously rainy time steps. The combination of *CNN* and *PC-Ph* is identical using *PC-Ph* with the same thresholds instead of *PC*. The results of six combinations are shown in Fig. 5. We named the combinations based on their initial product used for Step 1 as the other steps were identical for each combination e.g. *PC10-combined*

In total, we will evaluate two TSB methods (*RS* and *CNN*, eight ADB methods (derived from the two SEVIRI products *PC* and *PC-Ph* with four thresholds respectively), and six combinations of TSB and ADB methods.

~~In this study we apply a combination of different rain event detection variants, using different thresholds , for the TSB method *CNN* and the ADB method based on PC-Ph SEVIRI data. Each rain event detection variant is first calculated independently on raw CML.~~

3.2 CML data processing

Deriving rainfall estimates from CMLs is a delicate matter (Uijlenhoet et al., 2018). Various research groups developed individual CML processing methods that depend on e.g. the sampling strategy of data. The ~~reasoning of our approach is that dry time steps in a very wet (liberal) variant of a~~ most important aspects of data processing are briefly outlined hereafter while a more detailed description of the processing steps can be found in Graf et al. (2020). We removed default values and outliers that were outside the range [-10,40] dB for TSL and [-99,0] dB for RSL from the analysis. The total path loss along the CML (TL) was then calculated as TSL minus RSL. We interpolated gaps in TL time series of up to five minutes to obtain a more continuous data availability.

~~Flow chart for building a combination based on two different rain event detection methods. In this step-wise approach, which starts from a mean variant, only reliable wet (resp. dry) time steps derived from both methods are used to modify the mean variant based on a dry (resp. wet) variant.~~

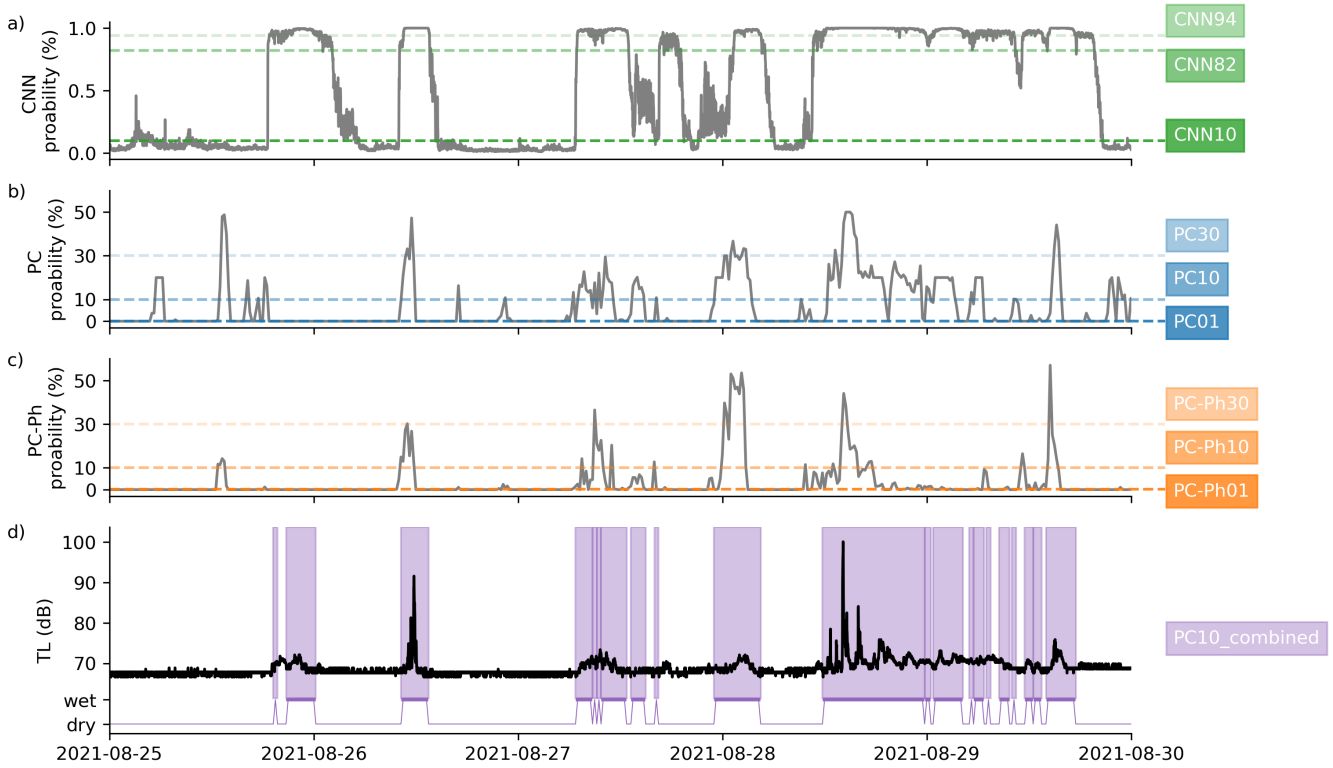

Figure 2. Example time series of five days showing the probability for rainy time steps from a) *CNN*, b) *PC*, and c) *PC-Ph* for the TL time series in d). The dashed lines in a), b), and c) represent the thresholds that separate the probabilities into rainy and dry time steps. Higher thresholds (*CNN94*, *PC30*, *PC-Ph30*) lead to a conservative rain event detection, classifying only time steps as rainy which have a high probability of being rainy. Vice versa, low thresholds (*CNN10*, *PC10*, *PC-Ph01*) lead to a liberal rain event detection, only time steps that are very likely dry are dry. Additionally, the combination of *PC10* with *CNN* and *PC* variants as rain event detection methods is shown in d).

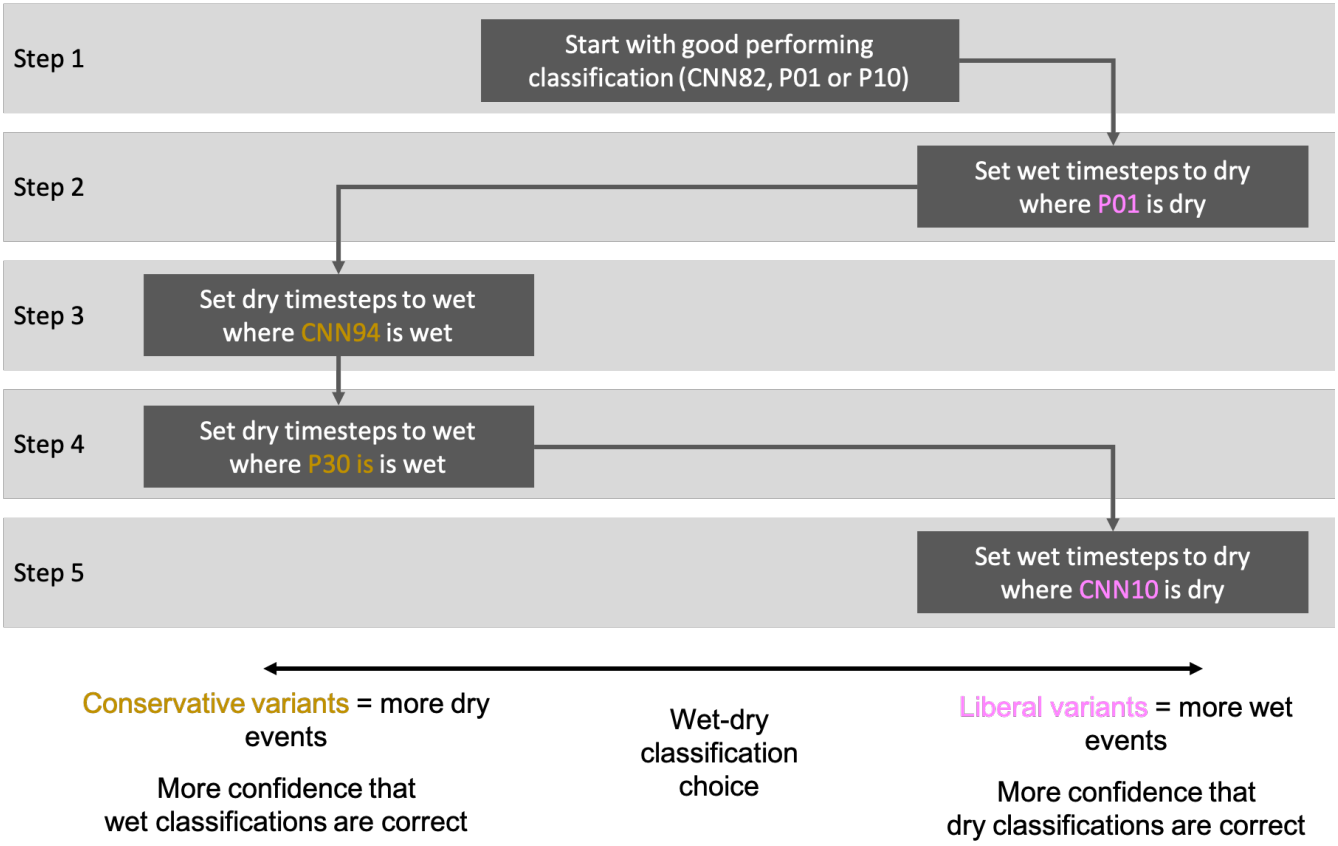

Figure 3. Flow chart for building a combination based on TSB and ADB rain event detection methods This step-wise approach starts with an already well-performing classification as a starting point. Then, during each step, time steps that have a high likeliness of being dry (or we) from a very liberal (or conservative) classification from the two methods used are added to the previous classification. If *PC* is used as a starting point, Step 2 and 4 also use *PC*. The same is true for *PC-Ph*. For **CNN82** as a starting point, we computed one variant with *PC* and one with *PC-Ph*. An algorithmic description of this flow chart can be found in Fig. A1.

For the rain event detection~~are most likely also dry in reality, and vice versa for wet predicted time steps in a very dry (conservative) variant. The~~ , we used two TSB methods, two ADB methods with four different thresholds as well as six combinations of TSB and ADB methods. As SEVIRI is only available with a 15-minute resolution, we resampled their classification to a 1-minute resolution with a forward-fill. We computed the following processing steps for each rain event detection method on a 1-minute basis to derive individual rainfall estimates. The baseline attenuation was dynamically identified from the preceding dry period of each rain event to derive the rainfall-induced attenuation along the path. The baseline was determined by the last dry time step of the TRSL time series and was set to be constant during the rain event~~detection of these highly reliable time steps is adopted and used to overrule the classification as wet or dry generated from a previously calculated default rain eventdetection, which has only a lower confidence.~~

~~Our procedure of combining the rain event detection variants consists of 5 individual steps shown in Fig. 2. In step 1, we choose a method (either CML-TSB or SEVIRI-ADB)that shows a good average performance as a first guess. Step 2 and 5 use wet variants of both data sets and provide additional high-confidence dry time steps. Step 3 and 4 derive additional high-confidence wet time steps from dry variants of both data sets. Since later steps can modify earlier classifications of wet and dry time steps, the priority of the individual steps increases.The order of steps 2-5 might be changed but it is recommended that two conservative and two liberal variants are applied. The chosen individual combinations are mentioned in Chapter 4.3.~~

. As Water on the CML antenna covers can lead to additional attenuation, this so-called wet antenna attenuation (WAA) effect has to be compensated for. We used the WAA correction from Leijnse et al. (2008). In this physical approach, the WAA depends on antenna cover properties (refractive index and thickness), microwave frequency, and rain intensity. We used the parameters given by Leijnse et al. (2008). The rainfall rate was derived from WAA-corrected attenuation using the k-R relationship. The parameter settings for the k-R relation were taken from ITU recommendations (ITU-R, 2005). From a purely practical point of view, we only evaluated data from the first of the two available sublinks.

For evaluation with the reference, the binary classifications from the TSB methods and all resulting rain rates were resampled to a 15-minute resolution. We believe that a resampling to 15 minutes and the fact that *RADKLIM-YW* is adjusted to rain gauges suffices to overcome the potential temporal mismatch between radar and ground-based CML observations. The evaluation with a 15-minute resolution might lead to worse results than one on an hourly basis. But for future applications of SEVIRI products for CML processing in regions with sparse reference data like Sub-Saharan Africa, this high temporal resolution is of advantage.

3.3 Statistical measures

The ~~Pearson correlation PCC~~ radar rainfall estimates at a 15-minute resolution serve as the ground truth for the computation of the scores listed below. For binary classification scores, the ground truth is considered wet if the path-averaged radar rain rate (r_{ref}) along the CML path is larger than 0.1mmh$^-$1. The Pearson correlation coefficient (PCC) is used to evaluate the ~~shape and temporal agreement of different precipitation time series. Focusing only on whether wet and dry time steps match between different data sets leads to categorical scores based on a~~ quality, in terms of the linear correlation, of different CML rainfall

estimates (r_{cml}) derived by using the proposed methods for rain event detection.

$$PCC = \frac{\sum(r_{ref} - \overline{r_{ref}})(r_{cml} - \overline{r_{cml}})}{\sqrt{\sum(r_{ref} - \overline{r_{ref}})^2}\sqrt{\sum(r_{cml} - \overline{r_{cml}})^2}}, \quad (1)$$

where the $\overline{r}$ indicates the mean of a quantity. The relative bias (RB, Eq. 2) is then used to measure an over-, or underestimation that cannot be derived from the PCC.

$$RB = \frac{\sum(r_{ref} - r_{cml})}{\overline{r_{ref}}}, \quad (2)$$

Binary classification scores are based on the confusion matrix (see ~~equation~~ Eq. 3):

$$\begin{pmatrix} TP & FP \\ FN & TN \end{pmatrix} = \begin{pmatrix} WET/wet & DRY/wet \\ WET/dry & DRY/dry \end{pmatrix} \quad (3)$$

with uppercase and lowercase denoting observed ~~event and prediction~~events and predictions, respectively. The ~~number of~~ correctly assigned wet time steps are called True Positives (TP) and the correctly assigned dry ones are True Negatives (TN). False Positives (FP) represent the number of time steps where rain event detections are incorrectly assigned wet time steps and False Negatives (FN) represent the number of incorrect dry time steps. The confusion matrix fully explains the performance of a classifier, but since the interpretation of four individual numbers is not straightforward, the computation of additional scores is necessary. The first simplification is to reduce it to the pair of True Positive Rate (TPR~~), the Positive Predictive Value (PPV), the Accuracy Correlation Coefficient (ACC) and the Matthews Correlation Coefficient (MCC) were selected from a large number of possible categorical statistical measures.~~

, Eq.4), that is the probability of a positive event being predicted positive, and False Positive Rate (FPR, Eq.5) also called false alarm rate. We include both scores since the importance of a high TPR or FPR may be weighted differently depending on the application of the CML rainfall rates.

$$TPR = \frac{TP}{(TP + FN)} \quad (4)$$

$$PPV FPR = \frac{TP}{(TP + FP)} \frac{FP}{(FP + TN)} \quad (5)$$

$$ACC = \frac{(TP + TN)}{(TP + FP + FN + TN)}$$

$$MCC = \frac{(TP \cdot TN - FP \cdot FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

~~The TPR is the probability that an actual positive (wet time step) will also test positive. For the PPV only the positive values (TP and FP) are used.It is applied to show how "wet" a certain data set is. ACC compares the true values to the total population. According to Polz et al. (2020) it depends on the class balance. With a precipitation probability of 5-10 % in Germany, the majority of time steps are dry. Due to this skewed distribution the main focus of~~ In this manuscript, we focus on improving the overall performance of rain event detection and thus use the Matthews Correlation Coefficient (MCC, Eq.6) which is more robust to influences of the ~~analysis is based on the MCC which is affected much less by unbalanced distributions~~skewed distribution of wet and dry classes (the ratio is roughly 1:20). The MCC is high ~~, if~~ only if the detection performance for both wet and dry classes ~~are well represented.~~

3.4 ~~Intensity classes~~

~~Despite the methodologically different rain event detections, the separation of dry and wet time steps is usually easier in the case of heavy precipitation since either the precipitation signal differs more clearly from the baseline CML signal or auxiliary precipitation data can more easily detect heavy precipitation.In order to confirm the assumption that heavy precipitation is easier to detect than light precipitation and to be able to evaluate~~ is high.

$$MCC = \frac{(TP \cdot TN - FP \cdot FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{6}$$

The classifier Accuracy (ACC, Eq.7) is used to analyze the performance for ~~different intensities, four additional intensity classes are applied in addition to the total precipitation. The classes for light, moderate and heavy precipitation are based on the (DWD, 2023) classification. "Light" precipitation is subdivided into two classes. This results in the five classes for "light1", "light2", "moderate", "heavy"~~ the different rainfall intensity classes and ~~"total" precipitation shown in Tab. 2. Accordingly, the lower threshold for precipitation in this analysis is 0.1 mm/h and values below are considered dry~~ gives the percentage of the time steps in the intensity class that were detected as wet (dry for the dry class).

~~Overview of rain intensity classes for evaluations **Title light1 light2 moderate heavy total**intensity mm/h0.1 – 1.0 1.0 – 2.5 2.5 – 10.0 > 10.0 >= 0.1~~

$$ACC = \frac{(TP+TN)}{(TP+FP+FN+TN)} \tag{7}$$

4 Results~~and Discussion~~

~~We compared raw SEVIRI products with RADKLIM-YW to assess its suitability as a ADB wet/dry indicator for~~ In this section, we first compare the ADB products derived from Meteosat-SEVIRI with the weather radar reference to assess their suitability as wet-dry indicators within the processing of CML data. ~~Then, we analyzed the performance of CML rainfall derived with a total of eight TSB and ADB rain event detection methods. Based on these results, we created six TSB and~~ Subsequently, we analyze the relative performance of ADB ~~combinations and evaluated their performance compared to individual TSB and ADB methods. Finally, we investigated the~~ and the combination of ADB and TSB methods with respect to the performance of established TSB methods. Then, we analyze the performance regarding different rain intensity classes and investigate the

influence of different rain event detection methods on the performance of individual CMLs. All presented rain event detection methods were used in combination with the CML processing routine described in Sec. 3.2 and the resulting data sets were compared to the path-averaged weather radar reference on a 15-minute basis. All scores computed for all methods and the full dataset are additionally shown in (see Tab. B1).

4.1 ~~Performance~~ Rain event detection performance of ~~raw~~ the *PC* and *PC-Ph* SEVIRI ~~products~~product

To assess the quality of ~~SEVIRI data it was directly compared to radar data. At this first step, no CML data was involved, but both data sets have been path-averaged~~ the SEVIRI products, they can be compared directly to the weather radar reference. To overcome the issue of their different grid sizes and to asses their quality for using them for CML processing, we compared SEVIRI and *RADKLIM-YW* data as path averages along the CML paths.

~~In Fig. 3a, TPR shows the percentage of wet *RADKLIM-YW* time steps per intensity class and precipitation amount, represented by SEVIRI data for different thresholds.The largest differences were obtained for small intensities. While for *P01* around 80 % of the wet~~ To compare the PCC of the different SEVIRI-based products and the applied thresholds, processing the CML rainfall rates according to Sec. 3.2 was done using the SEVIRI wet-dry indicator.

Fig. 4 compares the performance of the ~~*RADKLIM-YW*~~*PC* ~~time steps were correctly assigned wet, this proportion decreased to around 20 % for *P30*. For higher intensities, the proportion was over 90 % for *P01*~~ and ~~roughly 60 % for *P30*. In this way, 90 % of *RADKLIM-YW*'s precipitation amount was covered by *P01* and only 50 % by *P30*. Differences between day and nightexisted, but the general behavior over different rain intensity was similar.~~

~~According to TPR(blue column) in Fig. 3b, a maximum of 80 % of the wettime steps (according to radar data) are also classified wet by SEVIRI (~~*PC-Ph* products according to the TPR, FPR, MCC, and PCC scores. The highest TPR (0.83 during the day and 0.9 during the night) was achieved by the lowest threshold for *PC* and *PC-Ph* (*P01*~~). However, this corresponds to only 25 % (nighttime) and 40 % (daytime) of the total SEVIRI wet time steps (*P01*, PPV). Consequently, *P01* ~~), which also showed the highest FPR (0.14 day and 0.22 night). For *PC*, the TPR and FPR were increased during the night, but for *PC-Ph*, only the FPR was significantly increased during the night, except for the *P30* version which also showed a higher nighttime TPR. Both, the TPR and FPR decreased with increasing threshold which is due to the decreasing number of positive predictions. The ratio of $TP + FP$ and $TP + FN$ describes how many more timesteps a method predicted as wet, compared to the reference. Accordingly, *PC01* showed 195% (151% for *PC-Ph01*) more wet time steps than the radar data ~~. and *P30*~~*PC30* showed the opposite behavior with 36% (52% for *PC-Ph30*) less wet time steps than the reference (see Tab. B1).

The MCC shown in Fig. 4 which measures the overall classification performance showed that the *P10* threshold yields the best results for *PC* during day and night. ~~This was also reflected in the performance metrics ACC and MCC.~~ The best ~~performance based on MCC was achieved by~~ results for *PC-Ph* were achieved by the *P10* ~~as a compromise between coverage of wet radar time steps and a meaningful number of wet and dry time steps. MCC was 0.53 during the day and 0.42 at night. This difference can probably be attributed to the missing SEVIRI channels at night.Although the results between PC and PC-Ph were similar,~~

there were slight tendencies: PCC revealed a significantly better performance for PC-Ph than for PC. ACC showed similar tendencies, while the results for MCC were inconclusive.

a) TPR of SEVIRI data and *RADKLIM-YW* per intensity class and the resulting *RADKLIM-YW* precipitation amount represented by PC (left column) and PC-Ph (right column) for SEVIRI rain probabilities from *P01* to *P30* for day (top) and night (bottom). b) respective TPR without subdivision into intensity classes and the statistical performance measures PPV, ACC, MCC and PCC.

According to these analyses, threshold during daytime and by the *P01* threshold during nighttime. Fig. 7b) shows the accuracy of *P01**PC* and *P10*are promising candidates representing *RADKLIM-YW* precipitation. The direct comparison of SEVIRI data and radar data was only of limited significance since the combination of CML raw data signals and satellite information is important when using SEVIRI as a wet/dry indicator. So, the next step was to analyze CML data based on SEVIRI and the TSB rain event detection methods *RS* and *CNN*.

4.2 Evaluation of individual methods for rain event detection

Eight different rain event detection methods were used within one CML processing routine described in Chapter 3.1 and the resulting data sets were PC-Ph. The accuracy increases with a higher threshold whereas the relative bias shown in Fig. 7d) becomes more negative with higher absolute values suggesting an increasing underestimation compared to the reference *RADKLIM-YW*. First, they were evaluated based on the frequency of rainfall occurrence and total precipitation amount. This described the average behaviour, but did not allow to draw any conclusions on the temporal correspondence of CML data sets and radar data.For this purpose, statistical measures radar. This is due to an increasing amount of TN and FN predictions. When computing CML rainfall rates based on the confusion matrix were compared in the second part of this subsection.

4.1.1 Effect of rain event detection on total counts of wet time steps and precipitation

The distribution of rain intensities and total precipitation amount of eight rain event detection methods is shown in Fig.4. Percentage deviation of CML with the rain event detections *RS*, *CNN* as well as the SEVIRI probabilities represented by PC (left bars) and PC-Ph (right bars) *P01* till *P20* and *RADKLIM-YW* for counts per intensity class and precipitation amount for day (top) and night (bottom). Blue (resp. red) columns indicate overestimations (resp. underestimations) of CML data sets. Overall, all rain event detection methods underestimated the occurrence of light precipitation and overestimated the occurrence of heavy precipitation, albeit with a general lower frequency of occurrence for the latter. Since all data sets showed this tendency, this can be attributed more to the general performance of CML data than to *PC* and *PC-Ph* wet-dry labels, the rain event detection method.This behaviour probably resulted from both the WAA correction tending to correct too many small intensities (Graf et al., 2020), and a tendency of radar data to underestimate heavy rain intensities (Schleiss et al., 2020). The highest PCC was achieved using the *P01* threshold with values around 0.72. *P10* showed only marginally lower scores. When higher thresholds were applied, the performance decreased more severely. Overall, *PC* showed slightly higher scores than *PC-Ph* and the performance during daytime was equal to or better than the performance at nighttime.

According to this analysis, *P01* and *P10* were the most promising thresholds to apply for SEVIRI-basedADB data sets behaved

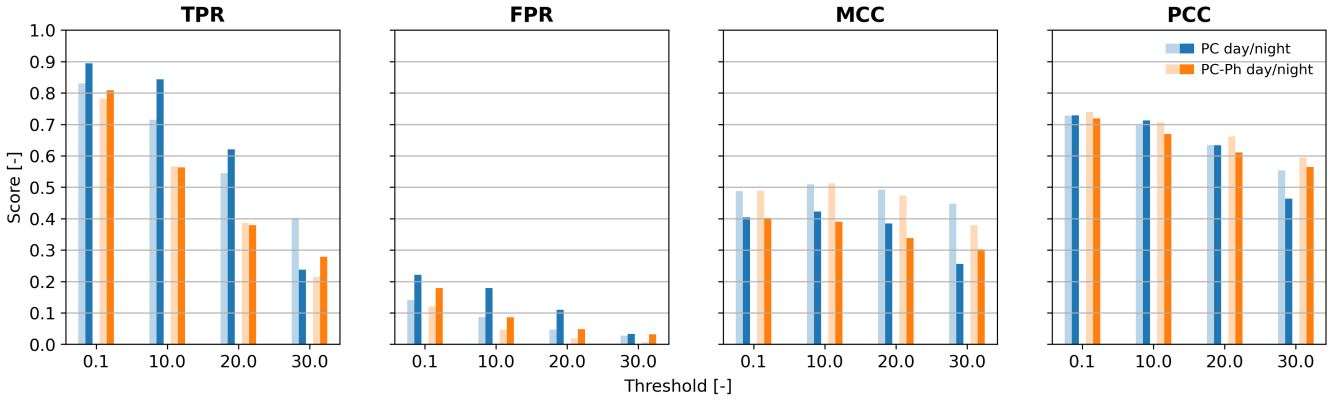

Figure 4. TPR, FPR, MCC, and PCC of the *PC* (blue) and *PC-Ph* (orange) products when compared to the radar reference. The results of each score are presented for different thresholds (x-axes) and split into day (light colors) and night (dark colors).

~~very similarly to the two TSB data sets~~ , pure ADB wet-dry detection. Both, ~~*RS*~~*PC* and ~~*CNN. The chosen probability level of SEVIRI data dominated the behaviour with a clear decrease of rainfall occurrences with higher precipitation probabilities. For ADB methods the night is slightly wetter than the day and a similar tendency occurred in the TSB data sets. Regarding precipitation amounts, the TSB data sets as well as SEVIRI P01 and PC10 (left column of p10) agreed very well with RADKLIM-YW~~*PC-Ph* show good classification and regression scores for these thresholds. Because of the bad PCC, *P30* is not suitable as a threshold for a standalone ADB method and we did not analyze its performance any further. However, we did use it for wet labels with high confidence in the combination algorithm. ~~*PC-Ph10* (right column of p10) and especially p20 tended to underestimate precipitation amounts. Based on these results, a performance similar to that of the TSB data sets was to be expected from the SEVIRI data.~~

4.1.1 ~~Performance measures based on the confusion matrix~~

~~Based on the confusion matrix, the MCC was chosen as a meaningful criterion for performance with regard to rainfall intensity, differences between the two SEVIRI products as well as day and night differences. Again, there was a subdivision into intensity classes.The analysis~~

4.2 Comparison of ADB (*PC* and *PC-Ph*), TSB, and combined rain event detection methods

Due to the increased number of products that are evaluated in this subsection, we limit the analysis to using only the MCC as a measure of classification performance since it gives the best summary of the confusion matrix in on scalar value. Fig.5 shows an analysis that was carried out for the ~~TSB data sets~~ two TSB methods *RS* (grey) and *CNN* ~~and for~~ (green) and for six ADB methods using the three lowest thresholds ~~of PC and PC-Ph~~ (.1, 10 and 20) ~~(Fig.5) . The lowest MCC was found for rain intensities below 2.5 mm/h (first two blocks), which illustrates the difficulty of rain event detection for CML data. This~~

was particularly important as these low intensities made up a large amount of all rainfall and therefore dominated the mean statistical measures. Both TSB and ADB approaches naturally show weaknesses for light rain. TSB methods have to distinguish between precipitation signal and noise and the ADB method with SEVIRI data has to distinguish between precipitation and non-precipitating clouds. For higher intensities all rain event detection methods were robust. of *PC* (blue) and *PC-Ph* (orange). The SEVIRI products, *PC* and *PC-Ph*, performed similarly well as the *CNN* and *RS* methods during daytime with MCC scores ranging from 0.49 to 0.51. The best performance and highest MCC values were observed for moderate rain where precipitation signals are clearly visible in the CML time series and hardly any convective and small-scale precipitation is involved, which is more difficult to detect.

Comparing both SEVIRI products, PC performed slightly better than PC-Ph for most intensities and precipitation probabilities. The pre-analysis in Chapter 4.1 showed different results. The reason for this deviant behaviour is that for the latter case, the combination of CML signals and the SEVIRI product was responsible for the performance. *P01* for instance showed too many wet time steps but provided very good results when applied as a wet/dry indicator. The probability threshold had a larger influence on the rain event detection and hence the performance of the product than the differences between PC and PC-Ph.

In general, *CNN* performed better than *RS* for all intensities as shown by Polz et al. (2020) for a different period. SEVIRI performed better for smaller probability thresholds. during daytime was achieved by *PC-Ph10* and the worst by *PC-Ph20*. During the nighttime, the CNN outperformed the ADB methods, while the standard deviation approach showed a similar performance to *PC* and *PC-Ph* except for the higher *P20* threshold which showed the worst performance. Probabilities of 20 % and above (not shown here) usually showed even worse resultsthan *RS*.In contrast, *P01* performed at least as well as *CNN*, usually even better.

Same CML data sets and classes as in Fig. 4 but showing the MCC compared to *RADKLIM-YW*. *PC10* will be applied for further evaluations and is marked in brown.

At night, the results of *RS* were worse than during the day. This was most notable for small intensities. One reason could be dew formation on the CML radomes antenna cover causing an increasing signal attenuation due to increasing number and size of droplets on the radomes. The *RS* method is only partially capable of distinguishing between such dewing events and precipitation. resulting in increased standard deviation which leads to false classification as wet. The *CNN* did not show differences between day and night because suggesting that it coped better with this issue. The SEVIRI-based *SEVIRI*-based products also showed slightly worse results for during nighttime, but this was probably more likely due to the lack of three SEVIRI *SEVIRI* channels in the VIS and NIR at night. The calculation of microphysical variables is mainly based on those channels. Their absence consequently limits the reliability of all derived precipitation products. Differences between day and night already became apparent for the same precipitation probability with usually wetter nights. Additionally, ADB methods such as SEVIRI's wet/dry indicators are more independent from noise and artefacts in the CML time series. Hence, they were less affected by dew than TSB methods.

SEVIRI-based rain event detection with probabilities up to 10 % provided a at least similar performance as the *CNN* method and showed significant improvements compared to the *RS* method.*PC10* (marked in brown) showed the best performance when

combining with the CNN method (described and shown in the next section) and is thus used as a new ADB method for all final results in the following sections.

4.3 Combinations

The main goal of this study is to improve the and as *CNN* during the daytime, but performed worse than *CNN* during the night. Fig. 6 shows an example time series of TL, TP, FP, and FN classifications, and rainfall rate of the *PC10* and *CNN* methods as well as the combined method *PC10-combined*. It can be seen that the choice of rain event detection method has an impact on the resulting rainfall rate. The combined method improves the classification performance and correlation of rainfall rates compared to the conventional methods.This can be achieved by combining the advantages of different methods and overcoming certain disadvantages in this way. The main principle is to use a robust rain event detection, which performed well over all analysed metrics and rain intensity classes, as a first estimate. Next, very conservative (liberal) rain event detection methods were used to derive time steps with high confidence of being wet (dry) and replaced the time step from the previous estimate. For example, wet time steps were taken from a conservative method which had only few but trustworthy wet time steps. Based on our data, the following five processing steps (see Fig. 2) and the order thereof turned out to provide best results: pure ADB and TSB methods.

1. Either *CNN*In addition to the individual TSB and ADB methods, Fig. 5 shows six different combined methods (purple and red colors) that are composed of either *cnn*, *PC01*, , or *PC10* , *PC-Ph01* as a starting point, and done for either *PC* or *PC-Ph10* served as a first estimate (typical individual method for rain event detection)

2. Additional dry time steps were derived from the very wet SEVIRI variant *PC01* or *PC-Ph01*.

3. Additional wet time steps were based on the very dry variant of *PC-Ph*. As before *PC* is shown on the left and *PC-Ph* on the right.

The main difference among the combinations is that *PC* data performed better than *PC-Ph* at daytime and similar at nighttime. *CNN* and *PC(-Ph)10* were the best performing TSB and ADB methods and were, therefore, compared to the combined products. *PC10-combined* showed the largest overall improvements with an MCC of 0.59 compared to 0.52 (*CNN*) and 0.52 (*PC-Ph10*) at daytime and 0.55 compared to 0.50 (*CNN*) and 0.41 (*PC10*) at nighttime. *PC10-combined* also showed the lowest relative bias of -2.1% compared to *CNN* *combined* with a threshold of 0.94.

4. Further wet time steps were derived from SEVIRI with a precipitation probability of at least 30 % (PC30 or PCPh30).

5. Finally, time steps were assigned dry, which were dry based on *CNN* with a threshold of 0.1 (very wet variant).

The combinations are either based on PC or PC-Ph to minimize the expense of data processing. If a SEVIRI data set is chosen as the first estimate, Step 2 has no effect on 2.6% and *PC01 combined* with -10.8% (see Fig. 4 d). The rel. bias was also lower than that of pure TSB or ADB methods. In the following, we concentrate on *PC10-combined* to evaluate the performance gain using a combination of ADB and TSB methods in more depth, for different intensity classes and individual CMLs.

PC10 showed the best performance when combined with the selection of wet and dry time steps. *CNN* method (described and shown in the next section) and is thus used as a new ADB method for all final results in the following sections.

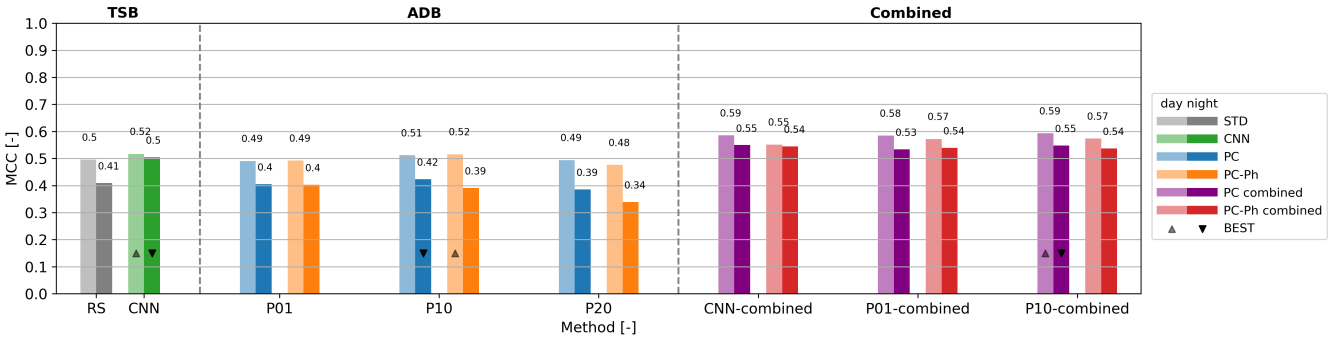

Figure 5. Matthews correlation coefficient (MCC) of pure TSB, ADB, and combined methods. *P01*, *P10*, and *P20* refer to the respective *PC* and *PC-Ph* thresholds. The results are split into day- (light colors) and night-time (dark colors). The best performing day and night products of each group are marked with a triangle.

~~Figure 6 shows how the TRSL (top) was interpreted by different rain event detection methods for an arbitrary CML and 5 days.The *CNN* led to an overestimation of wet time steps and to more precipitation than radar data showed. *PC10* was much dryer with a tendency of underestimation of wet time steps. PC10allwas a combination applying *PC10* as a first estimate in step 1 of building a combination. It can also visually be regarded as a combination of the other two rain event detection methods.~~

~~To assess the quality and~~

4.3 Performance of rain event detection methods for different rain intensity classes

Fig. 7 shows the accuracy and relative bias for the TSB, ADB, and combined methods. The relative bias is, for all individual intensity classes, a percentage of the average reference rainfall rate (mean over all classes) such that Fig. 7 b) is the sum of the ~~potential of combinations, the following analysis were performed: First, the best of six combinations were chosen based on the MCC. Then the selected combination, a TSBdata set and a ADBdata set were compared in-depth, based on differences of TP, FP and FN with rain intensity. Finally, the performance of these three data sets was evaluated with regard to individual CMLs.~~

4.3.1 ~~MCC of combinations~~

~~In order to evaluate the performance of combinations, the MCC of the TSB data sets and six different combinations are plotted~~ classes shown in Fig. ~~??, again the variant with PC on the left and with PC-Ph on the right. *call* means combination applying~~ 7 a).

It can be observed that for all methods the accuracy increased with increasing intensity and that the *CNN* ~~in the first step of building a combination and~~ performed better than *RS* for all intensities which confirms the results of Polz et al. (2020) for the same CML network, but a different period. Therefore, we omit ~~*p01allRS* and *p10all* the SEVIRI data with probabilities of 0.1 % and of 10 % respectively.~~

~~Example time series of a CML for 5 days. Raw TRSL (top) and *CNN*, *PC10* and *PC10all* (green line) are opposed to *RADKLIM-YW* (black line). Wet time steps which are assigned dry by CML (FN) are marked shaded red and dry time steps which are assigned wet (FP) are marked shaded blue.~~

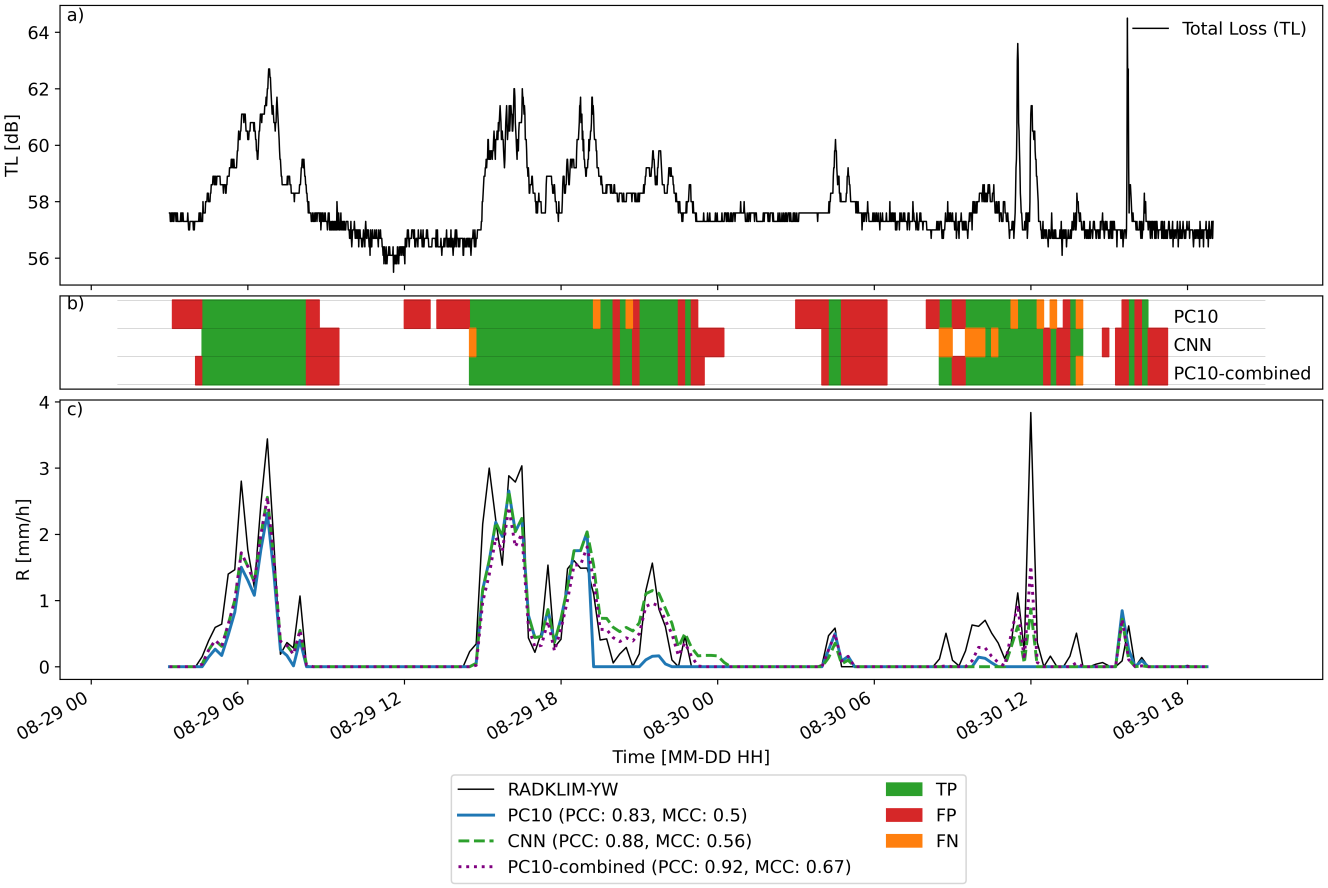

Figure 6. Example time series of total loss (TL) (a), classifications compared to the reference (b), and rainfall rates (c). The scores provided in the bottom legend are computed for data from the CML and period shown in this figure.

Same as Fig. 4 but for the combinations where *CNN* (*call*) and PC/PC-Ph with thresholds *P01* (*p01all*) and *P10* (*p10all*) serve as the first estimate. *PC10all* will be applied for further evaluations and is marked in brown.

The improvements by combining in the remaining results section. Overall, all rain event detection methods became obvious for light precipitation with an increase of the MCC between 0.3 and 0.1 for the lowest intensity classes. There was only minimal improvement for moderate precipitation, and no improvement for heavy precipitation.The main difference among the combinations is that PC data performs better than PC-Ph especially for lowest intensities and as a result of the large number of time steps there, also for the total class lead to an underestimation of light to heavy rainfall which was only partly compensated by an overestimation from false positives (dry class).

PC performed better than *PC-Ph* for light to moderate intensity and similar for heavy intensity, but had a lower accuracy for the dry class suggesting more false positives for *PC* when the same threshold was chosen.

Although more methods are shown, we will focus on the best TSB (*CNN*), ADB (*PC10*), and combined (*PC10-combined*) methods from here on to sharpen the analysis.

For the lowest class light1 we observed large differences in accuracy with 60.1% (*CNN*), 73.2% (*PC10*), and 77.8% (*PC10-combined*), while the accuracy was similar for the highest class heavy with 93.7% (*all*). Very similar results were obtained for the night.Again, PC performed slightly better than PC-Ph and combinations were superior to the applied TSB data sets.The combination *PC10all* (marked in brown)showed the largest overall improvements with a MCC of 0.62 compared to 0.57 (*CNN*)and 0.54 (, 91.0% (*PC10*), and 94.3% (*RSPC10-combined*)at daytime and 0.6 compared to 0.55 (. The TSB and combined methods showed a similar dry accuracy which was higher than for the ADB methods.

The positive relative bias due to FPs (Fig. 7c) dry class) is similar for *CNN*) and 0.47 (*RS*)at nighttime. Combinations are most effective for the very low intensities where the largest uncertainties of individual methods were observed. For individual methods, there is often a trade-off between few FN (liberal)or few FP (conservative)that might not be necessary when applying two data sets. Analyzing the same time steps or time series by different methods likely increases the reliability of a resulting rain event detection. In the following, we concentrate on PC10allto evaluate the performance of such a combination in more depth. (11.3%), *PC10* (8.9%), and *PC10-combined* (12.7%). The relative bias for the light, moderate, and heavy classes is negative (underestimation) for all methods but has a smaller absolute value for *PC10-combined*. For example, the bias for the light2 class is -7.1% (*CNN*), -8.4% (*PC10*), and -4.4% (*PC10-combined*). Overall, *PC10-combined* (-2.1%) shows a much lower bias than the TSB and ADB methods *CNN* (-11.5%) and *PC10* (-20.8%).

4.3.1 Performance of confusion matrix values vs. intensity

In order to Fig. 8 shows histograms of the occurrence and accumulated rainfall amount of *CNN*, *PC10*, and *PC10-combined* using logarithmic bin widths to be able to understand what distinguishes the individual methods for rain event detection, the distribution of TP, FP and FN is plotted against intensity and shown for frequency of occurrence (top) and total precipitation (bottom) in Fig. **??**. The left column show those distributions for the *CNN* methodvisualize differences for all intensity classes. *PC10-combined* showed the highest count of TPs and the lowest count of FNs. Despite this, the number of FPs was lower than for *PC10*. *PC10* showed a similar TP count as *CNN*, but also has the highest amount of FPs. The other columns reveal

the corresponding differences between *PC10* and *CNN* was the most conservative method with the lowest amount of FPs. The average rainfall amount per count per CML was different for TPs and FPs. TPs occurred for higher rainfall rates and, therefore this average was 0.43 mm (*CNN*), 0.39 mm (*PC10*), and 0.40 mm (*PC10-combined*). FPs occurred for lower rainfall rates and, therefore the average was 0.09 mm (middle)and between the combination *PC10all* and *CNN* (right)for day andnight. In summary, a typical TSB method (*CNN*), a new ADB method (*PC10*) and a combination thereof (*PC10* combined) and 0.07 mm (*PC10all**PC10*)were compared.

Top: The number of FP, TP, and FN time steps with a specific rain rate are compared for *CNN* (left) and the respective amount of rainfall (bottom). *diff_PC10-CNN* reveal the respective differences between *PC10* and *CNN* (middle) and *diff_PC10all-CNN* between *PC10all* and *CNN* (right). a) is daytime and b) is nighttime. The rain intensities for FP and TP are estimated by the CML, while the rain intensities for FN are taken from the reference.

According to Fig.**??**a (middle), . The average missed rainfall per FN, as measured by the radar reference, was 0.16 mm (*CNN*), 0.21 mm (*PC10*is drier than), and 0.15 mm (*CNNPC10-combined*with less TP.Improvements were achieved by a decrease in FP but also a similarly large increase in FN during the day showed up. The results for the combination were different. Here, the desired effect of a general improvement became apparent: both FP and FN decreased for all intensities.TP and thus precipitation were slightly increased.

At night,).

To analyze the confidence of individual methods for light, moderate, and heavy rainfall predictions, we computed the probability of a positive prediction to coincide with the reference (i.e. the ratio $\frac{TP}{TP+FP}$). For the light1 class, *PC10PC10-combined* is much wetter than during the day with an increase of TP and a decrease of FN compared to was the most confident method with 0.56 compared to 0.55 and 0.53 (*CNN* . For intensities below 0.5 mm/h, there was a massive increase of FP, but not for higher intensities. The combination *PC10all* also shows this SEVIRI issue, however, to a lesser extent. There are clear improvements for all intensities, except for intensities below 0.5 mm/h where FP was also increased compared to and *PC10*). For the light2 and moderate classes, *PC10* was the most, and *CNN* . Since nights were usually wetter based on SEVIRI data, a larger precipitation probability than 10 % to reduce FPs from SEVIRI might lead to overall improvements, but this was not tested in this investigationwas the least confident method. For the heavy class, *CNN* was the most, and *PC10-combined* was the least confident method.

4.4 Influence of chosen rain event detection method on individual CMLs

The results so far were based on averaged metrics over all CMLsmetrics that we computed using all CML data. The performance of individual CMLs might, however, differ from this mean behaviour behavior, and systematic differences between well and worse performing CMLs are possible for different rain event detection methodsCMLs that compare well or badly with the reference are possible. Therefore, scatterplots of different rain event detection methods and combinations against the *CNN* method for ACC, the MCC and PCC show the change of performance of each CML calculated for each CML individually are shown to compare *CNN* and *PC10* to *PC10-combined* in Fig. **??**. *PC10* and 9. For the majority of CMLs, the MCC and PCC could be improved by using *PC10all**PC10-combined* are compared with instead of the best TSB (*CNN*for day and night.

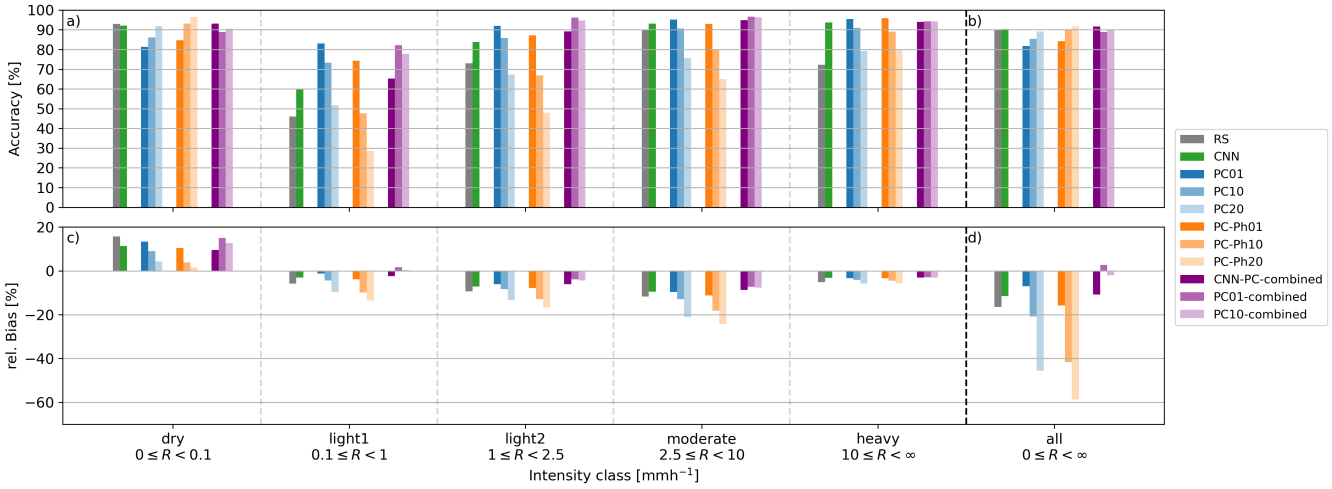

Figure 7. Accuracy (top) and relative bias (bottom) of selected TSB, ADB, and combined methods (see legend). Both scores are shown for different intensity classes (x-axis) as defined by the radar reference. The accuracy is the percentage of correctly detected rain events in this intensity class. The relative bias is computed as a percentage deviation of the mean of the full reference data such that the five leftmost classes add up to the 'all' class on the right.

~~The differences between~~) and ADB (*PC10*~~and *CNN* were small. No systematic differences were found using PCC. Based on ACC, the majority of CMLs are slightly improved by~~) method. The average of the MCCs using *PC10* ~~including the worst performing ones.Regarding MCC, the overall performance is equal for~~ (resp. *CNN*~~and~~) was increased from 0.44 (0.51) to 0.56. The CMLs with the worst MCC from *CNN* were improved most when using the combination. For PCC the improvement was smaller (from 0.75 (0.78) to 0.79) but still affected more than 80% of all CMLs. While *PC10-combined* compared to *PC10* improved the PCC by up to 0.36 for individual CMLs, the largest improvement from *CNN* to *PC10-combined* was 0.15. CMLs with the worst PCC could not be improved by using *PC10-combined*.

Compared to the overall PCCs of *PC10*, ~~only the worst performing ones were almost always improved here, too. A very similar picture emerged at night, with improvements in particular for the worst performing ones. These CMLs with low performance usually showed a high level of noise which is similar to the signal fluctuations caused by rainfall. Consequently,~~ *CNN*, and *PC10-combined* shown in Tab. B1, the ~~separation of precipitation signals is difficult for all TSB methods.The fact that~~ mean of all CML's PCCs is slightly higher for all three products. This means that individual CMLs have a higher linear correlation with the reference than the full dataset. Therefore, one can assume that individual CMLs show different biases that could not be compensated by any of the methods.

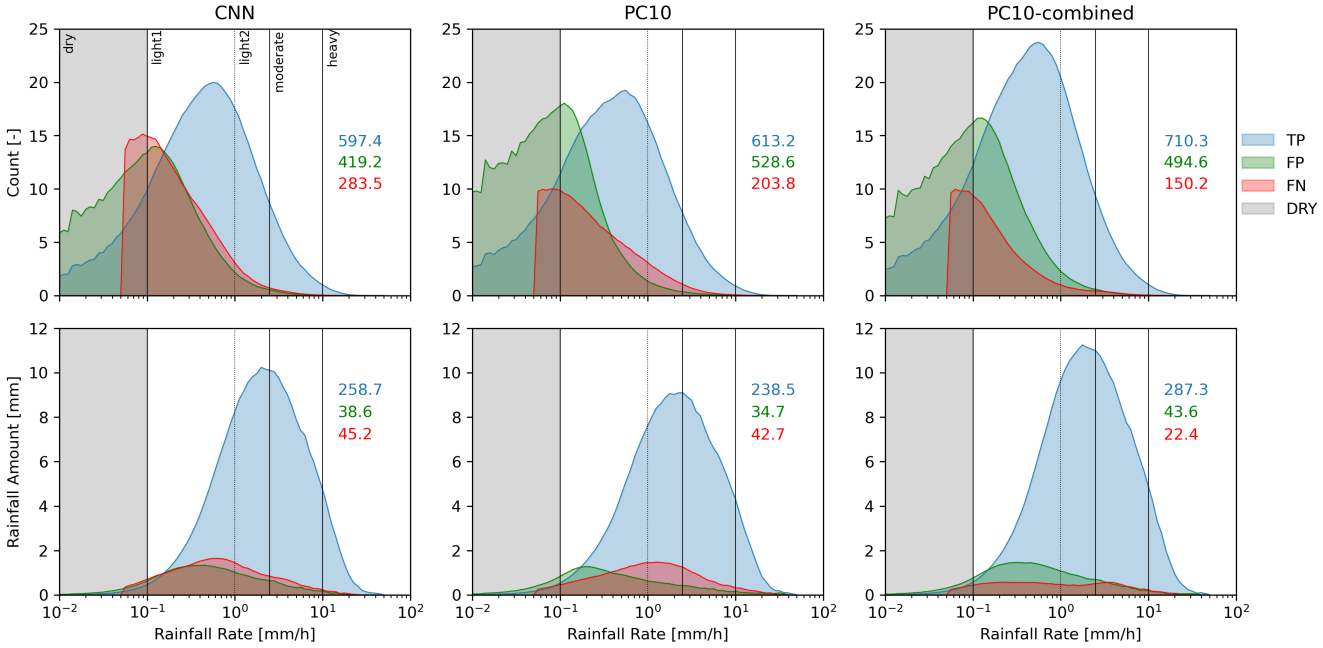

Figure 8. In the top panels, the number of FP, TP, and FN time steps with a specific reference rain rate are compared for *CNN* (left), *PC10* (middle), and *PC10-combined* (right). The respective amount of rainfall is shown in the bottom row. All quantities are an average per CML. The rain intensities for FP and TP are estimated by the CML, while the rain intensities for FN, where the CML rainfall rate is zero, are taken from the reference. The numbers in each panel are color-coded according to TP, FP, and FN and show the integrated amount for each curve.

5 Discussion

5.1 Suitability of PC and PC-Ph products as wet-dry indicators for CML data

The performance of *PC* and *PC-Ph* as wet-dry indicators for CML data was analyzed in Sect. 4.1 and compared to TSB methods in Sect. 4.2. The results showed that the classification scores were only slightly lower than for the TSB methods. In general, the probability threshold had a larger influence on the rain event detection than the differences between *PC* and *PC-Ph*. Both products showed better performance for smaller probability thresholds. The linear correlation (PCC) of derived CML rainfall rates with the reference was highest, when the threshold was lowest, which is due to the lower impact of FPs on PCC compared to FNs. This is surprising because a more liberal classification than using PC01 was not possible. Usually, extremely liberal or conservative methods are assumed to perform badly, which was the case for conservative thresholds *P20* and higher.

In summary, *PC* and *PC-Ph* products perform surprisingly well as wet-dry indicators for CML data. Since these ADB methods are independent ~~from the CML time series is responsible for this improvement~~ of the CML data, they can significantly improve the rain event detection step for noisy CMLs where erratic fluctuations hamper TSB methods. One limitation of the approach

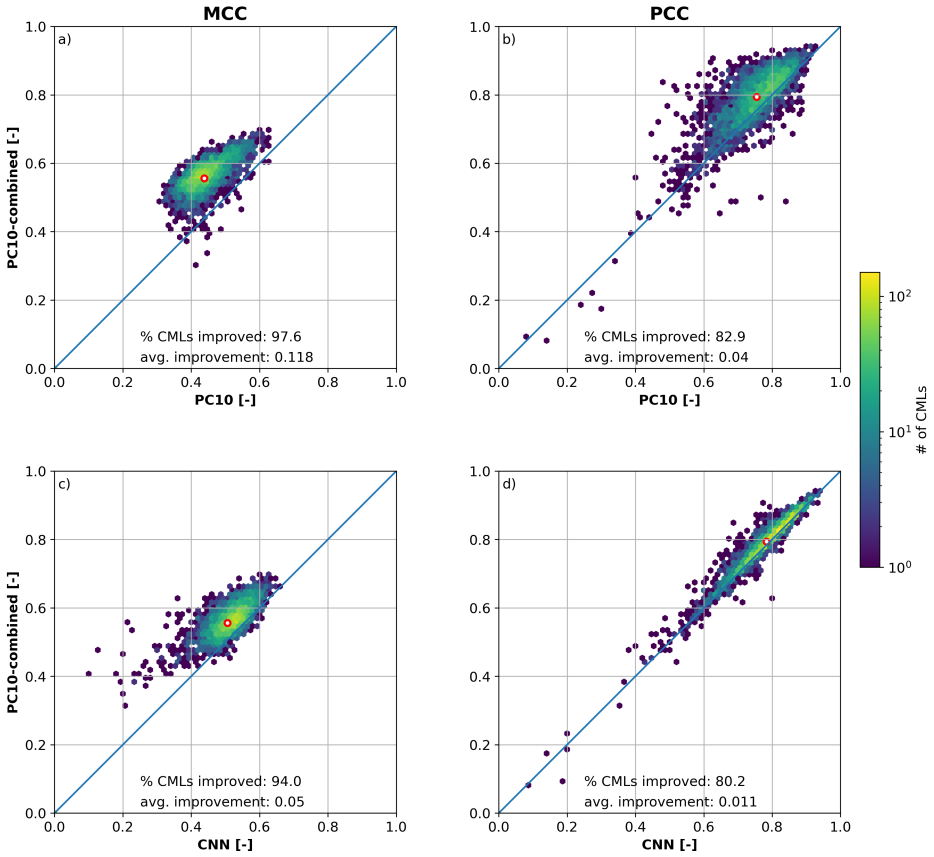

Figure 9. Scatter density comparison of the combination of ADB and TSB methods (PC10-combined) with pure ADB (*PC10*, top row) and TSB (*CNN*, bottom row) methods using the Matthews correlation coefficient (MCC, left column) and the Pearson correlation coefficient (PCC, right column). Each score is computed individually for each CML. The red point is the mean of the respective metrics of all CMLs per product.

based on SEVIRI data is that the temporal resolution can be more than one magnitude lower than the resolution of the CML data (e.g. 15 min. vs. 1 min.). However, the recently launched Meteosat Third Generation (MTG) satellites with the Flexible Combined Imager (FCI) will improve the temporal resolution from 15 minutes of MSG SEVIRI to 10 minutes.

5.2 Performance of a combination of TSB and ADB rain event detection methods compared to TSB-only and ADB-only methods

The combined method was able to outperform both, pure ADB, and TSB methods in detecting rain events. It was able to make a good trade-off between conservative and liberal methods used in the combination, which was shown by the superior MCC scores. Fig. 6 illustrated how such an improvement can be achieved, for example by reducing FNs that lead to a reset of the baseline at a higher level that ultimately leads to too small rainfall rates for *PC10*.

The choice of a starting method for the combination only had a small impact on the results, because most, but not all, initial predictions are overwritten in the combination process. The best MCC and lowest bias were achieved by *PC10-combined* which is why it was chosen for a more detailed comparison of scores for individual CMLs.

PC10-combined improved the MCC and PCC scores for the vast majority of all CMLs and only single outliers achieved better scores using the pure ADB or TSB method. Thus, we conclude that the combination algorithm is robust against varying CML behavior. A limitation of the proposed combination is that **CNN** relies on one-minute instantaneously sampled CML data. However, a similar combination of ADB with other TSB methods that can handle for example the common 15-minute "min-max" sampling should be easily applicable following the logic presented in 3.

5.3 Performance differences between day and night

We observed a notable classification performance drop during nighttime for *PC* and PC-Ph, but also for *RS*. For *PC* and *PC-Ph* this day and night difference is very likely due to the missing SEVIRI channels in the VIS range at night. The calculation of microphysical variables is mainly based on these channels. Their absence consequently limits the reliability of all derived precipitation products. According to Fig. 4), this effect significantly increases the rate of positive predictions.

The ~~combination *PC10all* showed clear improvements for ACC and MCC compared to *CNN* for daytime and nighttime~~*CNN* was able to perform equally well during nighttime, which leads to the assumption that signal fluctuations, caused for example by dew on the antenna, that disturb *RS* can be distinguished by a more sophisticated pattern recognition algorithm like *CNN*. The combined methods utilized the high confidence of the *CNN* and reduced the day and night difference.

However, the PCC results for *PC* and *PC-Ph* showed that a decrease in MCC during the night did not lead to a worse correlation of derived rainfall rates.

5.4 Performance for different rainfall intensity classes

Our results confirmed that the detection performance is much higher for moderate and heavy compared to light rainfall. This was already shown for TSB methods by Polz et al. (2020) and is now also confirmed for the ADB methods based on *PC* and *PC-Ph*. The TSB methods have to distinguish between rainfall signal and noise which can become similar for low rainfall intensities. The ADB method using SEVIRI data suffers from its indirect measurement principle and may have difficulties in distinguishing between precipitation and non-precipitating clouds, particularly for light rain.~~Almost all CMLs were improved by the combination compared to individual methods for rain event detection and especially the worse performing ones~~

Although *PC10* was more liberal, i.e. generally favoring positive predictions, the *CNN* was more confident in the prediction of heavy rainfall. The combined method increased the performance on low rain rate TPs compared to the TSB method, because of the higher confidence of the ADB method with only a small increase in the count and rainfall amount through FPs. Therefore, we confidently claim that *PC10-combined* was able to improve the detection performance for the dry and all positive intensity classes.

~~Scatter density of *PC10* (top) and *PC10all* (bottom) versus *CNN* for ACC (left), MCC (middle) and PCC (right) for a) day and b) night.~~

6 Conclusions

In this study, we aimed to address the questions of whether satellite-derived precipitation products are suitable indicators for ~~the~~ rain event detection in CML data with respect to day/night differences~~and rain intensities~~, different rain intensities, and whether combinations of TSB and ADB methods show an additional added value. We achieved this by using ~~PC and PC-Ph~~ *PC* and *PC-Ph* from MSG SEVIRI in an ADB rain event detection and compared the results to two TSB rain event detection methods. Then, we combined the most promising variants in such a way that the most confident detection method was used for any given time step.

The results clearly show that ~~PC and PC-Ph~~ *PC* and *PC-Ph* products from MSG SEVIRI can be used for the detection of rain events in CML attenuation time series. They performed ~~even better than RS, one of the most commonly used rain event detection methods and similar to the sophisticated CNN method~~almost as well as the TSB methods during daytime and worse than *CNN* during nighttime. Minor differences between the SEVIRI products ~~PC and PC-Ph~~ *PC* and *PC-Ph* exist, but the ~~applied~~ chosen threshold of precipitation probability dominated the overall ~~behaviour~~behavior. An improved ~~PC-Ph product has been~~ *PC-Ph* product is available since April 2022 potentially making its application in rain event detection even more attractive (personal communication with NWC SAF).

However, the performance of ADB methods based on SEVIRI ~~were worse~~ was lower at night than during the day due to the lack of the three SEVIRI channels (VIS and NIR). Since ~~CNN~~ *CNN* did not show a decrease ~~of~~ in quality at night, it would be logical to vary the rain event detection for day and night time. We did not aim for such a temporal variation ~~because the differences between CNN and ADB methods at night were negligible, to keep computational effort low and to~~ to avoid inconsistencies in the resulting time series. Compared to the ADB and ~~CNN, the RS method can only partly be recommended. Especially at nighttime, when dew formation on radomes is probably responsible for false positive rain event detection and therefore a decrease of quality.~~

CNN, the *RS* method does not show particular advantages, except for the easy application.

The quality of the rain event detection methods clearly depended on the rain intensity, with a better performance for moderate and heavy rain than for light rain. For flood forecasting light rain is often negligible but this is not the case for water balance analyses. Additionally, the increasing threat of droughts in ~~context to~~ the context of climate change also requires a high-quality representation of light rain. Low ~~precipitation intensities show the largest~~ rainfall intensities show a large potential for improvement and ~~the main~~ major differences in this study were also obtained there.

The effort to use ADB is larger than for TSB methods because the processing of ~~additional~~ additionally needed satellite data ~~sets can be time consuming. With stepwise combinations~~ can be time-consuming. However, the global availability of the data allows for the unified processing of CML datasets from different countries and we assume that extensive re-calibration is not needed. Stepwise combinations with TSB methods that need to be adjusted to the characteristics of the CML data, as presented here, ~~the effort increases accordingly~~do need re-calibration and increase the effort. The shown additional improvements by combinations are promising and justify the effort. ~~We expect added value of combinations also for other regions or different kinds of CML data sets. However, thresholds and the order of correction steps might vary, there. The~~ There is a multitude

of possibilities when combining different rain event detection methods~~is particularly interesting. Varying methods or~~ . Using methods like the nearby link approach or using different thresholds depending on data quality or rainfall intensity is also possible ~~.~~

in future applications.

In principle, combinations ~~from two~~ of multiple TSB methods are possible and may lead to improvements. However, we recommend applying a combination of TSB and ADB methods to exploit the advantages of both approaches: TSB methods are easy to apply and provide precise results where the separation of noise and ~~precipitation~~ rainfall signals is obvious. Whereas, ADB methods show a better performance for noisy or unstable CML time series due to their independence from the actual CML signal. Burkina Faso for example has only a few rainfall stations, but several hundred CMLs. Most CMLs outside the capital Ouagadougou are long (>20km) and use frequencies around 7 GHz. Their time series are quite noisy and show large fluctuations from other sources than rain. With abundant information from geostationary satellites and the methods presented in this work, we expect to extract useful ground-based rainfall information from such CMLs in this region with scarce rainfall data.

Data availability. CML data were provided by Ericsson Germany and are not publicly available. RADKLIM-YW was provided by the German Weather Service (DWD) and are publicly available.: https://opendata.dwd.de/climate_environment/CDC/grids_germany/5_minutes/radolan (last access: 28 July 2023; DWD CDC, 2023)

The PC and PC-Ph products were provided by Llorenç Lliso and José Alberto Lahuerta who are affiliated with NWC SAF. Recent data is shown at https://www.nwcsaf.org/web/guest/nwc/geo-geostationary-near-real-time-v2021 (last access: 20 December 2023) and long-term records must be requested individually from NWC SAF.

Appendix A: Additional figures

Appendix B: Additional tables

Author contributions. AW, CC, MG, and JP designed the study layout, and AW carried it out with the contribution of CC, MG, and JP. Data and related support was provided by LL and JAL. AW prepared the article with contributions from all co-authors. MG and JP prepared the revised article with contributions from CC.

Competing interests. The authors declare that they have no conflict of interest.

Algorithm 1
Combination of TSB and ADB methods. Example of CNN-PC-combined.

```
t ← CNN82
if PC01 is dry then
    t ← dry                          ▷ high confidence in dry predictions
if CNN94 is wet then
    t ← wet                          ▷ high confidence in wet predictions
if PC30 is wet then
    t ← wet                          ▷ high confidence in wet predictions
if CNN10 is dry then
    t ← dry                          ▷ high confidence in dry predictions
```

Figure A1. Algorithm describing the combination *CNN-PC-combined* of *PC* and *CNN* starting with *CNN82*. A given time-step of CML data that needs to be processed is denoted by *t*. The different wet-dry methods give a wet or dry prediction for this time-step as described in Sect. 3.1.1.

Acknowledgements. We thank Ericsson for their support and cooperation in the acquisition of the CML data. This research has been supported by the Federal Ministry of Education and Research (Grants 01LZ1904A-C and 13N16432), the Helmholtz Association (Grant ZT-0025) and the German Research Foundation (Grant CH-1785/1-2).

Table B1. Confusion matrix values TP, FP, TN, and FN as well as classification scores ACC and MCC and regression scores PCC and RB for all considered methods.

Group	Method	TP	FP	TN	FN	ACC	MCC	PCC	RB
Time series based (TSB)									
	RS	2086257	2761086	36939753	1487374	0.902	0.449	0.657	-0.165
	CNN	2511242	3125653	36575186	1062389	0.903	0.51	0.742	-0.115
Aux. data based (ADB)									
	PC01	3103560	7412469	32288370	470071	0.818	0.438	0.73	-0.07
	PC10	2809793	5525566	34175273	763838	0.855	0.452	0.707	-0.208
	PC20	2100523	3258161	36442678	1473108	0.891	0.423	0.636	-0.455
	PC30	1136635	1205697	38495142	2436996	0.916	0.35	0.509	-0.719
	PC-Ph01	2860243	6112184	33588655	**713388**	0.842	0.439	0.731	-0.159
	PC-Ph10	2029703	2703467	36997372	1543928	0.902	0.441	0.689	-0.417
	PC-Ph20	1376271	1408014	38292825	2197360	0.917	0.392	0.637	-0.587
	PC-Ph30	893173	**815028**	**38885811**	2680458	**0.919**	0.324	0.58	-0.711
Combined									
	CNN-PC-combined	2676762	2727541	36973298	896869	0.916	**0.566**	0.746	-0.108
	PC01-combined	**3122738**	4379019	35321820	450893	0.888	0.555	0.742	0.026
	PC10-combined	3010509	3762008	35938831	563122	0.9	**0.566**	0.743	-0.021
	CNN-PC-Ph-combined	2529058	2588598	37112241	1044573	0.916	0.548	**0.747**	-0.134
	PC-Ph01-combined	3003963	3988729	35712110	569668	0.895	0.553	0.743	**-0.017**
	PC-Ph10-combined	2656801	2865316	36835523	916830	0.913	0.554	0.746	-0.114

References

Atlas, D.: Radar in Meteorology - Battan Memorial and 40th Anniversary Radar Meteorology Conference, Amer. Meteor. Soc., 1990.

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, 1322 – 1331, https://doi.org/10.1175/1520-0450(1977)016<1322:PAAIRM>2.0.CO;2, 1977.

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., and Gerlach, N.: Projekt RADOLAN - Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer), Deutscher Wetterdienst, Hydrometeorologie, 2004.

Bruni, G., Reinoso, R., van de Giesen, N. C., Clemens, F. H. L. R., and ten Veldhuis, J. A. E.: On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution, Hydrology and Earth System Sciences, 19, 691–709, https://doi.org/10.5194/hess-19-691-2015, 2015.

Chwala, C. and Kunstmann, H.: Commercial microwave link networks for rainfall observation: Assessment of the current status and future challenges, WIREs Water, 6, e1337, https://doi.org/10.1002/wat2.1337, e1337 WATER-363.R2, 2019.

Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., and Kunstmann, H.: Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of southern Germany, Hydrol. Earth Syst. Sci., 16, 2647–2661, 2012.

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

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, 3859–3878, https://doi.org/10.5194/hess-21-3859-2017, 2017.

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, https://doi.org/10.1175/JHM-D-23-0015.1, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2023.

DWD: German Met. Service Glossary, German Meteorological Service, 'https://www.dwd.de/DE/service/lexikon/Functions/glossar.html?lv2=101812&lv3=101906', (last access: 28 July 2023), 2023.

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, 1827–1831, https://doi.org/10.1109/LGRS.2016.2614326, 2016.

Fu, S., Sonnenborg, T. O., Jensen, K. H., and He, X.: Impact of Precipitation Spatial Resolution on the Hydrological Response of an Integrated Distributed Water Resources Model, Vadose Zone J., 10, 25–36, 2011.

Giannetti, F., Moretti, M., Reggiannini, R., and Vaccaro, A.: The NEFOCAST System for Detection and Estimation of Rainfall Fields by the Opportunistic Use of Broadcast Satellite Signals, IEEE Aerospace and Electronic Systems Magazine, 34, 16–27, https://doi.org/10.1109/MAES.2019.2916292, conference Name: IEEE Aerospace and Electronic Systems Magazine, 2019.

Giro, R. A., Luini, L., Riva, C. G., Pimienta-del Valle, D., and Riera Salis, J. M.: Real-Time Rainfall Estimation Using Satellite Signals: Development and Assessment of a New Procedure, IEEE Transactions on Instrumentation and Measurement, 71, 1–10, https://doi.org/10.1109/TIM.2022.3165840, conference Name: IEEE Transactions on Instrumentation and Measurement, 2022.

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, 2931–2950, https://doi.org/10.5194/hess-24-2931-2020, 2020.

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, 100 883, https://doi.org/10.1016/j.ejrh.2021.100883, 2021.

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), pp. 149–153, https://doi.org/10.1109/SAM.2018.8448679, 2018.

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, 10 740–10 756, https://doi.org/10.1002/2017WR021015, 2017.

Hernanz, A., Lahuerta García, J. A., Calbet, X., and Rípodas, P.: Algorithm Theoretical Basis Document for the Precipitation Product Processors of the NWC/GEO, NWCSAF, https://www.nwcsaf.org/Downloads/GEO/2021/Documents/Scientific_Docs/NWC-CDOP3-GEO-AEMET-SCI-ATBD-Precipitation_v1.0.1.pdf, (last access: 28 July 2023), 2019.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T.: The Global Precipitation Measurement Mission, Bull. Amer. Meteorol. Soc., 95, 701–722, 2014.

ITU-R: Specific attenuation model for rain for use in prediction methods (Recommendation P.838-3), ITU-R, https://www.itu.int/rec/R-REC-P.838-3-200503-I/en, (last access: 28 July 2023), 2005.

Kreklow, J., Tetzlaff, B., Burkhard, B., and Kuhnt, G.: Radar-Based Precipitation Climatology in Germany—Developments, Uncertainties and Potentials, Atmosphere, 11, https://doi.org/10.3390/atmos11020217, 2020.

Kumah, K., Maathuis, B., Hoedjes, J., and Su, Z.: Near real-time estimation of high spatiotemporal resolution rainfall from cloud top properties of the MSG satellite and commercial microwave link rainfall intensities, Atmospheric Research, 279, 106 357, https://doi.org/10.1016/j.atmosres.2022.106357, 2022.

Kumah, K. K., Hoedjes, J. C. B., David, N., Maathuis, B. H. P., Gao, H. O., and Su, B. Z.: The MSG Technique: Improving Commercial Microwave Link Rainfall Intensity by Using Rain Area Detection from Meteosat Second Generation, Remote Sensing, 13, https://doi.org/10.3390/rs13163274, 2021.

Lahuerta García, J. A.: Scientific and Validation Report for the Precipitation Product Processors of the NWC/GEO, Agencia Estatal de Meteorología, https://www.nwcsaf.org/AemetWebContents/ScientificDocumentation/Documentation/GEO/v2016/NWC-CDOP2-GEO-AEMET-SCI-VR-Precipitation_v1.0.pdf, (last access: 28 July 2023), 2019.

Lahuerta García, J. A.: Algorithm theoretical basis document for the precipitation product processors of the NWC/-GEO, Agencia Estatal de Meteorología, https://www.nwcsaf.org/Downloads/GEO/2021/Documents/Scientific_Docs/NWC-CDOP3-GEO-AEMET-SCI-ATBD-Precipitation_v1.0.1.pdf, (last access: 28 July 2023), 2021.

Leijnse, H., Uijlenhoet, R., and Stricker, J. N. M.: Rainfall measurement using radio links from cellular communication networks, Water Resources Research, 43, https://doi.org/10.1029/2006WR005631, 2007.

Leijnse, H., Uijlenhoet, R., and Stricker, J.: Microwave link rainfall estimation: Effects of link length and frequency, temporal sampling, power resolution, and wet antenna attenuation, Advances in Water Resources, 31, 1481–1493, https://doi.org/10.1016/j.advwatres.2008.03.004, hydrologic Remote Sensing, 2008.

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, 3549–3563, https://doi.org/10.5194/amt-7-3549-2014, 2014.

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, 1101 – 1117, https://doi.org/10.1175/JHM-D-15-0190.1, 2016.

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

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

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, Water Resources Research, 47, https://doi.org/10.1029/2010WR010350, 2011.

Overeem, A., Leijnse, H., and Uijlenhoet, R.: Country-wide rainfall maps from cellular communication networks, Proceedings of the National Academy of Sciences, 110, 2741–2745, https://doi.org/10.1073/pnas.1217961110, 2013.

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, 8039–8065, https://doi.org/10.1002/2016WR019412, 2016.

Pollock, M. D., O'Donnell, G., Quinn, P., Dutton, M., Black, A., Wilkinson, M. E., Colli, M., Stagnaro, M., Lanza, L. G., Lewis, E., Kilsby, C. G., and O'Connell, P. E.: Quantifying and Mitigating Wind-Induced Undercatch in Rainfall Measurements, Water Resources Research, 54, 3863–3875, https://doi.org/10.1029/2017WR022421, 2018.

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, 3835–3853, https://doi.org/10.5194/amt-13-3835-2020, 2020.

Rafieeinasab, A., Norouzi, A., Kim, S., Habibi, H., Nazari, B., Seo, D.-J., Lee, H., Cosgrove, B., and Cui, Z.: Toward high-resolution flash flood prediction in large urban areas – Analysis of sensitivity to spatiotemporal resolution of rainfall input and hydrologic modeling, Journal of Hydrology, 531, 370–388, https://doi.org/10.1016/j.jhydrol.2015.08.045, hydrologic Applications of Weather Radar, 2015.

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, 3571–3584, https://doi.org/10.5194/hess-19-3571-2015, 2015.

Roebeling, R. A. and Holleman, I.: SEVIRI rainfall retrieval and validation using weather radar observations, Journal of Geophysical Research: Atmospheres, 114, https://doi.org/10.1029/2009JD012102, 2009.

Roebeling, R. A., Deneke, H. M., and Feijt, A. J.: Validation of Cloud Liquid Water Path Retrievals from SEVIRI Using One Year of CloudNET Observations, Journal of Applied Meteorology and Climatology, 47, 206 – 222, https://doi.org/https://doi.org/10.1175/2007JAMC1661.1, 2008.

Roversi, G., Alberoni, P. P., Fornasiero, A., and Porcù, F.: Commercial microwave links as a tool for operational rainfall monitoring in Northern Italy, Atmospheric Measurement Techniques, 13, 5779–5797, https://doi.org/10.5194/amt-13-5779-2020, 2020.

Schleiss, M. and Berne, A.: Identification of Dry and Rainy Periods Using Telecommunication Microwave Links, IEEE Geoscience and Remote Sensing Letters, 7, 611–615, https://doi.org/10.1109/LGRS.2010.2043052, 2010.

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, 3157–3188, https://doi.org/10.5194/hess-24-3157-2020, 2020.

Schmid, J.: The SEVIRI instrument, in: Proceedings of the 2000 EUMETSAT meteorological satellite data user's conference, Bologna, Italy, vol. 29, pp. 13–32, 2000.

Sevruk, B.: Rainfall Measurement: Gauges, chap. 35, John Wiley & Sons, Ltd, https://doi.org/10.1002/0470848944.hsa038, 2006.

Song, K., Liu, X., Zou, M., Zhou, D., Wu, H., and Ji, F.: Experimental Study of Detecting Rainfall Using Microwave Links: Classification of Wet and Dry Periods, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5264–5271, https://doi.org/10.1109/JSTARS.2020.3021555, 2020.

Steiner, M., Smith, J. A., and Uijlenhoet, R.: A Microphysical Interpretation of Radar Reflectivity-Rain Rate Relationships, J. Atmos. Sci., 61, 1114–1131, 2004.

Thoss, A.: Algorithm Theoretical Basis Document for Precipitating Clouds of the NWC/PPS, NWCSAF, https://www.nwcsaf.org/AemetWebContents/ScientificDocumentation/Documentation/PPS/v2014/NWC-CDOP2-PPS-SMHI-SCI-ATBD-4_v1_0.pdf, (last access: 28 July 2023), 2014.

Uijlenhoet, R., Steiner, M., and Smith, J. A.: Variability of Raindrop Size Distributions in a Squall Line and Implications for Radar Rainfall Estimation, J. Hydrometeor., 4, 43–61, 2003.

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

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

van de Beek, R. C. Z. ., Olsson, J., and Andersson, J.: Optimal grid resolution for precipitation maps from commercial microwave link networks, Advances in Science and Research, 17, 79–85, https://doi.org/10.5194/asr-17-79-2020, 2020.

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

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, 107–129, https://doi.org/10.1007/s10712-009-9079-x, 2010.

Wagner, A.: Spatiotemporal Variability of Precipitation : Measurements - Simulations - Limitations, doctoralthesis, Universität Augsburg, https://opus.bibliothek.uni-augsburg.de/opus4/files/38014/Wagner_Diss.pdf, (last access: 28 July 2023), 2018.

Wagner, A., Seltmann, J., and Kunstmann, H.: Joint statistical correction of clutters, spokes and beam height for a radar derived precipitation climatology in southern Germany, Hydrol. Earth Syst. Sci., 16, 4101–4117, 2012.

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, 1847–1859, https://doi.org/10.5194/amt-5-1847-2012, 2012.

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-AISH publication, pp. 323–329, 2012.

Winterrath, T., Brendel, C., Hafer, M., Junghänel, T., Klameth, A., Lengfeld, K., Walawender, E., Weigl, E., and Becker, A.: RADKLIM Version 2017.002: Reprocessed quasi gauge-adjusted radar data, 5-minute precipitation sums (YW), https://doi.org/10.5676/DWD/RADKLIM_YW_V2017.002, 2018.