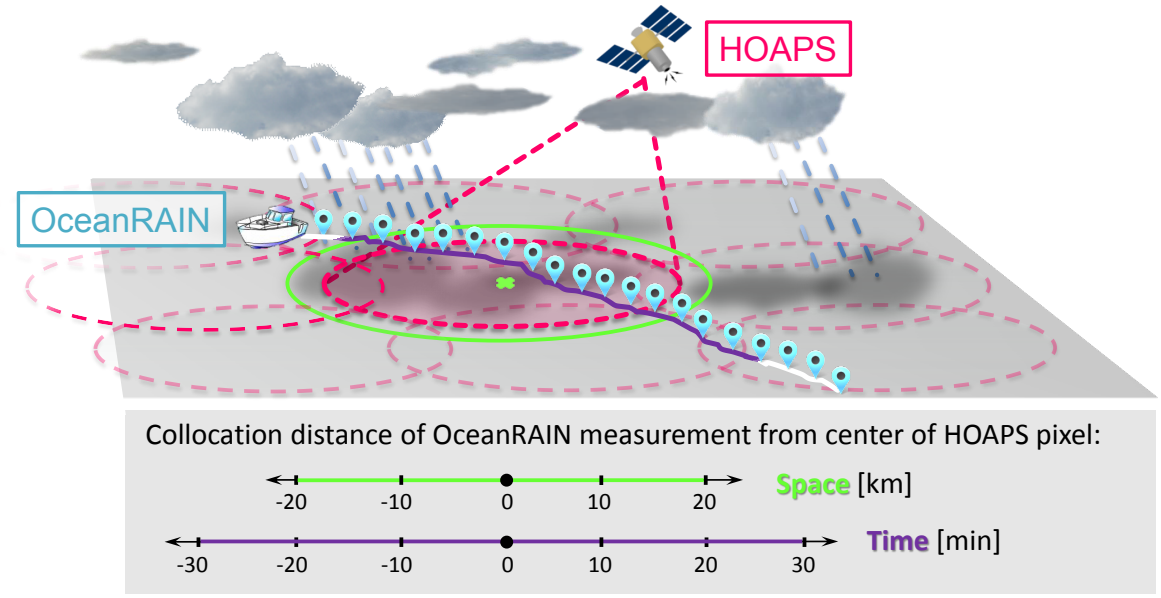
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Point-to-area validation of passive microwave satellite precipitation with shipboard disdrometers



Jörg Burdanowitz

Hamburg 2017

Berichte zur Erdsystemforschung
Reports on Earth System Science

188
2017

Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.

Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

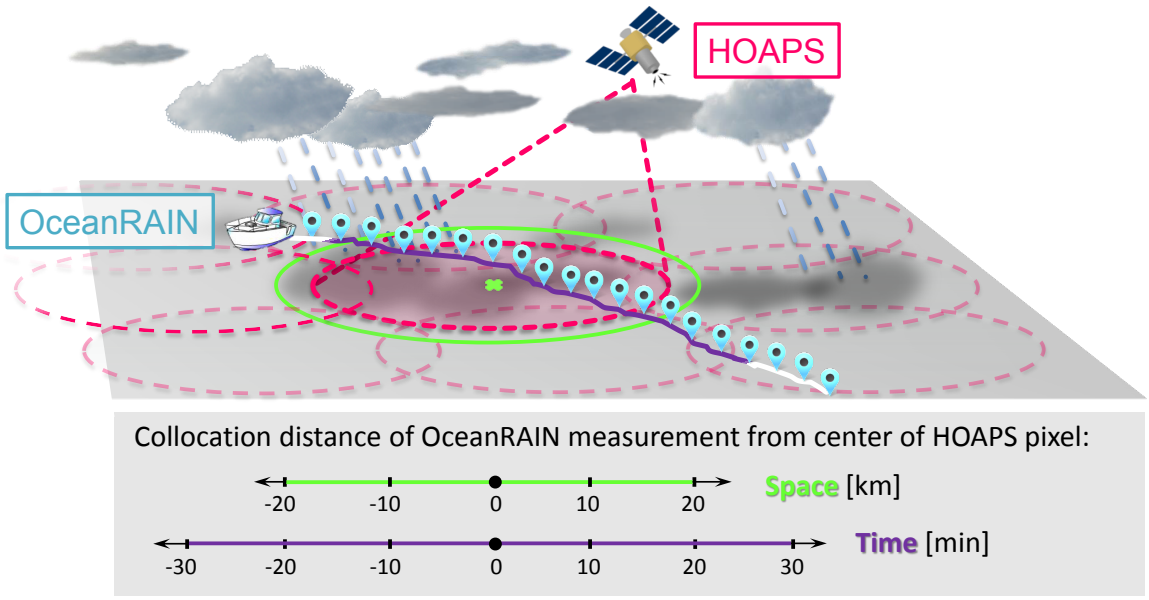
name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Point-to-area validation of passive microwave satellite precipitation with shipboard disdrometers



Dissertation with the aim of achieving a doctoral degree
at the Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg

submitted by

Jörg Burdanowitz

Hamburg 2017

Jörg Burdanowitz

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 19.12.2016

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Stefan Bühler
Dr. Stephan Bakan

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 188
Reports on Earth System Science / Max Planck Institute for Meteorology 2017

ISSN 1614-1199

Abstract

The high spatial and temporal variability of precipitation poses a substantial challenge to its global monitoring, particularly over the global oceans. The insufficient coverage of surface-based measurements adds to this problem. These data-sparse areas profit from passive microwave (PMW) satellite sensors, which essentially contribute to the estimation of global long-term precipitation estimates. However, the uncertainty of PMW satellite products demands a validation that requires high-quality surface-based precipitation reference data over the ocean. The Ocean Rainfall And Ice-phase precipitation measurement Network (OceanRAIN) collects high-resolution point-like precipitation data from long-term equipped RVs for the validation of low-resolution PMW satellite precipitation products. Deployed on research vessels (RVs), these optical disdrometers designed for ocean usage are suitable instruments to accurately record oceanic precipitation. This thesis addresses the emerging point-to-area problem and data-imposed validation constraints on the example of collocated PMW-based Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data (HOAPS) and OceanRAIN data.

The thesis presents three essential aspects for representative satellite validation: the precipitation phase (PP) distinction (rain, snow and mixed-phase), the point-to-area precipitation representation and the instrument/algorithm sensitivity. For the PP distinction, a newly developed algorithm is presented that predicts the PP using the air temperature, relative humidity and the 99th percentile of the particle diameter provided by the particle size distribution (PSD) of the disdrometer. Incorporating the PSD mainly adds to the high accuracy of 91 % for a rain–snow distinction. However, including the important mixed-phase precipitation yields 81 % in accuracy for the PP separation.

The point-to-area problem is addressed by applying two statistical adjustments derived from subtropical radar data to ensure that the along-track data represents the areal precipitation distribution in a PMW satellite pixel. The first statistical adjustment uses the average precipitation event duration to correct underrepresented precipitation rates by sampled short events (1–3 min) and overrepresented precipitation rates of long events (5–20 min). The second adjustment takes the median-normalized precipitation rate to mainly correct cases of strong precipitation overrepresentation on the ship track with respect to the area. Both adjustments reduce the sum of squared errors for the track–area precipitation difference by 86 %.

The HOAPS–OceanRAIN collocation reveals about 14 % misses from which more than 90 % relate to the low HOAPS algorithm sensitivity with a cut-off at 0.3 mm h^{-1}. HOAPS produces 3 % of false detections, which are merely caused by scenes when the RV of OceanRAIN spatially misses the precipitation falling within the HOAPS pixel. An exemplary case documents precipitation missed by an RV in OceanRAIN but detected by other PMW- and infrared-based satellite products. In the inner tropics, these HOAPS false detections reach 7 % where vigorous convection accounts for most of the precipitation accumulation. Treating false detections as hit cases reveals a good overall agreement in the average precipitation rate that HOAPS underestimates by less than 15 % except for the subtropics (30 %); however, the exact percentage of misclassified false detections remains unknown. Comparing average precipitation rates excluding zero precipitation, HOAPS underestimates the average precipitation by 13 % while the uncertainty obtained from resampling lies in the same range. The relatively high uncertainty points at the, so far, insufficient sample size of the OceanRAIN database to derive individual HOAPS uncertainty estimates. However, this thesis highlights the importance of the point-to-area problem for satellite validation and demonstrates a statistical adjustment for the shipboard precipitation reference data to spatially represent the satellite data.

Zusammenfassung

Die hohe zeitliche und räumliche Variabilität von Niederschlag stellt eine Herausforderung in der globalen Niederschlagserfassung dar. Besonders über den Ozeanen wird dieses Problem zusätzlich durch die geringe Dichte bodengebundener Niederschlagsmessungen vergrößert. Passive Satellitensensoren im Mikrowellenspektrum (PMW) tragen erheblich dazu bei, den globalen Niederschlag flächendeckend abzuleiten. Allerdings benötigt deren Validierung dennoch verlässliche, bodengebundene Referenzmessungen des Niederschlags. Das langfristig orientierte, schiffsbasierte Messnetz zur Erfassung von Regen und Eisphasenniederschlag (OceanRAIN) sammelt punktuelle ozeanische Niederschlagsdaten in hoher Qualität. OceanRAIN nutzt speziell für den Ozeaneinsatz konzipierte optische Disdrometer und liefert damit eine einzigartige Quelle zur Validierung von Niederschlag aus grob aufgelösten PMW-Sensor Satellitendaten. Bei der Validierung von PMW-Sensordaten werden in dieser Arbeit daher insbesondere das Punkt-zu-Fläche Problem sowie Datensatz-spezifische Limitierungen thematisiert. Dies geschieht am Beispiel des PMW-basierten "Hamburg Ozean- und Atmosphärenflüsse aus Satellitendaten“-Datensatzes (HOAPS) mithilfe von OceanRAIN.

Diese Arbeit beleuchtet drei wesentliche Aspekte, die für eine gelungene Satellitensensor-Validierung von Bedeutung sind: die Unterscheidung der Niederschlagsphase (PP), die unterschiedliche Repräsentation von Niederschlag in Punkt- und Flächenmessungen, und die Sensitivität des PMW-Sensors bzw. des Algorithmus schwache Niederschläge zu detektieren. Für die Erkennung der PP wurde ein neuer automatischer Algorithmus entwickelt, der die Parameter Lufttemperatur, relative Luftfeuchte sowie das 99. Perzentil des Niederschlagspartikeldurchmessers aus dem Partikelspektrum (PSD) des Disdrometers nutzt. Die Nutzung des PSDs trägt in hohem Maße zu einer Genauigkeit von 91 % in der Regen/Schnee Unterscheidung bei. Wird die Mischphase als wichtige, zusätzliche PP ermittelt, erreicht die Genauigkeit 81 %.

Der Verringerung des Punkt-zu-Fläche Problems zwischen HOAPS und OceanRAIN dienen zwei statistische Anpassungen des Referenzdatensatzes OceanRAIN, damit dieser besser die Fläche eines Satellitenpixels repräsentiert. Beide Anpassungen entstammen einer Untersuchung mit subtropischen Radardaten, die das Punkt-zu-Fläche Problem zwischen Schiffsspur und Satellitenpixel simulieren. Die erste daraus abgeleitete Anpassung nutzt die mittlere Niederschlagsdauer entlang der Schiffsspur, wobei die Niederschlagsrate für eine mittlere Dauer von 1–3 min unterschätzt wird, die Niederschlagsrate für eine Dauer von 5–20 min jedoch überschätzt wird. Die zweite statistische

Anpassung nutzt die mittlere Niederschlagsrate entlang der Schiffsspur normiert mit deren Medianwert und korrigiert insbesondere stark überrepräsentierten Niederschlag auf der Schiffsspur. Beide Anpassungen reduzieren die Summe des quadratischen Fehlers zwischen Spur- und Flächenniederschlagsrate um 86 %.
Nach der Kollokation von HOAPS mit OceanRAIN ergeben sich 14 % Fälle, in denen HOAPS den von OceanRAIN detektierten Niederschlag nicht erkennt. Die Ursache liegt in der niedrigen Sensitivität des HOAPS Algorithmus begründet, der keine Niederschlagsraten unter 0.3 mm h^{-1} auflöst und damit 90 % der nicht detektierten Fälle erklären kann. In 3 % der Fälle registriert HOAPS von OceanRAIN nicht registrierten Niederschlag, was überwiegend dadurch erklärt werden kann, dass das Schiff am Niederschlag innerhalb des Satellitenpixels vorbeifährt. Ein Fall der RV *Meteor* dokumentiert beispielhaft wie OceanRAIN Niederschlag nicht registriert, der von zwei anderen PMW-basierten sowie einem infrarot-basierten Satellitenprodukt jeweils detektiert wird. Besonders in den inneren Tropen treten von HOAPS angeblich falsch detektierte Fälle mit 7 % besonders häufig auf, da hier intensive, kleinräumig-konvektive Schauer besonders stark zum mittleren Niederschlag beitragen. Behandelt man all diese Fälle so als ob auch OceanRAIN Niederschlag detektiert hätte, ergibt sich eine gute Übereinstimmung in der mittleren Niederschlagsrate beider Datensätze, wobei HOAPS die Niederschlagsrate um weniger als 15 % unterschätzt, in den Subtropen jedoch um bis zu 30 %. Allerdings bleibt der wirkliche Anteil von Falschdetektionen unbekannt. Unter Ausschluss von Null-Niederschlagsraten unterschätzt HOAPS die mittlere Niederschlagsrate um 13 %, wobei die statistische Unsicherheit beider Datensätze ermittelt durch Bootstrapping im selben Bereich liegt. Diese relativ hohe Unsicherheit zeigt, dass eine individuelle Fehlerabschätzung für HOAPS mangels ausreichender OceanRAIN-Messungen noch nicht abschließend möglich ist. Dennoch unterstreicht diese Arbeit die Bedeutung des Punkt-zu-Fläche Problems und demonstriert eine Methode, wie der Referenzdatensatz an die räumliche Repräsentativität des zu validierenden Satellitendatensatzes angepasst werden kann.

Contents

Chapter 1

Introduction

Precipitation sets the basis for life on Earth by providing freshwater and shaping the ecosystem. Despite all technological developments, our modern society entirely depends on available freshwater as the most essential resource. While a lack of freshwater threatens life in the form of droughts, excessive precipitation endangers life by causing severe flooding. In everyday life, the heterogeneous distribution of precipitation demands global precipitation forecasts and monitoring that is crucial for many economic sectors such as agriculture, energy production, risk management or tourism, among others. Moreover, precipitation plays a key role in a changing climate as to what extent precipitation patterns shift (Trenberth et al., 2003; IPCC, 2014a).

1.1 Basic characteristics of precipitation

From a more scientific perspective, precipitation represents an essential climate variable (ECV) at the Earth's surface (Bojinski et al., 2014). Precipitation (P) and evaporation (E) form the freshwater flux that links the global water cycle to the Earth's energy budget through latent heat exchange (Trenberth et al., 2009). Over the ocean, the freshwater flux sources denote areas where the atmosphere gains energy from the ocean, whereas in sink regions the ocean gains energy from the atmosphere.

Compared to other ECVs, precipitation holds some special characteristics. First, the particles that constitute precipitation can occur in various forms of either liquid or solid water phase, or as a mixture of both. Second, precipitation varies dramatically in space and time. Third, the frequency distribution of precipitation particles as well as the precipitation rate follow a log-normal distribution instead of a Gaussian distribution. All these special characteristics enormously complicate the prediction and monitoring of precipitation.

The precipitation phase (PP) denotes the aggregation state of the falling precipitation particles. The PP is usually separated into liquid and solid particles, whereas the solid

particles can be further distinguished into precipitation types such as snow, hail or graupel. During PP transitions, liquid and solid particles can co-exist and the PP is referred to as mixed-phase. Mixed-phase precipitation usually occurs when solid particles fall through a relatively warm atmospheric layer that partially melts them. Likewise, a lower atmospheric inversion layer of colder temperatures than aloft can cause partial freezing of liquid particles. These conditions demonstrate that phase changes in the atmosphere are highly variable processes so that the PP possibly changes within minutes.

The high spatial and temporal variability of precipitation reflects the complex processes that interact in the Earth's atmosphere. These processes can be classified into different temporal and spatial scales that might or might not trigger precipitation. The concerned timescales range from few minutes (precipitating shallow Cumulus cloud; Stevens and Seifert, 2008; vanZanten et al., 2011) over several hours (heavily precipitating organized convective systems; Laing and Michael Fritsch, 1997) to multi-annual phenomena such as the El Niño Southern Oscillation (ENSO; Torrence and Webster, 1999). The spatial scales vary dramatically between the two classes of convective and stratiform precipitation (Stevens, 2005). The formation of convective precipitation requires strong vertical displacement of buoyantly unstable or forcedly lifted air parcels, whereas stratiform precipitation tends to form under generally weaker vertical air motion or as a remainder of older convective cells (Houze, 1997). Thus, the occurrence and intensity of convective precipitation can vary much stronger on spatial and temporal scales compared to stratiform precipitation.

The frequency distribution of precipitation particle sizes follows a non-Gaussian distribution. The tremendously higher number of small precipitation particles also puts its mark on the precipitation-rate distribution. Accordingly, light precipitation arise much more frequently compared to heavy precipitation. Despite the fact that low precipitation rates contribute most to the precipitation occurrence, high precipitation rates contribute strongest to the total precipitation amount. This non-Gaussian particle distribution highlights the need to resolve both extremes – the smallest and the largest particles. Whereas the resolution of small particles demands a high measurement sensitivity, the resolution of larger particles calls for a sufficiently long sampling period. Only considering both aspects can reflect the natural precipitation occurrence as well as precipitation amount.

1.2 The measurement of precipitation

As a consequence of the precipitation characteristics, the *ideal* instrument measures precipitation directly, at high sampling frequency, at high spatial resolution and covers a large area over a long period of time under uniform conditions. However, not a single instrument can sufficiently meet all these requirements at once. As a consequence,

very different instruments have been developed in order to measure precipitation while considering most of its special characteristics (Michaelides et al., 2009). Combining these different instruments offers extensive opportunities for global precipitation monitoring but poses new challenges at the same time. With the advent of the satellite-era, instruments to measure precipitation separate into surface-based and space-based instruments (Tapiador et al., 2012).

1.2.1 Surface-based instruments

Surface-based instruments that record the accumulated amount of precipitation are commonly called gauges. Various types of gauges exist as the simplest instrument to measure precipitation (Strangeways, 2004). All these gauges have in common that they are considered *direct* precipitation measurements because a bucket actually collects falling precipitation particles. Therefore, gauges often serve as precipitation reference at a certain location. When maintained regularly and properly arranged, gauge networks can help to validate precipitation satellite measurements (Sorooshian and Kuo-Lin, 2000; Yatagai et al., 2009; Sunilkumar et al., 2016). However, substantial sources of error limit the application area of gauges. In particular, wind-induced undercatch and solid precipitation let gauges strongly underestimate precipitation (Goodison, 1978; Yuter and Parker, 2001; Michelson, 2004). To a lesser extent, evaporative loss of collected precipitation, splashing drops or a partially blocked orifice of the gauge can diminish the precipitation collected. Though gauges may be the simplest direct instrument to measure precipitation, they are prone to large errors associated with wind and frozen precipitation.

Disdrometers represent the second type of surface-based instruments to measure precipitation pointwise. Instead of collecting precipitation amount, disdrometers count or image the particles so that the precipitation rate can be derived from the resulting particle size distribution (PSD). Three main types of disdrometers emerged: the impact disdrometer, the optical disdrometer and the imaging disdrometer. The most widely used impact disdrometer is the Joss-Waldvogel disdrometer (Joss and Waldvogel, 1967). Impact disdrometers sense the PSD by converting the mechanical energy of falling precipitation particles into equivalent electrical pulses. Video disdrometers sense precipitation particles with cameras but their operation is costly. Whereas video disdrometers record the PSD visually, optical disdrometers record the extinction of near-infrared (NIR) light induced by falling precipitation particles that is jointly converted into a PSD (Löffler-Mang and Joss, 2000). With the PSD, disdrometers provide essentially more information compared to ordinary gauges; however, the calibration and maintenance of disdrometers remains crucial for an accurate precipitation measurement.

Surface-based scanning weather radars are indirect precipitation measurements of an atmospheric column. They actively sense the surrounding atmosphere and record the

backscattered radiation in the form of the radar reflectivity Z. This reflectivity needs to be converted into a precipitation rate R by applying a Z-R relationship. The Z-R relationship depends on the PSD so that disdrometers can serve to accurately convert the radar reflectivity into a precipitation rate (e.g. Miriovsky et al., 2004; Lee and Zawadzki, 2005). However, no uniform Z-R relation exists because PSDs can change within minutes, which strongly affects the calculated precipitation rate. Depending on their operating frequency, the radar coverage ranges between a few tens of kilometers and several hundred kilometers in diameter. Radars are indispensable instruments to measure precipitation in a large area but require accurate calibration and validation.

1.2.2 Satellite-based instruments

Satellite sensors constitute the only instrument to measure precipitation on a global scale with short revisit time. A distinction is made between sensors deployed on polar-orbiting and geostationary satellites. Polar-orbiting satellites can sense the whole globe but have a relatively low revisit time, which limits its temporal resolution. In contrast, geostationary satellites hold a high temporal resolution but cover a limited but constant area. In addition, they operate about two orders of magnitude higher in space compared to the polar-orbiting satellites. Passive microwave (PMW) instruments have large antenna sizes and, consequently, cannot operate on geostationary satellites. However, plans exist to deploy microwave sounders on geostationary satellites (Lambrigtsen et al., 2006). The so-far deployed visible (VIS) and infrared (IR) sensors retrieve information mainly from the cloud top. Cloud top information is used to indirectly derive the precipitation rate, which can lead to erroneous detection of precipitation from non-convective high clouds with cold cloud tops (Kidd and Levizzani, 2011).

PMW sensors represent the second commonly used type of satellite instruments to estimate precipitation from space. They provide a more physically complete image of the atmospheric water content compared to VIS/IR satellite sensors (Levizzani et al., 2007). Whereas low-frequency channels serve to directly detect medium-to-large water droplets below the freezing level, high-frequency channels can infer smaller particles and specifically ice particles indirectly from scattering above the freezing level. As a downside, particularly the low-frequency channels have a coarse spatial resolution of several tens of kilometers in diameter. In contrast, active microwave sensors reach much higher spatial resolutions but with a very narrow swath width. Due to the low swath width, these spaceborne radars commonly serve as calibrator for PMW sensors or for case studies.

1.2.3 Combined precipitation products

All state-of-the-art precipitation satellite climatologies are based on PMW sensors that serve better to accurately quantify precipitation from space. However, IR sensors

present a suitable way to compensate for the low spatial and temporal resolution of PMW sensors. Some algorithms use IR sensors indirectly to propagate PMW images forward in time, e.g. the Climate Prediction Center MORPHing technique (CMORPH; Joyce et al., 2004). Some algorithms interpolate between consecutive PMW images using precipitation data derived from IR sensors, e.g. the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA; Huffman et al., 2007), the Global Satellite Mapping of Precipitation (GSMaP; Okamoto et al., 2005) or the Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman, 2015). Other products feed IR sensor data into a neural net algorithm that uses PMW estimates for comparison, e.g. the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN; Hsu et al., 1997). Some of the products such as TMPA or IMERG also include gauge data over land. However, ingesting data from different sources complicates the tracking of errors and uncertainties while making the product less independent for validation purposes. In contrast, single-sensor climatologies such as the Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data (HOAPS; Andersson et al., 2010b) provide independent precipitation estimates and allow calibration not only at the product level.

1.3 Precipitation validation over the ocean

In contrast to the global oceans, a great variety of ground-based instruments can be used to validate precipitation satellite products over land. These products ingest data from networks of gauges, disdrometers and radars, among others. The global oceans cover 71 % of the Earth's surface but are at the same time almost void of systematic long-term precipitation measurements (IPCC, 2014b; Maggioni et al., 2016). The lack of oceanic precipitation data poses a challenge to the validation of satellite retrievals over oceanic areas. As a consequence, coastal or island-based radars usually serve as reference for satellite retrievals over the ocean (e.g. Schumacher and Jr., 2000; Burdanowitz et al., 2015). However, these radar data sets are not entirely free from land influences such as emission of aerosol particles altering the precipitation efficiency of clouds (Twomey, 1977), or surface effects affecting the atmospheric dynamics (Dai, 2001). However, Bumke and Seltmann (2012) find no land–ocean difference in PSD over the relatively small Baltic Sea. These potential influences demand remote-ocean precipitation data to validate satellite retrievals over the global ocean.

Qualitative precipitation data over the ocean exists from weather reports of voluntary observing ships (VOS; Kent et al., 2010). Together with quantitative estimates of underway gauges on research vessels (RVs), merchant and cruise ships (Bumke et al., 2012), and tropical buoy gauge arrays (Hayes et al., 1991; Bourlès et al., 2008), these measurements complement recent ship-based precipitation databases such as the International Comprehensive Ocean-Atmosphere Data Set (ICOADS; Woodruff et al., 2011), the National Oceanography Centre Southampton Flux data set (NOCS; Berry

and Kent, 2011), the Shipboard Automated Meteorological and Oceanographic System (SAMOS; Smith et al., 2009) or OceanSITES (2016). However, these types of gauge data and weather reports contain large biases due to insufficient maintenance or instruments or unsuited instruments for open-ocean usage. These deficiencies disqualify them for most satellite validation applications.

As an alternative, special ship rain gauges (Hasse et al., 1998) as well as optical disdrometers find increasing usage onboard of RVs (Weller et al., 2008). In particular, optical disdrometers are considered a reference in precipitation measurements, provided that they are calibrated properly (Taylor, 2000). However, these instruments suffer from the same rough open-ocean conditions as conventional gauges. These conditions include frequently changing high wind speeds, sea salt, flow distortion by the ship's superstructure as well as freezing in polar regions. The optical disdrometer ODM470 has been specifically designed to measure under these open-ocean conditions (Großklaus, 1996; Clemens, 2002). A wind vane attached to a rotatable axis aligns the measurement volume of the disdrometer perpendicular to the wind direction while its cylindrical measuring volume is independent from the incidence angle of falling particles. These features qualify the ODM470 as the core instrument of the Ocean Rain And Ice-phase precipitation measurement Network (OceanRAIN; Klepp, 2015), the first systematic long-term shipboard precipitation measurement effort over the ocean.

Optical disdrometers of the same type as ODM470 have been effectively used before for validating reanalysis (Bumke, 2016) and satellite products (Bumke et al., 2012). Despite their great value, these studies neglect the different spatial representativeness of point-like ship measurements compared to areal satellite precipitation estimates. This difference, commonly referred to as "point-to-area problem", affects PMW satellite estimates with their large pixel size of about 50 km by 50 km. However, the effect of not adjusting point measurements to areal estimates has not been quantified so far for the PMW satellite validation using ODM470 data where this thesis draws on. These investigations pave the way for a systematic error analysis of HOAPS precipitation satellite data, similar to that of the near-surface specific humidity (Kinzel et al., 2016).

1.4 Research objectives

This thesis is targeted to unveil the essential factors that mostly influence the validation of precipitation satellite data using surface-based shipboard data. Identifying these factors allows to develop a methodology to reduce their influence on the validation of oceanic precipitation from the PMW satellite climatology HOAPS using the newly available OceanRAIN shipboard data set. The example of HOAPS and OceanRAIN should serve as guidance for other satellite validation applications in a larger scope.

Besides an accurate space–time matching, a meaningful validation requires to adapt the reference data set to the satellite data. This adaptation presumes to understand

the precipitation characteristics with respect to the data set specifications. Three main precipitation characteristics tremendously influence the validation procedure that are addressed successively in the thesis. First, the PP is an essential piece of information to correctly estimate the precipitation rate and diagnose errors thereof (Chapter 3). Second, the temporal and spatial variability of precipitation events sets high requirements to the data set's temporal and spatial resolution (Chapter 4). Third, the sensitivity of a data set's instrument and algorithm determines whether low precipitation rates can be resolved (Chapter 5). These low precipitation rates represent an important factor to get the precipitation occurrence correct caused by the non-Gaussian precipitation distribution with most frequently occurring low intensities. Finally, a sufficiently large data set guarantees an accurate representation of the precipitation rate to which rarely occurring high precipitation rates contribute most.

1.5 Outline

Chapter 2 contains a general description of the data sets used while most of the methodology is developed in the results of the thesis. Chapter 3 presents the automatic PP distinction algorithm to determine the PP for the HOAPS–OceanRAIN validation. Chapter 4 discusses the influence of the spatial representativeness of precipitation point measurements with respect to the typical spatial scales of PMW satellite data. From a study with island-based radar data, a statistical adjustment of the OceanRAIN data is developed. In Chapter 5, the statistical adjustment is applied to OceanRAIN. Furthermore, a focues lies on the space–time matching of HOAPS and OceanRAIN as well as the sensitivity to light precipitation. Chapter 6 summarizes the main results of the thesis, followed by concluding remarks and an outlook.

Chapter 2

Instruments and data

2.1 OceanRAIN

Since 2010, the Ocean Rain And Ice-phase precipitation measurement network (OceanRAIN) collects atmospheric data including precipitation rates on several RVs. Current permanent deployments include the German ships RV *Polarstern* (since June 2010), RV *Meteor* (since March 2014), RV *Sonne* (since November 2014). RV *Maria S. Merian* recorded data from October 2010 until June 2014. Focusing on consistent and high-quality precipitation data, this study employs data from all 4 mentioned German RVs. The backbone of OceanRAIN is the optical disdrometer ODM470, which is explained in detail in Sect. 2.1.1. Klepp (2015) describes the OceanRAIN data post-processing and quality-checking in detail. More information on OceanRAIN can be accessed online (OceanRAIN, 2015).

2.1.1 The ODM470

The ODM470 is an optical disdrometer to measure precipitation, manufactured by the German company Eigenbrodt GmbH & Co KG near Hamburg (Germany). The instrument consists of a NIR light-emitting diode (LED) at 880 nm and a photo diode receiver (Lempio et al., 2007). The IR-LED of the ODM470 is only activated once at least 8 particles per minute pass the active sensing area of the precipitation detector IRSS88 (Fig. 2.1, right) in order to increase IR-LED lifetime and exclude measurement artifacts caused by birds or other non-precipitation objects. The IRSS88 switches off the ODM470 after 1 min without any particle passing the IRSS88 active sensing area. The entire ODM470 system was developed in a way to minimize undesired influences by changing wind directions and high wind speed. The ODM470 sensitive optical volume has a cylindrical shape of 120 mm length and 22 mm in diameter. The cylindrical shape guarantees an independence from the incidence angle of the falling hydrometeors, which becomes crucial under high wind speeds and superstructure-induced local turbulence.

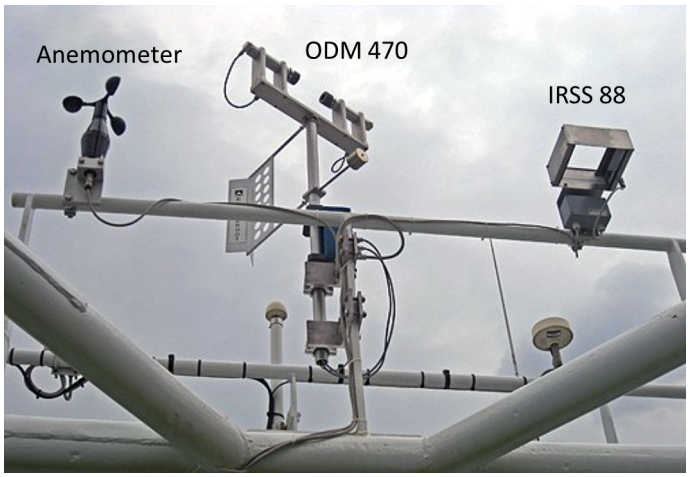


Figure 2.1: Automatic ODM470 measurement system including a cup anemometer (left), the optical disdrometer ODM470 (center), and the precipitation detector IRSS88 (right), all deployed in the highest main mast at about 43 m height onboard RV *Polarstern*.

Mounted on a pivotable axis, a wind vane ensures the optical volume to adjust perpendicular to the instantaneous local wind direction. The ODM470 mounting height typically ranges between 30 and 45 m height, depending on the RV's specifications (Fig. 2.1). This elevated deployment reduces influences on the measured precipitation by splashing wave water.

OceanRAIN ingests time and navigation data in the form of GPS coordinates from the RVs while the optical disdrometer system measures wind speed and the particle size distribution (PSD) of precipitation. In addition, the RVs collect numerous other meteorological information of the atmospheric state. These ancillary data samples aid to predict the precipitation phase (PP) using the automatic PP distinction algorithm introduced in Chapter 3.

During precipitation events, the falling hydrometeors attenuate the emitted NIR radiation, which decreases the voltage signal measured. The duration of the voltage drop determines the particle transit time, that is, the total time it takes a particle to pass through the optical volume of the disdrometer. From the amplitude of the detector voltage drop the cross-sectional area can be deduced, which determines the particle diameter. The measured particle diameters are split into 128 logarithmically distributed size bins, of which the smallest is less than 0.02 mm and the largest corresponds to the optical volume diameter of 22 mm. However, wind- or wave-induced ship vibrations passed on to the instrument might cause artificial signals that are not distinguishable from real precipitation, which is why particles below bin 14 (0.43 mm diameter) are not considered in OceanRAIN. This exclusion of small particles also removes sea spray particles from the PSD spectra. The remaining particles are accumulated and binned over 1 min. From the resulting PSD, the precipitation rate R (mm h^{-1}) or liquid water

equivalent (kg m^{-2} h^{-1}) after Großklaus (1996) can be calculated using

$$\mathrm{R} = 3600 \cdot \sum_{\mathrm{bin}=1}^{128} n(\mathrm{bin}) \cdot v(\mathrm{bin}) \cdot m(\mathrm{bin}), \tag{2.1}$$

where v(bin) in m s^{-1} denotes the particle terminal fall speed and m(bin) in kg the particle mass; both are parameterized. n(bin) in m^{-3} denotes the PSD density per bin class that is calculated following Clemens (2002) by considering the geometrical features, diameter d in m and length l in m, the sampling time t in s of the ODM470 as well as the sum of local wind speed U_{rel} in m s^{-1} relative to the ship movement measured by a cup anemometer, and the empirical terminal fall speed v(bin) in m s^{-1} as

$$n(\mathrm{bin}) = \frac{N(\mathrm{bin})}{l \cdot d \cdot t \cdot \sqrt{U_{\mathrm{rel}}^2 + \left[v(\mathrm{bin})\right]^2}}. \tag{2.2}$$

N(bin) is the number of measured particles per bin class, denoted as PSD. As explained, in Eq. (2.1) empirical relationships utilize the particle diameter D that strongly depends on the type of precipitation. Henceforth, we refer to precipitation phase (PP), which means either liquid precipitation (rain), solid precipitation (e.g., snow or graupel), or mixed-phase precipitation. For rain the drop mass m_{l} in kg, or liquid water content, and the particle terminal fall speed v_{l} in m s^{-1} are well known and calculated using Eqs. (2.3) and (2.4) from Atlas and Ulbrich (1974), respectively.

$$m_{\mathrm{l}} = 1000 \cdot \frac{4}{3}\pi \cdot (0.5D)^3 \tag{2.3}$$

$$v_{\mathrm{l}} = 9.65 - 10.3 \cdot e^{-600D} \tag{2.4}$$

For snow, the measured cross-sectional area differs from the required maximum dimension of the particle due to the non-spherical shape of snowflakes. This difference requires applying a transfer function. However, Lempio et al. (2007) found that the product of particle terminal fall speed and particle mass (liquid water equivalent) as a function of cross-sectional area is in the same order of magnitude for various frozen precipitation particle types. Hence, no transfer function between cross-sectional area and maximum diameter is required when using a spherical lump graupel assumption. The lump graupel assumption works well for frozen precipitation particles between 0.4 and 9 mm in diameter, whereas particles exceeding 9 mm in diameter rarely occur. Nevertheless, events with large particles introduce larger errors to the estimate in the same way as the retrieval quality may largely differ for individual snowfall events. Overall, no unique snowfall retrieval can be derived using optical disdrometers without recording the individual particle shape. Compared to a Geonor gauge, the optical disdrometer agreed well in most cases and overestimated a few light snowfall cases during the 1999/2000 winter period at Uppsala (Lempio et al., 2007). Following the lump graupel approximation by Hogan (1994), particle mass m_{s} (Eq. 2.5) and particle

terminal fall speed v_s (Eq. 2.6) are calculated empirically as

$$m_s = 1.07 \cdot 10^{-5} \cdot (100D)^{3.1}, \tag{2.5}$$

$$v_s = 7.33 \cdot (100D)^{0.78}. \tag{2.6}$$

Klepp et al. (2010) observed lump graupel being the most frequently occurring precipitation type over the cold-season Norwegian Sea during the LOFZY campaign. Battaglia et al. (2010) discuss several sources of error for a snow-measuring laser-optical PARticle SIze VELocity disdrometer (PARSIVEL) whereof those for particle shape and orientation, margin effects, and coinciding particles also apply to the ODM470. However, the PARSIVEL is more sensitive to influences by wind speed and wind direction on the falling precipitation particles because the PARSIVEL has a fixed non-pivotable horizontal optical sensing area.

For mixed-phase precipitation, OceanRAIN generally uses the snow retrieval (Eqs. 2.5 and 2.6) to calculate the precipitation rate because the absolute error of treating rain drops like snow particles, and thus underestimating the precipitation rate, results in a smaller error than vice versa. In more than 90 % of the precipitating cases from RV *Polarstern* in OceanRAIN the precipitation rate calculated with Eqs. (2.3) and (2.4) (theoretical rain rate) exceeds precipitation rate calculated with Eqs. (2.5) and (2.6) (theoretical snow rate) by a factor of 50 to 200. Accordingly, this large difference might cause large biases in the precipitation rate for misclassified PP. The influence of misclassified PP cases is discussed in Section 5.3.2 of Chapter 5. Correctly classified mixed-phase precipitation events might still strongly underestimate the precipitation rate if the instantaneous rain fraction sharply exceeds 0.5. The minute-aggregated fraction of liquid and solid particles cannot be identified by the ODM470 and would require ancillary data such as a video disdrometer. More details on the instrumentation can be found in Lempio et al. (2007) while algorithm features are explained in Klepp (2015).

2.2 The HOAPS satellite climatology

The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data (HOAPS; Andersson et al., 2010b; Fennig et al., 2012) set contains the parameters precipitation, evaporation as well as turbulent heat fluxes and other variables of the atmospheric state over the global ice-free oceans. HOAPS relies exclusively on PMW data from the Special Sensor Microwave Imager (SSM/I) and its successor with additional sounder (SSMIS), deployed onboard the Defense Meteorological Satellite Program (DMSP) satellite series. The last official release, HOAPS-3.2, is limited to the entire SSM/I record that ends in 2008. In this thesis, we employ the prolonged version HOAPS-3.3 of the precipitation parameter using the combined SSM/I and SSMIS Fundamental Climate Data Record (Fennig et al., 2015). Only the precipitation parameter in HOAPS-3.3 has been

processed until December 2015, kindly provided by the Satellite Application Facility on Climate Monitoring (CM SAF) of EUMETSAT and German Meteorological Service (DWD) for the German research initiative FOR1740 on the Atlantic freshwater cycle. The prolonged period of 1987–2015 obtains a larger overlap with the OceanRAIN data set that starts in 2010. This means, the HOAPS subset used in this thesis is exclusively based on SSMIS. Not incorporating further ancillary data, HOAPS remains an independent precipitation reference for re-analysis and model evaluation (Andersson et al., 2011).

To derive HOAPS precipitation estimates, a neural net algorithm has been developed. The neural net is trained with global-ocean data from DMSP F11, F13 and F14 and of ECMWF variational precipitation calculations of the year 1998. In that year, an El Niño event led to a wide range of measured brightness temperatures (Bauer et al., 2006). HOAPS excludes any data that has been sampled less than 50 km away from sea-ice or land areas to avoid contamination due to sudden changes in surface emissivity. The sea-ice mask excludes only grid boxes with an average sea ice fraction that exceeds 15 % for at least 5 consecutive days (Andersson et al., 2010b). The HOAPS precipitation estimates have been carefully quality-checked. Furthermore, an inter-calibration among the SSM/I and SSMIS instruments ensures a physically consistent retrieval (Andersson et al., 2010a).

HOAPS provides 3 different products with different resolutions: HOAPS-S, HOAPS-C and HOAPS-G. The scan-based HOAPS-S contains pixel-wise precipitation estimates of individual SSM/I and SSMIS swaths, respectively. This ungridded data provides precise information of time and location of the sensed precipitation scenes, usually unavailable in gridded products. The spatial resolution strongly depends on the frequency channels used in the HOAPS retrieval. The lowest SSM/I frequency channel of 19.35 GHz senses at the coarsest spatial resolution of about 43 km x 69 km. More information on the SSM/I and SSMIS scan geometry are given in Wentz (2013). Precipitation rates below 0.3 mm h^{-1} are excluded from the HOAPS-S data because the precipitation signal is no longer distinguishable from noise. HOAPS also provides gridded data in HOAPS-C (6-hourly) and HOAPS-G (monthly) with a spatial resolution of 0.5° by 0.5°. HOAPS 3.2 is freely accessible via http://www.hoaps.org from the CM SAF of EUMETSAT and DWD, HOAPS-3.3 is available upon request. The successor HOAPS4 is planned to be released in early 2017.

2.3 S-Pol radar data from RICO

During the wet season in winter 2004/05, a scanning S-band radar was deployed on the subtropical island of Barbuda (17.61 °N, 61.82 °W) for the Rain In Cumulus clouds over the Ocean (RICO) campaign. The S-Pol radar is introduced in Section 2.3.1) where features of data processing and quality checking are listed. Further, the S-Pol

radar data has been adjusted to the requirements of simulating ship tracks and satellite pixels, as explained later in Section 4.2.1 of Chapter 4.

2.3.1 The S-Pol radar

The S-Pol radar was developed by NCAR to provide a cost-effective portable weather radar that has been used in a number of field campaigns (Keeler et al., 2000). One of these campaigns was RICO during which the scanning S-Pol radar operated from 24 November 2004 to 25 January 2005 (Rauber et al., 2007). During RICO, the S-Pol performed surveillance scans with 10.68 cm wavelength at a 0.5 ° elevation angle, covering a 150 km domain in radius. Each scan was gridded onto a polar grid with a range resolution of 150 m, and has undergone extensive filtering to exclude anomalous returns from ground clutter. The minimum reflectivity threshold of 7 dBZ serves to exclude Bragg scattering (Knight and Miller, 1998). The measured radar reflectivity Z (dBZ) is converted into a rain-rate R (mm d^{-1}) by applying a Z-R relationship. Here, we apply a Z-R relation derived from PSDs that were measured during RICO aircraft flights (Snodgrass et al., 2009).

$$Z = 88 \cdot R^{1.52} \tag{2.7}$$

According to Eq. 2.7, the minimum reflectivity threshold of 7 dBZ corresponds to about 0.19 mm h^{-1}. However, for the application of simulating spatial effects from satellite sensors versus ship tracks, the absolute precipitation rates according to the chosen Z-R relationship play a minor role. The data is freely accessible online (UCAR/NCAR-EOL, 2011).

2.4 CMORPH

The Climate Prediction Center (CPC) MORPHing technique (CMORPH; Joyce et al., 2004) provides global precipitation estimates between 60°N and 60°S at fairly high temporal and spatial resolution. Additionally to coarser resolution products, CMORPH offers a resolution of 8 km in space and 30 min in time. CMORPH ingests precipitation estimates that have been derived from low earth orbiting (leo) microwave satellite sensors exclusively. These PMW sensors include SSM/I, AMSU-B, AMSR-E and TRMM Microwave Imager (TMI). From these PMW sensors, precipitation estimates are generated using algorithms of Ferraro (1997) for SSM/I, Ferraro et al. (2000) for AMSU-B and Kummerow et al. (2001) for TMI. However, as the spatial resolution of most of these PMW sensors remains fairly below 8 km, the gridded CMORPH precipitation estimates are spatially interpolated to obtain the high spatial resolution of 8 km.

To obtain a high temporal resolution of 30 min, the PMW-derived precipitation features are propagated forward in time using cloud system advection vectors (CSAV) from correlating consecutive images obtained entirely from geostationary IR satellite

sensors. Applying a time-weighted linear interpolation to these half-hourly CSAVs yields morphed half-hourly precipitation estimates (Joyce et al., 2004). Combining PMW and IR sensors takes advantage of the high spatio-temporal resolution of IR sensors and the relatively high-quality precipitation estimates from PMW sensors. The CMORPH data is available from December 2002 onward, also offering a near-realtime product called QMORPH. Further information and a link to directly access CMORPH precipitation data can be found on the NOAA CPC CMORPH website (Joyce and Janowiak, 2016).

2.5 GPM IMERG

The Global Precipitation Measurement (GPM; Hou et al., 2013) mission was launched in February 2014, whereas final precipitation data starts in April 2014. GPM builds on the concept of the widely used TRMM to center a core satellite that carries an advanced radar/radiometer system to measure precipitation from space while serving as a reference standard to unify precipitation measurements from a constellation of other satellites measuring precipitation. The GPM core observatory carries a multi-channel GPM Microwave Imager (GMI) as well as the first space-borne Ku/Ka-band dual-frequency precipitation radar (DPR). This setup makes the DPR more sensitive to light rain below 0.5 mm h^{-1} and snowfall compared to the TRMM-PR that was designed specifically to observe heavy tropical precipitation. Besides the DPR, the GMI is designed to accurately sense precipitation at various intensities using 13 channels that range from 10 to 183 GHz.

The Integrated Multi-satellitE Retrievals for GPM (IMERG; Huffman, 2015) is run twice in near-realtime, 4 (early) and 12 hours past observation (late or middle), and after the monthly gauge analysis (final). Independent of the 3 runs, IMERG is computed in half-hourly (3IMERGHH) and monthly (3IMERGM) resolution, whereas the latter is only available for final runs. The half-hourly satellite-gauge product is available in 0.1° x 0.1° resolution between 60°N and 60°S, beginning in mid-March 2014. We solely use the multi-satellite precipitation estimate with gauge calibration, final run (GPM 3IMERGHH v03). A technical manual describes the IMERG data and recent changes in detail (Huffman et al., 2015).

2.6 MSG-CPP

The METEOSAT Second Generation (MSG) Cloud Physical Properties (CPP) algorithm derives cloud, precipitation and radiation parameter from satellite sensors such as the Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard MSG (Roebeling et al., 2006). Besides CM SAF as one of the developers, KNMI offers real-time

and archived precipitation data, among other cloud and radiation parameters. Precipitation is calculated from cloud water path, cloud phase, cloud particle effective radius and cloud column height, all of which are taken from MSG-CPP. The combination of these cloud-microphysical properties strongly improves the generally low skill to separate non-precipitating clouds with small particles from precipitating clouds with larger particles or ice crystals (Roebeling and Holleman, 2009). All parameters exist for daytime only because the CPP algorithm relies on solar backscattered radiation. The MSG-CPP covers an area from -80 to 80°N and -50 to 50°E (MSG full disk), starting on 1 January 2004. More details can be accessed on the data set website (KNMI, 2011).

Chapter 3

An automatic precipitation phase distinction algorithm for optical disdrometer data[1]

3.1 Introduction

Rare and often low-quality gauge-based surface reference data sets challenge the in situ validation of oceanic precipitation as observed by PMW and active microwave satellite sensors (Taylor, 2000; Adler et al., 2012). Over land, radar and gauge-based precipitation monitoring networks cover a large fraction of the land surface for more than 2 decades, which qualifies them to validate precipitation satellite estimates (Schneider et al., 2014). However, the ocean surface lacks dense long-term in situ precipitation monitoring networks. Furthermore, existing coastal and island-based precipitation measurements may not fully represent oceanic precipitation because the measured particle size distributions (PSDs), precipitation rates, and accumulations may differ from those measured over the open ocean (Kidd and Levizzani, 2011). However, Bumke and Seltmann (2012) found no difference between PSDs over coastal areas and ocean. Most existing in situ oceanic precipitation data sets sample measurements from low-quality rain gauges on voluntary observing ships (VOSs; Kent et al., 2010) or buoy arrays (Weller et al., 2008). Many of these in situ ocean data sets include present weather observations but lack quantitative estimates of precipitation. The high-latitude ocean completely lacks precipitation measurements that sample solid and mixed-phase precipitation. However, recent and future precipitation satellite estimates demand high-quality in situ quantitative precipitation estimates including snowfall over the global ocean.

[1]This chapter has been published with minor modifications as:
Burdanowitz, J., C. Klepp and S. Bakan (2016): An automatic precipitation phase distinction algorithm for optical disdrometer data over the global ocean. Atmos. Meas. Tech., **9**, 1637-1652.

The large uncertainty in precipitation gauge measurements arises from the rough open-ocean conditions that complicate precipitation monitoring. Under high wind speed, standard rain gauges with horizontal catchment surfaces face a large undercatch (Yuter and Parker, 2001; Michelson, 2004). In the extratropics, mixed-phase and solid precipitation cause further difficulties strongly adding to the undercatch (Goodison, 1978) of horizontally oriented measuring surfaces. In contrast, optical instruments with a vertically oriented measuring surface such as disdrometers perform better at capturing precipitation under high wind speeds, though varying wind directions are challenging. Optical disdrometers are thus denoted as the reference in situ instrument to measure precipitation (Taylor, 2000).

To provide systematic high-quality in situ precipitation data over the ocean, the long-term Ocean Rainfall And Ice-phase precipitation measurement Network (OceanRAIN; Klepp, 2015) applies automatic optical disdrometers of type ODM470 that are deployed onboard sea-going RVs for operation in all climatic regions. The ODM470 was developed to measure under high wind speed and frequently varying wind directions. Its cylindrical measuring volume ensures being independent from the wind-driven incidence angle of the falling hydrometeors while a wind vane keeps the measuring volume perpendicular to the instantaneous wind direction. The ODM470 accuracy lies within a range of 3 % rain accumulation limited to rainfall at various wind conditions with respect to an improved ship rain gauge including side collectors on RV *Alkor* on the Baltic Sea (Bumke and Seltmann, 2012). Compared to an ANS410 WMO-reference rain gauge over land (Lanza and Vuerich, 2009), the ODM470 deviates by 2 % under low wind speed (Klepp, 2015). For snow, a predecessor of the current ODM470 agreed with the observer's log during the Lofoten Cyclones campaign (LOFZY; Klepp et al., 2010) in detecting snowfall events. More recent results for measuring snow and mixed-phase precipitation are expected soon from the WMO Solid Precipitation InterComparison Experiment (SPICE) at Marshall field site in Boulder (CO, USA), where the ODM470 was compared against a multitude of in situ solid precipitation instruments for more than 2 years. The ODM470 suits well to measure various types of precipitation under open-ocean conditions onboard sea-going RVs.

The deployment of the ODM470 on several RVs allows to sample OceanRAIN precipitation data from all climate zones including the cold-season high latitudes. This requires a precipitation-phase (PP) distinction between rain, snow, and mixed phase in order to provide correct precipitation rates for disdrometer-measured PSDs. The PP information usually originates from human observers' reports saved in the WMO present weather code *ww* (WMO, 2015). Efforts to automatize present weather observations impose high requirements on instruments such as present weather sensors. Automated present weather sensors encounter problems at temperatures around 0°C as well as for light precipitation and small particle sizes (Merenti-Välimäki et al., 2001). High wind speed also complicates the PP determination because the wind speed strongly interferes with the particle fall speed that solely carries the PP information. Thus, most studies to distinguish PPs limit the wind conditions to low wind speed or calm

conditions (Löffler-Mang and Joss, 2000; Yuter et al., 2006; Ishizaka et al., 2013). Only few studies apply more sophisticated instruments that use articulating particle size velocity (PARSIVEL) disdrometers to account for wind effects and thus directly derive the PP from the particle fall speed (Friedrich et al., 2013). More simply constructed instruments such as the ODM470 require ancillary data to determine the PP.

In OceanRAIN, we aim to replace the so far manual PP distinction method by an automatic algorithm for three main reasons. First, the manual method consumes a considerable amount of time and workforce because the 1 min precipitation data requires visual inspection of air temperature, present weather observations, and theoretical rain and snow rate. Second, the human-based PP decision based on visual data inspection lacks objectivity while the decision itself remains non-transparent to the user. Third, temporal gaps exist between the 3-hourly present weather observational timesteps, especially during nighttime adding to the uncertainty. Currently, no measures of this PP uncertainty are provided in the manual method. For these reasons, we present a new automatic PP distinction algorithm including a PP probability for OceanRAIN precipitation data that is also applicable to all other instruments sampling PSDs of precipitation.

The new PP distinction algorithm follows a statistical approach guided by many other studies that relate atmospheric predictors to the PP (Koistinen and Saltikoff, 1998; Fuchs et al., 2001; Dai, 2008; Froidurot et al., 2014). Most previous work focuses on PP distinction over land only, while we introduce a new PP distinction algorithm over the ocean. Dai (2008) compares ocean and land areas using a relatively coarse time step of few to several hours depending on availability of observations. In contrast, OceanRAIN offers atmospheric measurements at 1 min resolution while present weather observations are limited to 3-hourly timesteps during daytime only. These high-resolution ancillary data from the RV combined with PSD data from the optical disdrometer enable a more accurate and reliable PP distinction.

Section 3.2 introduces the manual PP distinction method, and the OceanRAIN data set in detail. Section 3.3 presents different atmospheric variable combinations and methods to predict the PP. In Sect. 3.3.1, one PP distribution distinguishes between two PPs, while in Sect. 3.3.2 one PP distribution distinguishes between three PPs. Section 3.3.3 introduces a newly developed method to predict three PPs using two PP distributions. Section 3.4 discusses the results by comparing with similar studies. Section 3.5 completes our investigations with a summary and concluding remarks.

3.2 Data and methods

The OceanRAIN database collects atmospheric data including precipitation rates on several RVs since 2010. From the current permanent deployments given in Section 2.1

of Chapter 2, RV *Polarstern* is chosen due to the high fraction of high-latitude precipitation. The core instrument of OceanRAIN, the optical disdrometer ODM470, is introduced in detail in Sect. 2.1.1. The following Section 3.2.1 explains how the manual method works to retrieve the PP that has been used so far in OceanRAIN. These manually determined PPs function as a benchmarking reference for the new automatic PP distinction algorithm.

3.2.1 The manual PP distinction

Though time-consuming, the manual PP distinction was so far employed to determine the PP that is required to calculate the precipitation rate. Because we apply this manual PP distinction data as reference to the new automatic PP distinction algorithm, a detailed explanation follows. If available, shipboard present weather observations stored in the WMO standard meteorological present and past weather code *ww* (WMO, 2015) are translated into the three PPs: rain, snow, and mixed phase according to Petty (1995), displayed in Table 3.1. However, the translation of *ww* codes into a PP partly

Table 3.1: Translation of WMO present weather codes *ww* (WMO, 2015) into the three PPs from Petty (1995), Froidurot et al. (2014), and OceanRAIN. *ww* codes printed in bold can be translated into multiple PPs in OceanRAIN depending on ancillary data. "Indet./hail" denotes indeterminate precipitation or hail used for classification in Petty (1995).

Source	Rain	Snow	Mixed phase	Indet./hail
Petty (1995)	21, 25, 50–55, 58–65, 80–82, 91–92	22, 26, 70–78, 85-86	23–24, 56–57, 66–69, 79, 83–84	20, 27–29, 87–90, 93–99
OceanRAIN	**20**, 21, 25, **29**, 50–67, 80–82, 91–92, **95**, **97**	**20**, 22, **26–27**, **29**, 70–79, 85–86, **87–90**, **93–95**, 96, **97**, 99	23–24, **26–27**, **29**, 68–69, 83–84, **87–90**, **93–95**, **97**	–
Froidurot et al. (2014)	58–65, 80–82, 91–92	70–79, 85–86	68–69, 83–84	–

differs between OceanRAIN and Petty (1995). While Petty (1995) assigns one single PP to each of the *ww* codes, OceanRAIN allows multiple PPs for a single ambiguous *ww* code (bold weather codes in Table 3.1). Instead, Petty (1995) lists ambiguous *ww* codes in a category called "indeterminable" (abbreviated "Indet." in Table 3.1) that,

however, includes no PP information anymore. For that reason we deviate from this procedure to retain as much PP information as possible. Another difference concerns *ww* codes for all kinds of freezing rain (i.e., rain at freezing temperatures) without snow that Petty (1995) classifies as mixed phase. Classifying freezing rain as mixed phase by applying the lump graupel assumption (Eqs. 2.5 and 2.6) leads to an underestimated precipitation rate. This underestimation arises because falling raindrops freeze only after passing the disdrometer's optical volume when hitting any obstacle, which is why we consider freezing rain cases in OceanRAIN as rain ($ww = 56, 57, 66, 67$). Likewise, we assign a snow flag to ice pellets ($ww = 79$) as well as mostly to hail ($ww = 89, 90$), graupel ($ww = 87, 88$), and combinations of both ($ww = 93, 94, 96, 99$). The aim is mainly to separate frozen (solid) from non-frozen (liquid) precipitation particles to account for differences in density and cross-sectional area that affect Eqs. (2.3) to (2.6) and hence the precipitation rate. In contrast, the study by Froidurot et al. (2014) concentrates exclusively on clear rain, snow, and mixed-phase observations (Table 3.1) by neglecting drizzle, freezing rain, and ice pellets, among others. In general, assigning the correct PP for a given ambiguous *ww* code requires visual inspection of PSDs and ancillary data collected onboard the RV.

The *ww* code from shipboard observations on RV *Polarstern* is available 3 hourly during daytime only. Nighttime conditions and PP changes between two consecutive 3-hourly observational time steps require ancillary data from the RV to derive the PP. By ancillary data we refer to atmospheric variables measured onboard the ship including the ODM470, such as air temperature, humidity, and precipitation rate. These ancillary data are available at a much higher resolution of 1 min compared to the 3-hourly observations. Initially, we assign the PP derived from the *ww* code directly to every single minute of precipitation that follows a 3-hourly observation as a first-guess information. If available, air temperature as one of the ancillary data serves to possibly correct this first-guess PP. For near-freezing air temperatures, the manual procedure calculates the precipitation rate after Eq. (2.1) for rain (Eqs. 2.3 and 2.4) and snow (Eqs. 2.5 and 2.6) assumption separately. Large differences between theoretical rain and snow rate can help to identify a plausible PP. However, if both theoretical rain and snow rate differ by much less than 2 orders of magnitude, their influence on the PP decision becomes negligible, which makes the PP more arbitrary. Accordingly, the manual method might be biased by the worker who decides for a PP and the observer on the RV. For these reasons, we aim at developing an automatic PP distinction algorithm at 1 min resolution that statistically derives a PP from atmospheric measurements.

3.2.2 OceanRAIN data from RV *Polarstern*

The manual PP estimation has been applied to more than 4 years of OceanRAIN data from RV *Polarstern* (11 June 2010–8 October 2014). This period consists of several expeditions to Arctic and Antarctic regions. In addition to the high latitudes, RV

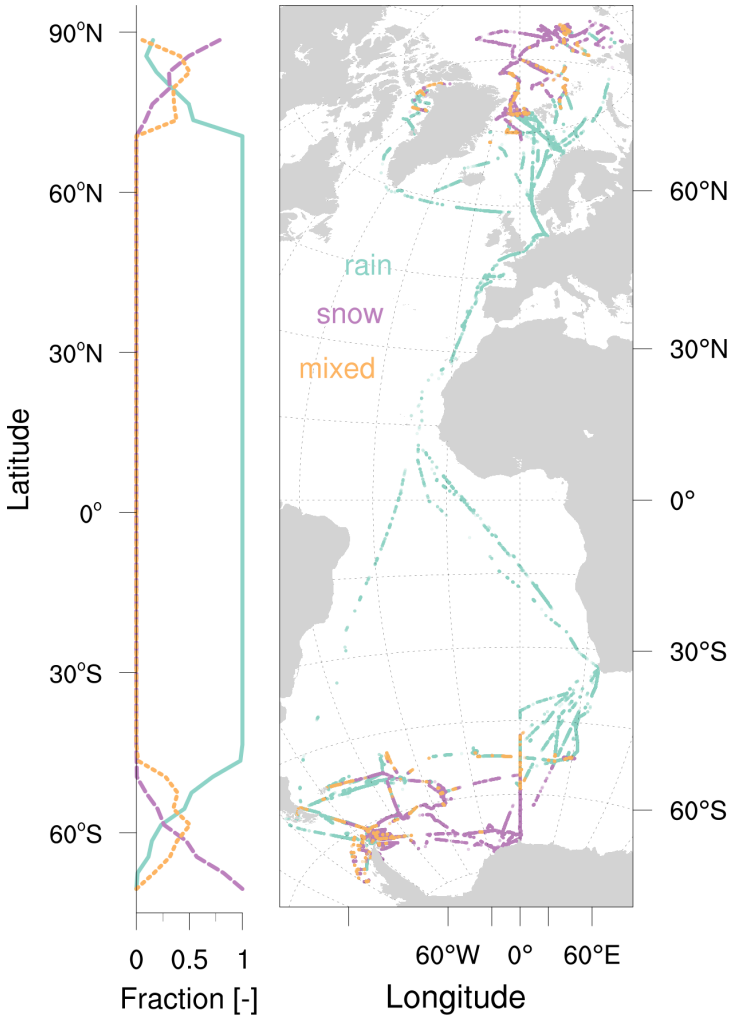


Figure 3.1: Map illustrates ship tracks from RV *Polarstern* ALL data (11 June 2010–8 October 2014), whereby dots denote minutes of occurring precipitation classified by the manual PP distinction (cyan: rain; orange: mixed phase; purple: snow). Harbor times and minutes without precipitation are not shown. Left side denotes the fraction of each PP averaged per latitude.

Polarstern collected precipitation data from the tropics and subtropics when crossing the equator in the Atlantic Ocean six times (Fig. 3.1). The whole measuring period amounts to more than 268 000 min of precipitation excluding periods of maintenance in harbors and instrument outages. Snow or mixed-phase precipitation occurred almost exclusively poleward of 45° S and 70° N, which largely depends on seasonality. During boreal warm season, RV *Polarstern* sailed over the northern hemispheric Atlantic Ocean and in the entire Arctic area, whereas during austral warm-season RV *Polarstern* cruised on the southern hemispheric Atlantic Ocean and at the Antarctic. As an exception, RV *Polarstern* spent the whole year 2013 including austral cold season in the Southern Hemisphere, which explains the multitude of mixed-phase and snow precipitation cases between 45 and 75° S that are not sampled at corresponding northern hemispheric latitudes. For the sake of polar research, RV *Polarstern* spends

Table 3.2: OceanRAIN data sets from RV *Polarstern* divided into sub-data sets that are used in the analysis. While RSM (rain, snow, mixed phase) and ALL (all data) include the mixed phase, RS (rain, snow) sub-data exclude mixed-phase precipitation. RSM and RS include only those minutes with at least 20 particles of precipitation falling at mid- or high latitudes at air temperatures around the freezing point (see Sect. 3.2.2). The no-rain fraction (rain fraction subtracted from 1) yields the fraction frozen precipitation meaning snow cases for RS and snow and mixed phase for RSM and ALL.

Name	Description	Size (min)	No rain (frac)
ALL	Complete data sample	268 340	0.57
RSM	Data sub-sample (incl. mixed phase)	165 632	0.61
RS	Data sub-sample (excl. mixed phase)	149 635	0.57

most research time in the polar regions, which results in a high time fraction of snow or mixed-phase precipitation of 0.57 (Table 3.2). Accordingly, precipitation occurred most frequently at temperatures around 0 °C and at high relative humidity (Fig. 3.2). The high time fraction of snow or mixed-phase precipitation around 0 °C makes RV *Polarstern* an extremely valuable data set for oceanic PP distinction analysis.

The whole RV *Polarstern* data set, denoted ALL (for all data), consists of about 268 000 min of precipitation. The ship's positions cover large areas of distinctly high or low temperatures where the PP assignment is trivial and does not help in developing the PP algorithm. Therefore, we reduce the complete RV *Polarstern* data set ALL to minutes of highest PP uncertainty (Table 3.2). Air temperatures below −6 °C and above 8 °C are excluded as well as ship locations between 45° S and 70° N latitude wherein virtually no solid or mixed-phase precipitation was observed within the 4-year period (Fig. 3.1). We exclude minutes with a total particle number of less than 20 particles because they cannot guarantee a meaningful PP distinction. These limitations leave a subset of data denoted RSM (for rain, snow, mixed phase) with 165 632 min of rain, snow, or mixed-phase precipitation. By that, the no-rain time fraction including snow or mixed-phase precipitation increases from 0.57 (ALL) to 0.61 (RSM). If we further exclude mixed-phase precipitation the gained sub-sample, denoted RS (for rain, snow), reduces to 149 635 min while the no-rain fraction decreases to 0.57 (Table 3.2).

Atmospheric variables measured onboard RV *Polarstern* include temperature-related (T, T_d, T_2h) and humidity-related variables (rH, T_d), air pressure (P), and, from the ODM470, wind speed (not used for analysis) and particle diameter (D). Instead of D, we use the 99th percentile of D, D_{99}, which is a measure for the maximum particle diameter measured within 1 min but excluding erroneously large particles possibly caused by particle coincidences, drip-off drops, or other artifacts. Table 3.3 lists all relevant variables from RV *Polarstern* and the ODM470. Note that all variables are

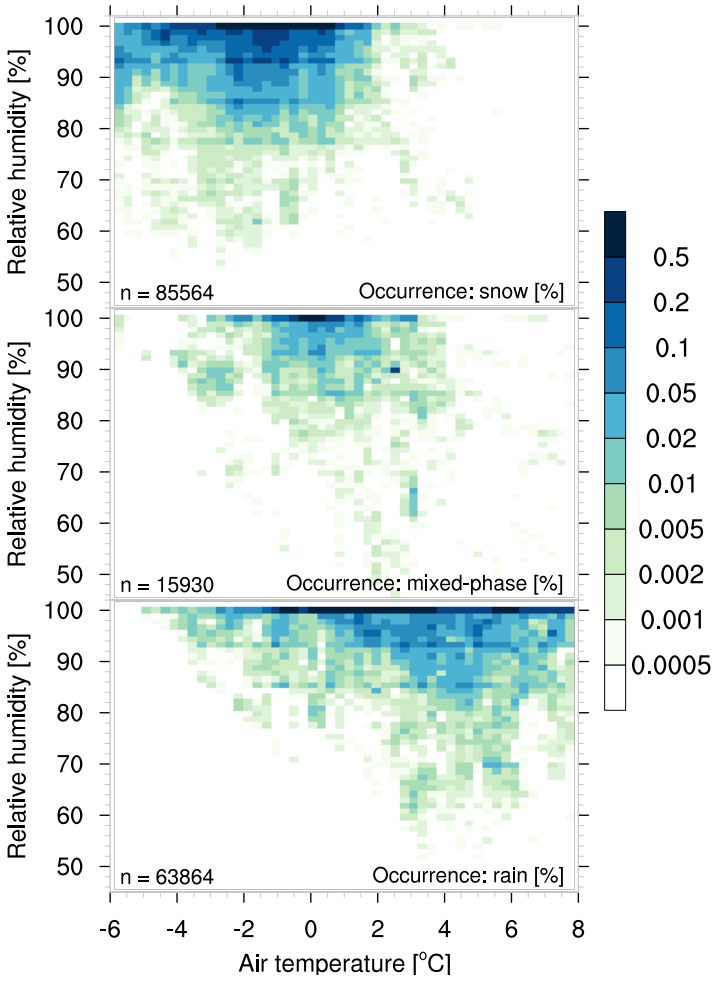


Figure 3.2: Two-dimensional histogram shows relative occurrence (%) for each PP (top: snow; middle: mixed phase; bottom: rain) after manual PP distinction from OceanRAIN RSM data set of RV *Polarstern*. n denotes the number of minutes used per PP (165 632 in total).

measured distinctly higher than 2 m above the surface at about 43 m in order to reduce interfering sea spray and splashing wave water.

3.3 The automatic PP distinction

3.3.1 One PP distribution to predict two PPs (2P1D)

This study aims at predicting the PP automatically by using available in situ atmospheric predictor variables (Table 3.3). While we first focus on predicting two PPs using one PP distribution (Sect. 3.3.1; 2P1D), we later apply one PP distribution to predict three PPs (incl. mixed phase; Sect. 3.3.2; 3P1D). Section 3.3.3 presents a novel approach that predicts three PPs applying two PP distributions (3P2D).

For the PP prediction we adopt a statistical model using logistic regression to relate the

Table 3.3: List of available meteorological predictor variables in OceanRAIN used in the logistic regression model to predict the PP.

Variable	Description	Unit	Source
T	Air temperature	°C	RV *Polarstern*
T_d	Dew point temperature	°C	RV *Polarstern*
T_2h	Air temperature 2 h prior to observation	°C	RV *Polarstern*
rH	Relative humidity	%	RV *Polarstern*
P	Sea-level air pressure	hPa	RV *Polarstern*
RR	Precipitation rate assuming rain	$\mathrm{mm\,h^{-1}}$	ODM470
D_{99}	99th percentile of particle diameter	mm	ODM470

available observed atmospheric variables (predictor variables) to the PP as suggested by Koistinen and Saltikoff (1998), henceforth KS98. The predictor variables are fitted against binary dependent variables to calculate the PP probability p(PP). Taken from the manual PP distinction data (Sect. 3.2.1), the binary dependent variables attain a rain probability p(rain) [frac] of either 0 (snow) or 1 (rain). Once fitted, p(rain) can attain any value between 0 and 1 depending on the predictor variables. p(rain) is calculated by

$$p(\mathrm{rain}) = \frac{1}{1 + e^{\alpha+\beta\cdot V_1+\gamma\cdot V_2+\ldots+\omega\cdot V_n}}, \tag{3.1}$$

whereby V_i represents the atmospheric predictor variables. α, β, γ, …, ω denote the regression coefficients that are determined by minimizing the sum of squared errors (nearest-neighbor method) with respect to the PPs from the manual PP distinction. Generally, we use the term PP probability, p(PP), representing both rain (p(rain)) and snow probability (p(snow)) if not stated differently. The snow probability is calculated as $1-p$(rain), excluding mixed phase for now in this simple model.

We calibrate various combinations of atmospheric predictor variables V_i (Table 3.3) for RS sub-data (Table 3.2) to find the combination that predicts best the PP. KS98 state that the combination of air temperature T and relative humidity rH, called T_rH, is suited best to predict the PP. For T_rH, Eq. (3.1) changes to

$$p(\mathrm{rain}) = \frac{1}{1 + e^{(\alpha+\beta\cdot T+\gamma\cdot \mathrm{rH})}}, \tag{3.2}$$

where the number of regression coefficients reduces to three. In lack of alternative reference data, we evaluate the calculated regression coefficients of RS sub-data using the same manually determined PPs as used for the model calibration. Nevertheless, we investigated the robustness of the regression coefficients using 100 realizations of only 50 % randomly chosen minutes of precipitation from the RS data set. The standard deviation of the 100 realizations rarely exceeds 10 % of the individual regression coefficients from the whole RS data set, which confirms the robustness of the calculated regression coefficients. If in the manual PP reference data set a minute of precipitation

is assigned rain, the statistical model by definition "agrees" for $p(\text{rain}) \geq 0.5$ while it "disagrees" for $p(\text{rain}) < 0.5$. For the rain/snow distinction four possible combinations exist – rain agreement, snow agreement, rain disagreement, and snow disagreement. For instance, rain disagreement means that the statistical model predicts rain that disagrees with the manual PP reference data indicating snow. Combined in a contingency table we choose four scores to evaluate how well the atmospheric predictor variable combinations serve to predict the PP in this statistical model.

First, the accuracy serves to evaluate the overall correctness of the predictor variable combinations with respect to the manual PP reference data set. The accuracy represents the sum of cases in which model and manually determined PP reference data of RS sub-data agree divided by the total number of minutes in RS sub-data. Ideally, the accuracy is close to 1.

Second, we consider the bias score defined as the ratio between the sum of disagreeing rain predictions and agreeing rain predictions to the sum of disagreeing snow predictions and agreeing rain predictions, all with respect to the manually determined PP reference data. Accordingly, a bias score of $b < 1$ represents an overprediction of snow by the model, whereas $b > 1$ represents an overprediction of rain by the model. However, the bias should be interpreted with caution because the manual reference data set might be biased itself. Thus, the bias rather carries the information in which direction the predicted PP deviates from the manual reference data.

Third, we determine the percentage of cases misclassified (PM). Misclassified means that predicted high-probability cases ($p > 0.95$) disagree with the manual PP reference data. For PM, the number of these misclassified cases is divided by the number of all RS cases. Ideally, PM is close to 0.

Fourth, the percentage of uncertain cases (PU) estimates how well the PPs are separated by the predictor variables used. PU represents the number of cases with $0.05 < p < 0.95$ divided by all RS cases. Accordingly, PU measures the fraction of cases that the model is unable to predict at a high level of certainty. The definitions of PM and PU follow the evaluation method in Froidurot et al. (2014).

The four performance scores are calculated for both 100 realizations of 50 % randomly chosen minutes of precipitation (black boxes and whiskers in Fig. 3.3) and for all minutes of RS sub-data (red stars). The percentiles (5th, 25th, 50th, 75th, and 95th) illustrate how strongly the RS data set scatters and whether differences among predictor variable combinations are significant ($p = 0.95$, $n = 100$).

The PPs calculated with the logistic regression model reach an accuracy of more than 88 % for combinations of atmospheric predictor variables that all include the air temperature T (Fig. 3.3). T carries the most straightforward PP information in most cases. Combining T with up to two other relevant predictor variables (connected by underscores) aids to assess their value in determining the PP. Table 3.4 displays the most important fitted regression coefficients for different combinations of predictor variables

Table 3.4: List of regression coefficients calculated with Eq. (3.1) by minimizing the sum of squared errors with respect to the manual PP reference data for two PPs using one PP distribution (2P1D; Sect. 3.3.1), three PPs using one PP distribution (3P1D; Sect. 3.3.2), and three PPs using two PP distributions (3P2D; Sect. 3.3.3). For 3P2D, the asterisk denotes the rain distribution that was derived setting the mixed phase to snow. KS98 denotes the coefficients recommended by Koistinen and Saltikoff (1998) derived over Finland.

Method	Variables used	Regression coefficients			
	(V_1_V_2_V_3)	α	β	γ	δ
KS98	T_rH	−22	2.7	0.2	–
2P1D	T_rH	−13.39	1.818	0.127	–
	T_rH_D_{99}	−10.83	1.780	0.118	−1.062
	T_rH_RR	−13.55	1.738	0.135	−0.325
3P1D	T_rH	−9.766	1.382	0.092	–
	T_rH_D_{99}	−8.364	1.364	0.090	−0.732
	T_rH_RR	−10.01	1.331	0.099	−0.204
3P2D	T_rH	−5.687	1.429	0.055	–
	T_rH*	−15.40	1.482	0.144	–
	T_rH_D_{99}	−4.794	1.467	0.056	−0.556
	T_rH_D_{99}*	−13.94	1.431	0.145	−0.959
	T_rH_RR	−5.888	1.412	0.060	−0.059
	T_rH_RR*	−13.95	1.382	0.136	−0.316

using the OceanRAIN sub-sample RS (2P1D) and the sub-sample RSM (3P1D and 3P2D).

Combining T with the air temperature 2 h prior to observation (T_{2h}) does not increase the accuracy of T (both 88.5 %). Other time intervals led to similarly small performance changes being in agreement with Froidurot et al. (2014). Accordingly, weather fronts associated with T drops do not seem to have an imprint on T_T_{2h} or they are currently underrepresented in the OceanRAIN data set. The air pressure P may have an impact on the PP at higher elevations due to lower air density (Dai, 2008). This, however, cannot explain the better accuracy of 89.2 % for T_P compared to T. Over the ocean, the additional skill in the predictor P might be caused by certain weather situations that favor either rain or snow, and are sufficiently sampled in the OceanRAIN data set. The relative humidity rH and the dew point T_d (not shown) reach about the same accuracy of 89.4 %. The addition of P and rH to T leads to a statistically significant ($p = 0.95$, $n = 100$) but only slight increase in accuracy compared to T alone.

With the 99th percentile of the particle diameter D_{99} and the calculated theoretical rain rate RR (Eqs. 2.3 and 2.4), physical properties of precipitation particles directly enter the PP distinction. This direct physical relation explains the notably higher accuracy of at least 90 % by T_RR, T_D_{99}, and other combinations containing RR and D_{99}. The similarly high performance of these three predictor combinations is driven by the particle diameter that mainly influences RR. Combinations of T, a humidity-related

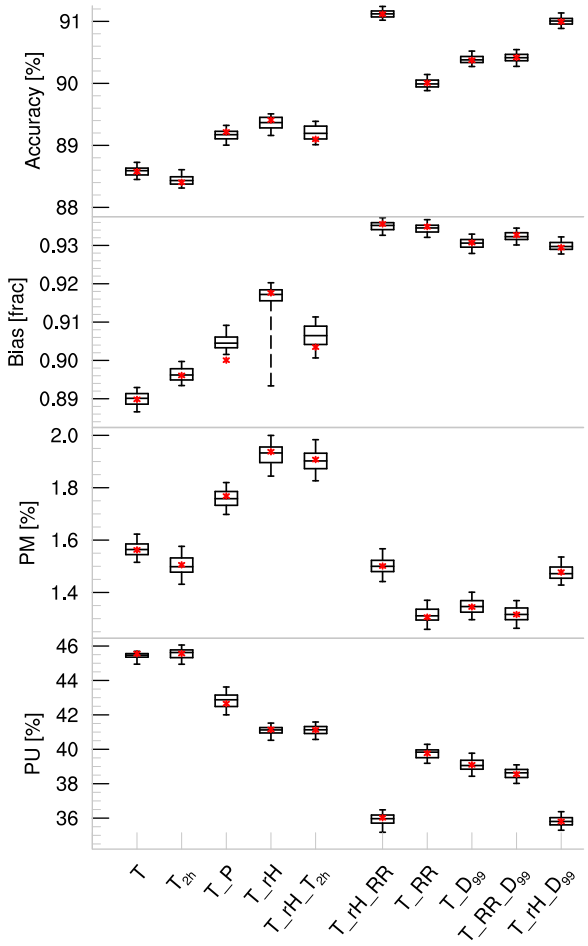


Figure 3.3: Box-and-whisker plot displays interquartile spread (black box: 25th, 50th, and 75th percentile) and lower (whisker: 5th percentile) as well as upper (95th percentile) extremes, calculated from 100 realizations of each 50 % randomly chosen minutes of precipitation from RS sub-data. Red stars denote the values for 100 % of RS sub-data. Accuracy (%), bias score (frac), percentage misclassified (PM: fraction of disagreeing cases with high certainty of $p > 0.95$ in %), and percentage unclassified (PU: fraction of uncertain cases of $0.05 < p < 0.95$ in %) serve as performance scores using the calculated coefficients in Table 3.4 against the manually determined PP reference data. Labels indicate variable combinations, whereby all combinations include T.

variable such as rH, and a diameter-related variable such as D_{99} reach the highest accuracy of more than 91 %. Combinations of four or five of the available atmospheric predictor variables such as T_rH_RR_D_{99} brought no noticeable further increase in accuracy (not shown). From the considered predictors, a combination of three out of the available predictor variables suits best to accurately distinguish between rain and snow.

The bias provides the ratio of rain cases predicted by the statistical model and observed rain cases from the manual PP reference data. All predictor variable combinations range between 0.89 and 0.94, which implies an underprediction of rain and hence an

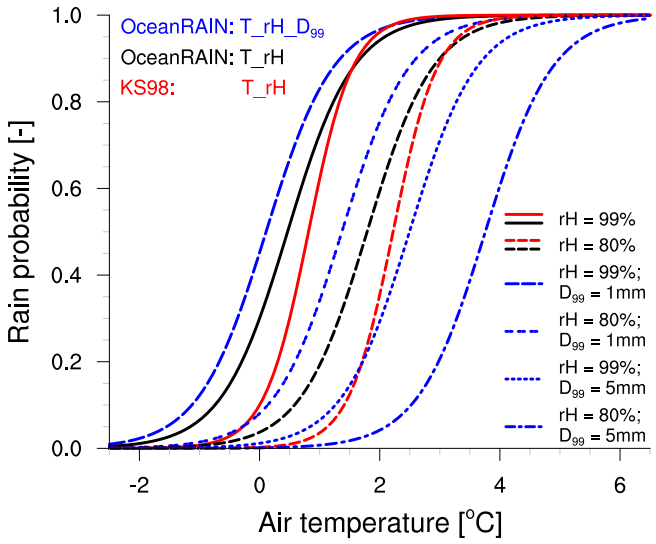


Figure 3.4: Rain probability using regression coefficients from Table 3.4 for OceanRAIN RS sub-data (2P1D) with the predictor variables T_rH (black), T_rH_D_{99} (blue) both fitted against OceanRAIN, compared to KS98-recommended coefficients for T_rH (red). Dashed lines (black, red) indicate a PP distribution where rH is set to 80 % while for solid lines it is set to 99 %. For T_rH_D_{99} (blue lines), D_{99} is set to either 1 or 5 mm in addition to rH.

overprediction of snow. Combinations that contain RR and D_{99} reach the smallest overprediction of snow, whereas T holds the strongest snow bias. The lower snow bias combined with the higher accuracy of predictor variables carrying particle diameter information highlights the need to include physically related variables in a statistical model to predict the PP.

Besides being accurate and unbiased, a small PP transition region of low PP certainty (low PU) combined with a low fraction of highly certain but misclassified PP cases (low PM) characterize a useful predictor variable combination. The PU decreases with increasing accuracy. Consequently, predictor variable combinations including rH and either D_{99} or RR reach the lowest PU of about 36 %. This low PU and thus fairly narrow PP distribution causes a slight increase in PM for T_rH_RR and T_rH_D_{99} (1.5 %) compared to T_D_{99}, T_RR, and T_RR_D_{99} (1.3 %). However, the positive effect of using RR or D_{99} outweighs the slightly negative influence of rH on PM. Consequently, the physical related predictor variables confirm their good performance in predicting the PP.

The T_rH coefficients that were calculated for Finland in KS98 and confirmed in Froidurot et al. (2014) over Switzerland reach an accuracy of 88.6 %, which is slightly lower than those coefficients optimized for OceanRAIN (89.4 %). A two-tailed t test confirms the difference to be statistically significant (p=0.99, n=100). The OceanRAIN-adapted coefficients exhibit a shallower rain/snow transition that results in a 0.8 °C lower temperature at p(rain) = 0.1 while both distributions equal at p(rain) = 0.9 (Fig. 3.4). Compared to OceanRAIN, the steeper rain/snow transition

against T fitted in KS98 holds a much lower PU of 24 % but to the expense of a much higher PM of 4 % and a snow bias of 0.8. Consequently, the coefficients from KS98 better predict most uncertain cases with T_rH but miss more extreme cases such as freezing rain. For the OceanRAIN data set, the PP prediction using the RS-fitted coefficients reflects better the OceanRAIN PP distribution compared to the KS98-fitted coefficients as indicated by the accuracy.

For $T\_\mathrm{rH}\_D_{99}$, the rain/snow transition shifts with T depending on D_{99}. While $D_{99} = 1\,\mathrm{mm}$ shifts the rain/snow transition to even lower temperatures by about 0.5 °C, $D_{99} = 5\,\mathrm{mm}$ shifts it towards higher temperatures by about 2 °C, both compared to T_rH derived from OceanRAIN RS sub-data. The shallower rain/snow transition of the PP distribution fitted for OceanRAIN compared to that over Finland is likely caused by more freezing rain cases sampled in OceanRAIN, which the KS98-fitted coefficients for T_rH cannot predict.

3.3.2 One PP distribution to predict three PPs (3P1D)

In a second step, we include mixed-phase precipitation into the algorithm because mixed-phase precipitation marks the transition from frozen to liquid particles and thus carries the highest uncertainty. We calculate the regression coefficients using the RSM sub-data including 165 632 min of precipitation measured onboard RV *Polarstern*. The three-phase distinction 3P1D fits $p(\mathrm{rain})$ against three PPs from the same manually determined PP reference data set as before. However, the calculated transition phase between snow with $p(\mathrm{rain}) = 0$ and rain with $p(\mathrm{rain}) = 1$ is interpreted as mixed phase, defined in the range of $0.3 \leq p(\mathrm{rain}) \leq 0.7$ after KS98. The approximated coefficients for predictor variable combinations V_i differ considerably from those calculated for the two-phase method 2P1D (see Table 3.4 in Sect. 3.3.1).

We evaluate the calculated PP probability against PPs from the manual PP reference data using RSM sub-data. Again, accuracy, bias, and PM serve as a measure of quality, while PU is no longer suitable for evaluation because the transition region of highest uncertainty between snow and rain represents mixed-phase precipitation. Overall, this three-phase method 3P1D yields an accuracy between 74 and 78 %, which corresponds to an absolute decrease of about 14 % compared to 2P1D (Fig. 3.5). To that large decrease in accuracy two reasons mainly contribute: (1) the manual PP reference data, acting as reference data, holds large uncertainties in the mixed phase, as well. The *ww* code represents snapshots of 3-hourly observations. Therefore, they hardly satisfy the need for minute-based observations because the mixed-phase rain/snow fraction can vary dramatically, both temporally and spatially. (2) KS98 assume the mixed-phase precipitation to occur in the transition region between rain and snow, which is true in most cases. However, several cases exist in which mixed-phase precipitation occurs at distinctly high or low air temperature (Fig. 3.2) and thus 3P1D misclassifies these cases.

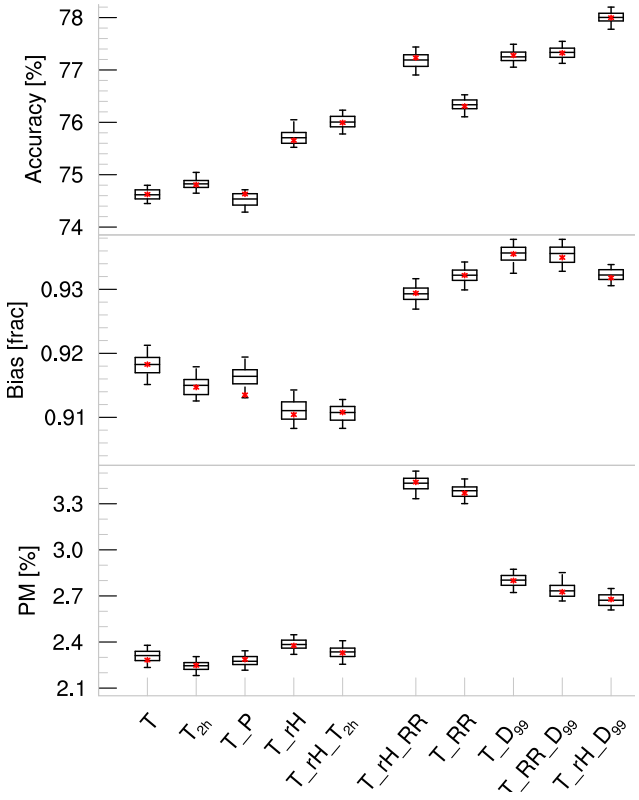


Figure 3.5: Performance of fit is shown for different combinations of atmospheric variables as in Fig. 3.3 for RSM sub-data. All variable combinations again include T.

Relative to each other, the individual variable combinations perform similar compared to 2P1D. T, $T\_T_{2\mathrm{h}}$, and T_P have the lowest accuracy of below 75 % (Fig. 3.5) and a bias below 0.92. The addition of rH significantly increases the accuracy by about 1 %, whereas T_rH_$T_{2\mathrm{h}}$, T_rH, and $T\_T_{\mathrm{d}}\_T_{2\mathrm{h}}$ (not shown) do not differ much from each other. The predictor variable combinations that include the diameter-related predictors RR and D_{99} lead to the highest accuracy of 76 up to 78 %. The highest accuracy of 78 % reached by T_rH_D_{99} represents a statistically significant performance increase to the remaining variable combinations in 3P1D, which contrasts to 2P1D where T_rH_RR does not perform significantly better than T_rH_D_{99}.

For the bias, predictor combinations including RR and/or D_{99} reach the least pronounced snow bias of about 0.93, whereas the remaining predictor combinations feature significantly lower biases, mostly below 0.92. In that respect, the bias of 3P1D resembles that of 2P1D (see Fig. 3.3 in Sect. 3.3.1) both in terms of magnitude and in the individual performance of the predictor variable combinations.

While the ranking of predictor variable combinations with respect to accuracy and bias looks very similar compared to 2P1D, PM tends to form three clusters. The first cluster comprises predictor variables without particle diameter information, holding the lowest PM of 2.2 to 2.4 %. The second cluster includes RR but not D_{99}, holding the highest PM (3.4 %). In the third cluster each predictor variable combination includes D_{99} but

performs better than the second cluster with PM of about 2.8 %. $T\_\mathrm{rH}\_D_{99}$ in the third cluster offers the best compromise in maximizing the accuracy while minimizing the fraction of misclassified cases.

In contrast to 2P1D, for 3P1D PM tends to scale with accuracy for many predictor variable combinations. While $T\_\mathrm{rH}\_D_{99}$ exhibits an about 0.5 % larger PM than T, the PM of $T\_\mathrm{rH}\_\mathrm{RR}$ and T_RR are even 1.1 % larger. A high PM indicates a clear disagreement between calculated PP and manually estimated PP. Note, however, that not in all of these clearly disagreeing cases the manual PP reference data necessarily contains the correct PP. Physically related predictor variables such as D_{99} can assist to unveil cases falsely classified by the manual PP estimation. For example, D_{99} is able to identify snow or mixed-phase cases, falsely classified as rain in the manual reference data. Except for the tropics, rain drops rarely exceed drop diameters of 5 mm (Bentley, 1904; Villermaux and Bossa, 2009). Larger drops mainly break up or collide with neighboring drops. D_{99} excludes coincidences of drops as well as artificial drops dripping off the instrument housing by discarding the uppermost percentile of measured drop diameters per minute. Therefore, particles classified as rain drops with D_{99}>5 mm very likely represent frozen particles, which means that they were falsely classified as rain (Fig. 3.6). Below 4 °C, 163 rain cases in RSM sub-data (about 0.25 %) are likely falsely classified. This could explain about half of the 0.5 % PM difference of $T\_\mathrm{rH}\_D_{99}$ to T in Fig. 3.5).

The T_rH coefficients calculated in KS98 reach an accuracy of 78.6 %, but PM amounts to 7.2 % misclassified cases (not shown), which is more than a factor of 2.5 higher than the PM of $T\_\mathrm{rH}\_D_{99}$. The shift towards higher air temperatures and the steeper rain/snow transition in the PP distribution using the coefficients recommended in KS98 (see Fig. 3.4) explain the large amount of misclassified cases. However, as stated before, the coefficients in KS98 derived over Finland cannot represent the temperature distribution of PPs in the OceanRAIN data set.

3.3.3 Two PP distributions to predict three PPs (3P2D)

The relatively low accuracy reached with the three-phase method after KS98 using one PP distribution (3P1D) motivates a novel investigation of how to further improve the PP prediction for three PPs. Instead of applying one PP distribution to determine rain, mixed-phase, and snow precipitation, we suggest to approximate two separate PP distributions for rain and snow (3P2D). These two individual PP distributions are derived analogous to the method for one PP distribution by assigning the mixed-phase PP differently – first set it to rain to calculate the snow PP distribution, then set it to snow to calculate the rain PP distribution. Subtracting the sum of both individually calculated PP distributions from 1 yields the PP distribution for mixed phase. In contrast to 3P1D, the separately calculated coefficients for rain and snow (Table 3.4) lead to individual distributions only connected via the mixed phase.

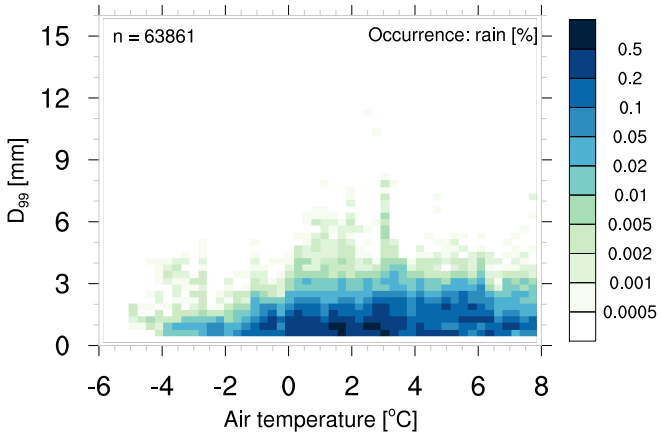


Figure 3.6: Two-dimensional histogram of temperature and the 99th percentile of the particle diameter for cases classified as rain by the manual PP estimation in RSM.

Analogous to 2P1D (Sect. 3.3.1), the accuracy represents the percentage of cases with $p(\mathrm{PP}) > 0.5$ that agree in their respective PP with the manual PP reference data. The bias represents the ratio between the sum of predicted rain cases and the sum of rain cases in the manual PP reference data. Please note that the bias definition remains unchanged for 3P2D that includes mixed phase compared to 2P1D. However, the additional PP distribution slightly modifies the calculation of PM and PU, illustrated in Fig. 3.7. PM represents the percentage of all certain cases ($p(\mathrm{PP}) > 0.95$; hatched area in Fig. 3.7) in which either one of the PPs disagrees with the manual PP reference data. PU as the percentage of uncertain cases ($0.05 < p(\mathrm{PP}) < 0.95$; shaded area) represents only those cases where all PPs are uncertain after definition. We introduce this limitation because if $p(\mathrm{PP}) < 0.05$ holds for at least one PP then we would not consider this PP uncertain anymore. Note that for mathematical reasons we cannot display $\mathrm{PM_{mix}} > 0$ and $\mathrm{PU} > 0$ in the same figure, which is why we set $\mathrm{PM_{mix}} > 0$.

This 3P2D method using two individual PP distributions reaches on average a higher accuracy compared to 3P1D (Fig. 3.8). Whereas T, $T\_T_{2\mathrm{h}}$, and T_P hold less than 78 % accuracy, $T\_\mathrm{rH}\_D_{99}$ reaches the highest accuracy of 81.2 %. As for 3P1D, the improvement is mainly caused by adding the predictor D_{99} that performs significantly better than when adding the predictor RR. Also the overprediction of snow by all predictor variable combinations with respect to the manually determined PP reference data stays the same in 3P2D. The physically related variables are least biased (about 0.93), which consistently highlights the improvement of including them in the predictor variable combination. However, for PM stronger differences among these physically related predictor variables arise. While T_RR holds the highest PM (about 2.3 %), $T\_\mathrm{rH}\_D_{99}$ reaches 1.9 % PM, which is on the order of the predictor variable combinations without RR and D_{99} (1.8 %). However, the physically related predictors reach again lowest PUs of about 38 % while T holds a PU of 51 %. In combination with the other scores we recommend $T\_\mathrm{rH}\_D_{99}$ followed by $T\_\mathrm{RR}\_D_{99}$ and T_D_{99} to most accurately predict the PP using two independent PP distributions.

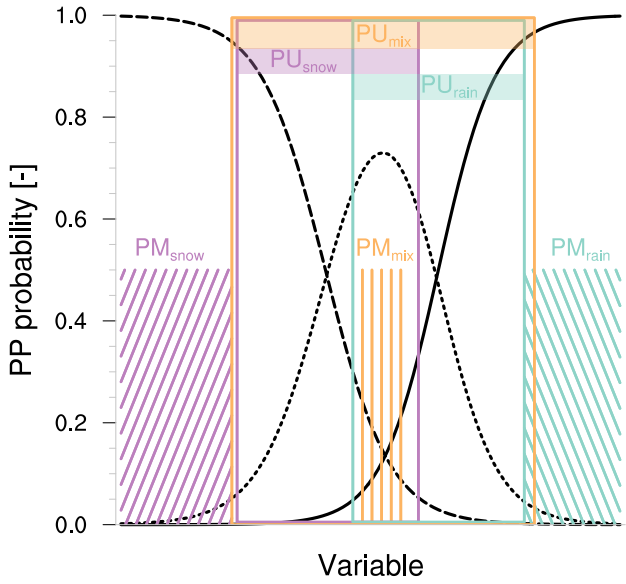


Figure 3.7: Graph illustrates the calculation of PU (framed) and PM (hatched) including snow (dashed/purple), mixed phase (dotted/orange), and rain (solid/cyan), analogous to Fig. 3 in Froidurot et al. (2014). PU divides the sum of cases with $0.05 < p(\mathrm{PP}) < 0.95$ for all PPs by the sum of all RSM cases. PM divides the sum of cases with $p(\mathrm{PP}) > 0.95$ for one of the PPs that disagrees with the manual PP estimation by the sum of all RSM cases. We set $\mathrm{PM}_{\mathrm{mix}} > 0$ because otherwise we could not display it in the same PP distribution (rH kept constant) with $\mathrm{PU} > 0$.

Compared to 3P1D after KS98, the PM decreases for 3P2D. This decrease in PM ranges between 0.5 and 1 % and thus highlights the improved performance of using two PP distributions instead of one to predict the PP. The lower PM and higher accuracy approve that the novel method applying two independent PP distributions better represents the PP distribution in OceanRAIN RSM.

To understand the better performance of 3P2D compared to 3P1D after KS98, we consider how the PP fraction is distributed with respect to T around the freezing point (rain/snow transition) in the manual PP reference data (Fig. 3.9). While the rain occurrence shows a relatively low skewness, the mixed-phase/snow distribution is slightly left-skewed. This higher skewness with a secondary maximum in the mixed-phase distribution at −3 °C (minimum in snow distribution) cannot be well represented by one PP distribution. One PP distribution is limited to match all three PP distributions at the same time that can only represent an average skewness. In that respect, deriving two independent PP distributions driven by mixed-phase precipitation better reflects the PP distribution of each PP individually with respect to the manual PP reference data in OceanRAIN RSM.

The question arises whether the left-skewed distribution of snow and mixed-phase precipitation in OceanRAIN sub-data RSM represents a feature of the oceanic PP distribution or if it simply reflects a currently insufficient length of the OceanRAIN time series.

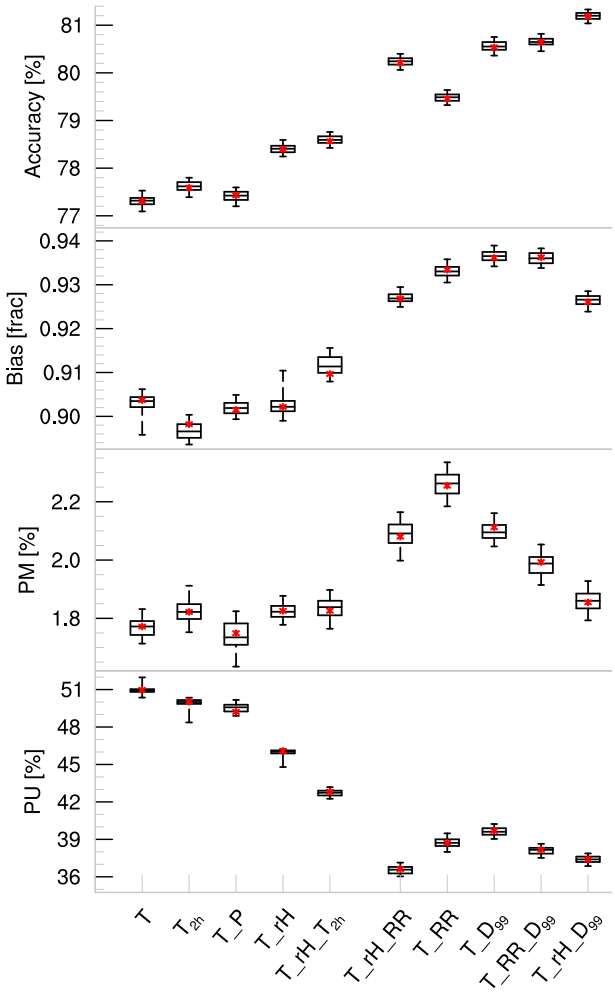


Figure 3.8: As Fig. 3.3 but for RSM including mixed phase, using two independent PP distributions (3P2D). The calculation of PM and PU differs from Fig. 3.3 as displayed and explained in Fig. 3.7.

Though the latter seems more likely, addressing this question comprehensively, however, remains beyond the scope of this study due to the limited available OceanRAIN data sample. Future studies could clarify this aspect by reanalyzing the constantly growing OceanRAIN database.

Nevertheless, differences remain due to the chosen PP distinction method. By discriminating three PPs, 3P1D and 3P2D enable a smoother rain/snow transition compared to 2P1D due to included mixed-phase precipitation (Fig. 3.10). At lower temperatures, 2P1D approaches the snow distribution of 3P2D, while at higher temperatures it approaches the rain distribution of 3P2D. In other words, the steeper rain probability distribution for 2P1D clarifies the slightly smaller unclassified range ($0.3 < p(\mathrm{PP}) < 0.7$) compared to 3P2D as seen in the percentage unclassified (compare Fig. 3.3 and Fig. 3.8).

D_{99} as additional variable in $T\_\mathrm{rH}\_D_{99}$ tends to shift the snow and rain distributions

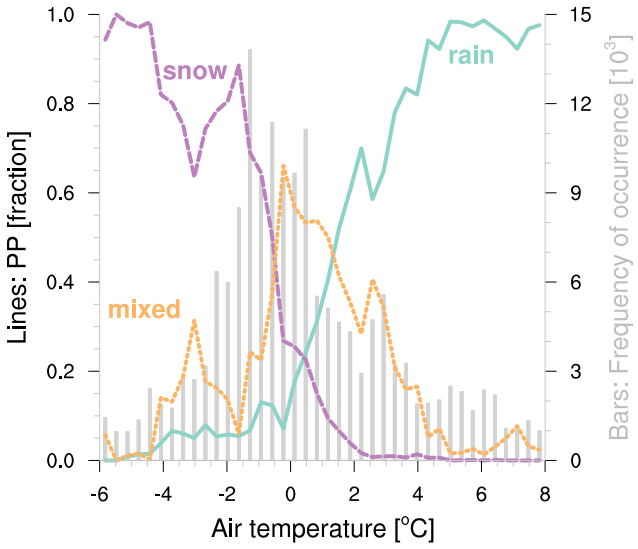


Figure 3.9: Lines show PP fraction for rain (solid, cyan), mixed phase (dotted, orange), and snow (purple, dashed) from OceanRAIN RSM (165 632 min) determined with the manual PP estimation against temperature. Gray bars represent the temperature frequency of occurrence (in 10^3).

to higher temperatures and apart of each other, which also resolves more extreme cases. This distribution shift with temperature follows a physical reason: large snow particles better withstand melting at high air temperatures than small snow particles. This physical information lacks in T_rH, which notably decreases its accuracy (cf. Fig. 3.8).

3.4 Discussion

After finding suitable methods for both the rain/snow distinction (Sect. 3.3.1) as well as for the rain/snow/mixed-phase distinction (Sect. 3.3.3) we compare the results to those of similar studies. For the rain/snow distinction over Switzerland using T_rH derived over Finland by KS98, Schmid and Mathis (2004) find a higher accuracy of 92.4 % compared to our calculated accuracy of 88.6 % when using the same KS98 regression coefficients $\alpha = 22$, $\beta = 2.7$, $\gamma = 0.2$. Schmid and Mathis (2004) find an overprediction of snow cases (bias 0.82), very similar to the OceanRAIN RS snow overprediction (bias 0.8) using the same coefficients derived by KS98. However, for fitting the regression coefficients to our data set (Table 3.4) we still obtain a slightly lower accuracy of 89.4 % calculated for T_rH and 91 % for $T\_\mathrm{rH}\_D_{99}$ while the low-bias decreases to 0.92 and 0.93, respectively. These performance improvements indicate, first, different conditions for PP transition over the ocean compared to Finland of KS98 while, second, the OceanRAIN data set is in relatively close agreement with the Swiss data.

With respect to two PPs, including the mixed phase decreases the accuracy to below 78 % while PM almost doubles. To elaborate on reasons for that accuracy decrease we

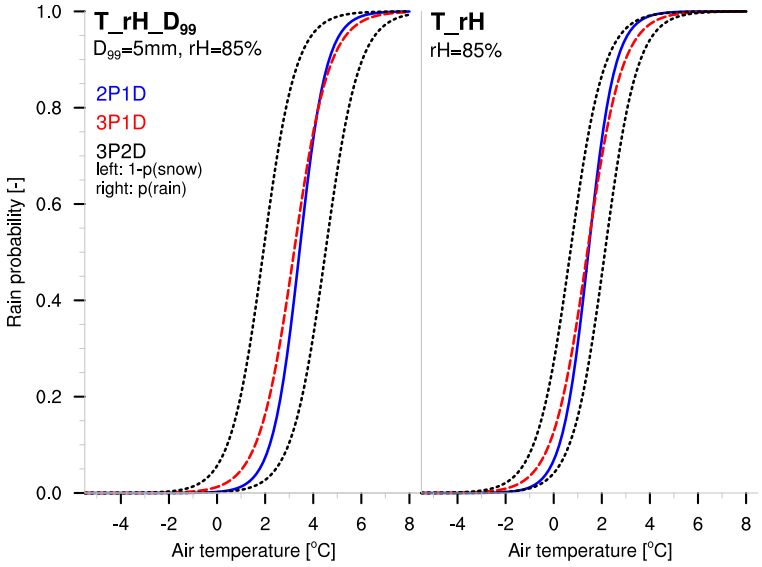


Figure 3.10: Air temperature versus predicted PP by the different methods: two PPs (2P1D; solid blue), three one-PP distributions (3P1D; dashed red), and three two-PP distributions (3P2D; dotted black). 3P2D consists of two curves (left: snow distribution as $1 - p(\text{snow})$; right: rain distribution as $p(\text{rain})$) for the calculated coefficients of $T\_\text{rH}\_D_{99}$ (left panel; $\text{rH} = 85\,\%$, $D_{99} = 5\,\text{mm}$) and T_rH (right panel; $\text{rH} = 85\,\%$).

consider a study of Gjertsen and Ødegaard (2005), who applied the same translation of *ww* codes into PPs for *ww* codes between 50 and 86. Using 3P1D, they find an accuracy of 86 % compared to Norwegian synoptic stations (6 months winter period) and 85 % compared to independent climatological stations over Norway (1 month). They obtain an overprediction of snow (bias of 0.92) and problems in predicting the PP of supercooled rain during prevailing temperature inversions. In OceanRAIN we find a similar overprediction of snow (bias T_rH: 0.91; $T\_\text{rH}\_D_{99}$: 0.93) with respect to the manual PP reference data in OceanRAIN. This overprediction of snow occurs predominantly around 0 °C that is the temperature range sampled most frequently (cf. Fig. 3.9). Hence, OceanRAIN is likely to face the same problems underpredicting rain when supercooled raindrops fall under prevailing temperature inversions. Further work is required in order to clarify whether we need additional ancillary data to reduce the bias or whether the logistic regression model is unable to provide a less biased PP prediction.

Assuming that mixed-phase precipitation causes most of the accuracy decrease between 2P1D and 3P1D as well as 3P2D, we consider the individual probability of detection (POD) for rain, snow, and mixed phase. For rain, the POD is calculated by dividing the number of agreeing rain cases by the number of all observed rain cases. For the POD of 3P1D using the KS98-fitted coefficients for T_rH for rain, snow, and mixed phase we find 0.92, 0.78, and 0.21 ($T\_\text{rH}\_D_{99}$: 0.92, 0.86, and 0.25). The respective PODs from Gjertsen and Ødegaard (2005) for the same settings reveal slightly different PODs of 0.81, 0.97, and 0.25. Whereas they obtain a notably higher POD

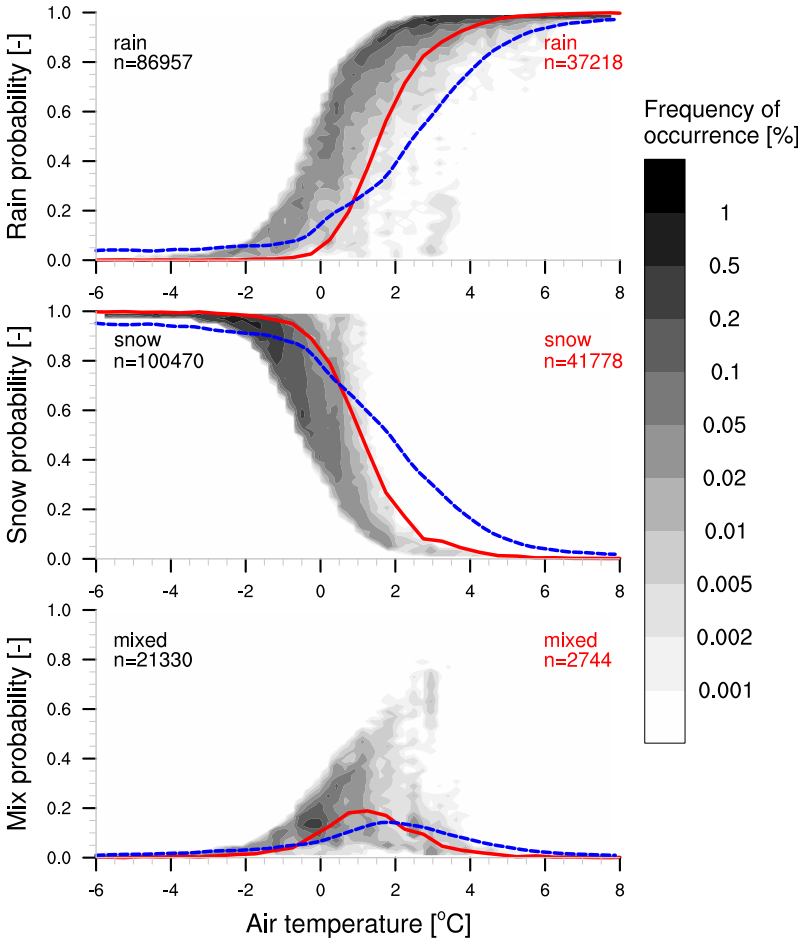


Figure 3.11: PP probability shown using the new 3P2D method with two individual PP distributions ($T\_\mathrm{rH}\_D_{99}$) as frequency of occurrence (%) in gray shades against air temperature according to PP reference data that separates rain, snow, and mixed phase in OceanRAIN ALL for more than 4 years of RV *Polarstern* data. Solid red lines represent the mean PP fraction from observations in the Swiss Alps (1991–2010) from Froidurot et al. (2014); dashed blue lines show mean PP fraction for oceanic ship data (DS464.0; 1977–2007) from Dai (2008).

for snow, the rain POD is lower compared to OceanRAIN. Nevertheless, mixed-phase precipitation confirms to carry the largest prediction uncertainty of all three PPs.

The variable combination $T\_\mathrm{rH}\_D_{99}$ distinguishes best rain, snow, and mixed-phase precipitation in OceanRAIN data. In comparison with PP fractions allocated into temperature bins from 30 years of Swiss Alps data from Froidurot et al. (2014), in most cases the PP transition in OceanRAIN occurs at lower temperatures (Fig. 3.11). However, the analysis by Froidurot et al. (2014), among other conditions, neglects all kinds of freezing rain ($ww = 56, 57, 66, 67$) that we assign to rain. Without these "cold rain" cases, the rain/snow transition shifts towards higher temperature that may in parts explain the temperature difference in Fig. 3.11. Additionally, the PP probability distribution in the OceanRAIN RV *Polarstern* data sample is biased by the high

number of temperatures around 0 °C that occur by a factor of 3 to 4 more often than temperatures between −4 and 4 °C (cf. Fig. 3.9), and relative humidity close to 100 %. These frequently sampled conditions put their mark on the average rain/snow transition by reducing the rain/snow transition temperature compared to the Swiss Alps where T and rH were sampled more homogeneously (Fig. 9 in Froidurot et al., 2014).

Despite the high number of available minutes with precipitation in OceanRAIN, the rather short time series on climatological timescales and the spatial distribution of along-track data limit the representativeness. However, a different rain/snow transition between land and ocean might be observable. Dai (2008) found a systematic land–ocean difference in the rain/snow transition between land and ocean in 30 years of NCEP ADP Operational Global Surface Observations (DS464.0; 1977–2007). Whereas over land, rain transitions into snow relatively quickly with increasing temperature ($-2 < T < 4$ °C), over ocean the transition zone is wider ($-3 < T < 6$ °C). Although the rain/snow transition zone within OceanRAIN appears to be wider compared to regression coefficients recommended by Koistinen and Saltikoff (1998) as seen in Fig. 3.4, the rain/snow transition in OceanRAIN compares better to the Swiss Alps data (Froidurot et al., 2014) than to the NCEP DS464.0 ocean data (Dai, 2008) that reveal a wider transition zone. In specific, OceanRAIN relatively closely agrees with the NCEP DS464.0 ocean data for $T < 0$ °C, whereas larger differences of > 1 °C occur in the range of $2 < T < 5$ °C. Two main reasons can explain the different rain/snow transition between OceanRAIN and NCEP DS464.0 ocean data by Dai (2008). First, *ww* codes used in the NCEP ocean data are subject to larger uncertainty compared to OceanRAIN. In contrast to the RV *Polarstern* onboard weather observatory by the German Meteorological Service, many VOSs such as cargo ships in NCEP DS464.0 ocean data have inadequately trained observers that might use certain *ww* codes preferentially, ships possibly avoid bad weather, or measurement quality may suffer from instrument biases (Petty, 1995). For these reasons, the wider rain/snow transition zone likely reflects a higher uncertainty of the NCEP DS464.0 ocean data compared to the OceanRAIN data from RV *Polarstern* or the Swiss Alps data. Second, RV *Polarstern* mainly sampled warm-season precipitation in the Atlantic Arctic and Antarctic regions with the exception of the austral cold season in 2013. In addition to that, the heterogeneous regional and seasonal sampling by RV *Polarstern* might have favored conditions under which inversions prevail that allow rain at fairly low temperatures but inhibit snow under relatively high temperatures.

While the sampling imbalance of RV *Polarstern* may indicate a restricted representativeness of PPs in OceanRAIN, the $T\_\mathrm{rH}\_D_{99}$ predictor variable combination recommended as the new PP distinction method for OceanRAIN well represents the observed PPs within OceanRAIN. The continuously growing time series of RV *Polarstern* among other RVs in OceanRAIN allows to recalibrate or refine the algorithm geographically for a longer time series with comprehensive statistical sampling in the future.

3.5 Summary and concluding remarks

We developed a novel automatic algorithm to distinguish the PP within OceanRAIN in situ precipitation data to introduce a statistical PP probability and to increase the data post-processing efficiency. The analysis focused on identifying the most suitable combination of available atmospheric predictor variables to determine the PP. For that purpose, we applied a simple logistic regression model suggested by Koistinen and Saltikoff (1998) that was shown to perform well over land. Previous studies mainly rely on air temperature T, relative humidity rH, air pressure P, and others to predict the PP. We investigated several of these atmospheric predictor variable combinations to obtain a PP probability. In particular, we tested the performance of the logistic regression model after Koistinen and Saltikoff (1998) for OceanRAIN using two (excl. mixed phase) and three PPs (incl. mixed phase) against the manually estimated observation-based PP in OceanRAIN. Besides increasing the efficiency in predicting the PP with an automatic method, we developed a novel three-phase method that uses two individual and independent PP distributions to predict the PP more accurately.

The study led to the following main results.

a. In OceanRAIN, the combination of air temperature T, relative humidity rH, and the 99th percentile of the particle diameter D_{99} (called $T\_\mathrm{rH}\_D_{99}$) predicts best the PP for all investigated methods.

b. Applying more than three of the chosen atmospheric predictor variables negligibly increases the accuracy in predicting the PP.

c. The two-phase method (2P1D) using the predictor variable combination $T\_\mathrm{rH}\_D_{99}$ and regression coefficients fitted to OceanRAIN reaches an accuracy above 91 % with a slight overestimation of snow cases for the mid- and high latitudes between -6 and 8 °C in the OceanRAIN data set with respect to the manual PP reference data including shipboard present weather observations.

d. A novel three-phase method using two individual PP distributions (3P2D) for rain and snow performs better than a three-phase method that relies exclusively on one PP distribution (3P1D after Koistinen and Saltikoff, 1998). As a reason, two individual PP distributions are capable of better representing unequally distributed or skewed PP distributions of atmospheric predictor variables as well as certain weather situations that might currently be over- or undersampled. Accordingly, this performance difference might decrease once the investigated 4-year OceanRAIN time series grows further while sampling biases vanish.

e. The OceanRAIN data using 3P2D reveal a wider rain/snow transition zone compared to data derived over Finland (Koistinen and Saltikoff, 1998). The rain/snow transition in OceanRAIN occurs at slightly lower temperatures compared to the

data from Finland as well as NCEP DS464.0 global ocean ship data (Dai, 2008). The difference in the rain/snow transition zone likely originates from heterogeneous spatial and seasonal sampling in OceanRAIN that is likely to decrease with an increasing OceanRAIN time series. In contrast, a higher quality of the derived *ww* codes in OceanRAIN compared to the average VOS suggests a higher certainty of the derived PPs. The Swiss Alps data (Froidurot et al., 2014) shows a similar rain/snow transition at slightly higher temperatures, likely caused by neglected cases of freezing rain, among others. Due to these differences we obtain the highest accuracy and lowest bias when applying regression coefficients fitted to the OceanRAIN data set instead of using recommended coefficients from the literature such as those from Koistinen and Saltikoff (1998).

f. The new PP distinction algorithm 3P2D including D_{99} as essential physical information identified several cases that were erroneously classified as rain within the manual PP estimation. Large particle diameters indicate that the PP should be classified as snow or at least mixed-phase precipitation instead of rain.

g. Mixed-phase precipitation carries the largest uncertainty of the three PPs and is most challenging to detect for the new algorithm with a probability of detection of up to 0.25 using the predictor variable combination T_rH_D_{99} and 3P2D.

Even though the newly developed automatic PP distinction algorithm strongly depends on the currently still limited OceanRAIN data set, remarkable improvements are made. First, a PP probability is provided on a minute basis that limits the number of highly uncertain cases requiring visual inspection of atmospheric variables. The PP probability further allows error characterizing other precipitation data sets such as satellite data using OceanRAIN precipitation rates to unveil systematic errors with respect to PP. Second, the PPs of a few critical cases could be corrected that were falsely classified by the manual method. Third, we give evidence that the particle diameter of the falling precipitation particles contributes valuable information to the PP separation and by that in a physical way significantly improves the algorithm accuracy. Fourth, the new PP distinction algorithm considerably speeds up the data processing within OceanRAIN, which is an important step towards a fast-growing global surface precipitation data set for validating and evaluating other oceanic precipitation data sets.

Chapter 4

Simulating the influence of spatial-scale differences in precipitation between HOAPS and OceanRAIN using S-Pol radar data

4.1 Introduction

The spatial-scale difference between satellite-measured precipitation and surface-reference precipitation data challenges the validation of precipitation retrievals using satellite sensors (Michaelides et al., 2009). Whereas satellite sensors in the visible (VIS) to IR spectrum retrieve precipitation rather indirectly from cloud properties such as cloud top temperature, PMW satellite sensors represent a rather direct measurement of precipitation over the atmospheric column (Kidd and Levizzani, 2011). However, PMW sensors sense a particularly large volume compared to IR/VIS sensors that is projected onto the earth's surface known as their footprint size. The footprint sizes of PMW sensors typically reach at least a few tens of kilometers in diameter, which strongly challenges the PMW sensor validation using high-resolution surface-based precipitation reference data (Tapiador et al., 2012).

Reliable global precipitation estimates from PMW sensors require validation and error characterization against surface-based precipitation reference data. Over land, radars are frequently applied for validation and error characterization purposes of satellite data because they cover a rather large area at high spatial resolutions of about 1 km (Habib and Krajewski, 2002; Chandrasekar et al., 2008; Amitai et al., 2012). The high spatial resolution of a radar can be easily brought to a coarse-grained grid in order to adjust to a PMW satellite sensor's footprint. Radar-derived precipitation rates are calculated from measured reflectivity Z using a Z-R relationship that can be calibrated

using disdrometer-derived PSDs (Tokay and Short, 1996; Bringi et al., 2009) and/or gauge-derived precipitation rates (Ciach and Krajewski, 1999).

HOAPS represents an ocean-only PMW data set that fully relies on SSM/I and its successor SSMIS (Andersson et al., 2010b). However, over the ocean usually neither radar data nor gauge data are available as constant reference data at fixed locations, except for a few buoys or island-based stations (Maggioni et al., 2016). In that respect, OceanRAIN offers new high-quality ship-based precipitation estimates measured on-board ships using optical disdrometers (Klepp, 2015). However, OceanRAIN data lacks a high radar-like spatial coverage. Therefore, a key issue arises from the different spatial resolutions of the instruments used in HOAPS and OceanRAIN. While the SSM/I and SSMIS satellite sensors used in HOAPS provide areal precipitation estimates inferred from measured brightness temperatures, OceanRAIN provides point-wise precipitation estimates aligned along ship tracks. This limitation in spatial coverage is stressed by the fact that precipitation naturally occurs spatially intermittent, and strongly varies on small spatio-temporal scales. In this chapter, we aim to understand how the spatial-scale difference between HOAPS and OceanRAIN affects a validation of HOAPS. If possible, we aim to adjust OceanRAIN statistically to HOAPS in order to account for the different spatial representativeness.

The quantification of the spatial-scale difference between precipitation from HOAPS and OceanRAIN calls for an independent high-resolution oceanic data set because effects of spatial representativeness are indistinguishable from other instrument- or model-related differences. These differences can include but are not limited to effects from different sensor sensitivity, retrieval limitations or measurement principle. As an independent surface-based precipitation measurement, island-based radars combine the advantage of large coverage with a high spatial resolution. The Polarimetric S-band radar (S-Pol) used during the Rain In Cumulus clouds over the Ocean (RICO; Rauber et al., 2007) campaign at the subtropical Island of Barbuda in the subtropical North Atlantic sets an excellent framework to study spatial-scale effects of precipitation for three reasons. First, Barbuda Island is located in the subtropical North Atlantic Ocean offering open-ocean conditions in upstream (eastward) direction of the prevailing trade wind flow, and merely marginal influence on precipitation formation by other small Caribbean islands. Second, the S-Pol radar offers a particularly high spatial resolution of a few hundred meters in range. Third, the S-Pol could sample frequent light rain events during the wet subtropical winter season (Nuijens et al., 2009; Burdanowitz et al., 2015). The high spatial resolution of the S-Pol combined with its sampled areal precipitation rates serve well to simulate the comparison in precipitation between the satellite view of HOAPS and the ship-board in-situ view of OceanRAIN, conducted in Chapter 5.

Many studies exist that focus on the validation of satellite data with either point-like precipitation reference data such as gauges (e.g. Li and Shao, 2010) or area-like data such as radar (e.g. Kubota et al., 2009). However, the individual measurements in

OceanRAIN arrange along variable tracks that entirely depend on the ship's movement. Generally, ship tracks appear to better represent an area compared to a single point measurement due to the enhanced spatial coverage. Only few studies focused on a point-to-area comparison by means of collocating ship-board disdrometer data and satellite or re-analysis data (Klepp et al., 2010; Bumke et al., 2012). Bumke (2016) simulated gridded areal data and ship data of precipitation using a coastal radar in Northern Germany. This chapter aims to advance the knowledge of simple point-to-area validation to the spatial representation of precipitation whereby each section of results addresses one of the following questions:

1. Are ship tracks suited to detect rainfall within the typical area of a PMW satellite pixel? (both simulated with S-Pol radar data; Sect. 4.3.1)

2. Can ship tracks correctly predict the average rain rate of an area, and can potential biases associated with different spatial representativeness be corrected using the S-Pol radar data? (Sect. 4.3.2)

3. How does the random uncertainty contribute to the overall error associated with the spatial representativeness? (Sect. 4.3.3)

We discuss the influencing factor of ship speed and direction relative to that of clouds (Sect. 4.3.4). Finally, direct conclusions are drawn for a PMW satellite data validation with ship-board disdrometer data using the example of HOAPS and OceanRAIN in Section 4.4.

4.2 Data and Methods

The entire chapter is based on S-Pol radar data that was deployed on the subtropical island of Barbuda for the RICO campaign for 2 months during winter 2004/2005. Section 2.3 describes the RICO campaign and the S-Pol radar in detail and lists features of data processing and quality checking. The following Section 4.2.1 introduces the methods used in order to adjust the S-Pol radar to the requirements of simulating ship tracks and satellite pixels.

4.2.1 The simulation framework

The polar grid of the S-Pol consists of 984 range gates and 540 azimuthal increments that represent 0.66 ° sectors of a circle. The arc length of these azimuthal increments increases with distance from the radar location. This inconstant spatial resolution could distort a spatial-scale comparison. To homogenize the spatial resolution, we brought the S-Pol radar data onto a Cartesian grid with a spatial resolution of about

0.4 km. Compared to the size of a HOAPS satellite pixel of about 50 km x 50 km, 132 x 138 S-Pol radar pixels fill each simulated HOAPS pixel. In other words, one single radar pixel covers about 0.006 % of a simulated satellite pixel and, thus, can be considered as a point measurement in a satellite pixel. In the scanned area around the radar, we choose 4 boxes of the typical size of a HOAPS pixel despite its rather circular shape (Fig. 4.1). These 4 boxes fulfill the following three conditions. First,

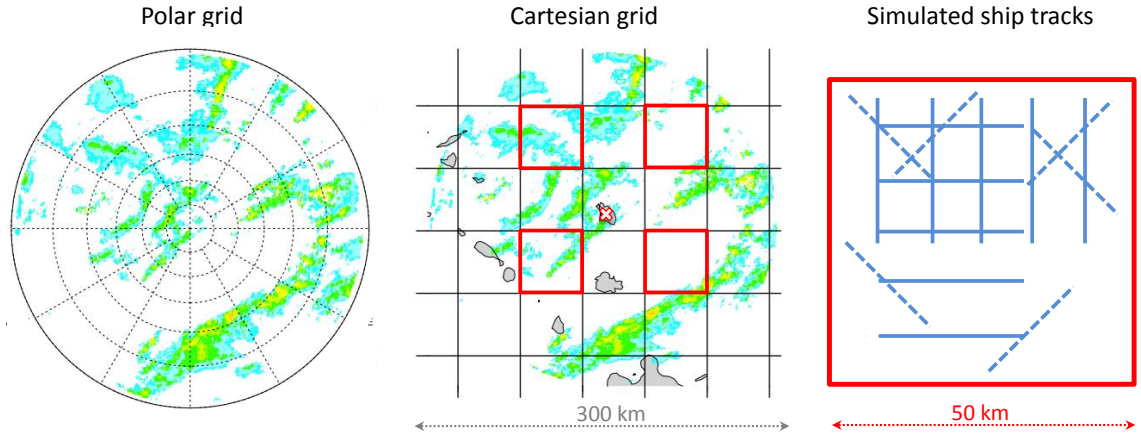


Figure 4.1: Overview of S-Pol radar data on original polar grid (left) and modified Cartesian grid (center). Red rectangles frame the 4 boxes chosen to simulate the satellite pixels. The enlarged red box (right) illustrates the 16 randomly chosen synthetic ship tracks each of 60 radar pixels in length (blue solid lines: 5 horizontal and 5 vertical; blue dashed lines: 6 diagonals) corresponding to a typical ship movement of 24 km h^{-1}.

they do not contain any islands to avoid land influences. Second, all boxes have about the same medium distance from the radar location. A not too large distance from the radar matters to reduce losses by radar beam overshooting above precipitating low clouds. Third, the 4 boxes are evenly distributed around the radar location to reduce influences by precipitating clouds that align with the prevailing easterly winds. For these reasons the location of the 4 boxes with respect to the radar suits well to study the effects of spatial-scale differences of precipitation observations over the subtropical ocean.

The whole RICO period spans 62 days (December 2004–January 2005) from which 3662 radar images are available from surveillance scans. For the 4 chosen boxes, this RICO sample results in overall 14 648 simulated satellite pixels. Into these simulated satellite pixels we place 16 arbitrary lines in order to simulate the ship tracks (Fig. 4.1 right). The 16 lines consist of 5 longitudinal, 5 latitudinal and 6 diagonal tracks with respect to the box. These simulated ship tracks within each of the 4 boxes cover the area of the box in a fairly homogeneous way. By this homogeneous orientation we rule out biases caused by prevailing cloud organization or cloud alignment within the box. Assuming that the ship moves at constantly high speed, we choose 24 km h^{-1} (approx. 13 kn) as a typical ship speed that corresponds to about 60 pixel in length or a distance

of 0.4 km per minute. This framework sets the ground for an independent statistical analysis that can serve as input for the HOAPS-OceanRAIN comparison.

Most of the RICO period was characterized by frequently occurring light rain events (Nuijens et al., 2009 and Burdanowitz et al., 2015). These conditions favor a comparison of precipitation along simulated ship tracks with areal satellite pixels because, first, the RICO period offers many suitable cases with observed precipitation. Second, the climate in the trades with its usually small-scale rain showers from shallow convective cumulus clouds challenges the realistic representation of precipitation by ship tracks within a satellite pixel. Thus, despite the limitation to 2 months, the collected S-Pol radar data sets a perfect framework to study the influence of the spatial-scale difference between line-oriented and areal estimates of oceanic precipitation under challenging meteorological conditions in the subtropics.

4.3 Results

The high spatio-temporal variability of precipitation complicates the detection and realistic representation for a series of precipitation measurements along a ship track. First, we investigate how well a ship track can detect precipitation in an area under the given conditions of RICO using the available S-Pol radar data (Sect. 4.3.1). Second, we analyze possible biases in observed rain rates from simulated ship tracks compared to simulated satellite-pixel rain rates that addresses the systematic error (Sect. 4.3.2). Third, the random error component is estimated (Sect. 4.3.3). Fourth, we discuss other factors influencing the spatial scale difference such as ship speed and orientation relative to clouds (Sect. 4.3.4). Finally, we aim at deriving a statistical adjustment for the OceanRAIN data with respect to the HOAPS satellite data in order to reduce the spatial-scale difference and quantify its uncertainty.

4.3.1 The detection of rainfall

A rainfall observation fulfills its purpose if the method above all proofs to reliably detect rain. The rain detection along a ship track mainly depends on two influencing factors.

First, raining cloud patterns associated with the underlying weather conditions set the distribution in which rainfall is spatially spread over a satellite pixel. The more homogeneously distributed over the area, the better an along-track observation can capture these rainfall patterns. The areal rainfall distribution measured by the S-Pol radar is given by the weather conditions during the RICO period that contained frequently occurring light rain showers from shallow cumulus clouds embedded in the trade-wind flow (Nuijens et al., 2014). On about 5 out of 60 days this typical trade-wind flow was disturbed by deeper convective systems that led to higher areal rain

rates sampled by the S-Pol radar (Burdanowitz et al., 2015). Overall, small showers of predominantly light rain from shallow trade cumulus clouds dominated the RICO period.

Second, the sampling of the measurements along the track determines how fine rainfall patterns can be resolved. Thus, the temporal sampling influences the along-track spatial resolution of measurements onboard a moving ship. A too coarse spatial resolution may result in rain events not being captured or dry patterns interpreted as rain, which distorts parameters such as the rain occurrence. In this simulation study, the sampling is defined by the S-Pol radar resolution and the simulation framework explained in Section 4.2.1. With a more than 100 times higher spatial resolution than the simulated "satellite pixel", the synthetic measurement sampling provided by the S-Pol radar meets the requirements set by the OceanRAIN-HOAPS comparison.

For the rain detection to check how often a simulated "ship track" misses rain with respect to a simulated "satellite pixel", the number of rainy "ship tracks" is compared to the number of rainy "satellite pixels". For both "ship tracks" and "satellite pixels", 'rainy' means that from the underlying radar pixels at least one holds a rain rate greater than 0 so that the average rain rate is also greater than 0. From now on, we refer to the simulated "satellite pixel" as "area" (A) and to the simulated "ship track" as "track" (T). Considering the 234 368 available cases (i.e., 3 662 radar images times 4 simulated "satellite pixels" times 16 simulated "ship tracks"), in less than 1 out of 5 cases the track and the area both detect rain while in more than 3 out of 5 cases the track misses the rainfall observed in the area (Tab. 4.1). The more than

Table 4.1: Contingency table lists relative occurrence (%) of rain-rate combinations for R_T and R_A from 234 368 available cases. In contrast to the left part of the table, the right part sets $R_A = 0$ if $C_A \leq 2\,\%$. Note that tracks are always located completely within the area so that $R_T = 0$ can only occur if $R_A = 0$ (no false alarms).

	All cases		$C_A > 2\,\%$	
	$R_T > 0$	$R_T = 0$	$R_T > 0$	$R_T = 0$
$R_A > 0$	17.3	61.2	14.2	17.6
$R_A = 0$	0	21.5	0	68.2

three times higher number of rainfall misses compared to rainfall hits points at a bad representation of areal rainfall by the tracks. However, the partial area covered with rain (area rain coverage: C_A) strongly influences the hit-miss ratio between track and area. Whereas a minimum area rain coverage of 0.01 % reveals a hit-miss ratio of 0.28, a minimum area rain coverage of 1 % leads to a hit-miss ratio of 0.6 (for $C_A > 10\,\%$: 2.4, Fig. 4.2). This relation indicates that the minimum area rain coverage scales with the hit-miss ratio of observable rain from tracks with respect to area. The hit fraction increases with increasing minimum area rain coverage because very small rain events tend to get excluded whereby these are most challenging to be captured on a track. An earlier study by Burdanowitz et al. (2015) revealed for the Northeast radar domain

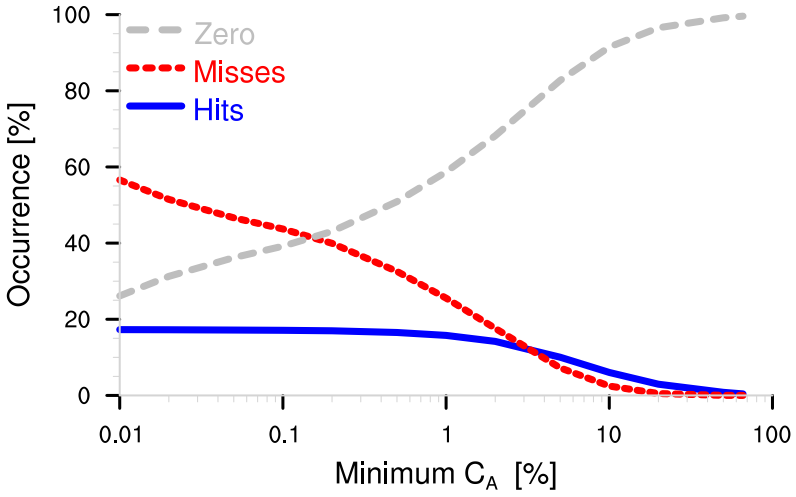


Figure 4.2: The frequency of occurrence (%) of hits (R_T>0 and R_A>0; blue solid line), misses (R_T=0 and R_A>0; red dotted line) and zero rain cases (R_T=0 and R_A=0; gray dashed line) for the S-Pol data as a function of the minimum area rain coverage C_A (%). The sum of hits, misses and zero rain always adds up to 100 % as in Table 4.1.

of the same S-Pol radar over Barbuda that the HOAPS scan-based data (HOAPS-S) can detect some of the rainfall in S-Pol match-ups with an area rain coverage between 1 and 2 %, whereas HOAPS certainly detects rainfall above 2 % area rain coverage by the S-Pol. Applying this threshold of $C_A > 2\,\%$ leaves 74 544 out of 184 016 rainy cases while the hit-miss ratio increases from 0.28 to 0.8 (Tab. 4.1 and Fig. 4.2). In 3 716 cases the along-track rain coverage C_T exceeds 2 % while the area rain coverage C_A does not exceed the minimum threshold of $C_A = 2\,\%$, in 5 930 cases vice versa (Fig 4.3). The regridded S-Pol radar resolution constrains C_T to a minimum of $\frac{1}{60}$ or about 1.6 %, determined by the number of 60 S-Pol radar pixels per track. As the actual area rain coverage C_A remains unknown in the HOAPS–OceanRAIN comparison (Chapter 5), other parameters need to be encountered, derived from the along-track OceanRAIN data.

Though C_T scales with C_A, more than half of all cases lie outside a range of $0.5 < \frac{C_T}{C_A} < 2$ (range enclosed by gray-dotted lines in Fig. 4.3). The variability of C_A with respect to C_T decreases with increasing rain coverage. For a low along-track rain coverage between 2 and 3 %, the possible area rain coverage values span almost the whole spectrum between 0.1 to 80 % area rain coverage if both catch rain. Thus, the along-track rain coverage holds the largest uncertainty to predict the area mean rain rate at low area rain coverage. The overall large variability in the ratio of C_A and C_T reflects the spatio-temporal variability of rainfall and disqualifies C_T as a suitable measure for C_A. Nevertheless, the area rain coverage strongly affects the ability of an along-track measurement to detect rain within an area.

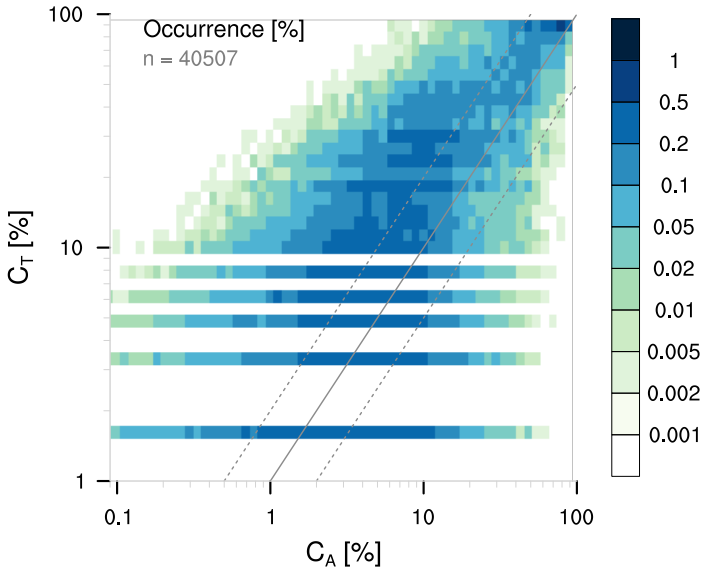


Figure 4.3: 2d-histogram with relative occurrence (%) of the along-track rain coverage C_T as a function of the area rain coverage C_A, both in %. Solid gray line denotes the 1-by-1 line, gray dotted lines the ratios of $\frac{C_T}{C_A} = 0.5$ and $\frac{C_T}{C_A} = 2$, respectively.

4.3.2 Representing rainfall: the systematic error

Once rainfall is detected within the area as well as along the track, the average rain rate might differ between both estimates. A difference in the average rain rates of track and area can be split into a systematic and a random error component. To quantify the systematic error component or bias, we concentrate on only those 40 538 (17 %) cases of all 234 368 cases where area and track contain at least one rainy S-Pol radar pixel (33 265 or 82 % of these rainy cases have an area rain coverage of at least 2 %). Due to mainly low absolute differences between the average along-track rain rate R_T and the average area rain rate R_A, the ratio $\frac{R_T}{R_A}$ serves to estimate how well R_T can represent R_A. R is composed of the product between the rain coverage C and the conditional rain rate D where D means the average rain rate of all rainy pixels ($R > 0$). Overall, $\frac{R_T}{R_A}$ strongly varies between more than 4 orders of magnitude (Fig. 4.4). The stronger the average along-track rain rate overestimates the average area rain rate ($\frac{R_T}{R_A} > 1$), the stronger a deviation in rain coverage contributes to that overestimation ($\frac{C_T}{C_A} > 1$), whereas $\frac{D_T}{D_A}$ rarely exceeds 5. However, most of these strongly differing cases ($\frac{R_T}{R_A} > 20$) vanish when applying the threshold for the minimum area rain coverage of $C_A = 2$ % (not shown). Setting a threshold for the area rain coverage does not only exclude cases in which tracks miss areal rainfall completely but also reduces the number of those cases in which tracks overestimate the rain area coverage with respect to C_A.

When average along-track rain rate and average area rain rate agree well ($\frac{R_T}{R_A} = 1$), the rain coverage is on average slightly overestimated along the track ($\frac{C_T}{C_A} > 1$), whereas the conditional rain rate is slightly underestimated along the track ($\frac{D_T}{D_A} < 1$), both compared

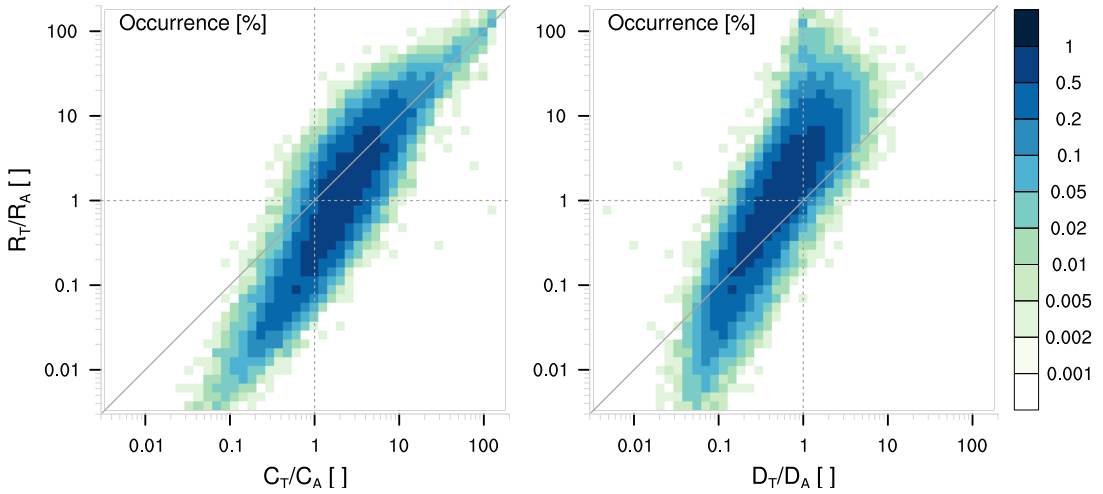


Figure 4.4: 2d-histogram with relative occurrence (%) of the average rain-rate ratio $\frac{R_T}{R_A}$ as a function of the rain coverage ratio $\frac{C_T}{C_A}$ (left), and the conditional rain-rate ratio $\frac{D_T}{D_A}$ (right). The solid gray line denotes the 1-by-1 line, gray dotted lines highlight the perfect ratio of 1 for $\frac{C_T}{C_A}$, $\frac{D_T}{D_A}$ and $\frac{R_T}{R_A}$, respectively.

to the area (Fig. 4.5). This tendency of the track to overestimate the area rain coverage while underestimating area rain intensity arises from, first, those not considered cases in which the track completely misses the rain observed in the area (Tab. 4.1 left: Misses). For low C_A, considering only hit cases artificially increases the probability that the track oversamples a rain event spatially. Second, the log-normal distribution of rain rates leads to low rain rates being sampled much more frequently than high rain rates. This non-Gaussian rain-rate distribution lets the track more likely undersample high rain rates observed in the area, resulting in a lower along-track conditional rain rate compared to the area.

For cases where the along-track rain rate slightly underestimates the area rain rate ($0.1<\frac{R_T}{R_A}<1$), the conditional rain-rate difference contributes stronger than the rain-coverage difference to the relative rain-rate difference $\frac{R_T}{R_A}$ (distance from 1-by-1 line in Fig. 4.4). For cases in which the along-track rain rate strongest underestimates the area rain rate ($\frac{R_T}{R_A}<0.03$), both the differences in area rain coverage and the conditional rain rate contribute about equally to the average rain rate difference. These cases of strongest R underestimation likely represent heterogeneously distributed rain showers in the area that are undersampled by the tracks. To better understand and account for such spatial-scale effects of rainfall heterogeneity we need to incorporate the spatial rainfall distribution more directly.

For advancing our knowledge of the rainfall variability within the area, we derive the individual rain event duration t_E that, in this study, actually represents a spatial length in radar pixel units. This length can be interpreted as an event duration when imagining a ship moving across the area underneath a raining cloud. This individual rain event duration as well as the number of identified rain events along a track N_E can be retrieved

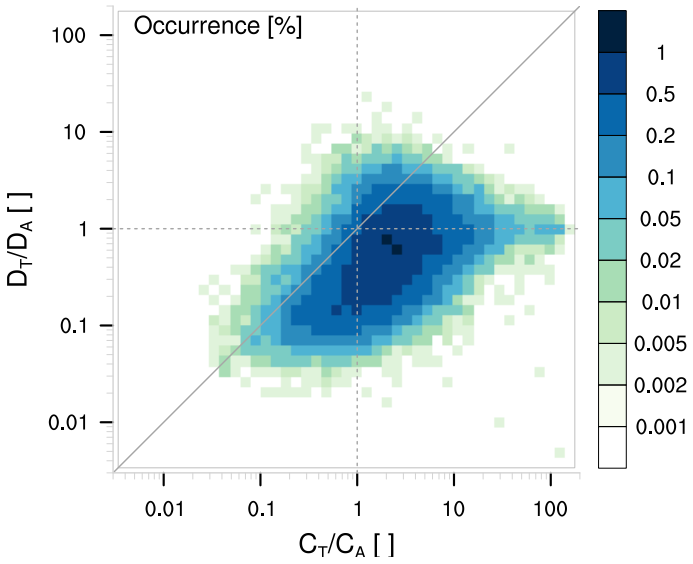


Figure 4.5: 2d-histogram with relative occurrence (%) of the conditional rain-rate ratio $\frac{D_T}{D_A}$ as a function of the area rain coverage ratio $\frac{C_T}{C_A}$. The solid gray line denotes the 1-by-1 line, gray dotted lines highlight the perfect ratio of 1 for $\frac{C_T}{C_A}$ and $\frac{D_T}{D_A}$, respectively.

from OceanRAIN for the HOAPS–OceanRAIN comparison. t_E contains additional information to a statistical variability parameter such as the standard deviation of the along-track rain rate. In specific, the standard deviation of the along-track rain rate does not account for the order in which the measurements/pixels are sampled but neglects that information. This means, the standard deviation could be equal for a case with one large rain shower and another case with several smaller rain showers for the same along-track rain coverage. In contrast, N_E and t_E carry part of the information of the rain distribution that is essential to predict the rainfall variability within the area.

Small rain showers dominate the rainfall distribution along the track of which one-pixel rain showers make up more than one fifths and small rain showers of 1 to 5 consecutive rainy pixels in length make up 60 % of all rain events (Fig. 4.6). Nevertheless, more than every hundredth rain shower reaches a size or duration of t_E>40 pixels along the track and, thus, covers at least two thirds of the track. In more than every second rain detection along the track, the track consists of only one single rain shower, whereas in 1 % more than 6 showers were detected along the track. These numbers indicate that small single rain showers dominate the rainfall occurrence along the track, which does not necessarily reproduce the area rainfall distribution as a whole.

For the goal to estimate whether or not the along-track rainfall well represents the areal rainfall, the individual rain event duration t_E alone does not suffice. However, dividing the sum of t_E by the total event number ($N_E = \sum n_E$) along the track reveals the more meaningful parameter named average rain event duration (T_E) that combines both the spatial rainfall structure along the track and the duration of each individual

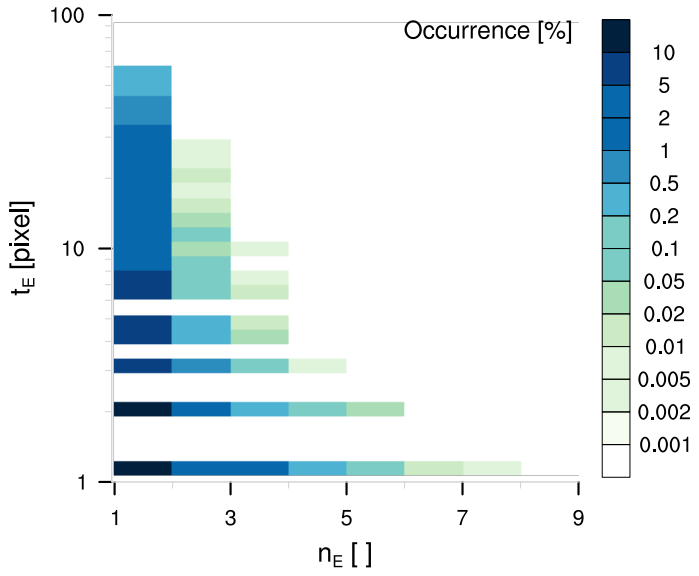


Figure 4.6: 2d-histogram with relative occurrence (%) of the individual rain event duration along the track (t_E) as a function of the number of identified rain events per duration class of t_E along the track (n_E) for $C_A > 2\,\%$.

rain event in a statistical way. The skill of T_E along the track to predict the average area rain rate R_A reaches its maximum around $3 < T_E < 6$ pixels where R_T equals on average R_A (Fig. 4.7, top). For the most frequently occurring cases of $T_E < 2$ pixels, the along-track rain rate R_T underestimates the area rain rate R_A on average by a factor of 3 to 6. In this range of low rain event duration or small-sized rain showers, the track on average likely underrepresents a rain shower in the area with respect to size. For T_E larger than 5 pixels, the track on average overestimates the rain rate in the area by up to a factor of 2. In these cases, the track overrepresents a rain shower compared to the area. Despite the variability in $\frac{R_A}{R_T}$, its dependence on T_E manifests in a rain rate that increases stronger along the track compared to the area with increasing T_E.

The difference in $\frac{R_A}{R_T}$ can be considered as a T_E-dependent bias that highlights the spatial-scale effects due to insufficient spatial sampling of the area by the track. The rain event duration can serve to remove this bias when fitting a bin-wise correction factor to yield $\frac{R_A}{R_T} = 1$ (Fig. 4.7, top left). The exponential least square fit of the form

$$\frac{R_A}{R_T} = 9.32 \cdot exp[-2.14 \cdot ln(T_E)] + 0.48, \tag{4.1}$$

reaches $R^2 = 0.98$ for the logarithmically binned mean bias of all 33 265 hit cases with $C_A > 2\,\%$. Bi-square weights help to minimize the influence of outliers. For T_E, we apply the natural logarithm in order to better fit the lower rain event duration that occurs most frequently (Fig. 4.6). Note that the fraction $\frac{R_A}{R_T}$ has been switched to fit an exponential curve. Multiplying R_T with the T_E-dependent correction factor from Eq. 4.1 reveals a bias-corrected along-track rain rate R_T^* (Fig. 4.7, bottom).

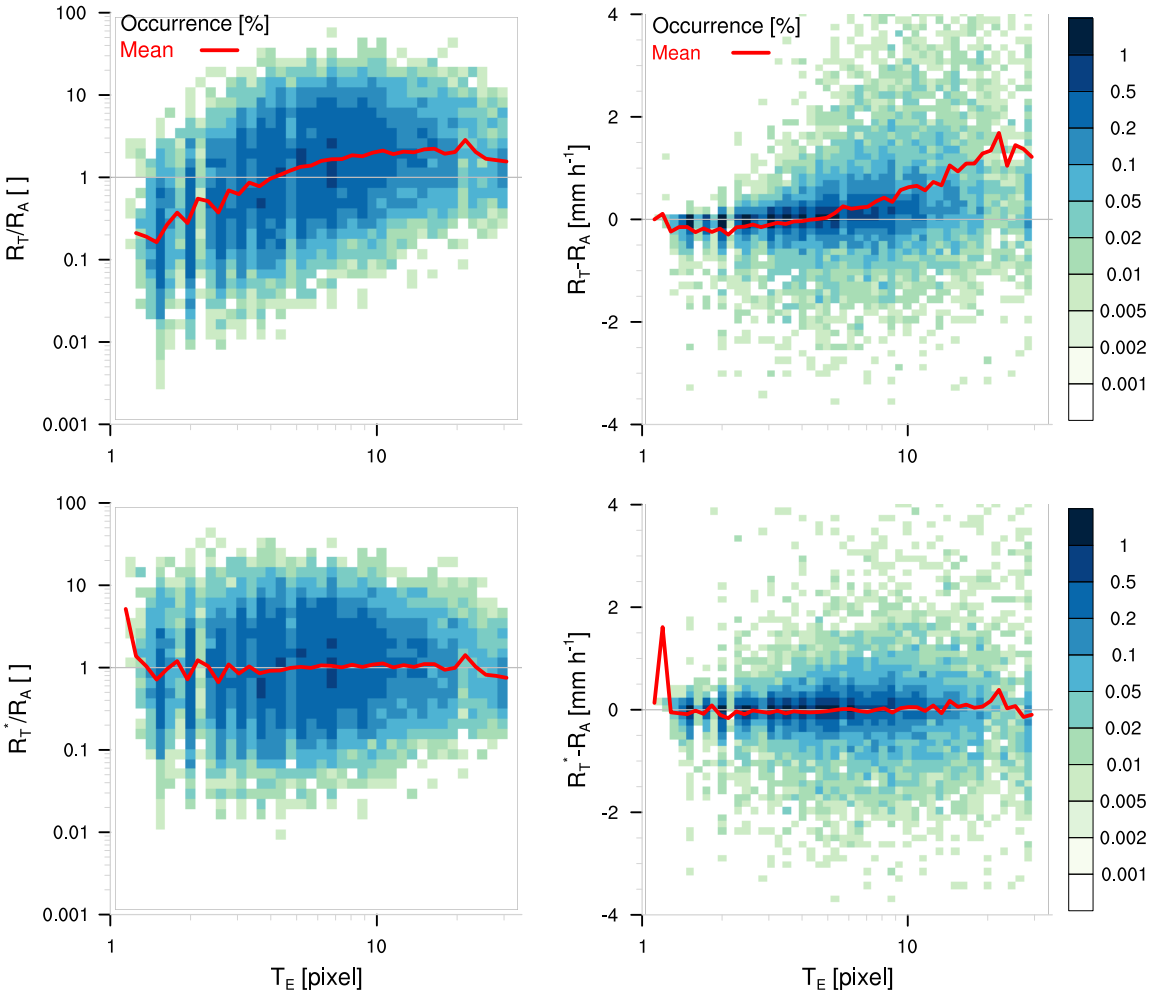


Figure 4.7: 2d-histogram with relative occurrence (%) of the rain-rate ratio $\frac{R_T}{R_A}$ (left) and the rain rate difference $R_T - R_A$ (right), both as a function of the along-track average rain event duration T_E for uncorrected R_T (top) and corrected R_T^* (bottom) for $C_A > 2\,\%$ using Eq. 4.1. Red lines mark the mean per bin, gray lines highlight where R_A equals R_T or R_T^*, respectively.

Compared to the average rain rate ratio $\frac{R_T}{R_A}$, the absolute rain rate difference between track and area $R_T - R_A$ provides a slightly different perspective on the scale-dependent rain rate. Whereas the underestimation of rain rates at low average along-track rain event duration lies below 0.3 mm h^{-1}, the overestimation at high average along-track event duration varies between 0.5 and 1.5 mm h^{-1} (Fig. 4.7, top right). Removing the bias with the same bin-wise correction factor from Eq. 4.1 as used for $\frac{R_T}{R_A}$ leaves a deviation not larger than 0.5 mm h^{-1} except for an outlier at $T_E \sim$1.2 pixel, likely due to insufficient sampling. Accordingly, the oversampling of rain events along the track compared to the area for $T_E > 10$ pixels has been corrected. The correction with respect to the average rain event duration along a track (Eq. 4.1) proofs to correct a substantial part of the spatial-scale bias due to limited spatial sampling along a track though the spread in absolute and relative rain-rate deviation from track to area stays

at a relatively high level.

The remaining high spread in rain-rate deviation between track and area might indicate that the T_E-correction (Eq. 4.1) could not remove the entire bias caused by the spatial-scale difference between track and area. This assumption is confirmed by the existing bias in $\frac{R_T^*}{R_A}$ with respect to R_T^* (Fig. 4.8, center left). However, the original bias in $\frac{R_T}{R_A}$ has been decreased applying Equation 4.1 while the overall spread of occurring along-track rain rates decreased compared to R_T before applying the T_E-correction (Fig. 4.8, top left). In particular, high R_T tend to overestimate R_A. In such cases, rain showers are overrepresented along the track while less rainfall occurs off the track elsewhere in the area. Similar to the T_E-correction, we aim to correct for that overrepresentation using an exponential fit with respect to $ln(R_T^*)$ of the form

$$\frac{R_A}{R_T^*} = 0.731 \cdot exp[-0.789 \cdot ln(\frac{R_T^*}{R_{T,50}^*})] + 0.306 \tag{4.2}$$

with a median value $R_{T,50}^*$ of 0.18 mm h^{-1}. Normalizing R_T with $R_{T,50}^*$ strongly reduces the dependence on absolute rain rates from the S-Pol. Otherwise, instrument-specific constraints such as the lower threshold for excluding Bragg scattering and clutter (Sect. 2.3.1) would put its mark on the corrected distribution of R_T^{**}. Obtaining a largely independent correction sets the ground to adjust the OceanRAIN data spatially to HOAPS satellite data.

The R_T^*-correction closely approximates the mean R_T^* bias of the 33 265 hit values ($R^2 = 0.99$) with a logarithmically spaced bin width b of $ln(b) = 0.29$. Applying this R_T^*-correction leaves almost exclusively along-track rain rates below 3 mm h^{-1}, which lies in the range of the maximum R_A of about 4 mm h^{-1} (Fig. 4.8, bottom). The R_T^*-correction removes most of the bias that remained after the T_E-correction.

Before the correction with Eq. 4.2, the track merely consisted of overestimated rain rates due to overrepresented rainfall along the track – a clearly statistical effect caused by the spatial-scale difference between area and track. However, the opposite effect of underrepresented rainfall along the track compared to the area remains uncorrected in R_T^{**} (cases of R_T^{**}–R_A<0 in Fig. 4.8, bottom right). This remaining part of the along-track rain-rate bias cannot be corrected in the same statistical way due to unavailable information of the area, such as C_A, D_A and R_A in the HOAPS–OceanRAIN comparison. For all hit cases, 8 % could be classified as underrepresenting rainfall on the track with R_T^{**}–R_A< -1 while 0.2 % overrepresent the rainfall with R_T^{**}–R_A> 1 compared to the area. These cases lead to an average underestimation of up to 0.5 mm h^{-1} by the track in parts of the spectrum (Fig. 4.8, bottom right), not reflected in the corrected relative rain-rate difference (bottom left).

The overall error can be quantified using the sum of squared errors (SSE). For the track, the area represents the reference, reflected in the equation

$$SSE = \sum_{i=1}^{33265} (R_{T,i} - R_{A,i})^2. \tag{4.3}$$

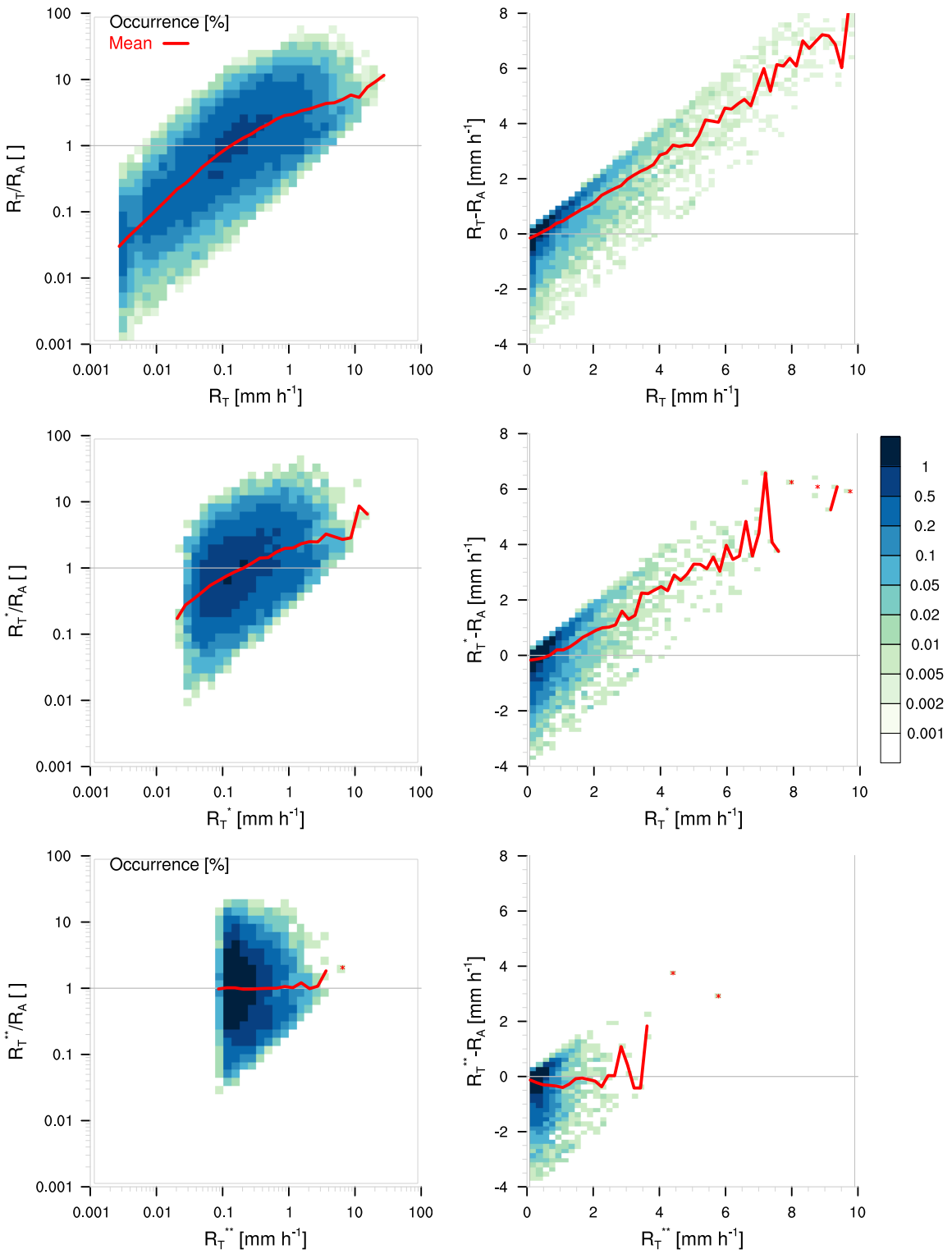


Figure 4.8: Same as Fig. 4.7 but as a function of R_T (top), R_T^* (center) and R_T^{**} (bottom). Note the linearly scaled axes on right.

SSE yields 37 675 mm^2 h^{-2} for the uncorrected R_T. The T_E-correction leads to a reduction of about 80 % to 7 678 mm^2 h^{-2} while both, T_E- and R_T^*-correction, reduce SSE by 86.7 % with respect to the uncorrected R_T.

4.3.3 Representing rainfall: the random error

The overall spread in the along-track rain rate R_T with respect to the area rain rate R_A represents the random error caused by the spatial-scale difference between track and area. However, this spread contains not only the random error but also the remaining systematic error that could not be corrected in a statistical way in Section 4.3.2. Thus, estimating the spread gives a good measure for the remaining overall uncertainty due to the spatial-scale difference between track and area.

The spread of the remaining overall uncertainty can be quantified using the standard deviation of the average rain-rate difference between track and area, $R_T - R_A$ (Fig. 4.9). The standard deviation normalized by R_T (i.e. the relative standard deviation, S_{rel})

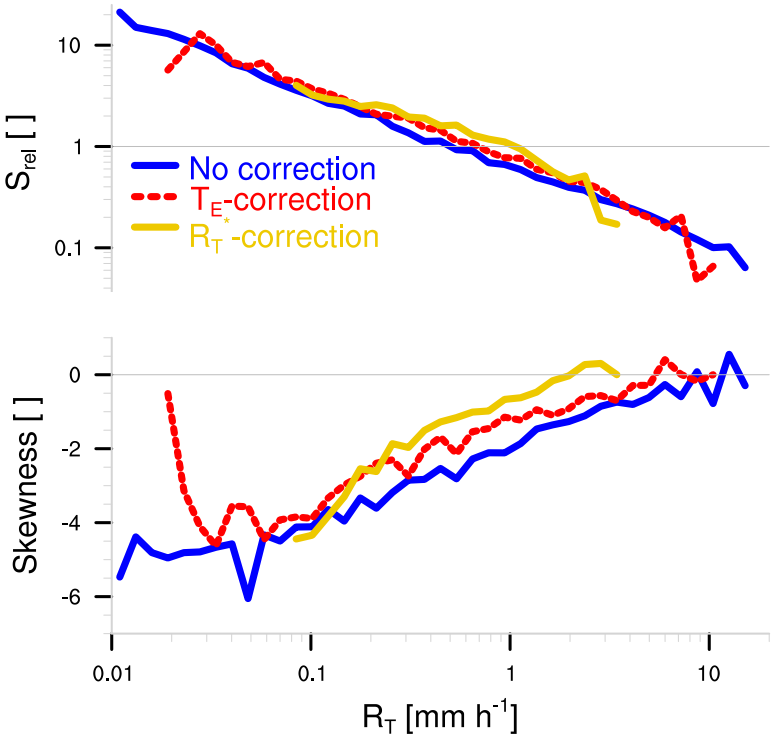


Figure 4.9: Standard deviation of the difference $R_T - R_A$ normalized by R_T (top) and the skewness of the difference $R_T - R_A$ (bottom) where negative values correspond to a left-skewed distribution ($R_T < R_A$), both as a function of the average along-track rain rate R_T for uncorrected R_T (blue), T_E-corrected R_T^* (red) and R_T^*-corrected R_T^{**} (yellow). The gray line indicates where the standard deviation of $R_T - R_A$ equals R_T (top) or where $R_T - R_A$ is normal-distributed (bottom), respectively.

decreases with increasing R_T. This means, the standard deviation of the difference

between R_T and R_A increases at a lower rate compared to the along-track rain rate R_T. For $R_T < 0.5$ mm h^{-1} the standard deviation exceeds R_T while it reaches up to 10 times lower values for $R_T > 10$ mm h^{-1}. However, at very low R_T of about 0.02 mm h^{-1}, the standard deviation exceeds the along-track rain rate by a factor of 10. In relative terms, low along-track rain rates hold a larger random uncertainty compared to high along-track rain rates. Overall, the standard deviation varies within one order of magnitude and does hardly exceed 1 mm h^{-1}. However, the random uncertainty holds relatively high values of 50 to 130 % of R_T for average rain rates between 0.3 and 1.3 mm h^{-1}, which emphasizes the large spread in rain rates measured along a track with respect to an area.

In this idealized study, the standard deviation of the difference between along-track and area rain rate characterizes the random uncertainty due to spatial-scale differences. This random uncertainty increases after applying the T_E- and R_T^*-adjustments exclusively for medium-intense rain rates between 0.3 and 1.3 mm h^{-1}. This means, the shrinking width in the rain-rate distribution coincides with an increase in standard deviation of up to 70 % for medium-intense rain rates. At the same time, the $R_T - R_A$ distribution becomes less left-skewed and, thus, more Gaussian (Fig. 4.9). In other words, the two corrections transform some of the systematic uncertainty to random uncertainty while the Gaussianity of the distribution increases particularly for higher rain rates (cf. Figs. 4.8 right).

4.3.4 Other interference factors

Most of the above mentioned results hold for conditions assumed to be constant in this idealized study. These assumed ideal conditions include that the ship moves on a straight track at constant speed of about 24 km h^{-1}. Any interference with wave water or other non-precipitation signals is ruled out here, but might clearly disturb the precipitation signal when measuring onboard of an actual ship. The influence of these erroneous precipitation particles cannot be considered in this idealized S-Pol study; however, we can consider a ship that moves at a lower or higher speed to study the influence on the representation of areal rainfall.

To realize a different ship speed in this idealized study we adapt the simulation framework as follows. Instead of lines with a constant length of 60 pixels, we change the length of the track running across the area. A length reduction corresponds to a decreased ship speed. For a halved ship speed the number of S-Pol radar pixels along that track needs to be incorporated two times in order to retain the same number of measurements. This procedure ensures a realistic representation of a ship that moves very slowly within a satellite pixel.

Under conditions of a slowly moving ship, we calculated the ability to detect rainfall for a ship speed reduced to about 12 km h^{-1} (Tab. 4.2) compared to 24 km h^{-1} as before (Tab. 4.1). In comparison, the slower moving ship holds a reduced percentage of hit

Table 4.2: Contingency table lists occurrence of rain detection (%) for track-averaged rain rates R_T and area-averaged rain rates R_A from 146 480 available cases from S-Pol radar with simulated shorter tracks of 30 pixels length representing a ship speed of about 12 km h^{-1}.

	All cases		$C_A > 2\,\%$	
	$R_T > 0$	$R_T = 0$	$R_T > 0$	$R_T = 0$
$R_A > 0$	10.69	67.82	9.11	22.69
$R_A = 0$	0	21.48	0	68.19

cases where both track and area detect rain from 17.3 to 10.7 % while the number of misses increases from 61 to about 68 %. Accordingly, the hit–miss ratio decreased from 0.28 to less than 0.16. This means, a slow-moving ship has a reduced probability of detecting rainfall in the area. More generally, the fraction of hits seems to scale almost linearly with the track length (Fig. 4.10). As a reason, a reduced track length loses the

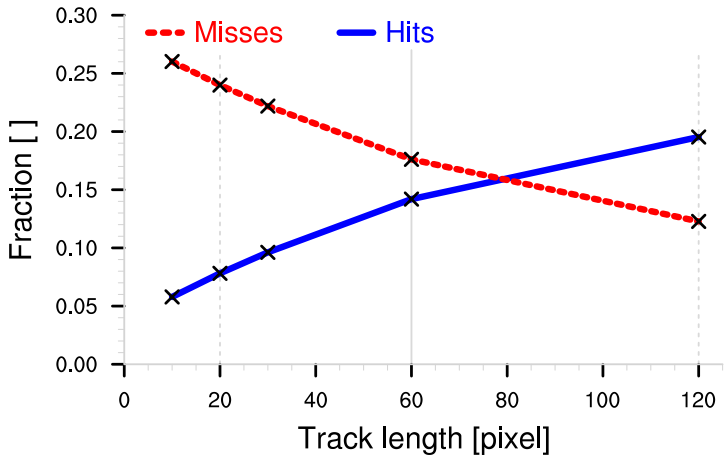


Figure 4.10: Hit (solid, blue line) and miss (dotted, red line) fraction as a function of track length (radar pixel) for all cases with a rain-covered area of $C_A > 2\,\%$. Solid gray line marks reference track length of 60 pixels (Tab. 4.1), whereas dotted-gray lines mark the track lengths displayed in Fig.4.11.

information content of the underlying spatial rainfall structure in the area that needs to be represented. The shorter a ship track becomes, the more the measurements transition from a line-like toward a point-like spatial coverage. A more point-like coverage increases the spread in the along-track rain rate R_T to predict the average area rain rate R_A for decreasing track lengths (Fig. 4.11). This increasing spread in R_T translates into a higher random uncertainty for the prediction of R_A. However, a reduced track length does not introduce a bias to R_T. Nevertheless, the skill to predict areal rainfall by means of along-track rainfall measurements onboard a ship decreases drastically with decreasing track length due to a decreasingly well represented spatial rainfall structure.

For simplicity, we neglected the cloud movement here, which also influences the areal

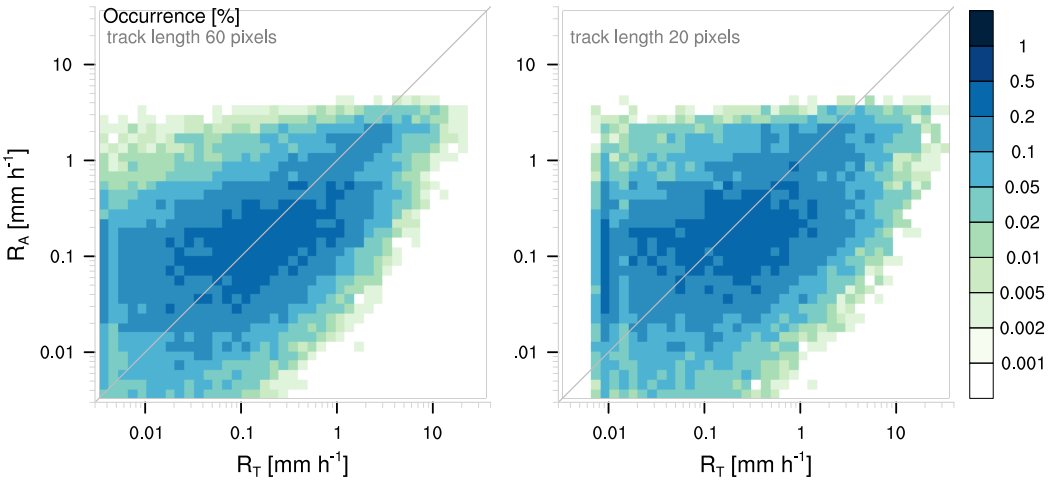


Figure 4.11: 2d-histogram with relative occurrence (%) of the area average rain rate R_A as a function of the along-track average rain rate R_T for the normal track length of 60 pixels (left) and for shorter tracks of 20 pixels (right). The 1-by-1 line is shown in gray.

rainfall representation (Fig. 4.12). The vectorial difference between the velocities of the moving ship and racking clouds determines how accurately a track can represent the spatial structure of rainfall. In other words, the longer the resulting velocity-difference vector from cloud and ship movement becomes the stronger do the atmospheric conditions observed by a satellite sensor put their mark on the resulting ship track.

The directions of ship and cloud movement play a key role for the velocity vector and thus the spatial rainfall representation. Ideally, ship and clouds move into opposite directions, which causes the maximum possible along-track coverage (case 1 in Fig. 4.12). In contrast, the superposition of clouds and ship that both move into similar directions at similar speeds would cause a point-like measurement coverage (case 3 in Fig. 4.12). This scenario would cause a bad areal rainfall representation by point-like measurements onboard a ship. For the estimation of areal rainfall, a point measurement represents the worst-case scenario to correctly capture the areal rainfall distribution.

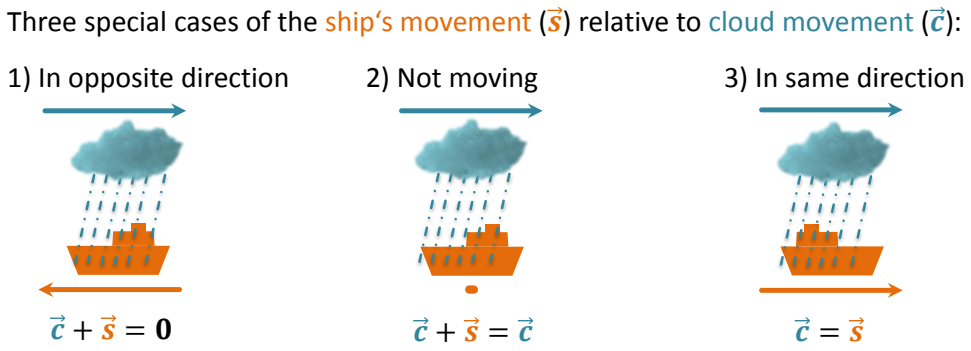


Figure 4.12: Schematic graph for the superposition of cloud velocity (blue arrow) and ship velocity (orange) for three special cases of orientation.

However, this case remains rather unlikely as not only speeds but also directions of clouds and ship need to agree with each other. For a resting ship (case 2), fast-moving clouds would cause a track-like coverage instead of a point-like coverage, which could still lead to a reasonable spatial representation of rainfall along the emerging track. Thus, the cloud movement holds an essential piece of additional information in order to assess the quality of how well a track can represent the spatial rainfall structure of an area.

4.4 Summary and concluding remarks

This chapter focused on quantifying and understanding the statistical error associated with spatial-scale differences between rainfall measured from space (satellites) and measured at the Earth's surface (ships). For this comparison we used S-Pol radar data from the RICO campaign in the subtropical Atlantic (winter 2004/05) in order to simulate rainfall measurements from satellite sensors versus track-like measurements onboard ships. In that respect, the simulated ship tracks have been tested for their ability to detect rainfall, correctly predict the rain rate and represent the spatial rainfall structure, all within a simulated satellite pixel. The following questions address these three topics, consecutively.

1. Are ship tracks suited to detect rainfall within the area of a typical passive microwave satellite pixel? (simulated with S-Pol radar data)

Our analysis suggests that a ship track can detect a part of the rainfall in an area. However, the detectability mainly depends on 3 factors. First, the area rain coverage of the rain events within the simulated satellite-pixel area sets the ground. Below an area rain coverage of 2 %, the ship track is more likely to miss rainfall in an area of the size of a PMW satellite pixel. Above 10 % area rain coverage, the ship track is 2.5 times more likely to detect rainfall within the area than to miss it. We recommend a minimum area rain coverage of 2 % for a HOAPS-OceanRAIN comparison, not only because a minimum hit–miss ratio of about 1 seems reasonable but also because HOAPS is unlikely to detect precipitation below a threshold of 2 % in area rain coverage.
Second, the ability of a ship track to capture rainfall depends on the spatial rainfall distribution in the area. A stretched cluster of rainfall poses a higher challenge to be detected compared to homogeneously distributed small showers or a widespread rainfall area.
Third, the ship speed relative to the cloud movement influences the ability to capture rainfall in the area. The slower the ship moves relative to the clouds, the shorter the relative track becomes and, thus, more unlikely it captures rainfall occurring in the area.

2. Can ship tracks correctly predict the average rain rate of an area, and can potential biases associated with different spatial representativeness be corrected using the S-Pol radar data?

As for the detectability of rainfall in an area, correctly predicting the area average rain rate depends on the same three factors: area rain coverage, homogeneity of spatial rainfall distribution and length of ship tracks relative to the cloud movement.
The along-track average rain rate consists of the product of area rain coverage and conditional rain rate. A strong overestimation of the area average rain rate by the track is driven by the strongly overestimated along-track rain coverage. In contrast, a strong underestimation of the area average rain rate along the track originates to equal portions from an underestimation in both along-track area rain coverage and conditional rain rate. However, for average along-track rain rates that well predict the average area rain rate, we find that the along-track rain coverage slightly overestimates the area rain coverage while the along-track conditional rain rate slightly underestimates the area conditional rain rate. This tendency of predicting slightly larger rain events of weaker intensity indicates a bias introduced by the different spatial-scale representation of rainfall between track and area.
Correcting for that bias requires a parameter that well scales with the bias. Whereas area rain coverage, conditional rain rate and rain event number are not suited, the average rain event duration well separates the bias. Using the average rain event duration, the representativeness error can be quantified and statistically corrected. Short along-track rain events more frequently underestimate the average area rain rate, whereas longer along-track rain events tend to slightly overestimate the average area rain rate. This purely statistical effect that can be removed by multiplying with a correction factor dependent on the average rain event duration. The correction decreases the systematic error introduced by the spatial-scale difference between ship track and satellite pixel, mainly due to partly detected rain events.
A remaining bias becomes visible when considering the rain-rate fraction between track and area with respect to the average along-track rain rate. As for the correction using the average along-track event duration, low along-track rain rates are increased while high along-track rain rates are decreased in order to statistically reduce the spatial-scale difference. This second bias correction mainly corrects for a strong rain-rate overestimation due to overrepresented rain events along the track. However, the second bias correction leaves a bias for underrepresented rain events along the track compared to the area. This remaining bias cannot be corrected in the same statistical way with the available parameters from the track.

3. How does the random uncertainty contribute to the overall error associated with the spatial representativeness?

The standard deviation of the average rain rate difference between simulated ship track and satellite pixel area represents a good measure to quantify the random uncertainty of precipitation caused by different spatial-scale data sets such as OceanRAIN and HOAPS. The change in standard deviation from uncorrected to the bias-corrected along-track rain rates experiences the strongest increase by about 70 % for rain rates of about 0.8 mm h^{-1} while below 0.2 and above 1.5 mm h^{-1} no noticeable changes appear. The increase in standard deviation for medium-intense rain rates implies that the corrections partly shift the systematic error into the random error. Overall, the standard deviation decreases in relative terms with increasing along-track rain rate. Above rain rates of about 1 mm h^{-1}, the standard deviation remains below the average along-track rain rate while the distribution of the track-area difference becomes almost entirely Gaussian. The more the rain-rate difference distribution approximates a Gaussian distribution the lower the remaining bias, revealing a noticeable improvement only for rain rates above 0.2 mm h^{-1}.

The results of this idealized study are derived from 2 months of radar data of a single radar located in the subtropical trade wind regime and, therefore, might change for other regions such as mid- and high latitudes. Similar studies should be undertaken in order to confirm the bias correction and analyze the random uncertainty with respect to climate regime and other influencing factors. Nevertheless, the results obtained from S-Pol radar are valuable for the HOAPS-OceanRAIN comparison. In addition, the quantified bias might already represent a large range of climate conditions due to varying rainfall patterns sampled throughout the RICO period. In particular, the prevailing climate in the subtropical North Atlantic fosters rainfall to occur from small shallow cumulus clouds, which are challenging to sample spatially along ship tracks.

The quantified statistical errors caused by spatial-scale differences can serve as valuable input for a satellite validation using surface-based observation of precipitation as ground truth. In specific, the estimated systematic and random uncertainty help to quantify errors in the precipitation parameter of the HOAPS satellite climatology using OceanRAIN ship data over the global oceans. Assuming the statistical bias due to spatial-scale differences between HOAPS and OceanRAIN remains constant with latitude, the correction derived from the S-Pol radar can be applied to remove the systematic error that complicates the direct comparison between track-like and area-like data sets. Additionally, the random error component derived from the S-Pol radar provides a first estimate of the remaining uncertainty due to spatial-scale differences between HOAPS and OceanRAIN. Further, the S-Pol radar study sets the ground to advance our knowledge of uncertainties associated with track-like precipitation observations that go beyond the traditional point-to-area problem in global precipitation observations.

Chapter 5

Validation of HOAPS precipitation estimates using OceanRAIN data

5.1 Introduction

Precipitation estimates from satellite data are crucial to monitor global precipitation and validate climate models. However, satellite data usage requires validation with surface-based reference data to quantify the uncertainty related to instrument characteristics, representativeness associated with different underlying areas and the instrument sensitivity, among others. These diverse sources of uncertainty need to be quantified and considered when it comes to model validation or long-term trend analysis. If possible, this overall uncertainty should be attributed to the specific sources of error in order to help understand the reliability and suitability of a precipitation data set as reference data for a certain purpose.

HOAPS represents such a long-term precipitation satellite data set over the global ice-free ocean for a time period of almost 30 years (Andersson et al., 2010b; Fennig et al., 2012). HOAPS performs well when compared to other satellite data sets in the subtropics (Burdanowitz et al., 2015) or against reanalysis data sets (Andersson et al., 2011). However, a rigorous error characterization exists only for the near-surface specific humidity parameter in HOAPS using a multiple triple collocation approach (Kinzel et al., 2016). The precipitation parameter so far lacks uncertainty estimates because systematic surface-based precipitation reference data has not been available over the global oceans. Since 2010, this gap is being filled by OceanRAIN that uses high-quality optical disdrometers onboard RVs (Klepp, 2015). OceanRAIN recently reached a reasonable size of quality-checked oceanic precipitation data that allows to start a systematic comparison with satellite data, such as HOAPS.

This chapter aims at investigating the uncertainty of HOAPS precipitation rates using OceanRAIN as precipitation reference database. The comparison involves three main aspects that need to be considered. First, even if adequately matched in time

and space, HOAPS and OceanRAIN do not necessarily estimate the precipitation rate for the exact same areal scene at the exact same time, condensed in the term "representativeness error". Previous studies discuss the representativeness effect between point-like ship measurements and areal HOAPS satellite measurements without an attempt of reducing it (Bumke et al., 2012 and 2016). In this thesis, the statistical adjustments are employed that were derived from the S-Pol radar study in Chapter 4. Second, OceanRAIN precipitation estimates hold an uncertainty due to potentially misclassified or abruptly changing PP, wind distortion at the RV's superstructure or measurement artifacts. Third, uncertainties might strongly be related to climate zones with distinct meteorological features such as the fraction of convective precipitation, PP, precipitation occurrence and the precipitation-rate spectrum. All these aspects need to be addressed as far as possible in order to validate the HOAPS precipitation parameter while discussing reasons for uncertainties.

5.2 Methods and data

For a most meaningful comparison between HOAPS satellite precipitation data and OceanRAIN surface-based precipitation data we use the longest available overlap of both data sets with a focus on high data quality and continuity. Section 2.1 describes the OceanRAIN database in detail as well as the optical disdrometer as its core instrument. Section 2.2 introduces the HOAPS satellite climatology and provides information on products and algorithm features. For a case-study, the Sections 2.4, 2.5 and 2.6 further introduce the satellite-based data sets CMORPH, IMERG and MSG-CPP. The applied methods to adjust OceanRAIN for a fair validation of HOAPS are given piece by piece in the following section of results.

5.3 Results

For a direct and fair comparison of HOAPS satellite data with OceanRAIN as precipitation reference data, both data sets ideally should match in most of their features. These features include but are not limited to temporal and spatial sampling as well as resolution, instrument sensitivity, measurement principle and retrieval. Due to the fact that HOAPS strongly differs from OceanRAIN in many of these aspects, a truly direct comparison of the exact same scene from both products seems virtually impossible to achieve. However, a series of methods can be applied to adjust OceanRAIN to HOAPS independently of HOAPS. These methods include a space–time matching (Sect. 5.3.1), a statistical adjustment of the spatial representation (Sect. 5.3.2) and an adaptation of the instrument sensitivity (Sect. 5.3.5).

5.3.1 Space–time matching of OceanRAIN and HOAPS

HOAPS and OceanRAIN strongly differ in their spatial sampling and their coverage. HOAPS provides quasi-global precipitation estimates over the ice-free ocean with a revisit time per DMSP satellite ranging between a few hours at the poles and a couple of days at the equator. In contrast, the OceanRAIN spatial sampling depends entirely on the RV's position whereby the routes strongly differ among the different RVs. The different spatial sampling can be addressed by a matching in space and time. The space–time matching represents the most essential requirement for a meaningful comparison of both data sets. Matching OceanRAIN and HOAPS both spatially and temporally is called collocation and ensures a minimum spatial distance as well as a minimum time lag between both precipitation measurements.

The number of valid collocations strongly depends on the chosen collocation boundaries (Fig. 5.1: green and purple). These boundaries mean the tolerance in time and

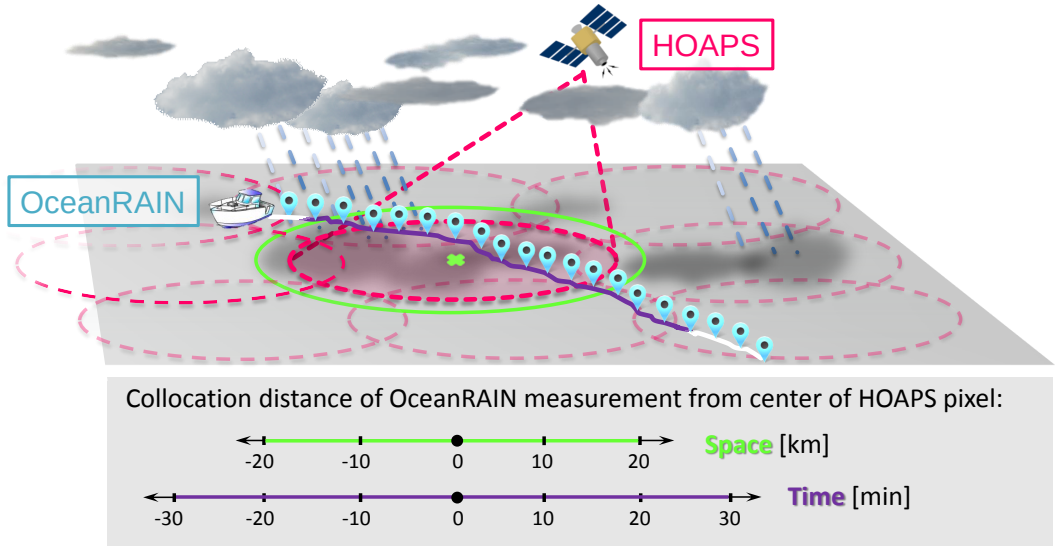


Figure 5.1: Schematic drawing illustrates the collocation procedure and the collocation boundaries set to ± 20 km in space (green) and ± 30 min in time (purple) as well as the track-averaging (blue markers) and HOAPS-S pixels (red-dashed circles).

space at which an observation of OceanRAIN is considered to be a valid collocation to an observation of HOAPS. The tolerance of the highly variable and intermittent precipitation parameter lies relatively low because precipitation events might change dramatically even within a few minutes in time or kilometers in space. Nevertheless, we need to choose relatively generous collocation boundaries in order to obtain a sufficient number of collocations for reasonable statistics. As a compromise of tight collocation boundaries and to obtain a decent number of collocations, for now, we allow a maximum distance of ± 20 km and a maximum time lag of ± 30 min, both measured from the center of a chosen HOAPS-S pixel.

The chosen collocation boundaries result in a maximum of 60 OceanRAIN measurements that can be collocated to one and the same HOAPS-S pixel according to the allowed 60 min time window. This maximum number illustrates that the boundary in time stronger limits the number of collocations compared to the boundary in space because of a typical ship speed far below 60 km h^{-1}, usually about 20 km h^{-1}. However, each of the 60 or less collocated OceanRAIN measurements along a ship's track represents a small sub-area of the 50 km by 50 km HOAPS-S pixel. To obtain a more similar spatial representation of the HOAPS-S pixel area we average over all collocated OceanRAIN measurements along the track that were collocated to the same single HOAPS-S pixel (cyan markers in Fig. 5.1). The track-averaging leads to a possible minimum precipitation rate of $\frac{1}{60} \cdot 0.01$ mm h^{-1} for OceanRAIN, which lies more than 3 orders of magnitude below the minimum precipitation rate of HOAPS, a discrepancy discussed later in Section 5.3.5.

The chosen collocation boundaries leave 24 990 collocations after OceanRAIN along-track averaging, displayed in Figure 5.2. From these number of all collocations, 3.5 % (885) represent scenes in which both OceanRAIN and HOAPS detect precipitation (hits), whereas in 14.2 % (3 552) HOAPS does not detect the precipitation detected by at least one of the along-track OceanRAIN measurements (misses). The opposite case that OceanRAIN did not detect precipitation that was detected by HOAPS appears in 2.8 % (692; false detection) while 79.5 % (19 861) remain as correct zeros. The percentage of hits, misses and false detections varies with latitude and season (Tab. 5.1).

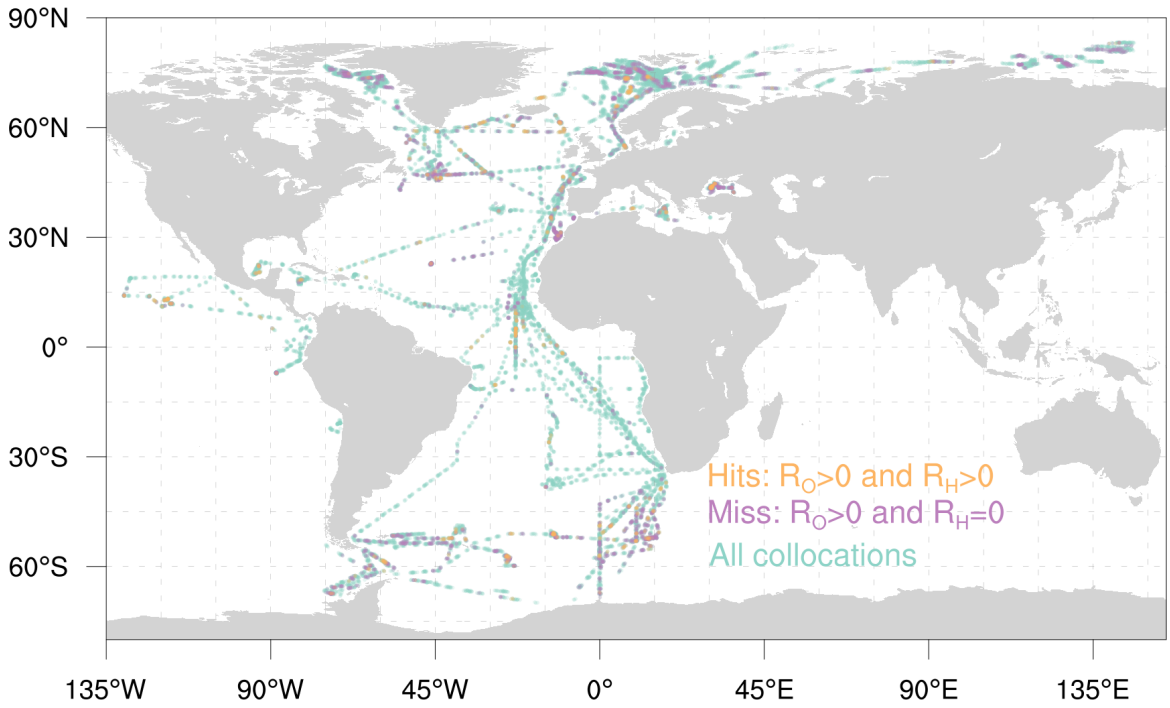


Figure 5.2: Map of all 24 990 HOAPS–OceanRAIN collocations from available long-term RVs in OceanRAIN from 06/2010–12/2015 (cyan dots), misses ($R_H = 0$ and $R_O > 0$; purple) and hits ($R_H > 0$ and $R_O > 0$; orange) using ±20 km and ±30 min as collocation boundaries.

Table 5.1: Contingency tables for all 24 990 collocations in % for the collocation boundaries of ±20 km and ±30 min per 20 ° latitude band. PP occurrence of rain, snow and mixed-phase precipitation (mix) given in % after the automatic algorithm presented in Chapter 3. n denotes total number of collocations per 20 ° latitude band.

Latitude	Occurrence [%]			PP occurrence [%]			n
		$R_H > 0$	$R_H = 0$	Rain	Snow	Mix	
70-90° N	$R_O > 0$	1.1	14	50.8	27.5	21.7	7 229
	$R_O = 0$	2.2	82.7				
50-70° N	$R_O > 0$	7.4	14.9	90.7	0	9.3	2 287
	$R_O = 0$	1.9	75.8				
30-50° N	$R_O > 0$	4.3	20.8	92.8	1.3	5.9	3 552
	$R_O = 0$	1.7	73.2				
10-30° N	$R_O > 0$	3.0	7.4	100	0	0	3 235
	$R_O = 0$	3.0	86.6				
10° N-10° S	$R_O > 0$	4.8	4.1	100	0	0	2 247
	$R_O = 0$	6.5	84.6				
10-30° S	$R_O > 0$	1.4	3.0	100	0	0	1 519
	$R_O = 0$	1.3	94.3				
30-50° S	$R_O > 0$	3.8	13.8	90.7	0	9.3	1 640
	$R_O = 0$	1.3	81.1				
50-70° S	$R_O > 0$	6.1	26.2	45.7	24.9	29.4	3 281
	$R_O = 0$	4.3	63.4				

Considering 20° latitude bands, the mid-latitudes (50° – 70°) followed by the equatorial band between 10° N and 10° S reach the highest percentage of hits ranging between about 5 and 7 %. Bumke et al. (2016) find somewhat higher hit percentages in the mid-latitudes and inner tropics, explained by higher collocation boundaries of 55 km and 45 min for ship tracks in the Atlantic for 2005–2008. In the subtropics, hit cases reach substantially lower percentages of 1.4 % in the southern and 3 % in the northern hemisphere, respectively. The difference between both hemispheres results not only from the prevailing climatic conditions but in particular from the locations of the sampled ship tracks. Whereas between 10° and 30° N, RV *Sonne* cruised through parts of the warm Pacific Ocean with occasional precipitation events, between 10° and 30° S, RV *Polarstern* went through the relatively cold and very dry southeast Atlantic Ocean. These spatial sampling differences within OceanRAIN influence the collocation statistics and need to be considered when drawing conclusions to the long-term precipitation mean.

The different sampling frequency between certain regions is accompanied by the different seasonal sampling. A different seasonal sampling mainly influences the PP. The twice-as-high fraction of snow and mixed-phase precipitation between 50° and 70° S compared to 50° and 70° N mainly reflects the exceptional overwintering period of RV

Polarstern in the southern hemisphere in August 2013. It should be noted that potentially more collocations from that southern-hemisphere austral winter stay would have been available but were flagged "missing" because HOAPS provides no data over the ice-covered ocean. Overall, snow or mixed-phase precipitation make about 50 % of the precipitation detected northward of 70° N and southward of 50° N while in-between their percentage does not exceed 10 %. The PP distribution strongly influences the precipitation rate because snowfall rates tend to lie below rainfall rates because larger particles more strongly contribute to the precipitation rate.

The contingency table from OceanRAIN and HOAPS per latitude band (Tab. 5.1) serves to derive how well HOAPS performs in detecting precipitation compared to OceanRAIN. The detectability of precipitation can be assessed calculating skill scores from hits, misses, false detections and correct zeros, as recommended by the International Precipitation Working Group (IPWG) and the WWRP/WGNE Joint Working Group on Forecast Verification Research (Tab. 5.2). These skill scores include accuracy, false alarm ratio, probability of detection, bias score, odds ratio and Heidke skill score, among others. The accuracy as a measure of overall agreement reaches highest values of more than 0.9 in the inner tropics, mainly caused by the high fraction of correct zeros exceeding 84 % (Fig. 5.3). The probability of detection better reflects the actual detected precipitation relative to missed precipitation. Exclusively in the equatorial band, HOAPS correctly detects precipitation more often than to miss it (Tab. 5.1).

Table 5.2: List of skill scores calculated from hits (h), misses (m), false detections (f), correct zeros (z) and the total number of collocations (n). More details and additional skill scores are listed on http://www.cawcr.gov.au/projects/verification.

Name	Calculation	Perfect score
Accuracy (ACC)	$\frac{h+z}{h+m+f+z}$	1
Probability of detection (POD)	$\frac{h}{h+m}$	1
False alarm ratio (FAR)	$\frac{f}{f+h}$	0
Bias score (BS)	$\frac{h+f}{h+m}$	1
Odds ratio (OR)	$\frac{h*z}{f*m}$	∞
Heidke skill score (HSS)	$\frac{h+z-h_r}{n-h_r}$ with $h_r = \frac{(h+m)(h+f)+(z+m)(z+f)}{n}$	1

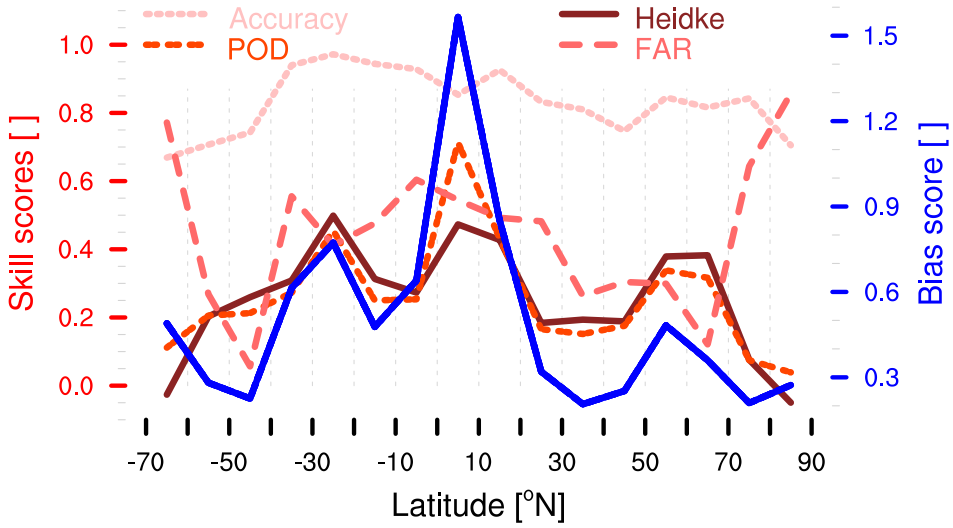


Figure 5.3: Skill scores for precipitation detection (explained in Tab. 5.2) in reddish colors (left axis) and bias score in blue color (right) of all HOAPS–OceanRAIN collocations as a function of 10° latitude bands.

In the northern-hemisphere mid-latitudes, HOAPS missed twice-as-much precipitation events while towards the poles four times more precipitation events are missed compared to the correctly detected precipitation events. In general, the lowest values coincide with regions in which more light precipitation or solid precipitation contribute to total precipitation that both are more challenging to detect for PMW sensors (Fig. 5.4). In reverse, the HOAPS false alarm ratio reaches highest values towards the poles as well as at the equator where more precipitation events are falsely predicted compared to those that were correctly predicted with respect to OceanRAIN. Lowest false alarm ratios occur in the mid-latitudes where neither solid precipitation nor light precipitation dominate (Tab. 5.1 and Fig. 5.4). Bumke et al. (2016) find more than a factor of 2 larger false detection percentages compared to this study. Whether false detections represent actual "falsely detected" precipitation is discussed in Section 5.3.5.

The bias score directly relates the missed precipitation to the falsely detected precipitation. Only at the equator, the falsely detected precipitation slightly exceeds the missed precipitation while in the extra-tropics the missed precipitation strongly exceeds the falsely detected precipitation. In addition to falsely detected or missed precipitation, the Heidke skill score relates the quality of precipitation predictions to that of random chance (Fig. 5.3: solid red line). The shape of the Heidke skill score resembles that of the probability of detection but its maximum at the equator does not strongly but slightly exceed the second maximum in the northern-hemisphere mid-latitudes of about 0.4. Towards the poles, the values decline to a minimum below 0.2. Globally, Heidke skill score values are positive and, thus, indicate a better skill of HOAPS to detect precipitation compared to random chance. Especially in the inner tropics and the northern-hemisphere mid-latitudes, HOAPS performs well in detecting precipitation.

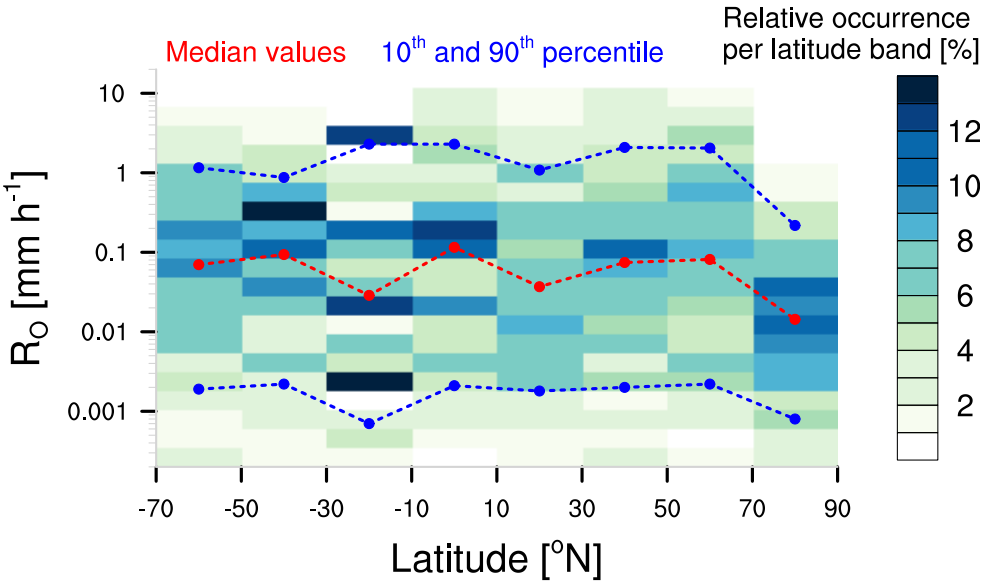


Figure 5.4: Relative occurrence in % for $R_O > 0$ mm h^{-1} (hits and misses: n = 4 437) as a function of 20° latitude bands. Relative occurrences per latitude band were normalized so that the sum per latitude yields 100 %. Markers indicate median (red) as well as 10th and 90th percentile (blue) per latitude band.

The initially set collocation boundaries of ±20 km and ±30 min were chosen in accordance with other studies that collocate satellite with ship-based data. Over the Baltic Sea, Bumke et al. (2012) estimated 17 km and 27 min as decorrelation lengths, which means the distance to which the correlation decreased to $\frac{1}{e}$. From these decorrelation lengths, Bumke et al. (2012) choose ±25 km and ±30 min for collocating precipitation derived from the same optical disdrometer onboard RV *Alkor* for the period 1995–1997. More specifically, Bumke and Seltmann (2012) estimate decorrelation lengths of 46–68 km for frontal and stratiform precipitation and about 18–46 km for convective precipitation. For snowfall in the Norwegian Sea, Klepp et al. (2010) applied more generous collocations boundaries of ±55 km and ±45 min with an accuracy of 0.99. The collocation boundaries of ±20 km and ±30 min seem to establish a global compromise between convective precipitation and more stratiform precipitation.

Despite the reasonable collocation boundaries, the question remains whether stricter boundaries in time and space would lead to a significantly higher relative amount of correctly detected precipitation by HOAPS. For space, we select ±15 and ±18 km as additional maximum distances of an OceanRAIN measurement from the center of a collocated HOAPS pixel. For time, we choose the more restrictive time lags of ±20 and ±25 min. Together with the initially set collocation boundaries of ±20 km and ±30 min, we obtain 9 possible combinations of spatial and temporal boundaries for collocation. For these 9 collocation boundaries, we consider the probability of detection, accuracy, odds ratio and the Heidke skill score to investigate whether stricter collocation boundaries cause significantly improved conditions to compare HOAPS to OceanRAIN

(for details see Tab. 5.2). All these skill scores were normalized by the maximum of all collocation boundaries per latitude band to neglect the effect of different climate regimes so that the maximum and perfect score per latitude band always equals 1 (Fig. 5.5). As a measure of robustness, the skill scores have also been calculated for 1 000 realizations of randomly chosen 50 % of all collocations. The distribution of these values for the sub data sets serves to estimate the uncertainty of a calculated skill score (illustrated as box-whisker in Fig. 5.5e for the accuracy). A skill score of a collocation boundary needs to lie outside the range of 2.5th and 97.5th percentile (whiskers) of the determined internal uncertainty (for $\alpha = 0.05$ shown as red boxes in Fig. 5.5a-d) to be statistically significant. As a secondary indicator for a weaker robustness, $\alpha = 0.5$ indicates when the skill score lies outside the box of the inter-quartile range (difference between 25 % and 75 % quantiles) and, thus, differs from at least half of the realizations (gray boxes in Fig. 5.5a-d). This means that improvements of non-framed fields are to be neglected. Thus, considering the internal uncertainty in the collocations constitutes an important tool to better rule out improved skill scores simply due to random chance.

The probability of detection indicates that the fraction of hits from all precipitation events detected by OceanRAIN varies by 15 to 20 % in the southern-hemisphere mid-latitudes and the inner tropics while staying about constant elsewhere. Especially in the inner tropics the lowest collocation boundaries of 15 km/20 min lead to the highest probability of detection while in the mid and high latitudes except for the northern polar regions more generous collocation boundaries are sufficient. However, the non-significant ($\alpha = 0.05$) improvements by stricter collocation boundaries do not yet justify an application of stricter collocation boundaries.

The accuracy experiences a markedly lower relative variability per latitude band compared to the the other skill scores, not exceeding 10 %. The main reason for the lower variability lies in the contribution of non-precipitating cases that make about 80 % of all cases and are not used in the calculation of the other skill scores. The non-precipitating cases induce a markedly lower variability per latitude, which causes an improvement for choosing stricter collocation boundaries to be significantly better everywhere but in the mid-latitudes. Particularly the polar regions and the northern-hemisphere subtropics indicate a statistically significant improvement for using 15 km/20 min as collocation boundaries, whereas the inner tropics and the southern-hemisphere mid-latitudes suggest to use 18 km/20 min and 18 km/25 min. The northern-hemisphere mid-latitudes and southern-hemisphere subtropics, however, do not justify using stricter boundary conditions, which rules out using different collocation boundaries on the global scale.

The odds ratio indicates the ratio of the probability that HOAPS correctly predicts precipitation compared to the probability that HOAPS incorrectly predicts precipitation. The best HOAPS performance in the polar regions as well as in the inner tropics occurs again for strictest collocation boundaries, though not significant. In contrast, in the northern-hemisphere mid-latitudes 20 km/30 min leads to the best performance. The Heidke skill score leads to a very similar distribution, but as the

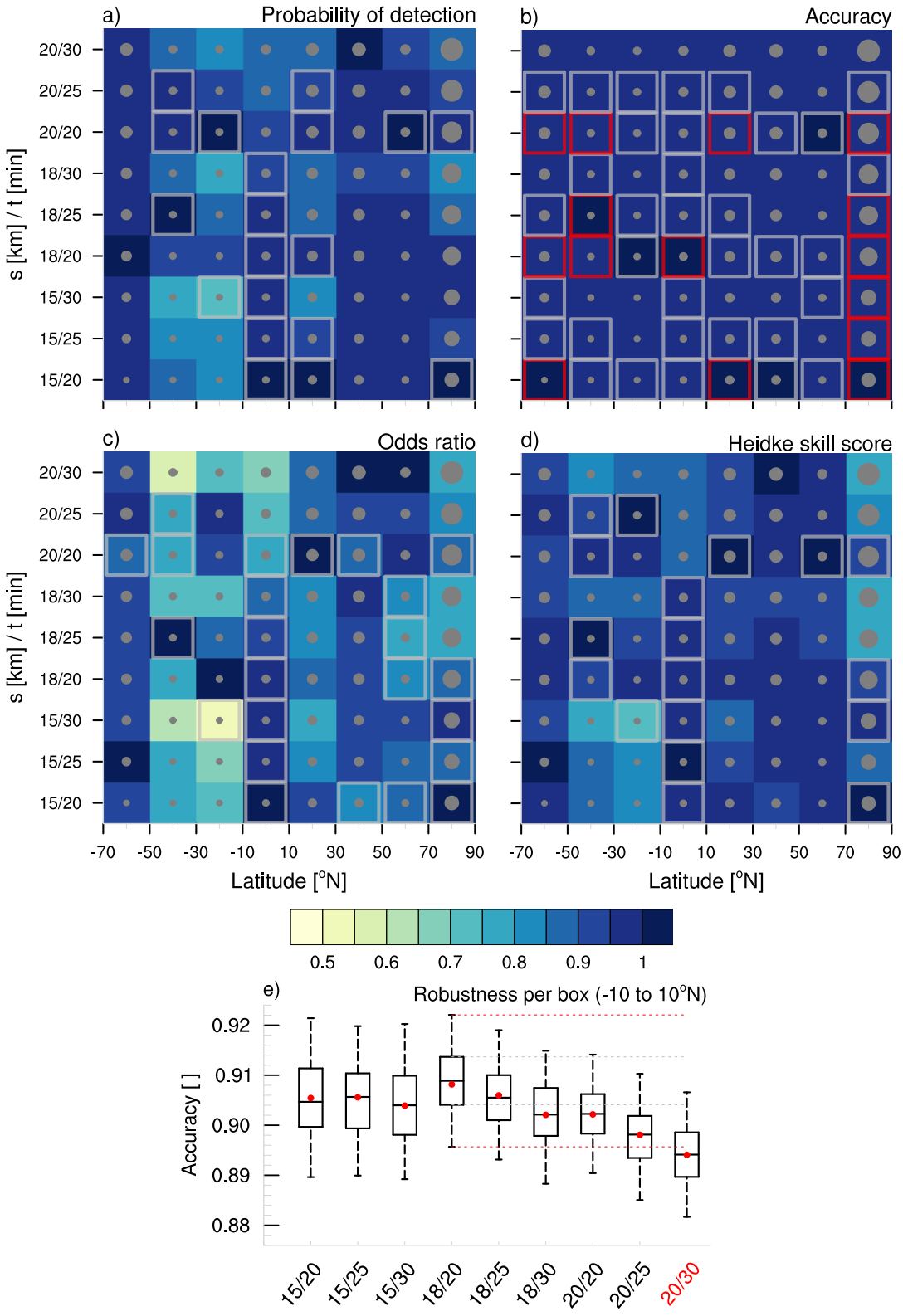


Figure 5.5: (a-d) Relative change with respect to maximum of skill scores for combinations of space/time collocation boundaries (format: km/min) as a function of 20° latitude bands. Size of gray dots in box denotes number of collocations relative to 7 229. Red (gray) boxes indicate a significant difference of $\alpha = 0.05$ (0.5) of a skill score to the reference 20km/30min with respect to internal uncertainty estimated from 1 000 realizations of a randomly chosen halved sub data set. (e) Graph exemplifies significance levels for accuracy at −10 to 10°N where boxes denote 25th and 75th percentile and whiskers 2.5th and 97.5th percentile.

odds ratio does not reveal significantly better collocation boundaries. However, to most extent the considered skill scores tend to recommend stricter collocation boundaries predominantly for the two climate extremes: inner tropics and polar regions. To robustly confirm these results, more HOAPS–OceanRAIN collocations are required to obtain statistically significant results. The tendency that particularly lower maximum distances of 15 and 18 km as collocation boundaries lead to an improved performance of HOAPS might be explained by the higher convective precipitation fraction in the tropics (Houze, 1997) and the tendency towards cellular cloud structures resulting from cold air advection over a warmer ocean in the polar regions (Klepp et al., 2010). The extent of convective precipitation events usually remains below that of stratiform precipitation. The inner tropics exhibit a particularly high degree of convective organization; though mid-latitudes reflect a similarly high occurrence of high R_O compared to the inner tropics (Fig. 5.4). These results tend to confirm earlier studies that find markedly lower correlation lengths for convective precipitation compared to stratiform precipitation (Bumke and Seltmann, 2012). A common definition of the term convective precipitation means precipitation formed in areas of vigorous upward air motion ($\omega > 1$ m s^{-1}) opposing to stratiform precipitation in areas of weaker upward or even downward motion ($\omega < 1$ m s^{-1}) after Atlas et al. (2000).

The collocation of HOAPS with OceanRAIN represents a major first step for a meaningful comparison of both data sets. In this first step both data sets were matched in time and space. However, despite the collocation they still represent different areas and the instruments of the underlying measurements are distinctly sensitive in resolving precipitation. These crucial differences between HOAPS and OceanRAIN are approached in the following sections.

5.3.2 Adjusting the spatial representativeness of OceanRAIN

The optical disdrometer used in OceanRAIN senses a substantially smaller area compared to the SSMIS used in HOAPS, which leads to a different spatial representativeness in both data sets. Whereas the OceanRAIN retrieval actively derives the precipitation rate from the measured precipitation PSD in a fairly small volume of less than 1 m^3, the SSMIS passively senses a fairly huge atmospheric column of about 50 km in diameter. Even after the along-track averaging of collocated OceanRAIN measurements in a HOAPS pixel, a largely different spatial representation remains that needs to be encountered when comparing HOAPS to OceanRAIN.

When precipitation along the OceanRAIN ship tracks is compared to the areal precipitation estimates from HOAPS, a considerable amount of cases from OceanRAIN inevitably over- or underrepresents precipitation events occurring in the HOAPS pixel (Fig. 5.6, top). From the S-Pol study we infer that precipitation events of short average along-track event duration T_E on average underestimate the precipitation rate in the area, whereas precipitation events of relatively high T_E on average overestimate

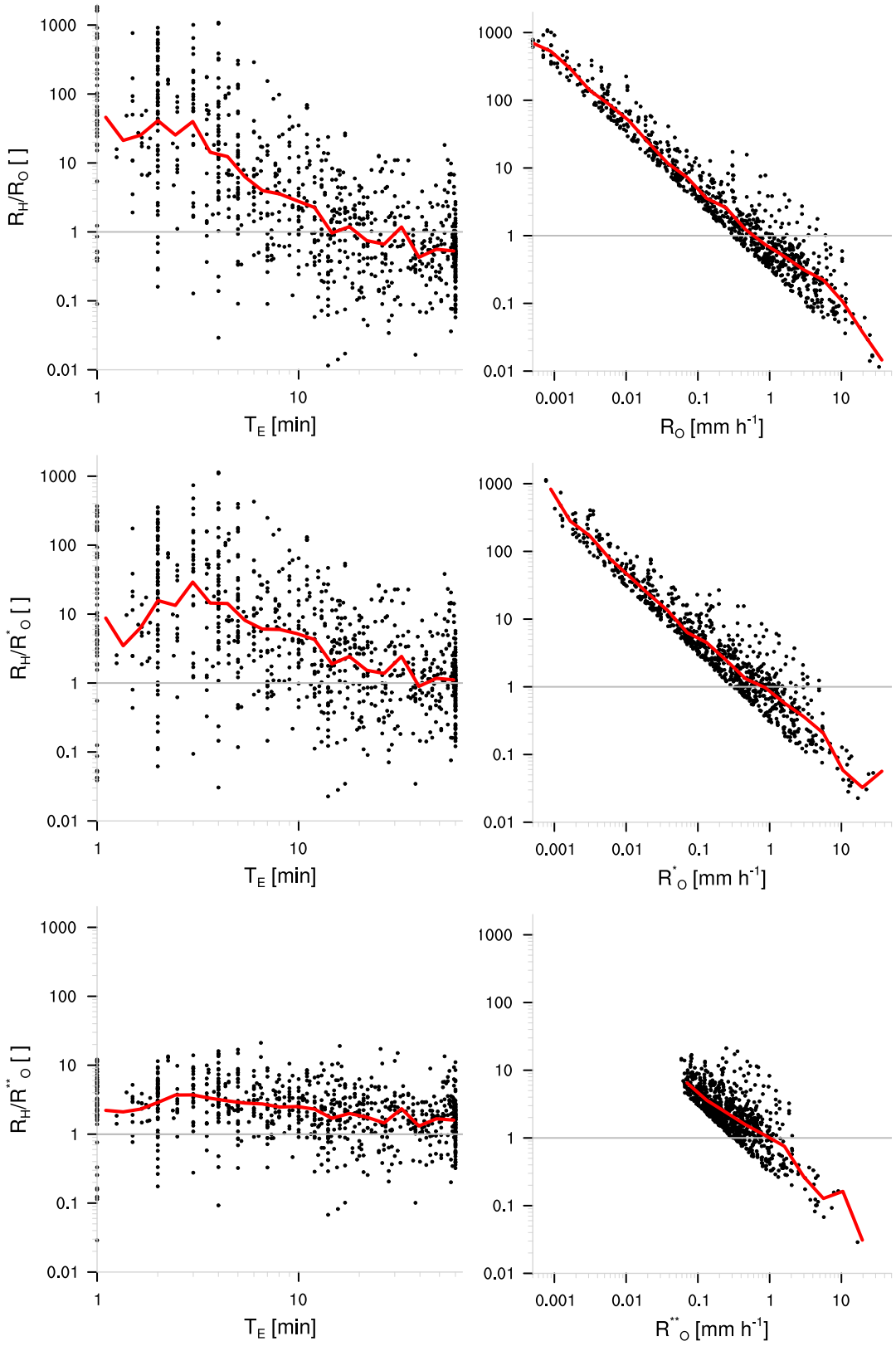


Figure 5.6: HOAPS–OceanRAIN precipitation-rate ratio $\frac{R_O}{R_H}$ as a function of the mean along-track precipitation event duration T_E (left) and the along-track average precipitation rate (right) as unadjusted R_O (top), T_E-adjusted R_O^* (center) and R_O^*-adjusted R_O^{**} (bottom), respectively. Solid red line indicates bin-wise mean.

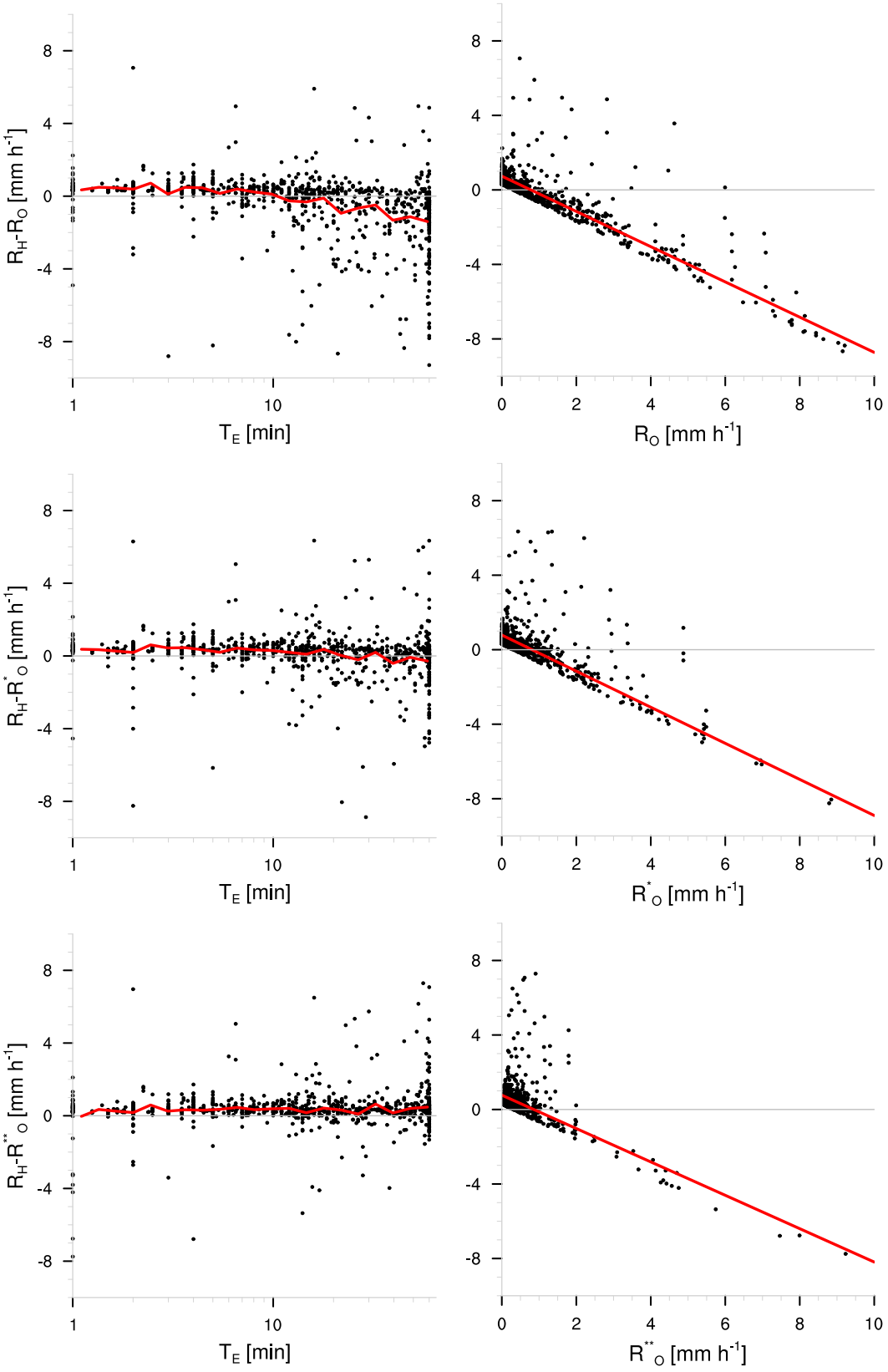


Figure 5.7: As Fig. 5.6 but for the HOAPS–OceanRAIN absolute precipitation-rate difference $R_H - R_O$.

the precipitation rate in the area, R_H here (compare with Figs. 4.7 and 4.8). HOAPS and OceanRAIN confirm this spatial effect where for short precipitation events of 1 to 3 min duration R_H exceeds R_O by a factor of 20 to 45 while for events of more than 20 min duration on average R_H remains under R_O. Note that the absolute difference of R_H and R_O for relatively long events with almost -1 mm h^{-1} stronger exceeds the absolute difference of relatively short events with about 0.5 mm h^{-1} (Fig. 5.7, top panel). The different spatial-scale representation of precipitation by OceanRAIN found in both relative and absolute differences of R_H and R_O cannot be corrected explicitly because the actual precipitation distribution in the area remains unknown. However, the spatial representativeness studied for the S-Pol radar in Chapter 4 allows to adjust OceanRAIN at least statistically to HOAPS.

The statistical adjustments derived from the S-Pol use the average duration per precipitation event T_E (Eq. 4.1; T_E-adjustment) and the adjusted precipitation rate of the track, R_O^* here (Eq. 4.2; R_O^*-adjustment). Note that T_E does not refer to the life cycle of an individual precipitation event but to the time it takes a ship to move through a shower, which is highly dependent on the actual track. Adjusting OceanRAIN according to T_E decreases the overestimation factor of $\frac{R_H}{R_O}$ from up to 45 toward factors ranging between 4 and 30 for $1 < T_E < 3$ min, whereas the occurrence of the maximum overestimation shifts from 1 to 3 min. For $T_E > 20$ min, the relative difference in R_H and R_O^* remains below 2.5 while the absolute difference strongly decreases leaving a difference of about 0.2 mm h^{-1} for highest T_E. The decreased HOAPS–OceanRAIN difference marks a first important step in adjusting the spatial representation of precipitation in OceanRAIN to that of HOAPS.

T_E certainly depends on direction and speed of the RV relative to the cloud movement as discussed in Section 4.3.4. Statistically, these influences introduce an additional uncertainty to the spatial-scale difference of HOAPS and OceanRAIN that is worth to be quantified in future studies. Here, we exclusively concentrate on the spatial representation of precipitation as measured along a ship track compared to a satellite pixel by neglecting the cloud movement. This limitation should not introduce a bias as OceanRAIN sampled a sufficiently long period of more than 5.5 years; nevertheless, it adds to the random uncertainty.

As for the S-Pol, the T_E-adjustment leaves large relative differences for $\frac{R_H}{R_O^*}$ ranging from 0.02 to about 1 000 (Fig. 5.6, center right). These more than 4 orders of magnitude of relative precipitation-rate difference indicate that OceanRAIN cannot yet statistically represent the same area as HOAPS. To achieve a better spatial representation of OceanRAIN we apply the second spatial adjustment derived from S-Pol using R_O^* (Eq. 4.2). The exponential fit from the S-Pol is transferable to OceanRAIN because the S-Pol along-track rain rate $R^*(T)$ has been normalized by its median value $R^*_{T,50}$. Without this normalization, the adjustment derived from the S-Pol would lead to a biased adjustment of OceanRAIN precipitation rates that were no longer independent from the S-Pol instrument features. The R_O^*-adjustment strongly reduces the relative

difference to values hardly exceeding $0.1 < \frac{R_H}{R_O} < 20$ (Fig. 5.6, bottom). As a second effect, the R_O^*-adjustment almost entirely removes the trend in the OceanRAIN–HOAPS difference with respect to T_E. However, on average there remains a difference between HOAPS and OceanRAIN after which R_H exceeds R_O^{**} by a factor of 2 to 3 or 0.3 mm h^{-1} (Fig. 5.7, bottom left). The on average larger precipitation rate of HOAPS compared to OceanRAIN reflects the adjustment of most of the spatial-scale difference between both data sets.

As expected, a statistical adjustment cannot fully account for the effect of spatial-scale differences between HOAPS and OceanRAIN for every single HOAPS–OceanRAIN collocation. Figure 5.8 provides a better overview of each individual collocation from the 885 hit cases before and after applying the statistical adjustments to OceanRAIN. Before the statistical adjustments, R_O spans a 3 orders of magnitude larger range compared to R_H, mainly related to the higher resolution and sensitivity of the disdrometer and the along-track averaging of OceanRAIN after collocation. Applying the statistical adjustments to OceanRAIN mainly increases R_O for cases with $R_O < 0.1$ while very high R_O are slightly decreased. Nevertheless, cases of strong disagreement such as $R_H \sim 0.1$ mm h^{-1} and $R_O \sim 10$ mm h^{-1} as well as vice versa experience marginal effects from the statistical adjustments and, thus, cause low Pearson correlation coefficients of 0.35 (logarithmic) and 0.11 (linear) after adjustments, respectively. The expectedly higher value for the linear regression in logarithmic space arises because precipitation follows a log-normal distribution that a simple linear regression cannot well approximate in linear space (Tian et al., 2013). Nevertheless, the overall correlation remains low because many hit cases disagree in the direct comparison of R_O^{**} and R_H. This means that on average, R_H does not well fit R_O^{**}. As a reason, the statistical adjustments leave some "extreme cases" in which the ship might have strongly undersampled precipitation in the HOAPS pixel domain ($R_H \gg R_O$) or in which the ship might have strongly oversampled precipitation in the HOAPS pixel domain ($R_H \ll R_O$). However, sampling solely cannot explain the relatively high number of strongly disagreeing cases.

Understanding the reasons for strongly disagreeing precipitation rates as depicted in Figure 5.8 is key to conduct a meaningful analysis and derive suggestions to improve the HOAPS data set. Thus, we investigate a number of factors that possibly could enhance or even induce differences between R_H and R_O^{**}.

First, a slow-moving ship decreases the track length over which precipitation rates are averaged in OceanRAIN. The decrease in track length causes a decrease in spatial representativeness (cf. Fig. 4.11). Short tracks due to a slow-moving ship (defined as $u < 5$ km h^{-1}) occur in 23 % of all collocations and in 31 % of all hit cases and, thus, cannot be neglected (Fig. 5.9d). With increasing tendency of HOAPS to exceed OceanRAIN precipitation rates ($\frac{R_H}{R_O^{**}} > 1$), the fraction of slow-moving ship cases starts to increase at a relative difference of 4 and almost doubles at 9 (0.57). Although this increase in low-speed fraction does not directly imply that a slow-moving ship leads to a lower R_O^{**} relative to R_H, the probability for that increases. For the opposite

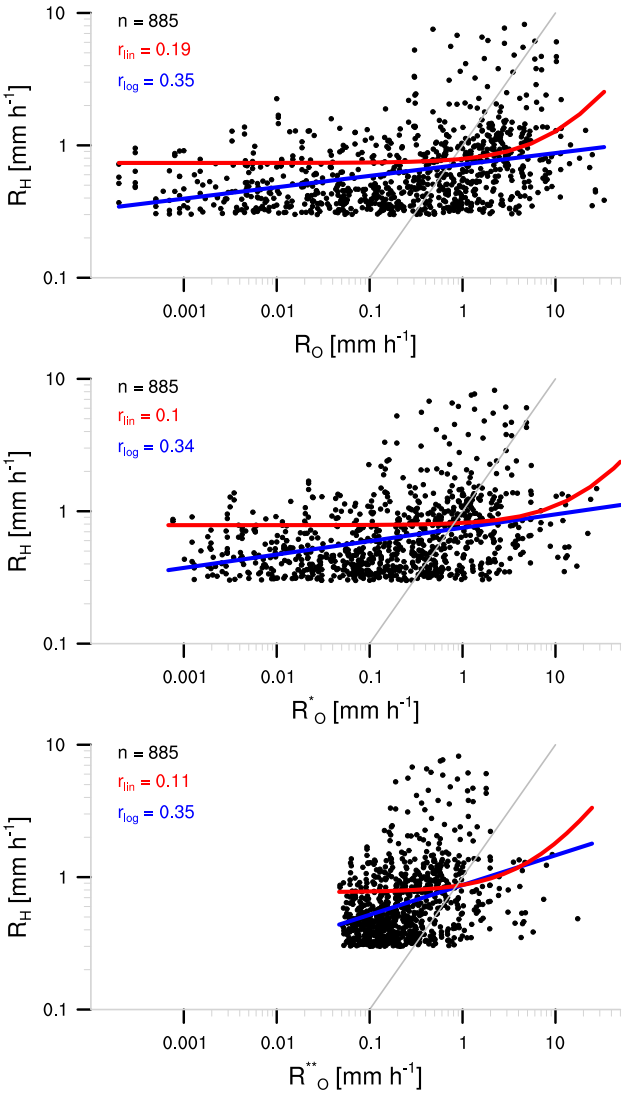


Figure 5.8: HOAPS precipitation rate R_H in mm h^{-1} as a function of OceanRAIN along-track averaged precipitation rate R_O (top), R_O^* (center) and R_O^{**} (bottom) in mm h^{-1}. Red line indicates a linear regression in lin-lin space with Pearson correlation coefficient r_{lin}, blue line a linear regression in log-log space with r_{log}; 1-by-1 line in gray. n gives number of underlying measurements (hits).

cases in which the precipitation rate of OceanRAIN exceeds that of HOAPS ($\frac{R_H}{R_O^{**}} < 1$), a doubled fraction of slow-moving ship cases occurs for relative differences below $\frac{1}{3}$ that decreases again for relative differences below $\frac{1}{6}$, which has two kinds of implications. First, the higher fraction of slow-moving ship cases for $R_H < R_O^{**}$ highlights that low ship speeds stronger influence cases in which HOAPS underestimates precipitation compared to those where HOAPS overestimates precipitation. As a consequence, a slow-moving ship might introduce a bias to the HOAPS–OceanRAIN comparison. Second, the maximum at $\frac{R_H}{R_O^{**}} = \frac{1}{6}$ indicates that other factors might have a stronger influence on cases where the precipitation rate of OceanRAIN more strongly exceeds that of HOAPS. Nevertheless, a decrease in representativeness of areal precipitation

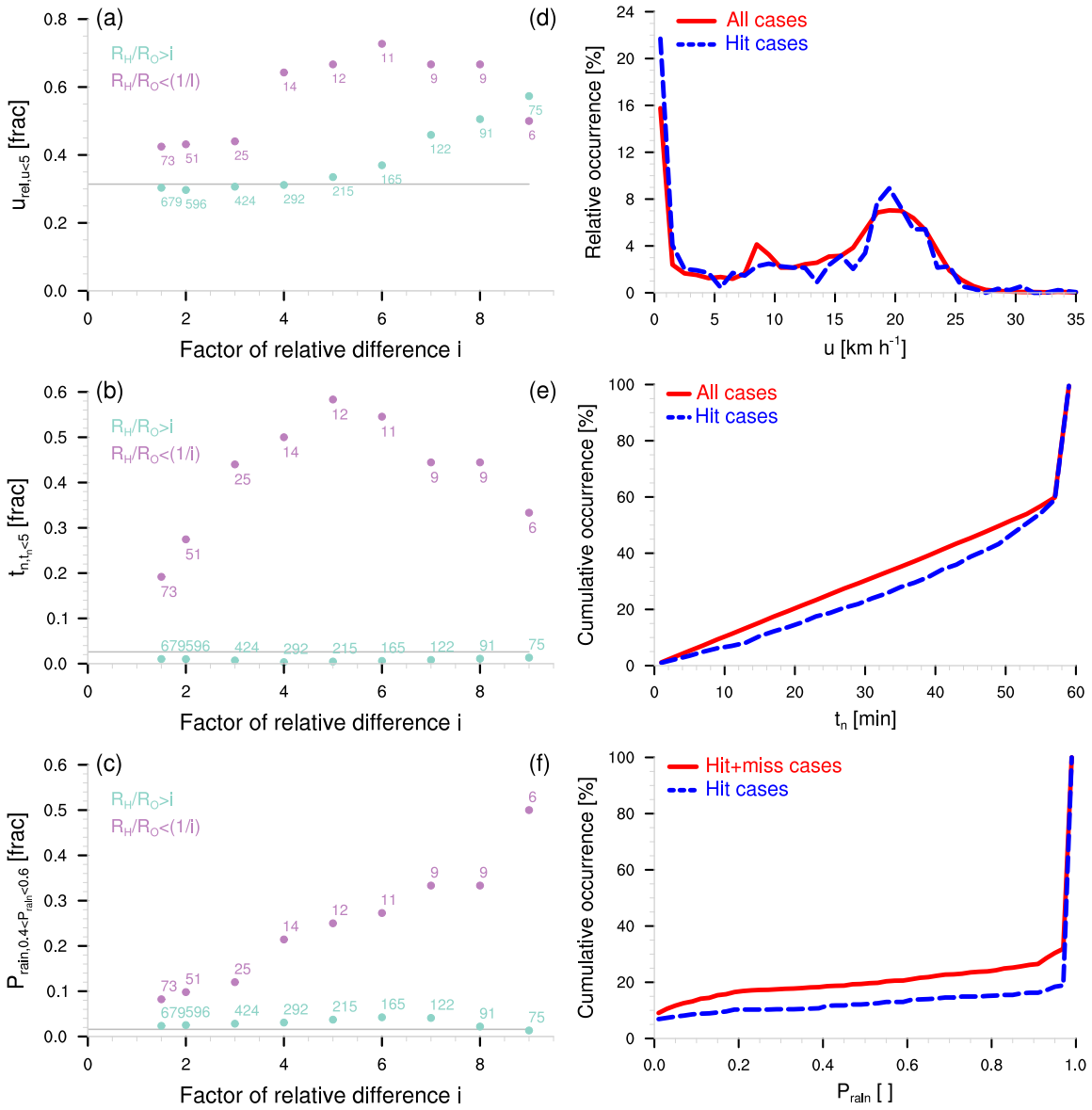


Figure 5.9: (*a-c*) Fraction of ship speeds below 5 km h^{-1} (*a*), OceanRAIN measurements per along-track average $t_n < 5$ min (*b*) and uncertain PP probabilities after using PP distinction algorithm $0.4 < P_{rain} < 0.6$ (*c*), each as a function of the relative difference range $\frac{R_H}{R_O} >$ i (cyan) and $\frac{R_H}{R_O} < \frac{1}{i}$ (purple) for extreme cases displayed in Figure 5.8 (bottom). Gray line indicates mean fraction for all hit cases. Numbers indicate total number of cases that fraction refers to. (*d-f*) Histogram of ship speed u (km h^{-1}) with a bin width of 0.5 km h^{-1} (*d*), t_n (min) with a bin width of 2 min (*e*) and P_{rain} with a bin width of 2 min (*f*), for hits (blue) and all cases (red), except for (*f*): hits and misses.

along a shortened ship track seems plausible, but also depends on the relative cloud movement as concluded in Chapter 4 (Sect. 4.3.4), which we neglected in this analysis.

Second, a low number of along-track OceanRAIN measurements t_n, collocated and then averaged to one HOAPS pixel, could have a similar influence as a slow-moving

ship for the same reason of reduced spatial representativeness of precipitation in a HOAPS pixel. In particular, very small numbers of OceanRAIN measurements per along-track average of less than 5 min can hardly represent the precipitation seen by a satellite sensor. Despite the fact that these cases rarely occur and in two thirds of all hit cases the collocated OceanRAIN measurements consist of at least 40 min, more than 2 % consist of less than 5 min (Fig. 5.9e). These cases of low OceanRAIN samples ($t_n < 5$ min) occur substantially more frequently if the precipitation rate of OceanRAIN exceeds that of HOAPS by at least a factor of 3, increasing the low-sample fraction to more than 25 % (Fig. 5.9b). Low OceanRAIN samples contribute by about 50 % to cases where R_O^{**} exceeds R_H by a factor of at least 4. For cases where R_H exceeds R_O^{**}, no noticeable change is observed in the low-sample fraction of OceanRAIN. The conclusion that low sample sizes strongly influence exclusively those cases where the OceanRAIN precipitation rate exceeds that of HOAPS seems plausible because with a decreasing number of measurements the sampled precipitation tends to overestimate the precipitation in the area.

Third, the PP distinction algorithm introduced in Chapter 3 holds an uncertainty. When the PP is misclassified, OceanRAIN precipitation rates might strongly deviate from the actual precipitation rate. These misclassified cases likely occur in a range of the most uncertain PP probability such as $0.4 < P_{rain} < 0.6$. This uncertain PP probability occurs in less than 2 % of all hit cases (Fig. 5.9f). The more R_O^{**} exceeds R_H, the higher the percentage of uncertain PPs gets with up to 50 % for $\frac{R_H}{R_O^{**}} = \frac{1}{9}$ (Fig. 5.9e). Overestimated R_O^{**} are usually caused by snow- or mixed-phase precipitation missclassified as rain. The monotonically increasing percentage of uncertain PP cases indicates that PP misclassifications can cause largest relative differences for the OceanRAIN precipitation rate exceeding that of HOAPS. In the opposite case of $\frac{R_H}{R_O^{**}} > 1$, rain cases misclassified as snow or mixed-phase precipitation cause hardly any increase in the fraction of uncertain PP.

An eyeball verification of the 14 cases that fulfill $\frac{R_H}{R_O^{**}} < \frac{1}{4}$ reveals that for 3 of them the automatic PP distinction algorithm diagnosed a change in PP from snow or mixed-phase to rain, which drastically increases the precipitation rate as a consequence of sampled relatively large particles. However, a disturbance by wave water cannot completely be ruled out. 6 other cases out of the chosen 14 cases consist of only 1 or 2 OceanRAIN measurements per HOAPS collocation that cause the strong overestimation of areal precipitation. From the remaining 5 cases, 2 cases have an average ship speed below 1 km h^{-1} from which 1 is additionally influenced by birds causing false signals. The remaining 3 cases seem plausible after thorough cross-checking with peak precipitation rates exceeding 50 mm h^{-1} (max. 367 mm h^{-1}). All these cases represent precipitation from clusters of vigorous convection, one in the inner tropics, another in the mid-latitude North Atlantic in the postfrontal sector of a cyclone and the third in the Mediterranean Sea. In all these 3 cases the ship likely overrepresents the precipitation in the HOAPS pixel. However, existing other HOAPS–OceanRAIN collocations need to proof that these 3 cases do not follow the rule but represent unusual cases.

5.3.3 How does precipitation vary spatially within a HOAPS satellite pixel as derived from other satellite products? – A case study

To gain insight into the spatial sub-pixel variability of precipitation, preferably independent high-quality precipitation data sets are required. For the reason that no completely independent high-quality precipitation product exists over the ocean, we choose other satellite data sets in order to derive the sub-pixel precipitation structure to conduct a case study. The data sets basically need to be available in high spatial resolution, provide precipitation estimates over the ocean and have a sufficient coverage to be matched in space and time to the HOAPS–OceanRAIN collocations. First, the MSG Cloud Physical Properties (CPP) algorithm serves as reference data to unveil the sub-pixel structure with a high spatial resolution of about 4 km (more details in Sect. 2.6). Precipitation from MSG-CPP has led to satisfying results compared to CMORPH and TRMM-PR data over Africa (Wolters et al., 2011) as well as GPCC and weather radar data over Europe (Roebeling et al., 2012). The POD reaches about 60 %, the FAR 40 % while the spatial correlation statistics yield 0.89 for 2 months of radar data over the Netherlands (Roebeling and Holleman, 2009). Second, CMORPH serves for comparison as a precipitation satellite data set that blends IR and PMW information (Sect. 2.4). Third, the IMERG product from the new generation GPM mission serves as additional source of information that uses PMW-IR data but gives preference to PMW data if available (Sect. 2.5). All 3 data sets fulfill the requirements of a higher spatial resolution than HOAPS, data availability for the HOAPS–OceanRAIN collocation domain and period as well as providing an independent data source (MSG-CPP), or at least additional data sources to SSMIS data and a different retrieval (IMERG and CMORPH).

Due to data availability, the above mentioned data sets largely constrain the selection of available cases. In specific, IMERG and CMORPH provide no data outside of 60°S to 60°N while the MSG-CPP full disk mainly covers the Atlantic Ocean and the algorithm depends on daylight conditions. These spatial and temporal limitations favor cases predominantly located in the tropics. The tropics offer a challenging environment for the along-track to satellite comparison because convective precipitation tends to strongly vary in time and space.

In a first case from 23 May 2015 in the morning hours, RV *Meteor* moved northeastward in the tropical North Atlantic, south of the Cap Verde Islands. At 9:31 UTC when DMSP-F18 overpassed the area, no precipitation has been measured by the OceanRAIN disdrometer onboard RV *Meteor* (Fig. 5.10). Few minutes before the overpass, the OceanRAIN disdrometer detected a rain shower that led to rain rates between 0.5 and 2 mm h^{-1} and up to 4 mm h^{-1} before the collocation period (gray line in timeseries of Fig. 5.10). The rain rates of OceanRAIN agree well with those from the MSG-CPP images at 9:00 UTC while the raining area well matches with that from RV

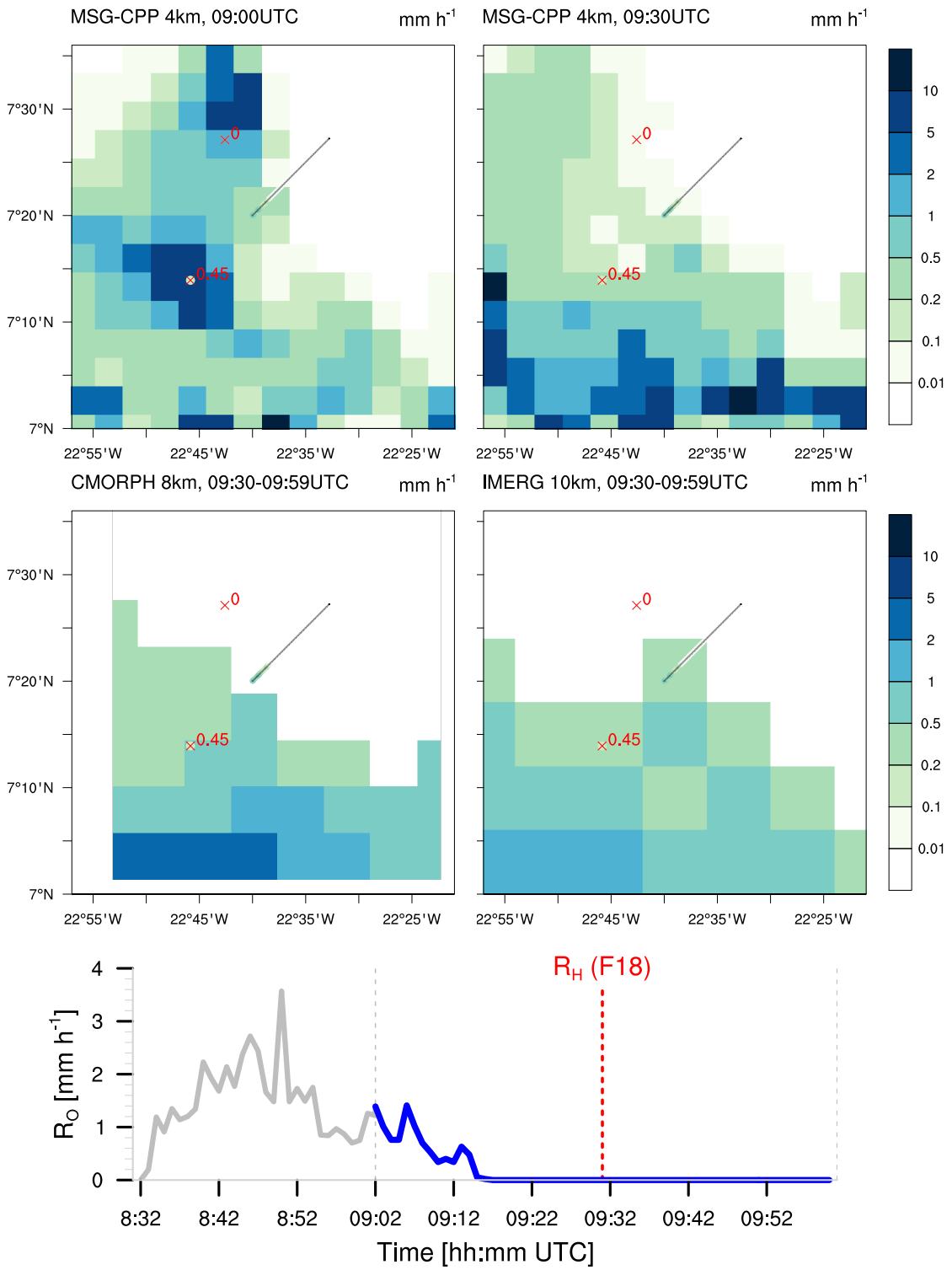


Figure 5.10: Maps of 23 May 2015 for MSG-CPP, CMORPH and IMERG precipitation rates compared to HOAPS (red numbers) and the OceanRAIN individual precipitation rates along the RV *Meteor* track (black dots and surrounding circles) in mm h^{-1}. Timeseries at bottom displays evolution of OceanRAIN precipitation rates (blue: within collocation period, gray: else) as a function of time (UTC), red-dotted line marks time of DMSP overpass (HOAPS), gray-dotted lines mark start and end of collocation period.

Meteor. According to the next MSG-CPP image from 9:30 UTC, the organization of the convective precipitation quickly changed. While the MSG-CPP image estimates an intense rain shower located next to the center of the HOAPS pixel at 9:00 UTC, at 9:30 UTC – the time of the nearest DMSP overpass – the area is almost devoid of rain with some signals hardly exceeding 0.2 mm h^{-1} west of the northern HOAPS pixel. If MSG-CPP is correct, HOAPS might have missed the relatively light rain because of its lower threshold of 0.3 mm h^{-1}. Nevertheless, the CPP algorithm almost entirely uses cloud top information that can be contaminated by Cirrus clouds in the outflow region of large convective systems. For the same reason, CPP might overestimate surface precipitation because evaporative loss of precipitation inside and below the cloud is so far neglected in CPP (Wolters et al., 2011). High resolution MODIS data from few hours past the observation (not shown) reveals large clusters of convection, surrounded by Cirrus that partially obscures the underlying atmospheric column. CMORPH and IMERG tend to agree with HOAPS in placing the precipitation and indicate a similar precipitation rate of about 0.5 mm h^{-1} in the area of the southern HOAPS pixel. As HOAPS, CMORPH and IMERG do not detect precipitation in the Northwest of the chosen domain, which could either indicate that all three data sets miss the precipitation or that it is overrepresented by CPP at 9:30 UTC. Overall the tropical scene reveals fairly good agreement between the 4 data sets for the area that RV *Meteor* passed through while there is no clear indication that the PMW-driven satellite data sets miss noticeable amounts of precipitation.

The second case from the evening hours of 26 May 2015 contains smaller convective clusters of rainfall in about the same tropical area as the first case (Fig. 5.11). According to the smaller extent of the showers, they are expectably more challenging to detect by PMW sensors. RV *Meteor* headed towards the south when the disdrometer sampled a rain shower during the last quarter of the collocation period (17:48–18:04 UTC), whereas the DMSP-F16 passed over the area at 17:34 UTC. West of RV *Meteor*'s track, the two collocated HOAPS pixels detected light precipitation of about 0.3 mm h^{-1} while east of the track another HOAPS pixel detects no rain. According to the OceanRAIN along-track precipitation, the HOAPS precipitation rates do not seem implausible, particularly in light of the beam-filling effect and the decreasing OceanRAIN along-track precipitation rates after the collocation period (not shown). However, MSG-CPP estimated peak precipitation rates exceeding 10 mm h^{-1} right before the DMSP overpass. Despite the small extent of the detected rain showers, HOAPS should have detected these high precipitation rates assuming that the MSG-CPP algorithm is correct. CMORPH and IMERG detect no precipitation at all in the area of the HOAPS pixels. As a consequence of ingesting the same SSMIS data from the same DMSP overpass as HOAPS, the differences between CMORPH and IMERG are exclusively driven by retrieval or algorithm differences. This implies that the HOAPS neural net algorithm outperforms both data sets with respect to rain detection proven by the surface-based OceanRAIN data for this scene. Whether HOAPS still underestimates precipitation rates against MSG-CPP remains hard to judge, although the more

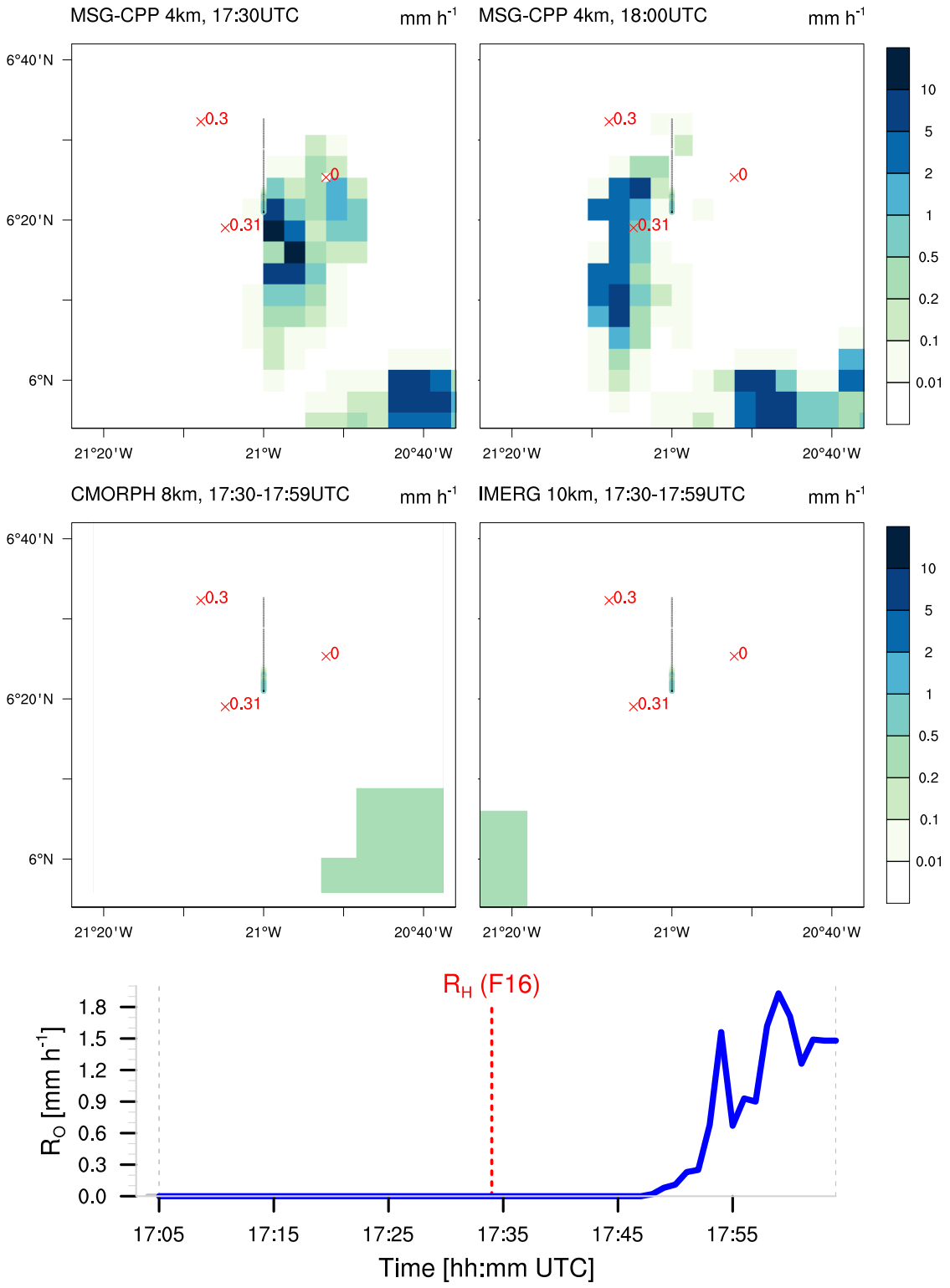


Figure 5.11: As Fig. 5.10 but for 26 May 2015 of RV *Meteor* moving southward.

than 5-times higher peak precipitation rates of MSG-CPP seem fairly high compared to the peak precipitation rates from OceanRAIN.

The third case from 19 September 2015 represents a rare case of convective precipitation in its early stage. Thus, the spatially limited rain showers conform to a typical scene in the subtropical trade region where clouds are usually kept shallow by a prevailing inversion. In this early stage of the precipitation formation process, the disdrometer onboard RV *Meteor* sampled 3 small showers in a row in almost the same tropical area as before (Fig. 5.12). The 3 sampled rain showers within the collocation period decrease in intensity from which the first shower reaches a peak rain rate close to 10 mm h^{-1} while the last shower peaks at 2 mm h^{-1}. MSG-CPP captured the second rain shower with a rain rate below 1 mm h^{-1} while OceanRAIN estimates somewhat higher rain rates of about 4 mm h^{-1}. This difference could be explained by beam-filling due to the very small-scale rain showers that might not completely fill a MSG-CPP pixel of 4 km in diameter. Two overpasses of DMSP satellites at 8:03 and 8:34 UTC lie within the collocation period, which results in 6 collocated HOAPS pixels. 5 out of these 6 collocated HOAPS pixels detect precipitation with rates ranging between 0.3 and 0.6 mm h^{-1}. These relatively low HOAPS precipitation rates seem very plausible when considering the scattered areal precipitation distribution, also reflected by the OceanRAIN along-track averaged rain rates between 0.8 and 1.5 mm h^{-1}, whereby the highest rates belong to the lowest numbers of collocated OceanRAIN measurements ($t_N = 30$). CMORPH estimates 0.2 to 0.5 mm h^{-1} for the southern three HOAPS pixels and up to 1 mm h^{-1} for the two northernmost HOAPS pixels in the scene, which agrees fairly well with HOAPS except for the southwestern edge at which CMORPH detects about 1 mm h^{-1} compared to no precipitation in HOAPS. IMERG agrees well in rain detection and rates with all other data sets in the northern part of the domain. However, IMERG completely misses the light precipitation from the scattered small-scale showers in the southern part of the domain. The missed light precipitation seems to confirm the tendency from case 2 that IMERG tends to miss light scattered precipitation that HOAPS can detect with evidence from OceanRAIN.

The fourth case from 24 August 2015 originates from RV *Meteor* heading in south-westward direction west of the coast of Portugal (Fig. 5.13). In the early evening hours, a cyclone approaches the domain from the Atlantic Ocean with its widespread warm sector and extensive precipitation. Compared to the previous 3 cases, this case is not expected to cause noticeable difficulties in precipitation detection for satellite data sets due to the large raining areas and the moderate rain rates. Within the collocation period (17:50–18:49 UTC), OceanRAIN constantly sampled precipitation rates between 1 and 5 mm h^{-1} until the rain stops at 18:42 UTC. At 18:19 UTC, DMSP-F17 passes over the area from which HOAPS estimates rain rates of 1.1 and 1.5 mm h^{-1}. These HOAPS precipitation rates agree well with the precipitation rates of OceanRAIN when considering the beam-filling that smears out the signal over the domain, which also contains areas without precipitation as indicated by OceanRAIN and MSG-CPP. MSG-CPP agrees fairly well with OceanRAIN in reflecting the peak

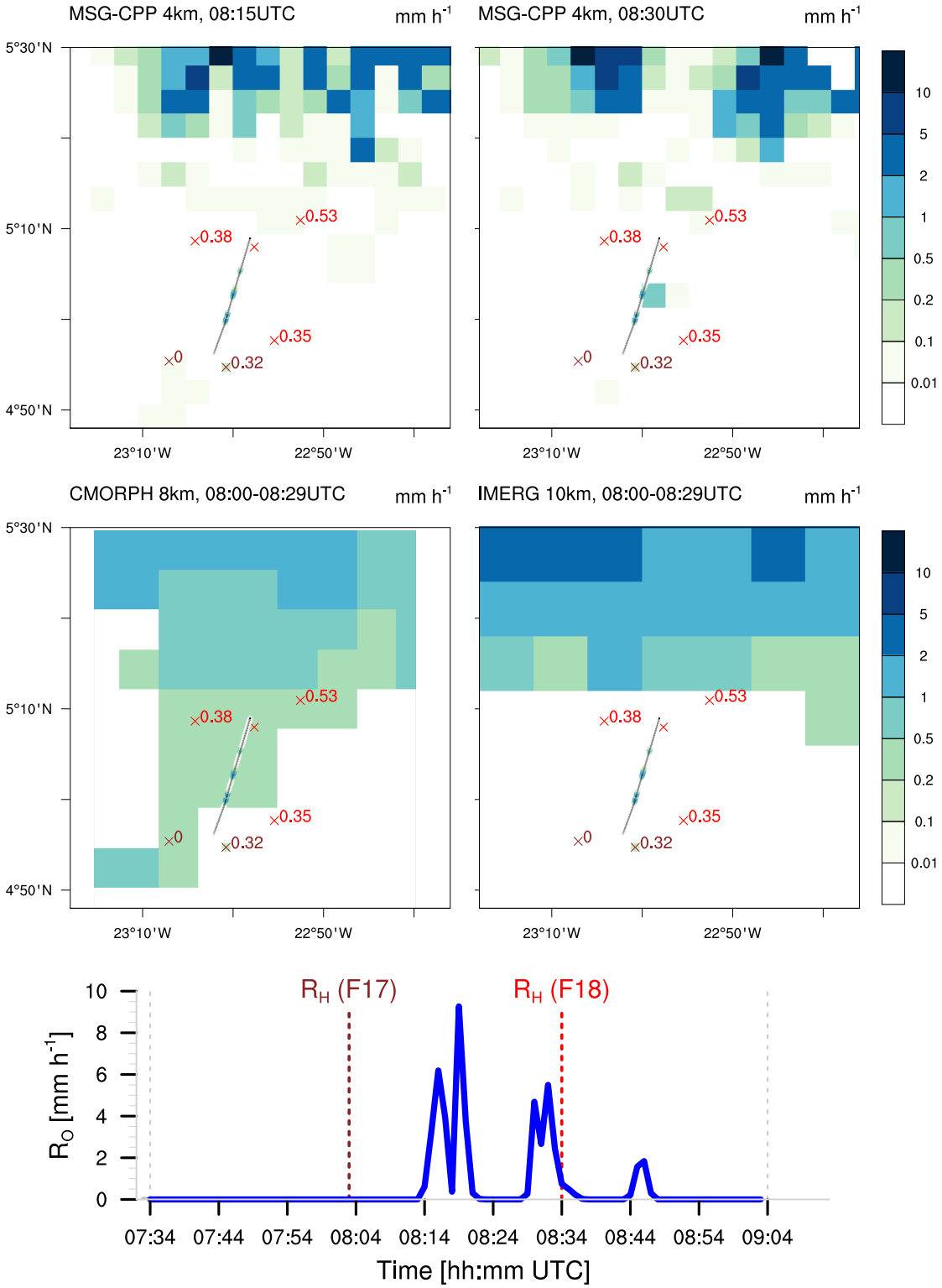


Figure 5.12: As Fig. 5.10 but for 19 September 2015 with RV *Meteor* moving northeastward and two collocated DMSP satellite overpasses of F17 (brown) and F18 (red).

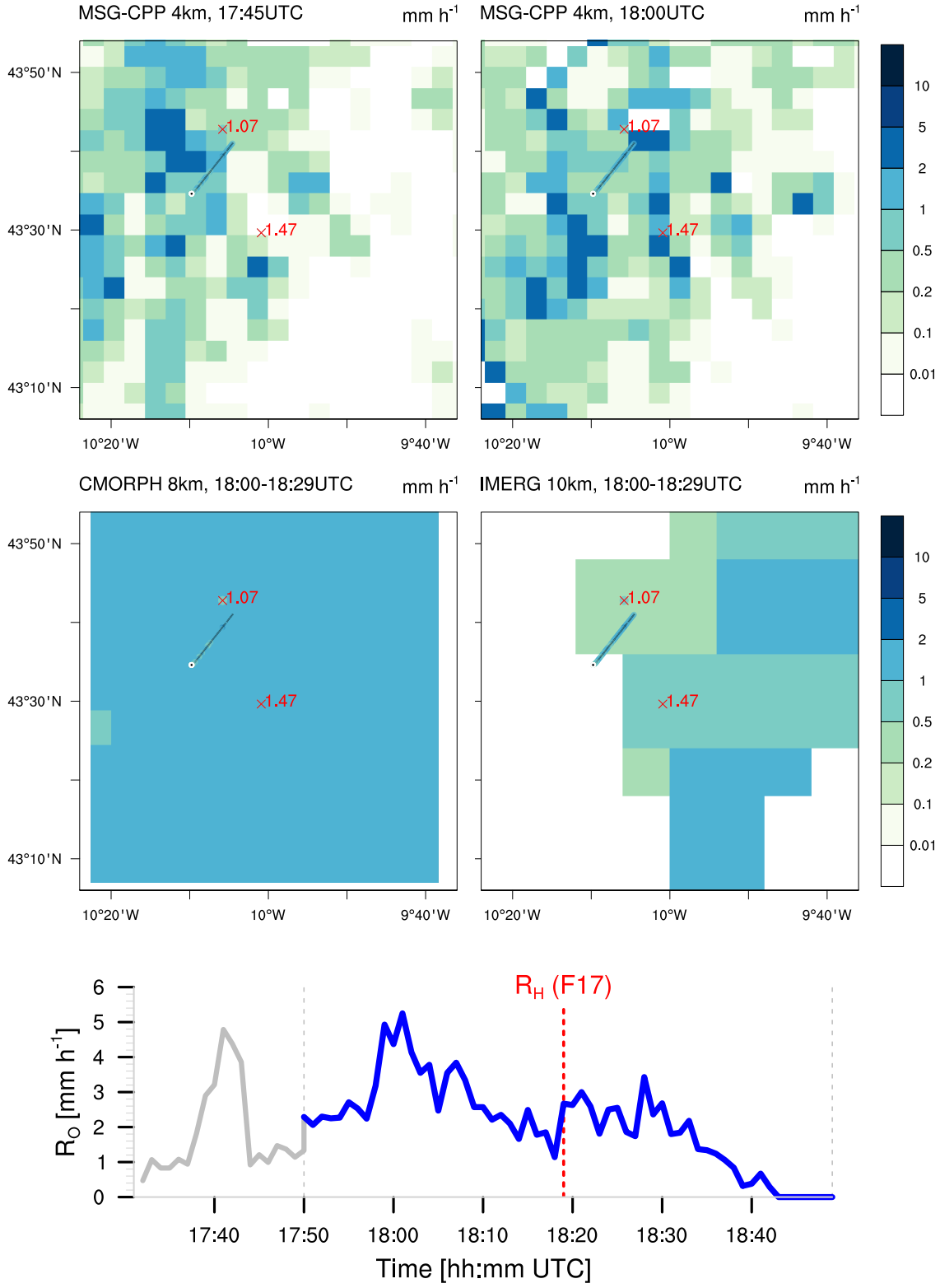


Figure 5.13: As Fig. 5.10 but for 24 August 2015 with RV *Meteor* moving southwestward.

around 18:00 UTC with up to 5 mm h^{-1} as well as the lighter precipitation between 17:45 and 17:55 of 1 to 2 mm h^{-1}. In absence of daylight, no later MSG-CPP images exist for this case. CMORPH estimates about the same precipitation rates as HOAPS between 1 and 2 mm h^{-1} reflecting almost no variability in its precipitation rate throughout the whole domain. IMERG estimates about a factor of 2 lower precipitation rates ranging between 0.2 and 1 mm h^{-1} compared to all other considered data sets. The ingested merged microwave-only product of IMERG estimates much higher rain rates (not shown). Apart from the deviations of IMERG, all considered data sets agree well in this fourth case of rather stratiform precipitation in the warm-sector of a cyclone.

In a fifth case from 30 November 2015, RV *Meteor* heads in northwestward direction close to the western coast of Central Africa (Fig. 5.14). In the afternoon hours, some small-scale convective precipitation lies over the Atlantic Ocean while somewhat larger convective clusters form towards the African coast that progress in southeastward direction. In contrast to the previously presented cases, OceanRAIN detects no precipitation event throughout the whole 60-minute collocation period (14:43–15:32 UTC). However, half an hour before the collocation period around 14:07 UTC, the OceanRAIN disdrometer onboard RV *Meteor* records a short but intense rain shower with peak precipitation rates of up to 18 mm h^{-1}. This rain event and the fact that the disdrometer was properly working during the whole collocation period rules out that OceanRAIN could have missed precipitation due to an instrument failure. The DMSP-F16 satellite passed over the area at 15:12 UTC and detected precipitation rates of about 0.6 mm h^{-1} within 2 neighboring HOAPS pixels collocated to OceanRAIN. The HOAPS precipitation estimate of 0.6 mm h^{-1} does not seem unrealistic from the perspective of MSG-CPP that estimates a cluster of rain showers in the east of the domain with peak precipitation rates between 2 and 5 mm h^{-1}. CMORPH and IMERG confirm that picture of precipitation in the eastern part of the domain with precipitation rates up to 1 mm h^{-1}, whereas IMERG places the edge of the cluster about 0.2° more eastward. Nevertheless, the scene exemplifies how an OceanRAIN ship track did not sample precipitation located in the HOAPS pixel, which represents a misclassified "false detection" by HOAPS.

The discussed cases have been chosen independent of the satellite data performance, whereas the data selection was strongly limited by the requirement to match all 4 data sets in both time and space. Despite the rather low number of cases presented in the case study, a relatively wide range of meteorological situations is covered. Nevertheless, the case study lacks high-latitude precipitation events with particularly frozen precipitation. For the chosen data sets, no data is provided for the high latitudes. However, these cases still pose a major challenge to current global precipitation satellite data sets, which opens a wide field for satellite validation, particularly over the global oceans.

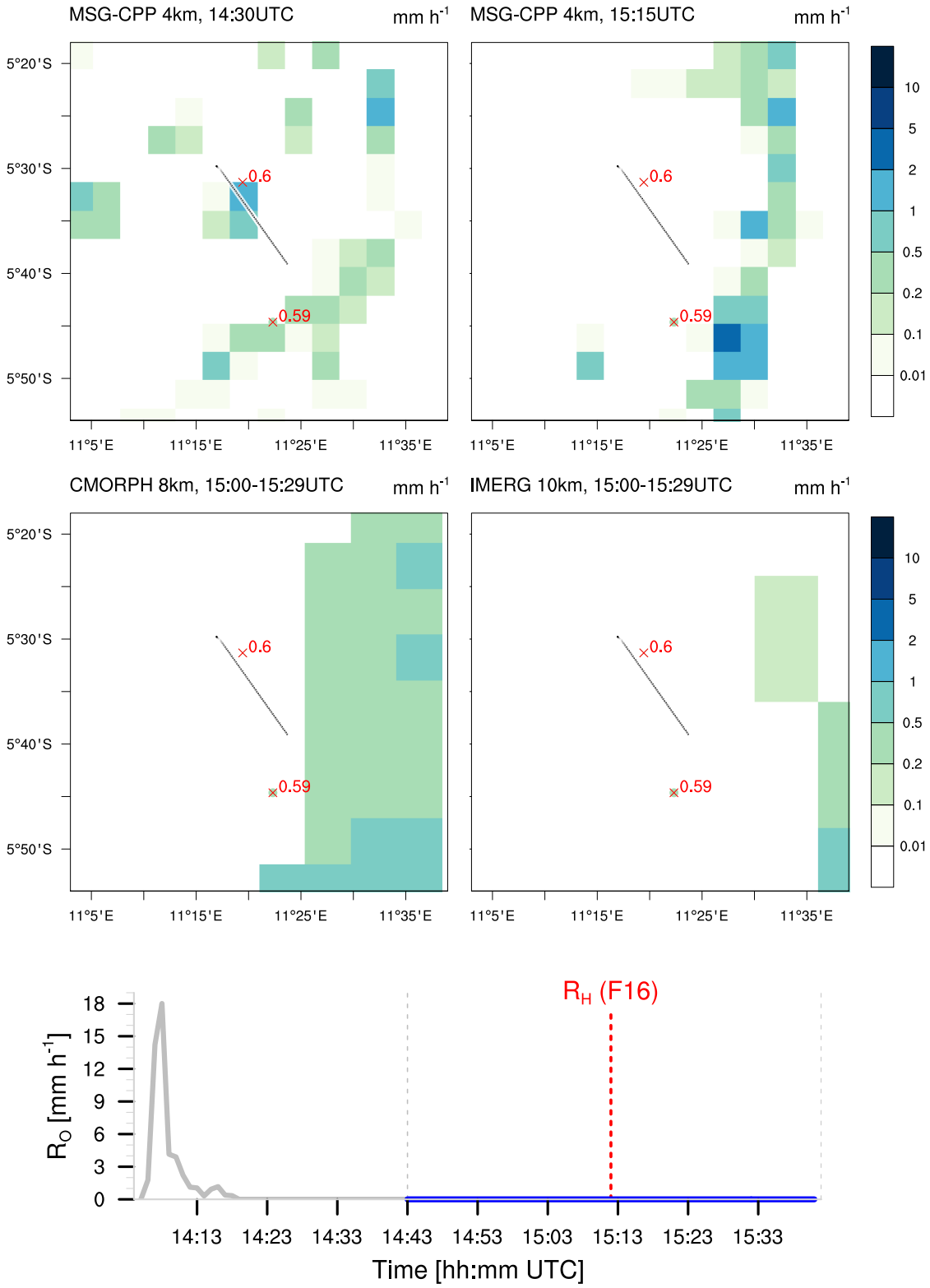


Figure 5.14: As Fig. 5.10 but for 30 November 2015 with RV *Meteor* moving northwestward. Note that $R_H > 0$ while $R_O = 0$ (false detection).

5.3.4 How do HOAPS and OceanRAIN compare on latitudinal average?

The detailed view on individual HOAPS–OceanRAIN collocations revealed two main outcomes. First, the case study exemplarily demonstrated that HOAPS suits to detect and accurately estimate precipitation in a number of cases. Second, even after statistically adjusting OceanRAIN to HOAPS, a number of cases does not fulfill the requirements of a fair comparison between both data sets. Most importantly, OceanRAIN needs to represent the largest possible area of the HOAPS pixel which needs both – a sufficiently long ship track and a sufficiently high number of measurements that form the along-track averaged precipitation rate of OceanRAIN. In particular, a very low number of collocated along-track measurements and a low ship speed do not fulfill these essential requirements in order to minimize artifacts distorting the OceanRAIN–HOAPS comparison. However, excluding all collocations with average ship speeds below 1 km h^{-1} would lead to a loss of about 25 % of all collocations. In neglecting the cloud movement, many of these 25 % of cases could still fairly well represent precipitation in the HOAPS pixel because fast-moving clouds can outweigh the effect of low ship speeds. Combining the ship speed with the number of collocated along-track OceanRAIN measurements implicitly considers this relation because the more measurements belong to the track, the higher the probability for a sufficient areal sampling of OceanRAIN. This means, we exclude collocations with a ship speed below 5 km h^{-1} that contain less than 30 OceanRAIN measurements. Additionally, cases of less than 5 OceanRAIN measurements per collocation are excluded because they are most unlikely to correctly represent the areal precipitation (cf. Fig 5.9b).

We exclude few critical cases with a highly uncertain PP probability of $0.4 < P_{rain} < 0.6$ that has been estimated by the automatic PP distinction algorithm. A misclassified PP might lead to strongly over- or underestimated precipitation rates. Excluding cases of uncertain PPs reduces the uncertainty in OceanRAIN precipitation rates because mixed-phase precipitation holds the highest uncertainty of all PPs. Besides the challenge in classifying mixed-phase precipitation, the OceanRAIN snow retrieval (Eqs. 2.5 and 2.6 in Sect. 2.1.1) that is used for mixed-phase precipitation ignores the actual fraction of snow particles and rain particles. Neglecting the rain–snow particle ratio adds to the uncertainty. Consequently, excluding cases of uncertain PP probability also reduces the uncertainty of precipitation rates at air temperatures close to the freezing point. Together with the thresholds for low ship speed and limited numbers of measurements, 1 453 out of 24 990 collocations (43 of 885 hits) are excluded from the following analysis.

From the case-by-case comparison of HOAPS and OceanRAIN, the influence of few cases stood out in which R_H and R_O^{**} clearly disagree. As depicted in Figure 5.9, many of these disagreeing cases could be removed applying the 3 thresholds to isolate samples of unfair conditions for comparison. These cases mainly involve collocations in which OceanRAIN oversampled precipitation with respect to the HOAPS pixel, whereas the

opposite situation of undersampled OceanRAIN precipitation is more challenging to unveil. As a main reason, the underlying precipitation distribution in the area remains unknown, which rules out to directly assess the spatial representativeness of OceanRAIN. Instead, we rely on the statistical adjustment of OceanRAIN, based on subtropical data from the S-Pol radar that, however, is incapable of correctly adjusting individual cases.

The statistical adjustments applied to OceanRAIN base on S-Pol radar data from the subtropical North Atlantic that can be interpreted as a "training data set" (Chapter 4). Thus, the applicability of the derived statistical adjustments to global ocean data does not seem obvious. However, the Caribbean area where the S-Pol was deployed during RICO experiences a prevailing trade inversion that keeps clouds shallow, frequently leading to light showers of warm rain (Nuijens et al., 2014). Embedded in the trade-wind flow these warm rain showers usually remain scattered and small (Nuijens et al., 2009). Thus, most of these showers are challenging to detect for PMW satellite sensors. In addition to the small isolated showers, some more extensive tropical depressions were sampled that led to more widespread showers and higher rain rates, as well. Accordingly, the S-Pol training data set comprises not only a variety of different precipitation regimes but, in particular, challenging cases for a track-to-area adjustment. Nevertheless, the training data lacks mid-latitude frontal passages as well as solid precipitation at all, that would need to be taken into account for a more robust adjustment. Generally, a rain-area coverage below 20 and above 80 % is sampled most frequently, whereas medium-covered areas are sampled least frequently (cf. Fig. 4.3). Overall, the S-Pol represents a suitable training data set for a statistical track–area adjustment of precipitation including most of the globally most frequently occurring precipitation regimes that range from scattered to clustered convective precipitation and more widespread stratiform precipitation.

The remaining cases should on average reflect differences between the climate regimes. Averaging all HOAPS precipitation rates R_H per 10° latitude band reveals average precipitation rates slightly below 1 mm h^{-1} for all hit cases, except for the southern-hemisphere mid-latitudes with a precipitation rate of 3 mm h^{-1} (black-dashed line in Fig. 5.15). Except for the poles and the southern-hemisphere mid-latitudes, R_O (solid gray line) exceeds R_H by about a factor of 2. Applying the T_E-adjustment to OceanRAIN decreases this difference because the lowering of high R_O outweighs the increase in low R_O, which mainly affects the average precipitation rate (cf. Fig. 5.7). Applying the second statistical adjustment in R_O^* and $R_O^{\circ *}$ (without the T_E-adjustment) amplifies this effect by further lowering the OceanRAIN average precipitation rate by a factor of 2 to 3. As a result, R_H exceeds R_O^{**} by about 1.5 up to 5 depending on the latitude. The large difference between R_H and R_O^{**} is likely to be explained by the different instrument sensitivities of optical disdrometer and SSMIS onboard the DMSP satellite (addressed in Sect. 5.3.5).

The uncertainty of the OceanRAIN precipitation rates is derived using a resampling

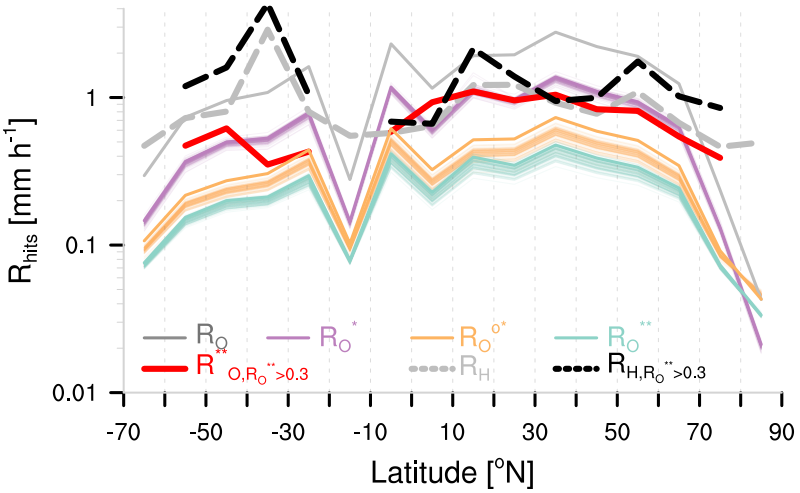


Figure 5.15: Average precipitation rate (mm h^{-1}) per 10° latitude band for all 885 hits for OceanRAIN R_O (gray), R_O^* (purple), R_O^{**} (cyan), R_O^{o*} (orange), $R_O^{**}(R_O^{**} > 0.3)$ (red) and HOAPS R_H (gray dashed) as well as $R_H(R_O^{**} > 0.3)$ (black dashed). Thin colored lines indicate uncertainty from 100 realizations of the randomly chosen halved sub data set.

of 100 randomly chosen sub data sets of half the size of all remaining 23 537 HOAPS–OceanRAIN collocations. The resulting spread indicates the precision of the calculated precipitation averages defined as the difference between the 25 % and 75 % quantiles. The inter-quartile range of these 100 realizations rarely exceeds 20 % of the actual value and reaches the highest spread for R_O^{**} after applying the R_O^*-adjustment. The relatively low uncertainty in OceanRAIN precipitation rates results from the extensive statistical adjustments and quality checking applied to the collocated OceanRAIN precipitation data. However, this internal uncertainty contains no information on the overall representativeness of the sampling per latitude, discussed in Section 5.3.7.

The hit cases partly represent the typical precipitation distribution per latitude with a relatively small peak at the equator and secondary peaks at or close to the mid-latitudes of both hemispheres. Towards the North Pole, precipitation rates decrease by about an order of magnitude compared to the northern-hemisphere high-latitudes. However, the distribution of precipitation entirely depends on the location of the ship tracks as well as their sampling frequency that strongly varies with latitude (Sect. 5.3.6). Including collocations of non-precipitating scenes includes the precipitation occurrence (Fig. 5.16 bottom). The higher number of non-precipitating cases over the subtropical ocean areas decreases precipitation rates mainly in the southern hemisphere, whereas the minimum in the northern hemisphere is much less pronounced. This uneven distribution depends on the location of most ship tracks, discussed in Section 5.3.6. OceanRAIN precipitation rates peak on both hemispheres at the mid-latitudes because the precipitation occurrence of OceanRAIN reaches 22 to 32 % compared to 4 to 10 % in the subtropics (Tab. 5.1). The influence of non-precipitating cases indicates that besides an accurate precipitation rate, the precipitation detection plays a key role for estimating the average precipitation rate per latitude.

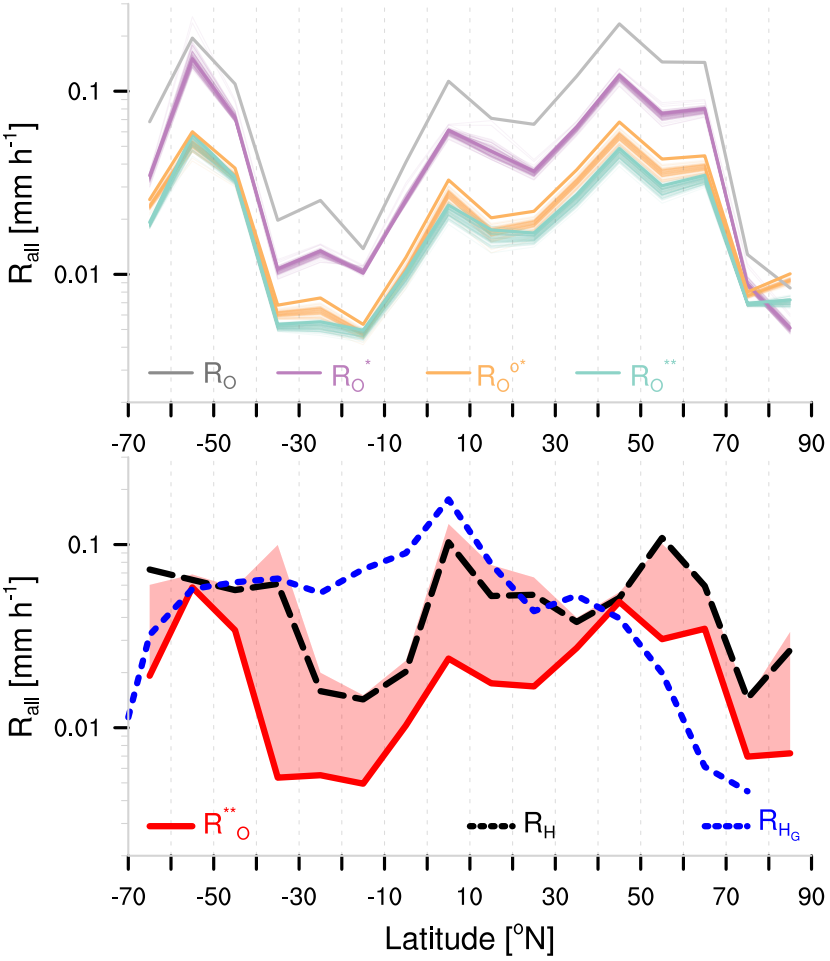


Figure 5.16: Average precipitation rate (mm h^{-1}) per 10° latitude band for all remaining 23 537 collocations for OceanRAIN R_O (gray), R_O^* (purple), R_O^{**} (cyan), R_O^{o*} (orange) in top panel. Bottom panel shows again OceanRAIN R_O^{**} (red) as well as HOAPS R_H (black dashed) and R_{H_G} (blue dotted) from 25 years of HOAPS-G data. Red-shaded area frames the uncertainty range of misclassified false detections by HOAPS with the upper edge counting all false detections as hits.

After applying the statistical adjustments to OceanRAIN (Fig. 5.16 top), R_H exceeds R_O^{**} by a factor of 2 to 4, except for the mid-latitudes with a markedly lower difference (bottom). These regions coincide with a low false alarm ratio (cf. Fig. 5.3). As discussed in Section 5.3.3 for the "30 November 2015" case, a relatively high fraction of false detections likely indicates scenes in which the RV of OceanRAIN "missed" the precipitation falling in the HOAPS pixel. Accordingly, many cases classified as false detections might represent scenes of actual precipitation detected by HOAPS. This situation becomes more likely for convective precipitation as illustrated by the case of "19 September 2015" (Fig. 5.12) that gives an impression how quickly convective precipitation can develop, change its intensity and be limited in its spatial extent. In addition to convective clustering, a ship track is less likely to detect precipitation when showers tend to be small or heterogeneously distributed over the HOAPS satellite

pixel. As these convective conditions frequently occur in most global oceanic regions (Short and Nakamura, 2000), a high fraction of cases classified as false detections would have been classified as hits if the ship took a "more appropriate" track through the HOAPS pixel to detect precipitation. This effect is even more pronounced for more generous collocation boundaries as used in Bumke et al. (2016). This study obtains false detection fractions of up to 20 % at the equator, likely explained by missed precipitation in the ship reference data set. To quantify the potential effect of spatial under-representation in OceanRAIN, we treat all false detections as hits (red shaded area in Fig. 5.16 bottom). The modified OceanRAIN average precipitation rate resembles that of HOAPS in almost all latitudes. The very good agreement indicates that likely most of the false detections point at insufficient spatial representativeness of OceanRAIN in the HOAPS pixel. In lack of the sub-pixel precipitation distribution we cannot correct for this effect. However, the attempted masking of this deficient spatial-scale representation suggests a good agreement between HOAPS and OceanRAIN. Nevertheless, the spatial representation of the reference data can only be statistically adjusted for detected precipitation; missed precipitation remains a known unknown in the validation of areal satellite data when using along-track shipboard data.

5.3.5 Adjusting the sensitivity of OceanRAIN to HOAPS

The different sensitivity to precipitation at the lower end of the spectrum remains as a key difference after statistically adjusting OceanRAIN to HOAPS because both adjustments do not depend on individual precipitation rates but were normalized. Thus, we take into account the lowest precipitation rates that OceanRAIN and HOAPS are able to resolve. The optical disdrometer used in OceanRAIN can resolve precipitation rates as low as 0.01 mm h^{-1}. In theory, much lower precipitation rates could be resolved with the optical disdrometer but sea spray and ship vibrations introduce large uncertainties particularly to tiny precipitation particles that mainly contribute to low precipitation rates (Klepp, 2015). In addition, there exists no practical utility to resolve precipitation rates below 0.01 mm h^{-1}. The HOAPS retrieval excludes measured precipitation rates below 0.3 mm h^{-1} that fall below a signal-to-noise ratio under which precipitation is not clearly distinguishable from sensor noise (Andersson et al., 2010b). The discrepancy of more than one order of magnitude between the lowest precipitation rates in HOAPS and those in OceanRAIN causes a systematic "overestimation" of light precipitation detected by HOAPS with respect to OceanRAIN after applying the statistical adjustments as mentioned in Section 5.3.2 that also affects average precipitation rates.

By excluding all $R_O^{**} < 0.3$ mm h^{-1}, OceanRAIN can be further adjusted to HOAPS to simulate a similar sensitivity and yield a more meaningful average precipitation rate. For the hit cases, the sensitivity-adjusted average precipitation rate resembles that of HOAPS predominantly in the northern hemisphere (solid red line in Fig. 5.15). However, in the southern hemisphere, large differences occur in particular at about

40°S. Nevertheless, these average precipitation rates are based on a relatively small sample of 207 cases that is left after excluding $R_O^{**} < 0.3$ mm h^{-1}, besides 142 misses, 639 false detections and 18 729 non-precipitation cases. In dry regions such as the southern-hemisphere subtropics none to very few cases are left, which explains the gap in Figure 5.15. For most OceanRAIN-sampled global regions, such a small number of remaining collocations does not suffice for a meaningful comparison of average precipitation rates. Despite the small amount of data left, HOAPS and OceanRAIN precipitation rates agree fairly well in the northern hemisphere.

The remaining small sample size for adjusted precipitation sensitivity in OceanRAIN certainly allows a statistical comparison of precipitation-rate percentiles. Instead of percentiles we rank all precipitation rates where either HOAPS or OceanRAIN detect precipitation (R_{rank}) to help identify how strongly certain precipitation rates contribute to the average precipitation rate calculated from $R < R_{rank}$ (Fig. 5.17). Excluding non-precipitating collocations makes the distribution independent from the ability to detect precipitation. This means, we calculate the precipitation-rate contribution of ranked precipitation rates for all cases in which one of the data sets detected precipitation for each data set individually (thick lines). For HOAPS, these cases include hits and false detections (h+f) while for OceanRAIN, these cases consist of hits and misses (h+m); hit cases are shown as thin lines.

Before the statistical adjustments of OceanRAIN, the average precipitation rate of HOAPS exceeds that of OceanRAIN by about 30 % when comparing exclusively those cases in which the respective data set detected precipitation. For hit cases, the average precipitation rate of HOAPS stays almost 40 % below that of OceanRAIN. This large difference between the OceanRAIN hits and misses points at the markedly lower precipitation rate in cases where HOAPS misses precipitation compared to those cases where HOAPS detects precipitation. Furthermore, these cases in which HOAPS misses precipitation predominantly contain low precipitation-area fractions. Applying the first and the second statistical adjustment to OceanRAIN does not essentially decrease the relative difference between the average precipitation rate of hits and misses to those of hits only. The only slightly decreasing relative difference after applied adjustments indicates that this effect does not merely result from spatial-scale differences but likely from the different instrument/retrieval sensitivities of HOAPS and OceanRAIN. However, the absolute average precipitation rates and, thus, the absolute difference strongly decrease so that OceanRAIN precipitation rates for hits and misses become more similar to each other induced by the statistical adjustment of spatial-scale differences. After excluding precipitation rates below 0.3 mm h^{-1}, the difference between hits and misses vanishes while the fraction of misses from all remaining 19 717 collocations decreases to about 1 %, confirming the assumption that a lower sensitivity in the HOAPS algorithm explains most of the misses. The good agreement of hits and misses from OceanRAIN approves that both kinds of adjustments – those due to spatial-scale differences and that for different precipitation sensitivity – serve to achieve a fair comparison between surface-based and satellite precipitation data.

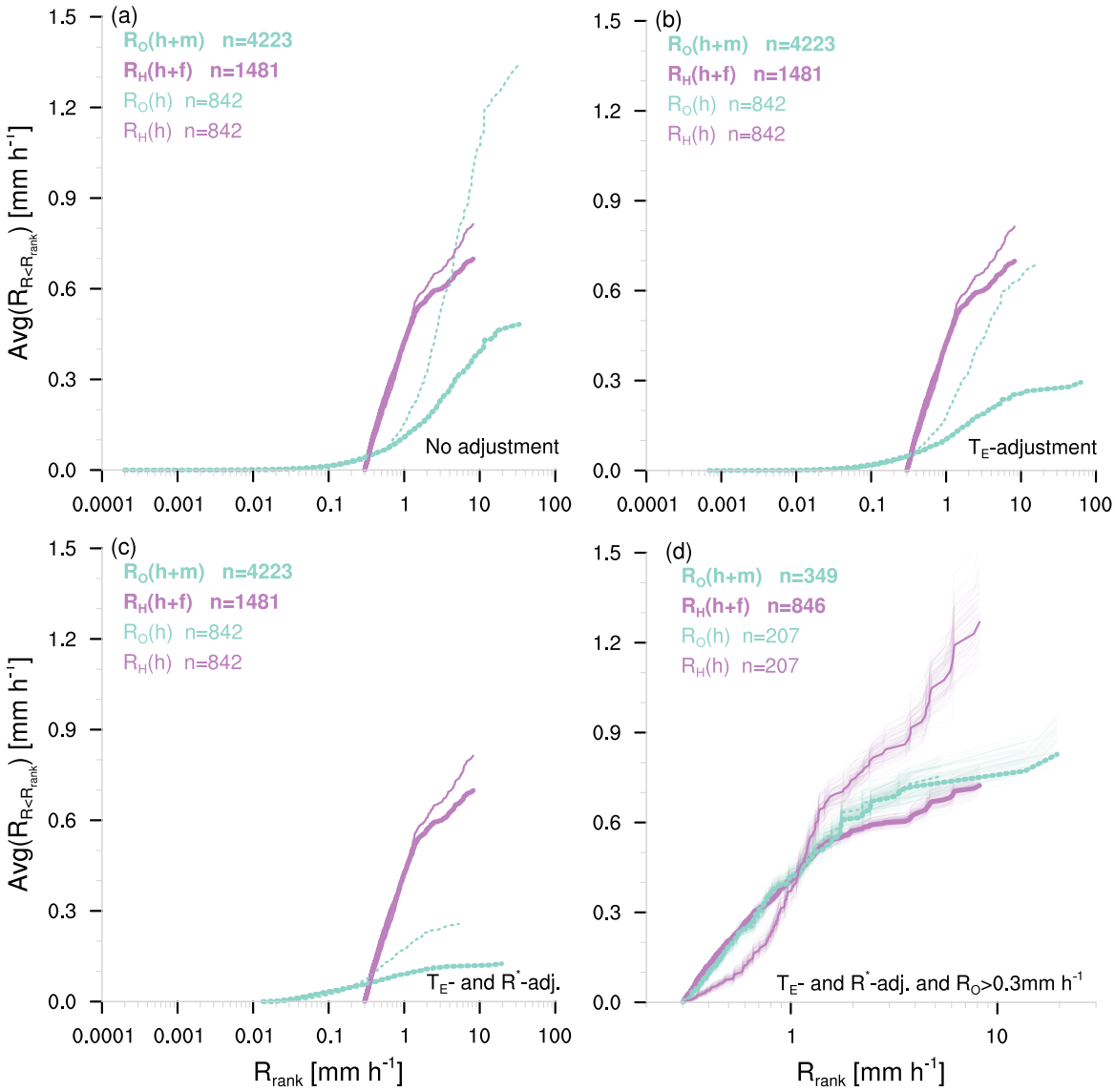


Figure 5.17: Average precipitation rate in mm h^{-1} as a function of their ranked precipitation rate from OceanRAIN (cyan lines) and HOAPS (purple) for OceanRAIN along-track averaged rain rate R_O (a), R_O^* (b), R_O^{**} (c) and $R_O^{**} < 0.3$ mm h^{-1} (d). Dotted lines indicate hit cases only, whereas solid lines include all cases in which the respective data set detected precipitation. Note the changed scaling in the abscissa of (d) that also contains 100 realizations of 50 % randomly chosen values of each sub data set as a measure of uncertainty.

The precipitation-rate distribution agrees well for HOAPS and OceanRAIN after applying the previously mentioned adjustments to OceanRAIN. In particular for $R < 2$ mm h^{-1}, both precipitation-rate distributions resemble each other. This means, precipitation rates below 1.3 mm h^{-1} lead to 0.5 mm h^{-1} average precipitation rate in both data sets. For $R > 2$ mm h^{-1}, the HOAPS average precipitation-rate distribution levels off earlier at about 3 mm h^{-1} but keeps increasing between 4 and 8 mm h^{-1}, which might also be subject to insufficient sampling. The highest OceanRAIN precip-

itation rates exceed 10 mm h^{-1} while the HOAPS precipitation rate stays below that value. Overall, this leads to an average precipitation rate of 0.72 mm h^{-1} for HOAPS and 0.83 mm h^{-1} for OceanRAIN, a 13 % underestimation of HOAPS that yet strongly depends on few intense precipitation cases and thus depends on sampling (Sect. 5.3.7).

The difference between HOAPS hits and false detections increases to 0.55 mm h^{-1} (43 %) after applying all adjustments while the difference used to be smaller before (0.12 mm h^{-1} or 14 %). This difference is mainly driven by the false detections in which HOAPS detects precipitation ($R_H > 0$) from which a certain percentage has not been detected by OceanRAIN ($R_O = 0$). A high fraction of false detections in Figure 5.16 (bottom) coincides with the largest differences in conditional precipitation rates of HOAPS and OceanRAIN (Fig. 5.18). The conditional precipitation rate includes all collocations per product in which precipitation is detected, excluding OceanRAIN precipitation rates below 0.3 mm h^{-1}. However, the significance of these results is strongly limited. First, the latitude averaged precipitation rates compare rather statistically than directly because different cases are included (numbers in Fig. 5.18). Second, the low number of remaining OceanRAIN hits and misses, e.g. of 12 cases between 35°S and 10°S, decreases the robustness of the comparison of these adjusted conditional precipitation rates. These latitudinal differences shed additional light on the uncertainty of the averaged precipitation rate of HOAPS and OceanRAIN. This internal uncertainty is estimated by resampling 100 randomly chosen halved subsets of HOAPS and OceanRAIN conditional precipitation rates (bootstrapping statistics). In the resulting distribution of the 100 realizations, the inter-quartile range is a robust measure of the standard deviation. OceanRAIN conditional precipitation rate holds an inter-quartile range of less than 15 %, whereas HOAPS remains below 10 % of the

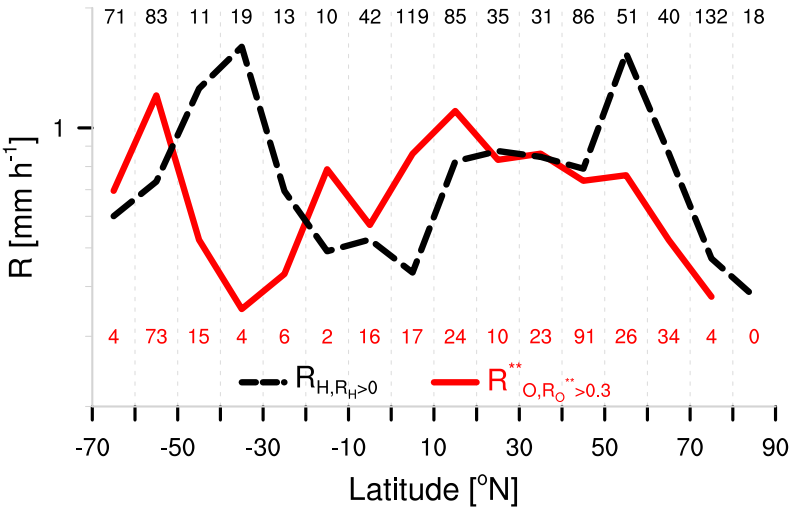


Figure 5.18: Average precipitation rate (mm h^{-1}) per 10° latitude band for all 349 precipitation detections of OceanRAIN with $R_O^{**} \geq 0.3$ (red solid) and 846 precipitation detections of HOAPS with $R_H > 0$ (black dashed). Numbers indicate the number of collocations used for the calculation per latitude.

respective conditional precipitation rate (Fig. 5.17d). The lower uncertainty in HOAPS likely results from the twice-as-large sample size. However, the estimated mean difference of 0.1 mm h^{-1} (13 %) between conditional precipitation rates lies within the same range of the estimated uncertainty. This similarity complicates to deduce reliable implications from the HOAPS–OceanRAIN validation.

5.3.6 Geographical distribution of precipitation differences

The contribution of individual precipitation rates to the average precipitation rate reveals that HOAPS and OceanRAIN mainly differ in precipitation rates that exceed 2 mm h^{-1}, whereas HOAPS cases classified as false detections reach markedly higher precipitation rates. These higher precipitation rates for false detections likely result from the large contribution by inner-tropical cases where on average higher precipitation rates occur. Bringing the HOAPS–OceanRAIN collocations on a 2° by 2° grid allows an insight into how the absolute precipitation-rate differences are distributed over the climate regimes of the Atlantic Ocean and adjacent polar regions (Fig. 5.19). Before adjusting OceanRAIN to HOAPS, the precipitation rates of OceanRAIN strongly exceed those of HOAPS in many regions, particularly in the tropics and mid-latitudes where overall most cases are sampled (Fig. 5.20). Nevertheless, in few boxes HOAPS precipitation rates strongly exceed those of OceanRAIN. After applying the statistical adjustments to OceanRAIN, the differences between both data sets markedly decrease while R_H exceeds R_O^{**} in the majority of grid-boxes. However, no clear differences between certain climate regions are noticeable, likely because of the very limited amount of data that rarely exceeds 5 measurements per grid-box.

Including the non-precipitating cases adds an order of magnitude more data that visualizes the most frequent ship tracks (see Fig. 5.20). These areas of most frequent ship presence include the tropical East Atlantic from which most of the case-study scenes are taken as well as the high-latitude Baffin Bay and the area south of Svalbard (Norway). Fewer collocated ship measurements come from the West Atlantic Ocean. The grid-box averaged difference in precipitation rate from OceanRAIN and HOAPS stays below 0.2 mm h^{-1} in most regions. In the inner tropics, HOAPS slightly tends to overestimate the average precipitation rate by about 0.4 mm h^{-1} around the equator. The tendency results from the fact that the overestimated precipitation occurrence with respect to OceanRAIN outweighs the underestimated conditional precipitation rates (Fig. 5.18). These cases highlight the influence of misrepresented false detections by HOAPS according to the increased number of clustered convective precipitation events in the tropics missed by RVs (discussed in Sect. 5.3.5). The number of grid-boxes in which OceanRAIN precipitation rates strongly exceed those of HOAPS has markedly decreased after applying the statistical adjustments to OceanRAIN.

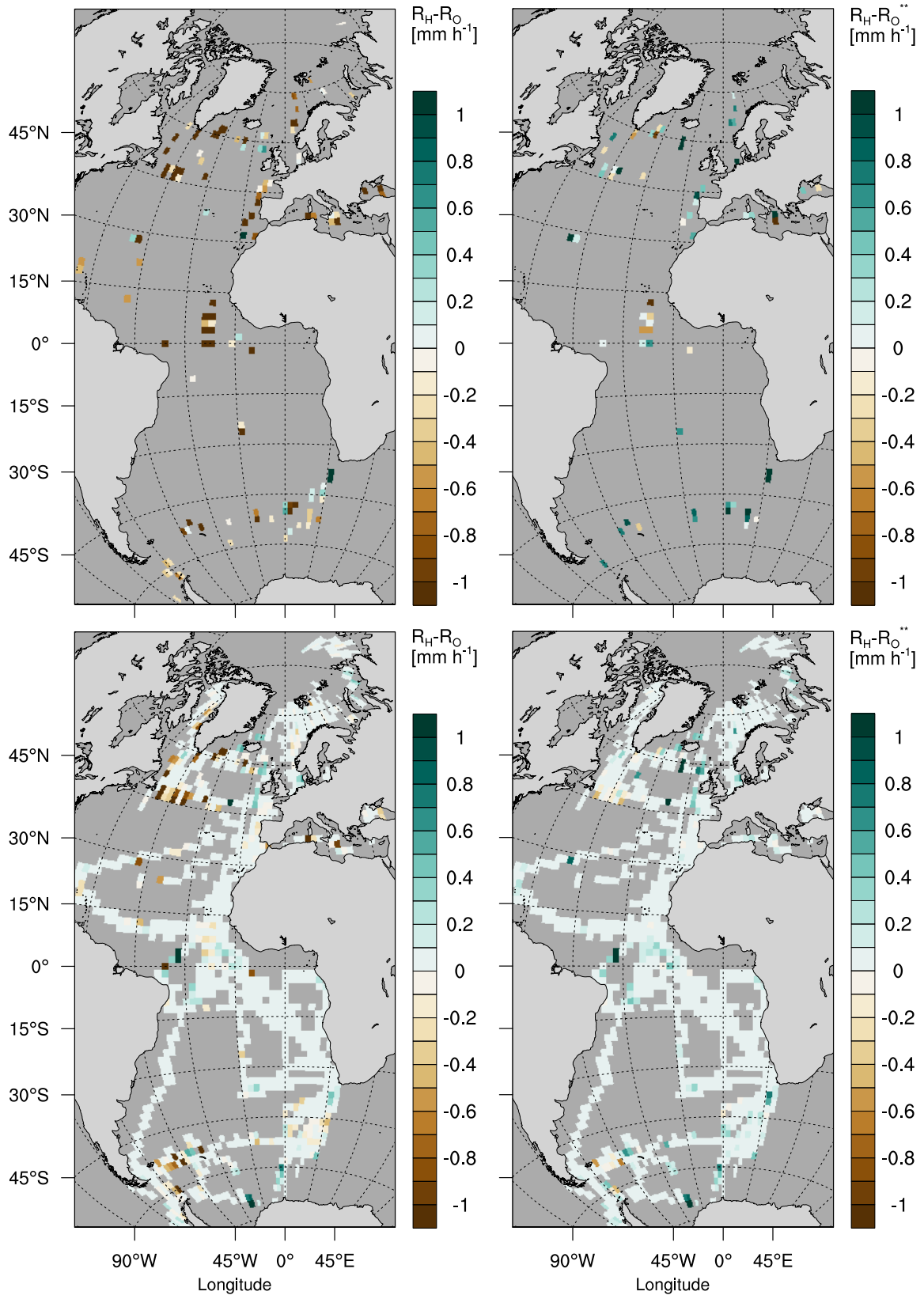


Figure 5.19: Maps of average precipitation-rate difference $R_H - R_O$ (left) and of $R_H - R_O^{**}$ (right) for hit cases (top) and all cases (bottom) in mm h^{-1} per 2° by 2° grid-box, all excluding $0 < R < 0.3$ mm h^{-1}.

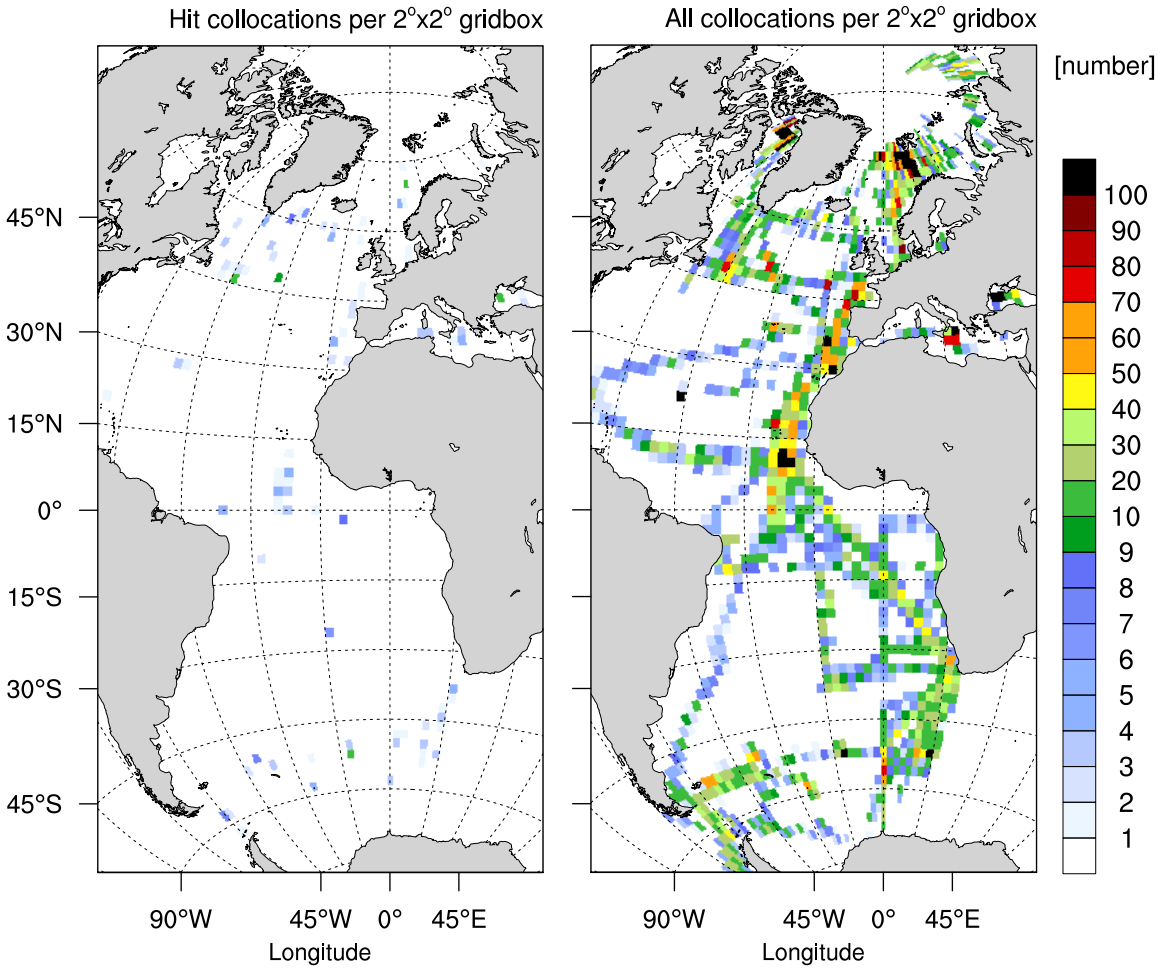


Figure 5.20: Maps depict number of cases per 2° by 2° grid-box used in Fig 5.19 for hit cases (left) and all collocations excluding $0 < R < 0.3$ mm h^{-1} (right).

5.3.7 OceanRAIN sampling: how long is long enough?

The global distribution of the HOAPS–OceanRAIN collocations entirely depends on the ship positions as well as the availability of HOAPS satellite data in consequence of available DMSP satellite overpasses. For such a heterogeneously sampled data set, the question of the minimum sample size cannot be addressed easily but depends on the objectives. A satellite validation requires the largest possible data sample size with respect to direct collocations. A climatological comparison, however, stronger depends on a homogeneous sampling with respect to seasons and multi-annual climate influences. The long data sampling is required in order to statistically cover the internal variability of precipitation in a certain region. Otherwise, recognized differences in precipitation over time cannot be attributed to long-term changes as long as the local precipitation variability remains unknown. For both kinds of comparisons, we discuss the status and requirements of the OceanRAIN data set.

In some regions of the Atlantic Ocean and the polar regions, the OceanRAIN database

has reached a reasonably large sample size. However, on a global scale, OceanRAIN cannot yet represent the long-term latitude-averaged precipitation rates of the gridded monthly HOAPS-G (blue-dotted line in Fig. 5.16, bottom). Largest differences appear in the southern-hemisphere subtropics and the northern high-latitudes and are almost entirely caused by the heterogeneous sampling of the RVs. Whereas most OceanRAIN RVs predominantly operated on the Atlantic Ocean, only few samples exist from the Pacific and Indian Ocean (not shown). This limitation implies that OceanRAIN in its current state does not serve for a globally representative comparison of oceanic precipitation. Achieving this ambitious goal would require at least an order of magnitude more RVs contributing to OceanRAIN, equally distributed over the global oceans. A higher number of RVs could ensure most homogeneous sampling in time and space that is essential to equally sample precipitation over all climate regions.

Besides spatial sampling, the temporal sampling plays an important role in a climatological comparison. In order to investigate the temporal sampling quality, we compare the collocated HOAPS-S precipitation rates R_H with those of the 25-year averaged precipitation rates of the monthly gridded HOAPS-G R_{H_G} (Fig. 5.21). Except for the

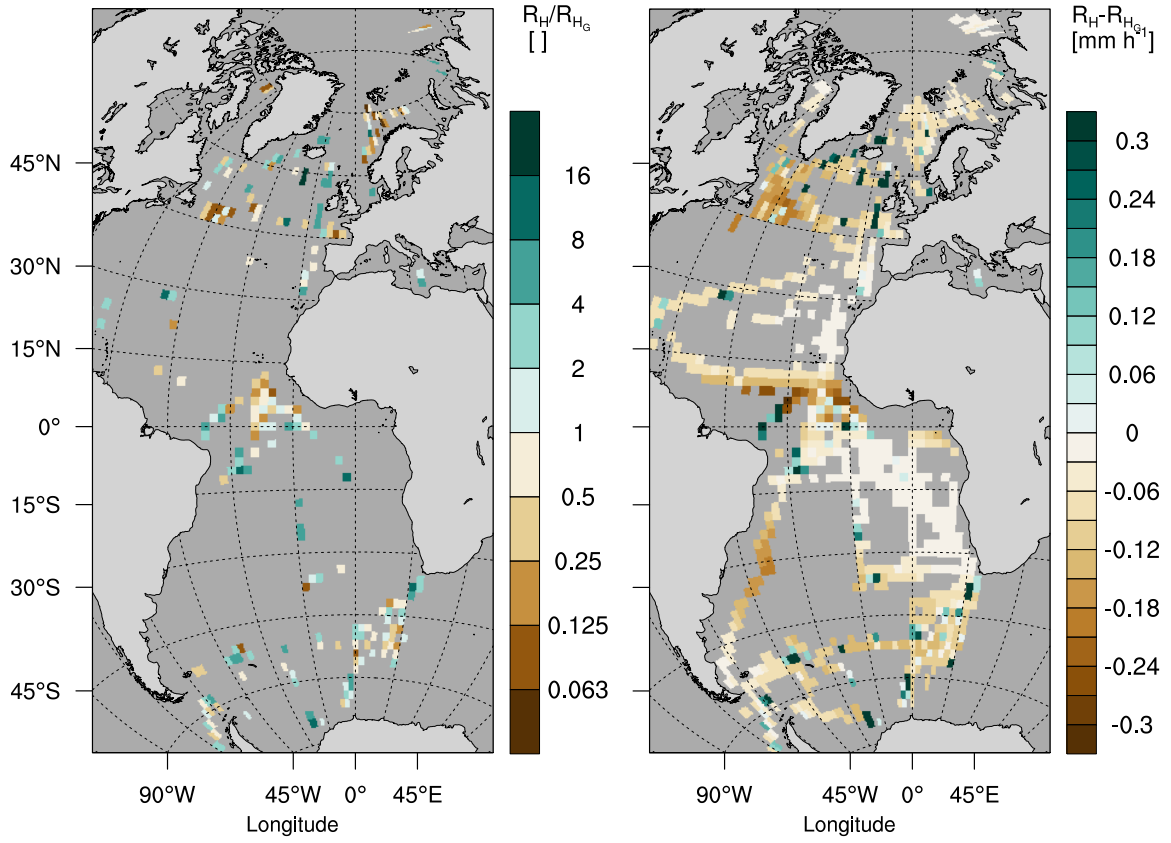


Figure 5.21: Map of the relative (left) and absolute (right) average precipitation-rate difference of HOAPS-S (R_H) and the climatological average over 25 years of HOAPS-G (R_{H_G}) for all 2° by 2° grid-boxes of collocated cases in mm h^{-1}. Note that non-precipitating boxes of R_H are excluded from precipitation-rate ratio on left.

same grid-boxes of particularly high HOAPS-S precipitation rates as in Figure 5.19 (bottom, right), widespread areas appear with on average slightly undersampled precipitation rates with respect to the long-term average of HOAPS-G. These undersampled areas include the tropical Atlantic northward the equator, the mid-latitude area south of Greenland and the southern-hemisphere West Atlantic. However, the southern-hemisphere West Atlantic HOAPS-S data is based on very few samples that did not experience any precipitation (Fig. 5.21 left). The undersampled area north of the equator and the oversampled area south of the equator in the tropical North Atlantic likely result from a seasonal bias because the majority of samples from RV *Polarstern*, RV *Maria S. Merian* and RV *Meteor* originates from the boreal spring months March, April and May (MAM) when the Intertropical Convergence Zone (ITCZ) with its peak precipitation lies farthest south. Most other of the sampled regions hold more patchy and nonuniform precipitation-rate differences that are shaped by short-term weather events during the RV sampling period rather than by the average climatic conditions. Predominantly, the mentioned areas coincide with areas of relatively low data sample size according to Figure 5.20. The limited data sample size also arises from the HOAPS ice-flag that strongly limits the number of available winter-season collocations at the poles. For all these given reasons, the seasonal sampling of the collocated HOAPS–OceanRAIN data sample does not yet suffice for a climatological comparison.

However, for a direct comparison as intended in this chapter, the HOAPS–OceanRAIN collocation data sample size is close to being large enough in order to robustly validate HOAPS satellite-retrieved data with OceanRAIN surface-reference data. As discussed in Section 5.3.5, the uncertainty of the collocated data sample lies in the same range of the difference in average conditional precipitation rates of HOAPS and OceanRAIN (cf. Fig. 5.17d). Besides a barely insufficient sampling, however, the excellent agreement in OceanRAIN and HOAPS at conditional precipitation rates below 2 mm h^{-1} indicates a good performance of the HOAPS precipitation algorithm as well as a sufficient sampling for low to medium-intensity precipitation rates. For the complete range of precipitation rates, a few years longer OceanRAIN data sample could underpin the outcome of this study while better representing the extremes in the precipitation rate.

In general, the natural variability of precipitation occurrence and precipitation intensity determine how long OceanRAIN would need to sample in a certain area to statistically fully represent the precipitation distribution. Thus, insufficient sampling obtains a potentially incomplete, partial image of the whole precipitation-rate distribution. The completion of this image requires long and continuous sampling that are key to statistically represent the rarely occurring intense precipitation events, which are the largest contributor to the average precipitation rate.

5.4 Summary and concluding remarks

This chapter focuses on directly comparing HOAPS-S precipitation satellite data to surface-based ship-board OceanRAIN as reference using a collocation method. For OceanRAIN, we applied the automatic PP distinction algorithm introduced in Chapter 3 and the adjustment to statistically reduce spatial-scale differences of the ship tracks to the precipitation estimates from the satellite sensors explained in Chapter 4. Most of this chapter focuses on adjusting OceanRAIN to HOAPS in order to assure a fair comparison between both data sets by reducing or neglecting the different spatial and temporal sampling, spatial representativeness and instrument/algorithm sensitivity between both precipitation data sets. A case study of few individual scenes compares HOAPS–OceanRAIN collocations against other state-of-the-art precipitation satellite products, followed by the actual comparison for latitude-averaged and gridded precipitation-rate differences as well as ranked precipitation-rate distributions.

The collocation of HOAPS and OceanRAIN serves to match both data sets for individual cases in space and time. As collocation boundaries we used ±20 km and ±30min, both with respect to the center of a HOAPS pixel and its overpass time of the respective DMSP satellite. These collocation boundaries conform to values chosen for similar ship-board disdrometer measurements in the literature (Bumke and Seltmann, 2012; Bumke et al., 2012). Stricter collocation boundaries seem plausible in regions dominated by convective precipitation but did not lead to significantly better skill scores except for the accuracy. However, the inner tropics and the polar regions experienced the largest tendency of better performing skill scores for ±20min as temporal boundary for collocation. Due to the very limited data sample size and the not consistently better performance of stricter collocation boundaries we keep ±20 km and ±30min.

The chosen collocation boundaries lead to a HOAPS-OceanRAIN sample of 24 990 cases of which 3.5 % agree on the presence of precipitation and 79.5 % on its absence. In 14.2 %, OceanRAIN detects precipitation not detected by HOAPS (miss) and in 2.8 % vice versa (false detection). The high fraction of HOAPS misses follows from the lower sensitivity to low precipitation rates because the HOAPS algorithm cuts off precipitation rates below 0.3 mm h^{-1} where precipitation is no longer distinguishable from noise. Excluding OceanRAIN precipitation rates below 0.3 mm h^{-1} after applying the spatial-scale adjustments, decreases the fraction of HOAPS misses from 14.2 % to about 1 % from 19 894 remaining collocations. This reduction clearly indicates that HOAPS well detects most precipitation in the collocated OceanRAIN measurements but cannot detect very light or small showers of precipitation as stated in Burdanowitz et al. (2015) for the subtropical North Atlantic.

The relatively high fraction of so called "false detections" by HOAPS with respect to OceanRAIN mainly results from the very limited ability by the OceanRAIN ship tracks to spatially represent the HOAPS pixel area. The highest false-detection fraction with respect to OceanRAIN appears in the inner tropics (10°S–10°N). The inner-tropic

latitude band frequently produces vigorous convection that is most challenging to be captured along a ship track. This aspect is not explicitly considered in the study of Bumke et al. (2016) to explain the high percentage of false detections by HOAPS of up to 20 % against comparable RV data of the equatorial Atlantic. In the tropical Atlantic region, RV *Meteor* missed precipitation along a 60-minute collocated track that was detected by HOAPS and other satellite products in an exemplary case on 30 November 2015. The demonstrated spatial limitation of the ship tracks in representing the areal precipitation distribution explains most of the "false detections" that, however, cannot be determined without ancillary information of that areal precipitation distribution.

Precipitation rates of cases where both HOAPS and OceanRAIN detected precipitation (hits), largely depend on the spatial-scale difference between both data sets. All collocated OceanRAIN precipitation rates according to the above mentioned boundaries are averaged (along-track average) to reach the largest possible coverage per satellite pixel. The averaging procedure inevitably leads to collocations in which OceanRAIN over- or undersamples the precipitation rate of the satellite pixel depending on the track and the areal precipitation distribution. The spatial sampling bias is reduced by adjusting OceanRAIN statistically, using the average duration of precipitation events and the median-normalized along-track averaged precipitation rate. Both statistical adjustments strongly narrow down the OceanRAIN precipitation-rate distribution but do not improve the weak correlation ($r = 0.35$ for linear regression in log-log space) in a case-by-case comparison of HOAPS–OceanRAIN collocations. Nevertheless, the relative and absolute differences between the HOAPS and OceanRAIN precipitation rates strongly decline. The weak correlation in the case-by-case comparison merely results from individual cases in which OceanRAIN represents the HOAPS area particularly poorly. Reasons for the poor representation include a relatively low number of OceanRAIN along-track samples per HOAPS collocation, very low ship speeds and misclassified PPs. However, excluding those critical collocations from the data sample does not only reduce the number of unrepresentative cases but predominantly excludes cases where the precipitation rate of OceanRAIN strongly exceeds that of HOAPS. However, as for the HOAPS "false detections", our study lacks information of the spatial precipitation distribution within the HOAPS pixel in order to unveil those collocations in which OceanRAIN underrepresents precipitation in the area. From the above mentioned 3 factors, only the fraction of low ship speed samples increases with increasing overestimation of HOAPS precipitation rates with respect to OceanRAIN while the other factors do not help to identify cases of underrepresented OceanRAIN precipitation.

The latitude-averaged HOAPS precipitation rate of all collocations exceeds that of OceanRAIN by a factor of 2 to 3. However, re-classifying cases of false detections by HOAPS to hit cases increases the OceanRAIN precipitation rate substantially. In the tropics, HOAPS underestimates the precipitation rate by 10 to 15 % according to the modified OceanRAIN sample with re-classified false detections. Though a re-classification of false detections into hits seems plausible, not all false detections might

have been caused by an OceanRAIN RV missing the precipitation. Therefore, the latitude-averaged precipitation rates should be interpreted with caution.

Excluding non-precipitation collocations, HOAPS tends to underestimate the average precipitation rate by 10 to 30 % in the inner tropics. In the mid-latitudes, HOAPS tends to overestimate the average precipitation rate by more than 50 %; however based on very few data samples. The one-sided exclusion of predominantly higher OceanRAIN precipitation rates compared to HOAPS could explain a part of the large difference. In the subtropics, HOAPS agrees well with OceanRAIN. Overall, considering the latitudinal average does not suffice in order to diagnose conclusive limitations in the HOAPS retrieval or algorithm.

Though no direct implications for the HOAPS performance can be derived from the latitudinal averages in different climate regimes, precipitation rates can be compared statistically. According to the ranked conditional precipitation-rate distribution for each data set (without non-precipitating cases), HOAPS and OceanRAIN closely resemble each other for conditional precipitation rates below 2 mm h^{-1}. For this comparison, OceanRAIN precipitation rates below 0.3 mm h^{-1} were excluded to reflect the lower sensitivity of HOAPS in detecting precipitation. Above 2 mm h^{-1}, OceanRAIN precipitation rates contribute slightly stronger to the average precipitation rate than those of HOAPS. In particular, the high conditional precipitation rates in OceanRAIN explain the about 13 % higher average conditional precipitation rate compared to HOAPS. Due to the relatively low sample size of HOAPS ($n = 846$) and OceanRAIN (346), the uncertainty lies in the same range. As a measure of robustness, statistical resampling of 100 realizations yields an inter-quartile range of less than 15 % for OceanRAIN and of less than 10 % for HOAPS. These uncertainties highlight the need for longer sampling in order to statistically better resolve the high precipitation rates that mainly contribute to average precipitation rates.

The average HOAPS precipitation rate for hit cases exceeds that of the false detections by a factor of 2. The large difference between both distributions highlights the need to either correct for the fraction of falsely identified "false" detections or consider the effect in the HOAPS–OceanRAIN comparison. A lack of information on the areal precipitation distribution within the HOAPS pixel only leaves the option to estimate the effect of collocations in which the ship track misses precipitation detected by HOAPS. Stricter collocation boundaries in regions dominated by small-scale convective precipitation could help to reduce the number of cases in which an RV misses precipitation and assigns a false detection to a satellite measurement. The adaptation of collocation boundaries goes to the expense of a reduced along-track sample size per collocation and, thus, requires careful checking of the conditions under which a validation is conducted.

Chapter 6

Summary and conclusions

This thesis investigates the point-to-area validation of precipitation derived from passive microwave satellite sensors using shipboard disdrometer data on the example of the HOAPS satellite climatology and the OceanRAIN ship database. The main focus lies on the careful representativeness adjustment of the OceanRAIN surface reference data to the characteristics of HOAPS in order to obtain a fair comparison. The adjustments, summarized hereafter, mainly relate to the special precipitation characteristics: precipitation phase (PP) changes, high spatio-temporal variability and the non-Gaussian frequency distribution of precipitation rates.

6.1 Summary of main results

The PP represents an important piece of information for an accurate retrieval of the precipitation rate from the disdrometer-measured PSD. We developed a novel PP distinction algorithm that ingests various parameters from ancillary data available from RVs in OceanRAIN. A combination of air temperature, relative humidity and the 99th percentile of the particle diameter from the disdrometer turns out to predict best the PP compared to a manually determined PP that is mainly based on 3-hourly present weather reports from 4 years of RV *Polarstern*. The combination of these 3 meteorological parameters reaches an accuracy of 91 % for the distinction of rain and snow using one single PP probability distribution. However, using two independent PP probability distributions allows to distinguish between rain, snow and mixed-phase precipitation. This novel approach of using two independent PP probability distributions reaches an accuracy of 81 % with a slight tendency to overpredict the occurrence of snowfall (bias score: 0.93). Applying the newly developed automatic PP distinction algorithm to OceanRAIN considerably speeds up the data processing and sets the ground to provide an independent and efficient surface-based precipitation reference database.

The high variability of precipitation in both time and space poses a particular challenge to the HOAPS validation due to largely different temporal and spatial resolutions

of both data sets. The influence of the resulting point-to-area effect is studied independently using 2 months of S-Pol radar data from the subtropical island of Barbuda during the RICO campaign. During RICO, the S-Pol radar mainly sampled small-scale shallow-convective precipitation as well as some more widespread precipitation events (Burdanowitz et al., 2015). In the S-Pol domain, we define 4 areas in which we place several randomly chosen ship-like tracks of 60 radar pixels in length corresponding to about 24 km. These simulated ship tracks detect precipitation in more than 14 % while missing it in less than 18 % of the scenes when neglecting scenes of below 2 % areal precipitation coverage. In contrast to the synthetic ship tracks put into the S-Pol radar domain, the detection in reality is mainly hampered by low ship speed, heterogeneous areal precipitation distribution as well as the detection sensitivity of the PMW sensor for particularly low area coverage. On average, ship tracks tend to slightly overestimate the spatial extent of precipitation events while underestimating their intensity. This bias can be quantified and corrected by defining the average precipitation event duration T_E (sum of event durations per track divided by their number). On average, low T_E coincide with strongly undersampled events while medium to high T_E coincide with oversampled events, which can be expressed in a statistical adjustment function of exponential type. Applying the T_E-adjustment reduces the systematic bias by about 80 % with respect to the sum of squared errors. However, a remaining bias in the relative precipitation-rate difference provokes a second adjustment using the along-track precipitation rate itself, normalized by its median. The normalization ensures an independence from the radar data when applying this second statistical adjustment to reduce the precipitation-rate bias. After the statistical adjustment, a small bias remains for precipitation rates below 1.5 mm h^{-1} that the track underestimates by less than 0.2 mm h^{-1} with respect to the area. This remaining bias can, however, not be corrected statistically. After both statistical adjustments, the random error of the precipitation-rate difference in form of the standard deviation equals the track precipitation rate at about 1 mm h^{-1} and decreases in relative terms for higher precipitation rates. The bias adjustments increase the standard deviation by up to 70 % at about 0.8 mm h^{-1} indicating a shift of the systematic error into the random error. Simultaneously, the distribution of the track–area precipitation-rate difference progressively approximates a Gaussian distribution, particularly at high precipitation rates, which indicates the remaining bias to be largest for low precipitation rates.

The generally non-Gaussian frequency distribution of precipitation rates reflects a substantially higher occurrence of low precipitation rates compared to high precipitation rates with twofold implications. First, the high occurrence of low precipitation rates requires a high measurement sensitivity in order to obtain a realistic precipitation occurrence. Second, the low occurrence of high precipitation rates calls for sufficiently long sampling in order to resolve the extreme events that mainly contribute to the average precipitation rate. Despite the low sensitivity of the HOAPS algorithm to resolve precipitation rates less or equal to 0.3 mm h^{-1}, in 75 % of all hit cases, the OceanRAIN precipitation rate remains below that threshold. The relatively low sensi-

tivity of HOAPS explains most of the low probability of detection of 0.2–0.4 in regions where light rain and snowfall dominate the precipitation occurrence. The percentage of false detections by HOAPS exceeds 5 % only in the inner tropics. A certain fraction of false detections likely represents misclassified hit cases when the RV missed the precipitation in a HOAPS pixel. A case study with other satellite data sets could give evidence for such a case. However, these cases strongly impact the average precipitation rate: the OceanRAIN precipitation rate of hit cases is underestimated while the HOAPS precipitation rate of hit cases is overestimated. For that reason, HOAPS strongly exceeds average precipitation rates when considering hit cases only. Comparing hits and misses of OceanRAIN with hits and false detections of HOAPS yields a fairer comparison. For precipitation rates of at least 0.3 mm h^{-1}, HOAPS underestimates the precipitation rate by about 13 %, mainly due to underestimated high precipitation rates. However, the uncertainty estimated by resampling lies in the same range as the difference between HOAPS and OceanRAIN. Together with the stronger contribution of high precipitation rates to the estimated HOAPS–OceanRAIN difference, more collocations are required for a more robust statement on the performance of HOAPS. The derived statistical adjustments prove to be useful for the validation of satellite data using along-track ship data as reference; however, they cannot improve the representativeness of every single collocation and might even exacerbate scenes well represented by the RVs. Nevertheless, applying both statistical adjustments paves the way towards a more reliable point-to-area validation by correcting for extreme over- or underrepresentation of precipitation along a ship track.

6.2 Implications

This thesis investigated the influence of spatial-scale differences on the validation of precipitation from passive microwave satellite data using ship-based disdrometer measurements. The following implications follow from the obtained results.

Implications for HOAPS and other satellite data sets

HOAPS performed well in detecting precipitation in most oceanic regions. The amount of false detections is rather small so that HOAPS can be used for precipitation detection with the limitation of a reduced sensitivity to very light precipitation. The precipitation rates are on average in the range of the observations with a tendency to underestimate precipitation in the inner tropics. As for light precipitation, this is a result of the relatively low spatial resolution of about 50 km. Studies that aim for investigating extreme precipitation or small-scale precipitation features might want to choose a multi-satellite data set with higher temporal resolution such as the Global Satellite Mapping of Precipitation (GSMaP; Okamoto et al., 2005) at hourly resolution, the CMORPH (Joyce et al., 2004), the GPM IMERG (Huffman, 2015) at half-hourly resolution, or

the daytime-only MSG-CPP (Roebeling et al., 2006) at 15 min resolution, respectively. For light precipitation, Cloudsat (Stephens et al., 2002) is recommended with its very high sensitivity for snow and drizzle. Owing to the independence from other data sets and model assumptions, and the long SSMI/SSMIS record, HOAPS serves better for climate model validation and long-term precipitation monitoring.

Implications for OceanRAIN

The novel PP distinction algorithm to distinguish between rainfall, snowfall and mixed-phase precipitation achieves a speedup in the OceanRAIN data processing. At the same time, the algorithm reduces the dependence on human observers and adds physical information on the PP, which can change frequently in-between two 3-hourly observations. Nevertheless, a manageable number of uncertain cases requires visual inspection to detect and correctly assign PP changes. We recommend inspecting cases with a PP probability between 0.4 and 0.6. For application over land, the PP distinction algorithm would need to be re-calibrated against land-based reference data. However, the algorithm can easily be transferred to other oceanic precipitation products.

In this thesis, OceanRAIN demonstrates its utility in validating precipitation data over the ocean. Besides other PMW-based satellite data sets, IR/VIS satellite products such as MSG-CPP could profit from a validation with OceanRAIN. A number of studies exist that validate MSG-CPP over land (e.g. Roebeling and Holleman, 2009; Wolters et al., 2011) but MSG-CPP has not yet been validated over the ocean. The case study of this thesis considers a number of cases in which the majority of precipitation detections seems promising; however, an in-depth validation could assess the ability of MSG-CPP to detect and correctly quantify precipitation in various climate zones. Such a study is easier to conduct due to the markedly higher spatial and temporal resolution of MSG-CPP compared to PMW satellite products.

Implications for the sample size of validation data

The sample size of precipitation reference data sets is crucial to derive meaningful implications for the validated satellite data set. Therein, OceanRAIN represents a suitable base for ocean precipitation validation using high-quality disdrometers. The OceanRAIN sample size suffices in order to derive a tendency for the performance of a satellite data set as well as to conduct high-quality case studies. Nevertheless, the current OceanRAIN sample size lies on the border to become useful for the broader validation of precipitation satellite products. The sample size and distribution of the OceanRAIN database does not yet satisfy the requirements for deriving global mean precipitation estimates. In particular, more data samples need to be collected from the Indian and the Pacific Ocean while sufficient extreme precipitation events should be contained. In these regions, RV *Sonne* (Pacific Ocean), RV *Investigator* (Southern

Pacific and Indian Ocean) and RV *Roger Revelle* (tropical Pacific) collect data available in OceanRAIN after the end of the current HOAPS version of December 2015. Due to the lack of other high-quality systematic precipitation data sets over the ocean, the unique OceanRAIN database is worth to be continued and extended.

Implications for the point-to-area problem

For precipitation, the main challenge in the point-to-area problem actually lies in the adjustment of the reference data set in order to represent the satellite data set without distorting the reference data. For this purpose, the statistical adjustments developed in this thesis use precipitation event duration and median-normalized precipitation rate as data-independent parameters from an independent radar data set. That the statistical adjustments are based on one single island-based radar in the subtropics represents one of the major shortcomings. It remains open to which extent radar data from other climate zones with frontal and frozen precipitation would lead to different results. However, the basic statistical feature of over-/undersampling of precipitation from short/long events should remain the same. A second aspect not considered in this thesis arises from the neglected cloud movement that mainly influences the spatial representativeness of the ship track. The resulting vector of ship velocity and cloud movement could act as a measure of quality for the spatial representativeness of an along-track collocation. In this respect, variable collocation boundaries could significantly reduce the number of precipitation events in the satellite pixel that are missed by the RVs. Specifically in the inner tropics, ± 20 min seems plausible as a maximum time lag from a satellite overpass and ± 18 km as a maximum distance from the center of a HOAPS pixel, respectively. As an alternative, a convective–stratiform distinction could assist to decide for reasonable collocation boundaries using a lookup table.

6.3 Final remarks

In conclusion, this thesis reveals a good performance of HOAPS both in detecting and quantifying precipitation. At the same time, these results are a first step towards an error decomposition of the HOAPS precipitation parameter. As of now, the uncertainty following from the limited number of HOAPS–OceanRAIN collocations remains too high to conclude robust error estimates or implications for the HOAPS algorithm development. Besides this direct outcome, a valuable algorithm could be developed to predict the precipitation phase using atmospheric parameters available from most ship measurements. For the satellite product validation, careful statistical adjustments have been derived from independent radar data that help reducing the different spatial-scale representativeness of precipitation between ship tracks and PMW satellite pixels. These results provide valuable information for the upcoming validation of precipitation

satellite products within the GPM community as well as for IR and potential future geo-microwave sensor products.

References

Adler, R. F., Gu, G., and Huffman, G. J. (2012). Estimating climatological bias errors for the global precipitation climatology project (GPCP). *J. Appl. Meteor. Clim.*, 51(1):84–99.

Amitai, E., Petersen, W., Llort, X., and Vasiloff, S. (2012). Multiplatform Comparisons of Rain Intensity for Extreme Precipitation Events. *IEEE Transactions on Geoscience and Remote Sensing*, 50(3):675–686.

Andersson, A., Bakan, S., and Graßl, H. (2010a). Satellite derived precipitation and freshwater flux variability and its dependence on the North Atlantic Oscillation. *Tellus A*, 62(4):453–468.

Andersson, A., Fennig, K., Klepp, C., Bakan, S., Graßl, H., and Schulz, J. (2010b). The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data – HOAPS-3. *Earth Sys. Sci. Data*, 2(2):215–234.

Andersson, A., Klepp, C., Fennig, K., Bakan, S., Graßl, H., and Schulz, J. (2011). Evaluation of HOAPS-3 Ocean Surface Freshwater Flux Components. *J. Appl. Meteor. Clim.*, 50(2):379–398.

Atlas, D. and Ulbrich, C. (1974). The physical basis for attenuation-rainfall relationships and the measurement of rainfall parameters by combined attenuation and radar methods. *J. Rech. Atmos.*, 8(1–2):275–298.

Atlas, D., Ulbrich, C. W., Marks, F. D., Black, R. A., Amitai, E., Willis, P. T., and Samsury, C. E. (2000). Partitioning tropical oceanic convective and stratiform rains by draft strength. *J. Geophys. Res.: Atmospheres*, 105(D2):2259–2267.

Battaglia, A., Rustemeier, E., Tokay, A., Blahak, U., and Simmer, C. (2010). PARSIVEL Snow Observations: A Critical Assessment. *J. Atmos. Ocean. Tech.*, 27(2):333–344.

Bauer, P., Lopez, P., Benedetti, A., Salmond, D., and Moreau, E. (2006). Implementation of 1d+4d-Var assimilation of precipitation-affected microwave radiances at ECMWF. I: 1d-Var. *Quart. J. Roy. Meteor. Soc.*, 132(620):2277–2306.

Bentley, W. (1904). Studies of raindrops and raindrop phenomena. *Mon. Weather Rev.*, 32(10):450–456.

Berry, D. I. and Kent, E. C. (2011). Air–Sea fluxes from ICOADS: the construction of a new gridded dataset with uncertainty estimates. *Int. J. Clim.*, 31(7):987–1001.

Bojinski, S., Verstraete, M., Peterson, T. C., Richter, C., Simmons, A., and Zemp, M. (2014). The Concept of Essential Climate Variables in Support of Climate Research, Applications, and Policy. *Bull. Amer. Meteor. Soc.*, 95(9):1431–1443.

Bourlès, B., Lumpkin, R., McPhaden, M. J., Hernandez, F., Nobre, P., Campos, E., Yu, L., Planton, S., Busalacchi, A., Moura, A. D., Servain, J., and Trotte, J. (2008). The Pirata Program: History, Accomplishments, and Future Directions. *Bull. Amer. Meteor. Soc.*, 89(8):1111–1125.

Bringi, V. N., Williams, C. R., Thurai, M., and May, P. T. (2009). Using Dual-Polarized Radar and Dual-Frequency Profiler for DSD Characterization: A Case Study from Darwin, Australia. *J. Atmos. Ocean. Tech.*, 26(10):2107–2122.

Bumke, K. (2016). Validation of ERA-Interim Precipitation Estimates over the Baltic Sea. *Atmosphere*, 7(6):82.

Bumke, K., Fennig, K., Strehz, A., Mecking, R., and Schröder, M. (2012). HOAPS precipitation validation with ship-borne rain gauge measurements over the Baltic Sea. *Tellus A*, 64(0).

Bumke, K., König-Langlo, G., Kinzel, J., and Schröder, M. (2016). HOAPS and ERA-interim precipitation over the sea: validation against shipboard in situ measurements. *Atm. Meas. Techn.*, 9(5):2409–2423.

Bumke, K. and Seltmann, J. (2012). Analysis of measured drop size spectra over land and sea. *ISRN Meteor.*, 2012:1–10.

Burdanowitz, J., Nuijens, L., Stevens, B., and Klepp, C. (2015). Evaluating light rain from satellite- and ground-based remote sensing data over the subtropical North Atlantic. *J. Appl. Meteor. Clim.*, 54(3):556–572.

Chandrasekar, V., Bringi, V. N., Rutledge, S. A., Hou, A., Smith, E., Jackson, G. S., Gorgucci, E., and Petersen, W. A. (2008). Potential Role Of Dual-Polarization Radar In The Validation Of Satellite Precipitation Measurements: Rationale and Opportunities. *Bull. Amer. Meteor. Soc.*, 89(8):1127–1145.

Ciach, G. J. and Krajewski, W. F. (1999). On the estimation of radar rainfall error variance. *Adv. in Water Resources*, 22(6):585–595.

Clemens, M. (2002). Machbarkeitsstudie zur räumlichen Niederschlagsanalyse aus Schiffsmessungen über der Ostsee. *Berichte aus dem Institut für Meereskunde an der Christian-Albrechts-Universität Kiel*, (327):1–50.

Dai, A. (2001). Global precipitation and thunderstorm frequencies. part II: Diurnal variations. *J. Climate*, 14(6):1112–1128.

Dai, A. (2008). Temperature and pressure dependence of the rain-snow phase transition over land and ocean. *Geophys. Res. Lett.*, 35(12):L12802.

Fennig, K., Andersson, A., Bakan, S., Klepp, C.-P., and Schröder, M. (2012). Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data - HOAPS 3.2 - Monthly Means / 6-Hourly Composites. *Satellite Application Facility on Climate Monitoring (CM SAF)*.

Fennig, K., Andersson, A., and Schröder, M. (2015). Fundamental Climate Data Record of SSM/I / SSMIS Brightness Temperatures.

Ferraro, R. R. (1997). Special sensor microwave imager derived global rainfall estimates for climatological applications. *J. Geophys. Res.: Atm.*, 102(D14):16715–16735.

Ferraro, R. R., Weng, F., Grody, N. C., and Zhao, L. (2000). Precipitation characteristics over land from the NOAA-15 AMSU sensor. *Geophys. Res. Ltr.*, 27(17):2669–2672.

Friedrich, K., Higgins, S., Masters, F. J., and Lopez, C. R. (2013). Articulating and stationary PARSIVEL disdrometer measurements in conditions with strong winds and heavy rainfall. *J. Atmos. Ocean. Tech.*, 30(9):2063–2080.

Froidurot, S., Zin, I., Hingray, B., and Gautheron, A. (2014). Sensitivity of precipitation phase over the swiss alps to different meteorological variables. *J. Hydrometeor.*, 15(2):685–696.

Fuchs, T., Rapp, J., Rubel, F., and Rudolf, B. (2001). Correction of synoptic precipitation observations due to systematic measuring errors with special regard to precipitation phases. *Phys. Chem. Earth Pt. B*, 26(9):689–693.

Gjertsen, U. and Ødegaard, V. (2005). The water phase of precipitation–a comparison between observed, estimated and predicted values. *Atmos. Res.*, 77(1-4):218–231.

Goodison, B. E. (1978). Accuracy of canadian snow gage measurements. *J. Appl. Meteor.*, 17(10):1542–1548.

Großklaus, M. (1996). *Niederschlagsmessung auf dem Ozean von fahrenden Schiffen*. PhD thesis, Institut für Meereskunde, Christian-Albrechts-Universität Kiel, Kiel, Germany.

Habib, E. and Krajewski, W. F. (2002). Uncertainty analysis of the TRMM ground-validation radar-rainfall products: Application to the TEFLUN-B field campaign. *J. Appl. Meteor.*, 41(5):558.

Hasse, L., Grossklaus, M., Uhlig, K., and Timm, P. (1998). A ship rain gauge for use in high wind speeds. *J. Atmos. Ocean. Tech.*, 15(2):380–386.

Hayes, S. P., Mangum, L. J., Picaut, J., Sumi, A., and Takeuchi, K. (1991). TOGA-TAO: A Moored Array for Real-time Measurements in the Tropical Pacific Ocean. *Bull. Amer. Meteor. Soc.*, 72(3):339–347.

Hogan, A. W. (1994). Objective Estimates of Airborne Snow Properties. *J. Atmos. Ocean. Tech.*, 11(2):432–444.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2013). The Global Precipitation Measurement (GPM) Mission. *Bull. Amer. Meteor. Soc.*, pages 701–722.

Houze, R. A. (1997). Stratiform Precipitation in Regions of Convection: A Meteorological Paradox? *Bull. Amer. Meteor. Soc.*, 78(10):2179–2196.

Hsu, K.-l., Gao, X., Sorooshian, S., and Gupta, H. V. (1997). Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.*, 36(9):1176–1190.

Huffman, G. (2015). GPM merged radiometer global .1 x.1 deg half-hourly precipitation retrieval – NASA Goddard Earth Sciences Data and Information Services Center.

Huffman, G. J., Bolvin, D. T., and Nelkin, E. J. (2015). Integrated multi-satellite retrievals for GPM (IMERG) technical documentation. `https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf`. [accessed on 29 Sep 2016].

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., and Stocker, E. F. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor.*, 8(1):38–55.

IPCC (2014a). Introduction. In Intergovernmental Panel on Climate Change, editor, *Climate Change 2013 - The Physical Science Basis*, pages 119–158. Cambridge University Press, Cambridge.

IPCC (2014b). Observations: Ocean Pages. In Intergovernmental Panel on Climate Change, editor, *Climate Change 2013 - The Physical Science Basis*, pages 255–316. Cambridge University Press, Cambridge.

Ishizaka, M., Motoyoshi, H., Nakai, S., Shiina, T., Kumakura, T., and Muramoto, K.-i. (2013). A new method for identifying the main type of solid hydrometeors contributing to snowfall from measured size-fall speed relationship. *J. Meteor. Soc. Jpn. Ser. II*, 91(6):747–762.

Joss, J. and Waldvogel, A. (1967). Ein Spektrograph für Niederschlagstropfen mit automatischer Auswertung (A spectrograph for rain drops with automatical analysis). *Pure Appl. Geophys.*, 68:240–246.

Joyce, R. and Janowiak, J. (2016). NOAA CPC morphing technique (CMORPH). `http://www.cpc.noaa.gov/products/janowiak/cmorph_description.html`. [accessed on 21 Aug 2016].

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). Cmorph: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, 5(3):487 – 503.

Keeler, R., Lutz, J., and Vivekanandan, J. (2000). S-pol: NCAR's polarimetric doppler research radar. volume 4, pages 1570–1573. IEEE.

Kent, E. C., Ball, G., Berry, D. I., Fletcher, J., Hall, A., North, S., and Woodruff, S. (2010). The Voluntary Observing Ship (VOS) Scheme. pages 518–528. European Space Agency.

Kidd, C. and Levizzani, V. (2011). Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.*, 15(4):1109–1116.

Kinzel, J., Fennig, K., Schröder, M., Andersson, A., Bumke, K., and Hollmann, R. (2016). Decomposition of Random Errors Inherent to HOAPS-3.2 Near-Surface Humidity Estimates Using Multiple Triple Collocation Analysis. *J. Atmos. Oceanic Technol.*

Klepp, C. (2015). The oceanic shipboard precipitation measurement network for surface validation — OceanRAIN. *Atmos. Res.*, 163:74–90.

Klepp, C., Bumke, K., Bakan, S., and Bauer, P. (2010). Ground validation of oceanic snowfall detection in satellite climatologies during LOFZY. *Tellus A*, 62(4):469–480.

Knight, C. A. and Miller, L. J. (1998). Early radar echoes from small, warm cumulus: Bragg and hydrometeor scattering. *J. Atmos. Sci.*, 55(18):2974–2992.

KNMI (2011). MSG cloud physical properties (CPP) algorithm. `http://msgcpp.knmi.nl/mediawiki/index.php/MSGCPP_product_description`. [accessed on 28 Sep 2016].

Koistinen, J. and Saltikoff, E. (1998). Eperience of customer products of accumulated snow, sleet and rain. *COST75 Advanced Weather Radar Systems*, pages 397–406.

Kubota, T., Ushio, T., Shige, S., Kida, S., Kachi, M., and Okamoto, K. (2009). Verification of High-Resolution Satellite-Based Rainfall Estimates around Japan Using a Gauge-Calibrated Ground-Radar Dataset. *J. Meteor. Soc. Japan*, 87A:203–222.

Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., Ferraro, R., Petty, G., Shin, D.-B., and Wilheit, T. T. (2001). The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors. *J. Appl. Meteor.*, 40(11):1801–1820.

Laing, A. G. and Michael Fritsch, J. (1997). The global population of mesoscale convective complexes. *Quart. J. Royal Meteor. Soc.*, 123(538):389–405.

Lambrigtsen, B. H., Brown, S. T., Dinardo, S. J., Gaier, T. C., Kangaslahti, P. P., Tanner, A. B., Piepmeier, J. R., Ruf, C. S., Gross, S. M., Musko, S., and Rogacki, S. (2006). GeoSTAR: a microwave sounder for geostationary applications. page 63610K.

Lanza, L. G. and Vuerich, E. (2009). The WMO Field Intercomparison of Rain Intensity Gauges. *Atmos. Res.*, 94(4):534–543.

Lee, G. W. and Zawadzki, I. (2005). Variability of Drop Size Distributions: Time-Scale Dependence of the Variability and Its Effects on Rain Estimation. *J. Appl. Meteor.*, 44(2):241–255.

Lempio, G. E., Bumke, K., and Macke, A. (2007). Measurement of solid precipitation with an optical disdrometer. *Advances in Geosciences*, 10:91–97.

Levizzani, V., Bauer, P., and Turk, F. J. (2007). *Measuring precipitation from space EURAINSAT and the future*. Springer, Dordrecht, the Netherlands.

Li, M. and Shao, Q. (2010). An improved statistical approach to merge satellite rainfall estimates and raingauge data. *J. Hydrology*, 385(1-4):51–64.

Löffler-Mang, M. and Joss, J. (2000). An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. Ocean. Tech.*, 17(2):130–139.

Maggioni, V., Meyers, P. C., and Robinson, M. D. (2016). A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. *J. Hydrometeor.*, 17(4):1101–1117.

Merenti-Välimäki, H.-L., Lönnqvist, J., and Laininen, P. (2001). Present weather: comparing human observations and one type of automated sensor. *Meteor. Appl.*, 8(4):491–496.

Michaelides, S., Levizzani, V., Anagnostou, E., Bauer, P., Kasparis, T., and Lane, J. (2009). Precipitation: Measurement, remote sensing, climatology and modeling. *Atm. Res.*, 94(4):512–533.

Michelson, D. B. (2004). Systematic correction of precipitation gauge observations using analyzed meteorological variables. *J. Hydrol.*, 290(3-4):161–177.

Miriovsky, B. J., Bradley, A. A., Eichinger, W. B., Krajewski, W. F., Kruger, A., Nelson, B. R., Creutin, J.-D., Lapetite, J.-M., Lee, G. W., Zawadzkt, I., and Ogden, F. L. (2004). An experimental study of small-scale variability of radar reflectivity using disdrometer observations. *J. Appl. Meteor.*, 43(1):106 – 118.

Nuijens, L., Serikov, I., Hirsch, L., Lonitz, K., and Stevens, B. (2014). The distribution and variability of low-level cloud in the north atlantic trades: Distribution and variability of low-level cloud in the trades. *Q. J. R. Meteor. Soc.*, 140(684):2364–2374.

Nuijens, L., Stevens, B., and Siebesma, A. P. (2009). The environment of precipitating shallow cumulus convection. *J. Atmos. Sci.*, 66(7):1962–1979.

OceanRAIN (2015). The Ocean Rain And Ice-phase precipitation measurement Network (OceanRAIN). `http://oceanrain.org`. [accessed on 29 Sep 2016].

OceanSITES (2016). OceanSITES. `http://www.oceansites.org`. [accessed on 1 Oct 2016].

Okamoto, K., Ushio, T., Iguchi, T., Takahashi, N., and Iwanami, K. (2005). The global satellite mapping of precipitation (GSMaP) project. volume 5, pages 3414–3416. IEEE.

Petty, G. W. (1995). Frequencies and characteristics of global oceanic precipitation from shipboard present-weather reports. *B. Am. Meteor. Soc.*, 76(9):1593–1616.

Rauber, R. M., Ochs, H. T., Di Girolamo, L., Göke, S., Snodgrass, E., Stevens, B., Knight, C., Jensen, J. B., Lenschow, D. H., Rilling, R. A., Rogers, D. C., Stith, J. L., Albrecht, B. A., Zuidema, P., Blyth, A. M., Fairall, C. W., Brewer, W. A., Tucker, S., Lasher-Trapp, S. G., Mayol-Bracero, O. L., Vali, G., Geerts, B., Anderson, J. R., Baker, B. A., Lawson, R. P., Bandy, A. R., Thornton, D. C., Burnet, E., Brenguier, J.-L., Gomes, L., Brown, P. R. A., Chuang, P., Cotton, W. R., Gerber, H., Heikes, B. G., Hudson, J. G., Kollias, P., Krueger, S. K., Nuijens, L., O'Sullivan, D. W., Siebesma, A. P., and Twohy, C. H. (2007). Rain in shallow cumulus over the ocean: The RICO campaign. *Bull. Amer. Meteor. Soc.*, 88(12):1912–1928.

Roebeling, R. A., Feijt, A. J., and Stammes, P. (2006). Cloud property retrievals for climate monitoring: Implications of differences between Spinning Enhanced Visible and Infrared Imager (SEVIRI) on METEOSAT-8 and Advanced Very High Resolution Radiometer (AVHRR) on NOAA-17. *J. Geophys. Res.*, 111(D20).

Roebeling, R. A. and Holleman, I. (2009). SEVIRI rainfall retrieval and validation using weather radar observations. *J. Geophys. Res.*, 114(D21).

Roebeling, R. A., Wolters, E. L. A., Meirink, J. F., and Leijnse, H. (2012). Triple Collocation of Summer Precipitation Retrievals from SEVIRI over Europe with Gridded Rain Gauge and Weather Radar Data. *J. Hydrometeor.*, 13(5):1552–1566.

Schmid, W. and Mathis, A. (2004). Validation of methods to detect winter precipitation and retrieve precipitation type. *12th SIRWEC Conference, Bingen, Germany 2004*.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theor. Appl. Climatol.*, 115(1-2):15–40.

Schumacher, C. and Jr., R. A. H. (2000). Comparison of radar data from the trmm satellite and kwajalein oceanic validation site. *J. Appl. Meteor.*, 39(12):2151–2164.

Short, D. A. and Nakamura, K. (2000). TRMM Radar Observations of Shallow Precipitation over the Tropical Oceans. *J. Climate*, 13(23):4107–4124.

Smith, S. R., Rolph, J. J., Briggs, K., and Bourassa, M. A. (2009). Quality-Controlled Underway Oceanographic and Meteorological Data from the Center for Ocean-Atmospheric Predictions Center (COAPS) - Shipboard Automated Meteorological and Oceanographic System (SAMOS).

Snodgrass, E. R., Di Girolamo, L., and Rauber, R. M. (2009). Precipitation characteristics of trade wind clouds during RICO derived from radar, satellite, and aircraft measurements. *J. Appl. Meteor. Climatol.*, 48(3):464–483.

Sorooshian, S. and Kuo-Lin, H. (2000). Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, 81(9):2035.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'Connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., Mitrescu, C., and CloudSat Science Team, T. (2002). The CLOUDSAT Mission And The A-TRAIN: A New Dimension of Space-Based Observations of Clouds and Precipitation. *Bull. Amer. Meteor. Soc.*, 83(12):1771–1790.

Stevens, B. (2005). Atmospheric moist convection. *Annual Review of Earth and Planetary Sciences*, 33(1):605–643.

Stevens, B. and Seifert, A. (2008). Understanding macrophysical outcomes of microphysical choices in simulations of shallow cumulus convection. *J. Meteor. Soc. Japan*, 86A:143–162.

Strangeways, I. (2004). Improving precipitation measurement. *Int. J. Clim.*, 24(11):1443–1460.

Sunilkumar, K., Narayana Rao, T., and Satheeshkumar, S. (2016). Assessment of small-scale variability of rainfall and multi-satellite precipitation estimates using measurements from a dense rain gauge network in Southeast India. *Hydrol. Earth Sys. Sci.*, 20(5):1719–1735.

Tapiador, F. J., Turk, F., Petersen, W., Hou, A. Y., García-Ortega, E., Machado, L. A., Angelis, C. F., Salio, P., Kidd, C., Huffman, G. J., and de Castro, M. (2012). Global precipitation measurement: Methods, datasets and applications. *Atm. Res.*, 104-105:70–97.

Taylor, P. (2000). Intercomparison and validation of ocean-atmosphere energy flux fields—final report of the joint WCRP/SCOR working group on Air-Sea fluxes. page pp. 325. WCRP-112, WMO/TD-1036.

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., and Wu, H. (2013). Modeling errors in daily precipitation measurements: Additive or multiplicative?: Modeling Errors in Precipitation Data. *Geophys. Res. Lett.*, 40(10):2060–2065.

Tokay, A. and Short, D. A. (1996). Evidence from Tropical Raindrop Spectra of the Origin of Rain from Stratiform versus Convective Clouds. *J. Appl. Meteor.*, 35(3):355–371.

Torrence, C. and Webster, P. J. (1999). Interdecadal Changes in the ENSO–Monsoon System. *J. Climate*, 12(8):2679–2690.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B. (2003). The Changing Character of Precipitation. *Bull. Amer. Meteor. Soc.*, 84(9):1205–1217.

Trenberth, K. E., Fasullo, J. T., and Kiehl, J. (2009). Earth's Global Energy Budget. *Bull. Amer. Meteor. Soc.*, 90(3):311–323.

Twomey, S. (1977). The Influence of Pollution on the Shortwave Albedo of Clouds. *J. Atm. Sci.*, 34(7):1149–1152.

UCAR/NCAR-EOL (2011). RICO: S-band Polarimetric (S-Pol) Data DORADE Format, Version 1.0. `https://data.eol.ucar.edu/dataset/87.008`. [accessed on 29 Sep 2016].

vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., Cheng, A., Couvreux, F., Jiang, H., Khairoutdinov, M., Kogan, Y., Lewellen, D. C., Mechem, D., Nakamura, K., Noda, A., Shipway, B. J., Slawinska, J., Wang, S., and Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during RICO. *J. Adv. Mod. Earth Syst.*, 3(2):M06001.

Villermaux, E. and Bossa, B. (2009). Single-drop fragmentation determines size distribution of raindrops. *Nat. Phys.*, 5(9):697–702.

Weller, R. A., Bradley, E. F., Edson, J. B., Fairall, C. W., Brooks, I., Yelland, M. J., and Pascal, R. W. (2008). Sensors for physical fluxes at the sea surface: energy, heat, water, salt. *Ocean Sci.*, 4(4):247–263.

Wentz, F. J. (2013). SSM/I Version-7 Calibration Report. In *report number 011012*, page 46pp. Rem. Sens. Syst., Santa Rosa, CA, USA.

WMO (2015). Code tables and flag tables associated with BUFR/CREX Table B. `http://www.wmo.int/pages/prog/www/WMOCodes/WMO306_vI2/LatestVERSION/WMO306_vI2_BUFRCREX_CodeFlag_en.pdf`. [accessed on 29 Sep 2016].

Wolters, E. L. A., van den Hurk, B. J. J. M., and Roebeling, R. A. (2011). Evaluation of rainfall retrievals from SEVIRI reflectances over West Africa using TRMM-PR and CMORPH. *Hydrol. Earth Syst. Sci.*, 15(2):437–451.

Woodruff, S. D., Worley, S. J., Lubker, S. J., Ji, Z., Eric Freeman, J., Berry, D. I., Brohan, P., Kent, E. C., Reynolds, R. W., Smith, S. R., and Wilkinson, C. (2011). ICOADS release 2.5: extensions and enhancements to the surface marine meteorological archive. *Int. J. Clim.*, 31(7):951–967.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I., and Hamada, A. (2009). A 44-Year Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *SOLA*, 5:137–140.

Yuter, S. E., Kingsmill, D. E., Nance, L. B., and Löffler-Mang, M. (2006). Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteor. Clim.*, 45(10):1450–1464.

Yuter, S. E. and Parker, W. S. (2001). Rainfall measurement on ship revisited: The 1997 PACS TEPPS cruise. *J. Appl. Meteor.*, 40(6):1003–1018.

List of Acronyms

CPP	Cloud Physical Properties
CMORPH	Climate Prediction Center Morphing
CM SAF	Satellite Application Facility on Climate Monitoring
CSAV	Cloud System Advection Vector
DMSP	Defense Meteorological Satellite Program
DPR	Dual-frequency Precipitation Radar
DWD	Deutscher Wetterdienst (German Weather Service)
ECV	Essential Climate Variable
ENSO	El Niño Southern Oscillation
FAR	False Alarm Ratio
GMI	GPM Microwave Imager
GPM	Global Precipitation Measurement
GSMaP	Global Satellite Mapping of Precipitation
HOAPS	Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data
ICOADS	International Comprehensive Ocean-Atmosphere Data Set
IMERG	Integrated Multi-satellite Retrievals for GPM
IPCC	Intergovernmental Panel on Climate Change
IPWG	International Precipitation Working Group
IR	Infrared
ITCZ	Intertropical Convergence Zone
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Royal Netherlands Meteorological Institute)
KS98	Koistinen and Saltikoff (1998)
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	METEOSAT Second Generation
NCAR	National Center for Atmospheric Research
NCEP ADP	National Centers for Environmental Prediction Automated Data Processing
NOAA	National Oceanic and Atmospheric Administration
NOCS	National Oceanography Centre Southampton
OceanRAIN	Ocean Rain And Ice-phase precipitation measurement Network
ODM	Optical Disdrometer
PARSIVEL	Particle Size Velocity

PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PM	Percentage Misclassified
PMW	Passive Microwave
POD	Probability Of Detection
PP	Precipitation Phase
PSD	Particle Size Distribution
PU	Percentage Unclassified
RR	Rain Rate
RS	Rain, Snow
RSM	Rain, Snow, Mixed-phase
RV	Research Vessel
RICO	Rain In Cumulus clouds over the Ocean
SAMOS	Shipboard Automated Meteorological and Oceanographic System
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SPICE	Solid Precipitation InterComparison Experiment
SSM/I	Special Sensor Microwave Imager
SSMIS	Special Sensor Microwave Imager Sounder
TMI	TRMM Microwave Imager
TMPA	TRMM Multi-satellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission
VIS	Visible
VOS	Voluntary Observing Ships
WMO	World Meteorological Organization
WWRP/WGNE	World Weather Research Programme / Working Group on Numerical Experimentation

List of Figures

List of Tables

Acknowledgment

First of all, I would like to express my deep gratitude to my supervisors, Stephan Bakan and Christian Klepp, who dedicated a tremendous amount of time and patience and who strongly supported me in developing my scientific research ideas.

I would like to thank Stefan Bühler for fruitful discussions and his helpful scientific advice. Furthermore, I appreciate his diligence in keeping track of my scientific progress and occurring problems during a number of constructive panel meetings. My gratitude goes to Jens Hartmann and Lars Kaleschke for their willingness and dedicated time to evaluate my work.

Special thanks goes to Karsten Fennig from CM SAF for carefully processing the extended HOAPS time series to obtain a larger sample of collocations. I thank Axel Andersson from DWD for his valuable comments discussing the results of my work. I well appreciate the extra time and effort that Nicole Albern put in carefully quality-checking and processing the OceanRAIN database. Thanks to Carsten Benecke and Andrea Dahl for their very personal technical support on the optical disdrometer. I thank Paul Kucera for helpful comments.

I thank Julian Kinzel, Katharina Meraner and Anne Häseler for their help in proofreading parts of my manuscript. Thanks to Jochem Marotzke and Dallas Murphy for their very inspiring scientific writing course. I would also like to express my gratitude to Antje Weitz, Connie Kampmann and Wiebke Böhm from the IMPRS office who were always friendly and of great help to solve organizational problems. I thank the IMPRS as a whole for the diverse course program offered as well as the provided travel money, which allowed me to participate in international conferences such as the International Precipitation Working Group workshop in Tsukuba or the GEWEX Global Watercycle Conference in The Hague.

Certainly not to forget, I deeply appreciate the time and company spent not only during countless lunch breaks together with Rohit, Eduardo, Manita, Sirisha, Kathi, Matthias, Markus and many others. Among other delicious restaurants, Kaalia was always worth a visit with its authentic Indian street kitchen, thanks to Wiebke and Ulf. Special thanks goes to Dirk Olonscheck for starting a table tennis team at the institute together with me. Thanks to everyone who supported us, it was and still is a lot of fun to play with you guys!

Mein besonderer Dank gilt meinen Eltern, die mir stets den größtmöglichen Rückhalt und gleichzeitigen Freiraum gegeben haben. Ich danke euch von ganzem Herzen. Nicht zuletzt danke ich ganz besonders dir, Nicole, für deine Zuneigung, dein Vertrauen und deine Unterstützung nicht nur in der finalen Phase meiner Doktorarbeit.

Thanks to everyone who supported me or my research during the time of my PhD!

Aus dieser Dissertation hervorgegangene Vorveröffentlichungen

List of Publications

Burdanowitz J., L. Nuijens, B. Stevens, and C. Klepp, 2015: Evaluating light rain from satellite- and ground-based remote sensing data over the subtropical North Atlantic, J. Appl. Meteor. Clim., 54, **3**, 556–572.

Burdanowitz J., C. Klepp and S. Bakan, 2016: An automatic precipitation phase distinction algorithm for optical disdrometer data over the global ocean. Atmos. Meas. Tech., **9**, 1637–1652.