A machine learning based 24-h-technique for an area-wide rainfall retrieval using MSG SEVIRI data over Central Europe

kumulative Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Geographie
der Philipps-Universität Marburg
vorgelegt von

Meike Kühnlein
aus Nürnberg

Marburg/Lahn, April 2014

Vom Fachbereich Geographie
der Philipps-Universität Marburg als Dissertation
am 23. April 2014 angenommen.

Erstgutachter: Prof. Dr. Thomas Nauß (Marburg)
Zweitgutachter: Prof. Dr. Peter Chifflard (Marburg)
Drittgutachter: Dr. Vincenzo Levizzani (Bologna)

Tag der mündlichen Prüfung: 18. November 2014

Preface

At the end of my thesis I would like to thank all those people who supported and accompanied me on my way to realise this work.

First and foremost, I would like to express my special gratitude to my supervisor Thomas Nauss for his invaluable guidance, constant support and encouragement. Moreover, I am deeply grateful that he made my research stay at the Centre for Atmospheric Research, University of Canterbury in Christchurch, New Zealand possible.

I am greatly thankful to Tim Appelhans for many scientific discussions, his generous assistance and enourmous helpful comments on the manuscripts. He has been a source of enthusiasm and encouragement during the last three years and made the commuting between Nürnberg and Marburg entertaining. Moreover, many thanks for opening me the door to my rewarding research stay in New Zealand.

I am very thankful to Boris Thies for his support and valuable comments on the manuscripts. Discussions and fruitful exchanges with Christoph Reudenbach are gratefully acknowledged. I thank Alexander Kokhanovsky for providing helpful comments on the first manuscript.

I thank my colleagues at the working group Environmental Informatics at Philipps-University of Marburg for the kind working atmosphere and their support ranging from little everyday matters to being open-minded discussants in scientific discourse. The very thought of our lunch or coffee breaks in our kitchen make me still laugh.

I gratefully acknowledge the financial support by the German Research Foundation (DFG) within the Priority Program 1374 “Biodiversity Exploratories”.

Last but not least, I would like to express my dearest gratitude to my parents who have provided the most supportive environment one could wish for. Moreover, I thank my sister, brother, granny and friends for the constant support and encouragement that I received.

Meike Kühnlein
Marburg, April 2014

Contents

List of Figures

List of Tables

List of Acronyms and Symbols

ACT	Advective Convective Technique
$\mathbf{a}_{ef}$	Cloud Effective Radius
$\mathbf{a}_{efSslalom}$	Cloud Effective Radius retrieved from SLALOM using MSG SEVIRI data
$\mathbf{a}_{efMslalom}$	Cloud Effective Radius retrieved from SLALOM using Terra/Aqua MODIS data
$\mathbf{a}_{efM06}$	Cloud Effective Radius retrieved from MODIS M06
2B-TAU	CloudSat Cloud Optical Depth Product
Cb	Cumulonimbus
CGMS	Coordination Group for Meteorological Satellites
CP	Cloud Phase
CPR	Cloud Profiling Radar
CREW	Cloud Retrieval Evaluation Workshop
CP	Cloud Phase
CSI	Critical Success Index
CST	Convective Stratiform Technique
CTH	Cloud Top Height
CTT	Cloud Top Temperature
CWP	Cloud Water Path
DFG	Deutsche Forschungsgemeinschaft (engl. German Research Foundation)
DWD	Deutscher Wetterdienst (engl. German Weather Service)
ECMWF	European Center for Medium-Range Weather Forecasts
ECST	Enhanced Convective Stratiform Technique
EOS	NASA's Earth Observing System
ETS	Equitable Threat Score
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites

FAR	False Alarm Ratio
GEO	Geostationary Orbit
GOES	Geostationary Operational Environmental Satellite
GPCP	Global Precipitation Climatology Project
GPI	GOES Precipitation Index
HKD	Hansen Kuipers discriminant
HSS	Heidke Skill Score
IR	Infrared
IPWG	International Precipitation Working Group
ISCCP	International Satellite Cloud Climatology Project
KNMI	Koninklijk Nederlands Meteorologisch Instituut (engl. Royal Netherlands Meteorological Institute)
LAADS	NASA's Level 1 and Atmosphere Archive and Distribution System
LEO	Low Earth Orbit
LUT	Look-up table
LWC	Liquid Water Content
MAE	Mean Absolute Error
MCC	Mesoscale Convective System
ME	Mean Error
ML	Machine Learning
MODIS	Moderate Resolution Imaging Spectroradiometer
MODIS 021KM	NASA MODIS Level 1B Calibrated Radiances - 1km
MODIS 03	NASA MODIS Geolocation - 1km
MODIS 06	NASA MODIS Cloud Product; MYD06 for Aqua MODIS and MOD06 for Terra MODIS
MSE	Mean Square Error
MSG	Meteosat Second Generation
MTG	Meteosat Third Generation

mtry	Number of predictor variables randomly sampled at each split
NASA	National Space Agency
NIR	Near Infrared
Ns	Nimbostratus
ntree	Overall number of trees in a RF
OOB	Out-of-bag
PFD	Probability of False Detection
POD	Probability of Detection
PR	TRMM Precipitation Radar
R	Correlation Coefficient
R^2	Coefficient of Determination
RADOLAN RW	Radar-based precipitation product of the DWD
RF	Random Forests
RMSE	Root Mean Square Error
Rsq	Coefficient of Determination
SACURA	Semi-Analytical Cloud Retrieval Algorithm
SEVIRI	Spinning-Enhanced Visible and Infrared Imager
SIST	Solar-Infrared Infrared Split Window Technique
SLALOM	Cloud Retrieval Algorithm (Simple Approximations for Cloudy Media)
SPACES	Science Partnerships for the Assessment of Complex Earth System Processes
SSM/I	Special Sensor Microwave Imager
τ	Cloud Optical Thickness
τ_{2B-TAU}	CloudSat 2B-TAU products
$\tau_{Sslalom}$	Cloud Optical Thickness retrieved from SLALOM using MSG SEVIRI data
$\tau_{Mslalom}$	Cloud Optical Thickness retrieved from SLALOM using Terra/Aqua MODIS data

τ_{M06}	Cloud Optical Thickness retrieved from MODIS M06
TOA	Top of the Atmosphere
TRMM	Tropical Rainfall Measurement Mission
UTC	Universal Time co-ordinated
VIS	Visible
VISST	Visible Infrared Solar-Infrared Split-Window Technique
WCRP	World Climate Research Program
WGDMA	Working Group on Data Management and Analysis
WMO	World Meteorological Organisation
WV	Water Vapour

Chapter 1

Motivation, Aims and Outline

1 Motivation, Aims and Outline

1.1 Why area-wide quantitative precipitation?

Water affects all aspects of human life, moreover, the availability of fresh water is vital to life on Earth. The availability of fresh water is regulated by the water cycle: water vapour condenses, forming cloud droplets from which precipitation may fall and therefore there is a water flux from the atmosphere to surface. Precipitation is a key-process of the global water cycle and, together with the energy cycle, helps to regulate the climate system. Its measurement and monitoring is important. Too little precipitation causes droughts while too much especially over short time span may put life and property at risk.

From a scientific point of view, area-wide information on precipitation distribution and its quantity at high spatio-temporal resolution is of increasing importance for all climatological, hydrological and ecological studies (New et al., 2001; Trenberth et al., 2003). This information also can improve the reliability of very-short change and now-casting applications (e. g. flood prediction and monitoring). Moreover, precipitation has an economic value since it plays a key role in activities such as agriculture and water resource management (Kidd et al., 2009; Thornes et al., 2010).

Despite its great importance, the correct spatio-temporal detection and quantification of this key parameter of the global water cycle is still associated with large uncertainties and particularly challenging with conventional observing systems. This is mainly due to the fact that precipitation is characterised by high variability in space, time and intensity (New et al., 2001). Convective precipitation events can be extremely local and result in rainfall rates above 50 mm/h that can last from less than 30 minutes to over 6 hours while stratiform events are usually characterised by low rain rates which can often last for several days and cover areas up to 1000 km by 1000 km.

Reliable precipitation estimates can be considered a fundamental contribution to many scientific and socio-economical studies, especially in areas where no in-situ information about precipitation is available. Ideally, information on precipitation is available:

- At high spatial resolution, in order to capture the high variability in space.
- At high temporal resolution, in order to allow for continuous precipitation monitoring.

1.2 Area-wide quantitative precipitation estimation - why satellite?

Traditionally, precipitation is observed locally by conventional instruments such as rain (or snow) gauges and where available, weather radar systems. However, the most obvious drawbacks of station-based precipitation measurements are their spatial incoherences, their uneven regional and global distribution and their highly variable density with some regions having a relatively dense network (e. g. Germany) while others have only a few or no gauges (e. g. Russia, Central Africa and Tibetan Plateau). Over oceans, gauges are almost non-existent apart from only a few island stations. Ground-based weather radars are expanding over Europe, Japan and North America to provide spatial measurements of precipitation (100 km from the radar). They are located over land and generally concentrated in regions which are also well covered with rain gauges. Both, gauges and radars can not fill the gap of missing precipitation measurements.

In this context, satellite imagery are the obvious data source to fill this gap. According to the requirements shown above, a satellite system with high spatial and temporal data resolution is needed. The latter criterion can only be fulfilled by geostationary (GEO) systems. There are already a variety of satellite-based rainfall retrieval techniques for the detection of precipitating cloud areas and the assignment of rainfall rates developed for GEO platforms (see valuable overviews by e. g. Kidd and Levizzani, 2011; Kidd and Huffman, 2011; Prigent, 2010; Thies and Bendix, 2011). Caused by the poor spectral resolution of the first satellite sensor systems, many of the techniques rely on a relationship between cloud top temperatures measured in an infrared (IR) channel and rainfall probability/intensity (e. g. Adler and Mack, 1984; Arkin and Meisner, 1987; Vicente et al., 1998). These methods are based on the assumption that cold clouds are related to convection and are likely to produce rain. As a consequence, cold cloud tops are assumed to be connected with high rainfall intensities. The variety of existing IR-based techniques benefit from a degree of simplicity and the day and night availability of the IR data enables the estimation of rainfall 24 hour day. Moreover, these methods perform reasonable well over areas where rainfall is mainly controlled by deep convection. However, they are less favourable at higher latitudes, where precipitation originates from both convective and advective-stratiform systems (e. g. Levizzani et al., 1990; Negri and Adler, 1993; Pompei et al., 1995; Amorati et al., 2000; Adler et al., 2001; Ebert et al., 2007; Früh et al., 2007). Advective-stratiform precipitating clouds are characterised by relatively warm and spatially homogeneous cloud top temperatures that do not differ significantly from raining to non-raining regions. Stratiform rain is usually not associated with high rainfall rates, but it often covers large areas and is characterised by long duration and therefore contributes to a significant portion of the rainfall in a region. As a result, retrieval techniques based solely on IR cloud top temperature lead to

an underestimation of the detected rain area and to uncertainties concerning the assigned rainfall rate in such cases (e. g. Ebert et al., 2007; Früh et al., 2007).

With the upcoming of new generation satellite systems, several authors suggested to use optical and microphysical cloud parameters derived from the now available multispectral data sets to improve rainfall retrievals (e. g. Ba and Gruber, 2001; Rosenfeld and Gutman, 1994; Rosenfeld and Lensky, 1998; Thies et al., 2008b,a,d; Nauss and Kokhanovsky, 2006, 2007; Roebeling and Holleman, 2009; Kühnlein et al., 2010; Feidas and Giannakos, 2010, 2012). They were able to show that cloud areas with a high optical thickness and a large effective particle radius possess a high amount of cloud water and are characterised by a higher rainfall probability/intensity than cloud areas with a low optical thickness and a small effective particle radius.

However, there is still a great deficit regarding the identification and quantification of precipitating cloud areas in the mid-latitudes especially in connection with extra-tropical cyclones in a continuous manner (day, twilight, night) resulting in a 24 hour estimation. In particular, all existing optical retrievals have considerable problems during twilight due to difficult illumination conditions. Moreover, the quantification of rainfall is still largely unsolved during all times of the day.

Precipitation processes leading to different rainfall intensities within extra-tropical cyclones are very complex (see section 2.1). So far, most of the retrieval techniques use parametric approaches to relate cloud properties to precipitation (e. g. Adler and Negri, 1988; Cheng and Brown, 1995; Kühnlein et al., 2010; Levizzani et al., 1990; Thies et al., 2008d). This approach requires the specification of the underlying parametric tests and the related conceptual models. It is rather straight-forward if only a few input variables are considered. However, when faced with a very complex and non-linear cloud top to precipitation relationship and a variety of precipitation processes as existing in extra-tropical cyclones, this is likely to be beyond the skill of parametric techniques.

In this context, machine learning algorithms have been successfully adopted to remote sensing and rainfall applications (e. g. Capacci and Conway, 2005; Rivolta et al., 2006; Grimes et al., 2003; Giannakos and Feidas, 2013) which show that the latter might be suitable to overcome these limitations. In particular, random forests (Breiman, 2001), a new and promising machine learning algorithm, which offers a number of features that makes it well suited for use in rainfall remote sensing applications, will likely help to benefit from the potential offered by current GEO systems (descriptions of machine learning in general and of the features of random forests are given in section 2.3). Especially, the new European GEO system Meteosat Second Generation (MSG) with its payload Spinning Enhanced Visible and InfraRed Imager (SEVIRI, Aminou, 2002; Schmetz et al., 2002) offers the spectral resolution which is needed to infer information about cloud properties and the potential for an improved precipitation estimation (Levizzani et al.,

2001b; Levizzani, 2003). Its relatively high spatial (3 km x 3 km at sub satellite point) and especially temporal resolution (15 minutes) permits a quasi-continuous area-wide observation of the rainfall, which is essential in order to address the deficiencies mentioned above.

1.3 Aim of this work

As mentioned in section 1.1, quasi-continuous area-wide information on precipitation at high temporal and spatial resolution can be considered a fundamental contribution to many applications ranging from now-casting to climatological purposes. Following the identified deficiencies of existing rainfall retrievals, the hypothesis of this work is:

- Based on spectrally and temporally adequate MSG SEVIRI data, it is possible to devise a process-related and quantitative precipitation scheme for daytime, twilight and night-time conditions suitable for precipitating clouds in connection with extra-tropical cyclones using the machine learning algorithm random forest.

Addressing the needs and requirements identified, the aims of the present work are:

- To develop a coherent daytime, twilight and night-time technique for the process-related and quantitative estimation of precipitation in connection with extra-tropical cyclones in the mid-latitudes.
- To develop this technique based on MSG SEVIRI data.
- To develop this technique using the machine learning algorithm random forest.

A methodology of this kind needs reliable information on precipitation for its development and validation. Germany was chosen as study area as the ground-based radar network of the German Weather Service (DWD) provides exactly these information. Moreover, this region can be regarded as sufficiently representative for mid-latitude precipitation formation processes since it is dominated by frontally induced precipitation processes in connection with extra-tropical cyclones.

1.4 Outline

From the aims of this work follows the structure of this thesis presented in figure 1.1, starting with the conceptual design in **chapter 2**. In section 2.1, dominant precipitation processes in the mid-latitudes that are mostly connected to extra-tropical cyclones are introduced. Then, in section 2.2, an overview over the

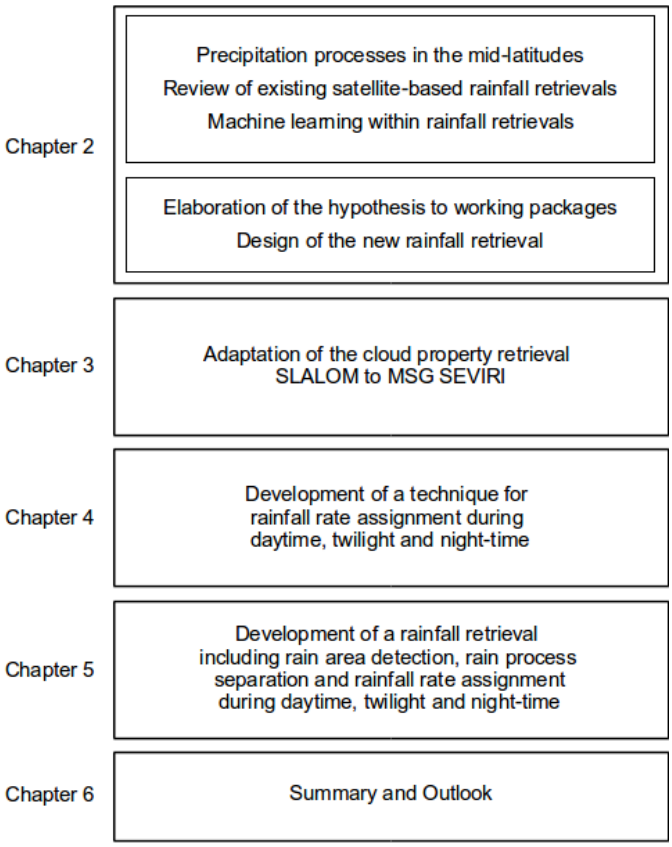

Figure 1.1: *Structure of this work.*

current status of existing satellite-based rainfall retrieval techniques is given to elucidate the need for a novel rainfall retrieval technique and to highlight the weaknesses and shortcomings to be addressed in the newly developed technique. After this, the advantages of machine learning algorithms as fundamental algorithms within rainfall retrievals are described in section 2.3. Based on this, the elaboration of the hypothesis to working packages as well as the conceptual design of the newly developed technique is outlined afterwards in section 2.4.

Chapters 3, 4 and 5 describe the core of the study with the implementation of the working packages (section 2.4). First, the adaptation of the SLALOM (SimpLe Approximations for cLOudy Media) algorithm, a semi-analytical cloud property retrieval technique, to MSG SEVIRI is presented in 'Kühnlein, M., T. Appelhans, B. Thies, A. A. Kokhanovsky, and T. Nauss, 2013: An evaluation of a semi-analytical cloud property retrieval using MSG SEVIRI, MODIS and CloudSat, Atmospheric Research, 122, 111–135' (**chapter 3**). Then, the development of a technique for rainfall rate assignment during daytime, twilight and night-time is introduced in 'Kühnlein, M., T. Appelhans, B. Thies, and T. Nauss, 2014: Improving the accuracy of rainfall rates from optical satellite sensors with machine learning - A random forests-based approach applied to MSG SEVIRI, Remote Sensing of Environment, 141, 129–143' (**chapter 4**). Finally, the development of the overall rainfall retrieval including rain area detection, rain process

separation and rainfall rate assignment during daytime, twilight and night-time is embedded in 'Kühnlein, M., T. Appelhans, B. Thies, and T. Nauss, 2014: Precipitation estimates from MSG SEVIRI daytime, night-time and twilight data with random forests, Journal of Applied Meteorology and Climatology, 53, 2457–2480' (**chapter 5**).

Chapter 6 provides a summary of the previous chapters and a short outlook.

The presented work was part of the Priority Program 1374 "Biodiversity Exploratories" funded by the German Research Foundation (DFG). It is also integrated into the Ecosystem Informatics PhD program funded by the University of Marburg. Furthermore, the work contributed to Eumetsat's cloud retrieval evaluation workshop (CREW, see http://www.icare.univ-lille1.fr/crew/) with the objective to evaluate cloud retrieval algorithms using satellite data from passive imager instruments. The work also concurs with the main objective of the International Precipitation Working Group (IPWG). The IPWG is a permanent working group of the coordination group for Meteorological Satellites (CGMS) and the World Meteorological Organisation (WMO) dealing with operational and research satellite-based quantitative precipitation measurement issues and challenges.

Chapter 2

Conceptual Design

2 Conceptual Design

This chapter introduces the conceptual design of the new rainfall retrieval technique based on MSG SEVIRI data for mid-latitudes and puts it in the context of existing techniques. In the first section, dominant precipitation processes in the mid-latitudes that are mostly connected to extra-tropical cyclones are introduced. Then, an overview over the current status of existing satellite-based rainfall retrieval techniques is given. After this, the advantages of machine learning algorithms as fundamental algorithms within rainfall retrievals are described. Based on this, working packages required to verify the hypothesis and the retrieval's design are presented afterwards.

2.1 Precipitation processes in the mid-latitudes

The mid-latitude weather is under the constant influence of extra-tropical cyclones forming along the polar front and moving from west to east driven by westerly winds. In doing so they transport tropical-subtropical warm air poleward and artic or antarctic cold air toward the equator. On their way, typical extra-tropical cyclones usually follow a definite life cycle from cyclogenesis to maturity.

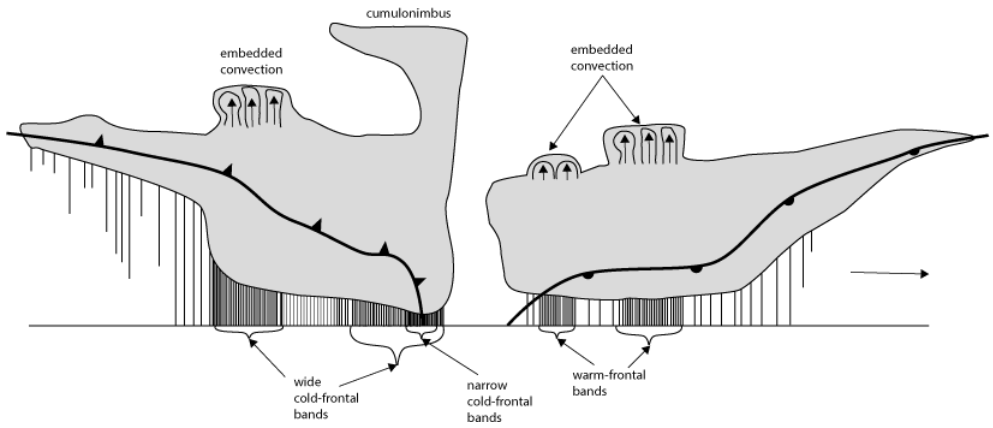

Figure 2.1: *Schematic cross-section through an extra-tropical cyclone (adapted from Houze (1993) after Matejka et al. (1980)). Vertical hatching below cloud bases represents precipitation. The density of the hatching corresponds qualitatively to the precipitation rate.*

Figure 2.1 shows a schematic vertical cross-section through an extra-tropical cyclone denoting the associated cloud and precipitation patterns in the mature phase with the leading warm- and trailing cold front. As a result of widespread rising processes, the general frontal precipitation field is formed by advective-stratiform precipitation areas which are characterised by light precipitation intensities. Within these precipitation areas, so-called rainbands are embedded. The warm-frontal and wide cold-frontal rainbands are enhancements of the advective-

stratiform precipitation while the narrow cold-frontal rainbands represent lines of intense convection (Houze, 1993). The narrow cold-frontal rainbands are the result of ascent of stable or just slightly unstable warm air masses which flow in an upward direction predominantly parallel to the cold front with a slight bend to the cold sector (Browning and Roberts, 1996) forced by colder and drier air subsiding on the cold side of the front forming the cold front. A zone of strong, concentrated convergence and abrupt upward motion arise by reason of the cold front in connection with the warm moist air masses resulting in a narrow precipitating cloud band with high convective rainfall rates. This forced ascent can be strong. However, the precipitating clouds remain limited in vertical and horizontal extent to the zone of lifting forced by the sharp front (Houze, 1993). If the air forced up by the cold front were unstable, high convective cells or mesoscale convective systems (MCCs) of deep thunderstorms would be caused. Usually these convective showers are embedded or adjoining a large region of stratiform precipitation (Houze, 1993; Barth and Parsons, 1996). As precondition for such convective-stratiform precipitation from Nimbostratus in connection with convective cores, ice particles in the upper part of the cloud together with aggregation and riming processes are necessary for the formation and growth of rain particles.

Unlike the narrow cold-frontal rainbands, the most active region of upward motion of wide cold-frontal rainbands (see figure 2.1) is aloft, above the front. These active regions are characterised by enhanced mean ascent and contain shallow convective generating cells. As precipitation particles generated in these cells fall trough the layer of frontal clouds, they enhance the general advective-stratiform precipitation rate at lower levels by the “seeder-feeder” mechanism (Houze, 1993). The layer of the generating convective cells in the upper part of the cloud is termed “seeder zone”. The increased buoyancy in this zone enables the development and accelerated growth of ice particles. As these ice particles fall through the cloud they increase greatly in mass by depositional (and possibly riming) growth in the middle cloud parts (“feeder-zone”). The ice particles can become larger and heavier than rain droplets which in turn leads to higher fall velocities and higher rainfall rates and therefore to enhanced rainfall intensities within the advective-stratiform precipitation area (Rutledge and Hobbs, 1983; Houze, 1993).

In contrast to the cold-frontal bands, wide warm-frontal rainbands occur in the forward part of the cloud shield of the developing cyclone where warm advection dominates. Both are similiar in dimension. An intensification of the advective-stratiform precipitation is caused by generating cells, located above the rainbands. They move over the lower-level rainbands and seed them with ice particles. The dynamical mechanism responsible for the enhanced vertical air motions in wide warm-frontal rainbands seem to be similar to wide cold-frontal rainbands. However, this has not yet been proven (Heymsfield, 1977; Houze, 1993).

While figure 2.1 shows the vertical pattern of clouds and precipitation associated with a mature extra-tropical cyclone, figure 2.2 summerises typical features of the area-wide horizontal pattern of the latter. Warm- and cold-frontal rain-

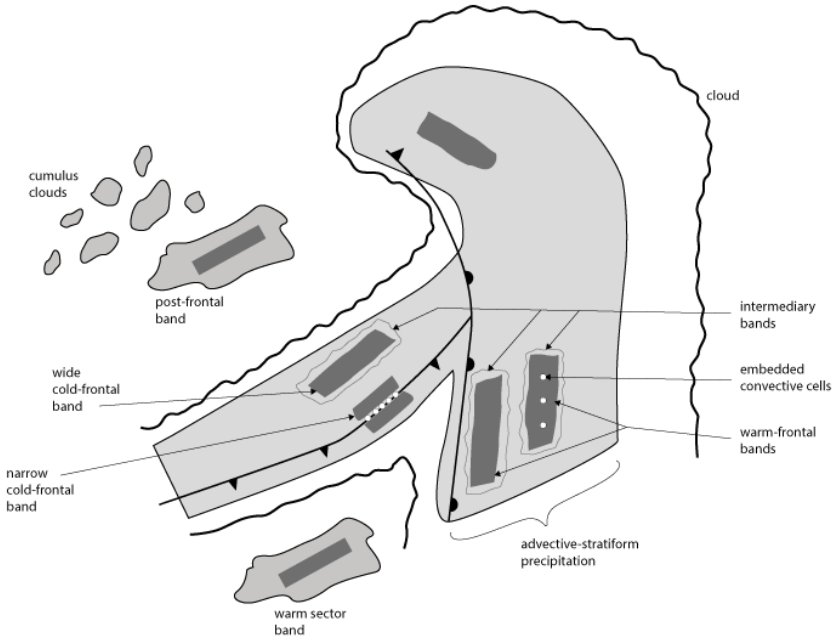

Figure 2.2: *Idealisation of the cloud and precipitation pattern associated with a mature extra-tropical cyclone. Light grey areas indicate advective-stratiform precipitation, dark grey areas represent enhanced advective-stratiform precipitation areas characterised by convective precipitation processes and white circles show embedded convective cells (adapted from Houze (1993) after Matejka et al. (1980)).*

bands can be clearly identified within the main envelope. Warm sector and post-frontal rainbands are located outside the envelope and often consist of lines of convective showers or thunderstorm. All rainbands are linked to moderate to high rainfall intensities (dark grey cloud areas). However, it also becomes obvious that a major part of the precipitating cloud areas are not necessarily connected to the warm-frontal and cold-frontal rainbands (light grey cloud areas). This precipitation in the main envelope is primarily advective-stratiform which is usually not associated with high rainfall rates, but it often covers large areas and the generally slower forward motion allow precipitation to last longer. Therefore advective-stratiform precipitation contributes to a significant portion of the rainfall in a region.

Further detailed descriptions of precipitation processes within extra-tropical cyclones are given in Houze (1993) and Pruppacher and Klett (1997). A more detailed description about the life cycle of an extra-tropical cyclone can be found in Barry and Chorley (2010). Kraus (1995) gives an overview about the different theories for the development of extra-tropical cyclones.

2.2 Satellite-based approaches to precipitation detection

With the advent of meteorological satellites in the 1970s, scientists developed techniques to estimate precipitation from observations from optical sensors available on GEO platforms. Since then a large number of satellite sensors have been developed and launched to observe the atmosphere with many of these specifically designed for precipitation monitoring.

According to the sensor characteristics, the variety of retrieval methods can be separated into primarily optical methods, microwave methods and multi-sensor techniques which utilise a combination of both.

Passive microwave measurements are directly sensitive to precipitation. However, because of their technical possibilities, so far microwave sensors (e. g. SSM/I (Special Sensor Microwave Imager), Hollinger et al., 1990) are limited to the use on Low Earth Orbiting (LEO) satellites. Observations from LEO satellites are usually only available twice a day (e. g. Terra/Aqua MODIS (Moderate Resolution Imaging Spectroradiometer)) for a given location and a given satellite. This highlights that LEO observations leave substantial gaps all over the globe and thus, a quasi-continuous rainfall monitoring in near-real time cannot be achieved. In this context, optical sensors onboard GEO satellites form an alternative to fill the information gap by providing area-wide data with up to 15 minutes sampling. This permits a quasi-continuous monitoring of rainfall distribution and rainfall rate in near-real time. For this reason and the fact that microwave sensor based observations are not used within this study, only optical rainfall retrievals are considered in the following overview.

For further information about the full spectrum of existing satellite-based rainfall retrievals, the reader is referred to Kidd and Levizzani (2011), Kidd and Huffman (2011), Prigent (2010), Stephans and Kummerow (2007), Ferraro (2007), Anagnostou (2004), Levizzani (2003), Scofield and Kuligowski (2003), Levizzani et al. (2001a), Adler et al. (2001), Kidd (2001) and Kidder and Vonder Haar (1995). An overview of existing retrieval techniques based on passive microwave sensors can be found in Joyce et al. (2004), Weng et al. (2003), Kummerow et al. (2001), Petty (1995) and Wilheit et al. (1994). Iguchi et al. (2000) and Ferreira et al. (2001) give explanations of the Tropical Rainfall Measuring Mission (TRMM) precipitation radar (PR).

In general, optical rainfall retrievals consist of two components. First, the rain area is identified. Then, a rainfall rate is assigned to the identified rain areas. The first techniques were restricted by the spectral resolution of the sensors, which only allowed concepts that rely on a relationship between IR cloud top temperature (CTT) and the rainfall probability and intensity. In the thermal IR the temperature of cloud tops, and therefore their heights can be identified. Through the assumption that colder/taller clouds are more likely to rain and to produce higher rainfall intensities than warmer/smaller clouds, IR imagery may

provide a simple rainfall estimate. The statistical relationship between IR CTT and rainfall is established using rainfall measured by ground-based rain gauges or radars. However, the CTT to rainfall relationship is indirect, with significant variations in the relationship occurring between rain systems and climatological regimes as well as during the lifetime of rainfall events (Kidd and Huffman, 2011).

The simplest and probably most widely used IR technique is the GOES Precipitation Index (GPI, Arkin, 1979; Arkin and Meisner, 1987). Precipitating clouds are delineated based on the threshold of a CTT of 235 K and a constant rainfall rate of 3 mm/h is assigned to these raining pixels. This can be considered to be appropriate for tropical rainfall over areas of 2.5° by 2.5°. Kerrache and Schmetz (1988) transferred the GPI to Meteosat. Menz and Zock (1997), Ba and Nicholson (1998) as well as Todd et al. (1995) used the GPI successfully over eastern Africa. The World Climate Research Programme (WCRP) uses the GPI within the Global Precipitation Climatology Project (GPCP) to estimate global precipitation from a few days to one month (Arkin and Ardanuy, 1989; Arkin et al., 1994; Huffman et al., 1997). Other IR-based techniques such as the autoestimator technique developed by Vicente et al. (1998) uses the GOES 10.7 μm band to compute real-time precipitation amounts based on a power-law regression algorithm. Similiar techniques use the Autestimator (Vicente et al., 2002) and Hydroestimator (Kuligowski, 2002). A neural network approach, which combines numerical weather model information with the IR satellite imagery to derive daily rainfall values was developed by Coppola et al. (2006). Anagnostou and Krajewski (1999a), Anagnostou and Krajewski (1999b), Grimes et al. (1999) and Todd et al. (1999) are further examples of IR techniques.

The variety of existing IR-based techniques benefit from a degree of simplicity. Furthermore, the day and night availability of the data enables the estimation of rainfall in a quasi-continuous manner resulting in a 24 hour prediction. However, IR techniques are only appropriate for convectively dominated precipitating cloud areas which are characterised by a large vertical extension and a cloud top rising up into the atmosphere resulting in cold CTT. A look at figure 2.2 shows that a major part of the precipitating cloud areas in connection with extra-tropical cyclones are not necessarily connected to convection which is only present in the narrow cold-frontal rainbands.

As a consequence, based on the studies from Adler and Mack (1984), Adler and Negri (1988) developed the Convective Stratiform Technique (CST) for convective systems in the tropics where CTT are related to rainfall rate and rain area through a one-dimensional cloud model. First, convective areas are determined by local minima in the IR temperature which are considered as potential precipitating areas. Then, temperature thresholds are used to determine potentially precipitating, stratiform cloud areas around the convective cores. Non-precipitating cirrus is eliminated by a radar-based discriminant function. Precipitation is then assigned to convective areas by means of the cloud model. To the detected stra-

tiform precipitating areas a fixed stratiform rainfall rate of 2 mm/h is assigned. The CST has become a widely used and intensively validated technique. Bendix (1997) and Bendix (2000) showed that the CST can be used successfully for the detection of rainfall processes in the tropics and subtropics while the transfer of the method to extra-tropical regions still reveal deficiencies (e. g. Levizzani, 2000; Marrocu et al., 1993; Pompei et al., 1995; Ebert et al., 2007). As a result of this, the Enhanced Convective Stratiform Technique (ECST) for convective systems in the mid-latitudes was developed by Reudenbach (2003). Inclusion of the water vapour (WV) channel provided a more reliable differentiation between different types of precipitating cloud areas. Reudenbach et al. (2007) developed the Advective Convective Technique (ACT) which allows the differentiation of convectively dominated precipitation areas into the following sub areas: convective cores (Cumulonimbus, Cb), convective-stratiform precipitating Nimbostratus clouds (Nb in connection with Cb, see section 2.1), and enhanced advective-stratiform precipitating areas (embedded shallow generating cells together with the "seeder-feeder" effect, see section 2.1).

However, another look at figure 2.2 represents that also retrievals such as ECST and ACT do not consider all different precipitation areas within extra-tropical cyclones. They still show considerable deficiencies regarding the detection of widespread advective-stratiform precipitation areas which are not connected to convective cores and not characterised by cold CTT (Früh et al., 2007; Ebert et al., 2007).

Continued technical enhancements of satellite sensor systems accompanied by an improved spectral resolution enabled the computation of optical and microphysical cloud properties such as cloud effective radius and cloud optical thickness (e. g. Nakajima and Nakajima, 1995; Platnick et al., 2003; Kokhanovsky et al., 2003, 2005; Roebeling et al., 2006; Nauss and Kokhanovsky, 2011). Several studies have shown, that using optical and microphysical cloud parameters derived from multispectral data of new generation satellite systems can be used to overcome the drawbacks of IR and advanced IR methods at mid-latitudes (e. g. Ba and Gruber, 2001; Lensky and Rosenfeld, 2003a,b; Nauss and Kokhanovsky, 2006, 2007). These techniques also allow to consider advective-stratiform clouds and therefore are able to detect all different precipitation areas of an extra-tropical cyclone (see figure 2.2) which led to an improvement in estimation of rainfall in the mid-latitudes.

Nauss and Kokhanovsky (2006, 2007) documented the potential of the use of optical and microphysical cloud properties for rain area identification using data from the LEO system Terra MODIS. They were able to show that cloud areas with a high optical thickness and a large effective particle radius possess a high amount of cloud water and are characterised by a higher rainfall probability than cloud areas with a low optical thickness and a small effective particle radius. However, the low temporal resolution of Terra MODIS and the restriction to

daytime data still impedes deficiencies. Based on the potential shown by Nauss and Kokhanovsky (2006, 2007), Thies et al. (2008b) and Thies et al. (2008a) introduced new methods for rain area delineation using cloud properties retrieved from MSG SEVIRI data during daytime and night-time. Using MSG SEVIRI data now enables the rainfall area detection at high temporal resolution. In this context, Thies et al. (2008d) showed the possibility to separate areas of differing precipitation processes and rainfall intensities within the rain area by means of MSG SEVIRI cloud properties. The day and night technique for precipitation process separation and rainfall rate differentiation relies on information about the cloud top height, the cloud water path and the cloud phase in the upper parts of the cloud. It is based on the assumption that areas with high cloud water path (i.e. large enough combination of the optical thickness and the effective particle radius) and ice particles in the upper parts posses high rainfall intensities. Moreover, several authors successfully used optical and microphysical cloud parameters for an improved rain process differentiation (Feidas and Giannakos, 2012; Giannakos and Feidas, 2013) and rainfall rate assignment (Kühnlein et al., 2010; Roebeling and Holleman, 2009) during daytime.

Even if the encouraging results concerning rain area detection, rain process separation and rainfall rate assignment along with the enhanced information content on cloud properties at high spectral, spatial and temporal resolution point to a quite promising potential of current and upcoming GEO systems as basis for reliable rainfall retrievals, there is still a great deficit regarding the detection of rain areas and assignment of rainfall rates in the mid-latitudes in connection with extra-tropical cyclones especially in a continuous manner which enables a 24 hour prediction at high temporal resolution. In particular, all existing optical retrievals that are based on optical and microphysical cloud parameters are restricted to daytime and night-time conditions and do not cover twilight conditions (see e. g. Kidd and Levizzani, 2011; Prigent, 2010). Moreover, the continuous quantitative determination of rainfall, especially light rain, is still largely unsolved.

2.3 Machine learning in rainfall retrievals

Most of the rainfall retrieval techniques introduced in section 2.2 take account of a specific precipitation process (e. g. IR CTT as a measure of cloud height and therefore convection) and relate the corresponding cloud properties to precipitation. For this purpose, the underlying conceptual model and the related parametric tests are specified and usually only a few input variables are considered. The approaches used are called parametric approaches (e. g. Adler and Negri, 1988; Cheng and Brown, 1995; Kühnlein et al., 2010; Levizzani et al., 1990; Thies et al., 2008d). These approaches are rather straightforward and easy to interpret, but their invariable inherent assumptions limit their applicability.

As depicted in section 2.1, precipitation associated with an extra-tropical cy-

clone is manifold. Differing rainfall intensities can be expected from rainbands dominated by different rainfall processes. In order to consider these different rainfall processes a number of input variables are necessary. As spectral channels are used as surrogates for these cloud parameters, also different illumination conditions have to be taken into account which beyond this, also change throughout the year (e. g. twilight hours).

Nonetheless, the aims stated in section 1.3 can only be achieved if all rainfall processes can be reproduced within the new rainfall retrieval. This however is likely beyond the skill of parametric tests and the related conceptual models. When faced with high dimensional and complex data, machine learning (ML) algorithms might be suitable to overcome the mentioned limitations. In this context, the recent developments in parallel computing with machine learning offer new possibilities in terms of training and predicting speed and therefore make the usage of real time systems realisable.

ML algorithms are broadly defined as computational methods using experience to improve performance (Mohri et al., 2012), e. g. to make accurate predictions. Here, experience refers to the past information or training data which is available to the methods. While parametric approaches define a sequence of instructions that have to be carried out to transform the input to output, machine learning approaches use computing systems with learning capability to infer and generalise dependencies from the past information or training data to extract automatically the algorithm for the prediction. Such systems adapt themselves to changing conditions and do not need an underlying conceptual model. The success of a learning algorithm strongly depends on the data used. This is the reason why ML are data-driven methods combining fundamental concepts in computer science with ideas from statistics.

ML algorithms feature exactly the required characteristics which are needed to make use of a) the knowledge acquired during the last 40 years but limited by the use of parametric approaches and b) excellent training data basis given by MSG SEVIRI data and ground-based radar network observations.

The goal of learning algorithms that can learn from their experience has attracted researchers from many fields, e. g. computer science, engineering, mathematics, physics, neuroscience, cognitive science, chemistry. Out of this research has come a wide variety of learning algorithms that have been successfully deployed in a variety of applications, including e. g. speech recognition, computational biology applications, medical diagnosis, text classifications (Alpaydin, 2010; Mohri et al., 2012). Such applications correspond to a wide variety of learning problems which can mainly be classified into classification, regression, clustering, ranking, dimensionality reduction or manifold learning. ML algorithms differ according to the types of training data available, the order and method by which training data is received and the test data used to evaluate the learning algorithm (Mohri et al., 2012). According to these characteristics, ML

algorithms are classified into supervised, unsupervised, semi-supervised, on-line, reinforcement and active learning algorithms as well as transductive inference algorithms. The most common learning scenario is supervised learning. Here, the algorithms are trained on labelled examples and make predictions for all unseen points. If the label is discrete, then the task is called classification problem, otherwise, for continuous labels, it is called regression problem (Alpaydin, 2010; Mohri et al., 2012).

ML algorithms have already been successfully adopted to rainfall applications. Most of these techniques use neural network algorithms to link the input information to the rainfall estimates (e. g. Hong et al., 2004; Capacci and Porcù, 2009; Sorooshian et al., 2000; Tapiador et al., 2004; Grimes et al., 2003; Hsu et al., 1997; Bellerby, 2004; Coppola et al., 2006; Capacci and Conway, 2005; Rivolta et al., 2006; Behrangi et al., 2009). In recent years, supervised machine learning techniques which use ensembles of classifications or regressions have received increasing interest (e. g. random forests, neural network ensembles, bagging and boosting, see Friedl et al., 1999; Krogh and Vedelsby, 1995; Rodriguez-Galiano et al., 2012; Ruiz-Gazen and Villa, 2007; Steele, 2000). They are based on the assumption that a whole set of trees or networks produce a more accurate prediction than a single tree or network (Dietterich, 2002).

A new, powerful and promising ensemble classification and regression technique is random forests (further referred to as RF, Breiman, 2001). RF is widely applied in disciplines such as bioinformatics (e. g. Cutler and Stevens, 2006) and ecology (e. g. Cutler et al., 2007; Mota et al., 2002; Prasad et al., 2006), and has also proven its performance in some land-cover classifications using hyper spectral and multispectral satellite data, radar and lidar data (Ghimire et al., 2010; Guo et al., 2011; Pal, 2005; Rodriguez-Galiano et al., 2012).

Indeed the utilisation of RF in climatology is rare, but it offers a number of features that makes RF well suited for use in rainfall remote sensing applications:

- RF runs efficiently on large data sets.
- RF can easily be parallelised.
- RF does not require the specification of an underlying data model.
- RF offers the ability to capture non-linear association patterns between predictors and response.
- RF offers the flexibility to perform several types of statistical data analysis (regression and classification).
- RF is relatively robust to outliers and noise.

These features likely offer the possibility to take account for high dimensional and complex data expected when dealing with cloud-top properties to precipitation

relationships. This is the reason why RF is chosen as the fundamental prediction algorithm in the novel rainfall retrieval.

2.4 Detailed elaboration of the hypothesis into working packages

The hypothesis stated in section 1.3 was that a daytime, twilight and night-time scheme for the process-related and quantitative estimation of precipitation in connection with extra-tropical cyclones in the mid-latitudes can be developed based on MSG SEVIRI data using the machine learning algorithm random forest. The overview presented in section 2.2 shows that existing satellite-based rainfall retrievals for mid-latitudes show deficiencies in the following areas:

- No operationally applicable scheme is available for rain area detection and rain process separation during twilight conditions.
- No operationally applicable process-related rainfall rate assignment scheme is currently available for daytime, twilight and night-time.

Up to now, these drawbacks have very much limited the usability of satellite-based rainfall estimation in the mid-latitudes.

In order to derive a rainfall retrieval for process-related and quantitative estimation of precipitation in the mid-latitudes using MSG SEVIRI daytime, twilight and night-time data, the hypothesis is split into working packages (WP) for the examination. Based on the review of precipitation processes within extra-tropical cyclones (see section 2.1) and machine learning approaches (see section 2.3), the following working packages are differentiated:

(WP1) The transfer and adaptation of the cloud property retrieval SLALOM to MSG SEVIRI in order to retrieve cloud optical properties such as cloud effective radius (a_{ef}) and cloud optical thickness (τ) that are needed for improved satellite-based rainfall estimation in mid-latitudes (e. g. identification of advective-stratiform clouds, see section 2.1). SLALOM was developed by Nauss and Kokhanovsky (2011) for Terra MODIS data and is characterised by fast computation speed and good accuracy, which motivates its utilisation for this study.

(WP2) The investigation of the potential of the machine learning algorithm random forest for satellite-based rainfall rate assignment. Therefore, a technique for rainfall rate assignment based on MSG SEVIRI daytime, twilight and night-time data with random forests as fundamental prediction algorithm is developed. Based on the precipitation processes in connection with extra-tropical cyclones (see section 2), rainfall rates are assigned to previously identified advective-stratiform and convective precipitating areas by means of individual RF models. Since this WP focuses on an improved assignment of rainfall rates based on random forests, the derivation of the rain process

is realised based on observations from the radar network. To estimate the response variable (rainfall rates), RF needs a set of predictor variables which must consider the different mid-latitudinal precipitation processes within extra-tropical cyclones. Hence, satellite-based information on the cloud top height, cloud top temperature, cloud phase and cloud water path are chosen as predictor variables.

(WP3) Development of a new coherent daytime, twilight and night-time rainfall retrieval to retrieve rainfall rates of precipitation events in connection with extra-tropical cyclones in the mid-latitudes based on MSG SEVIRI. The technique is realised in three steps: (i) Precipitating cloud areas are identified. (ii) The areas are separated into convective and advective-stratiform precipitating areas. (iii) Rainfall rates are assigned separately to the convective and advective-stratiform precipitating areas. These three parts are realised sequentially by means of individual RF models. Again, combined satellite-based information on cloud top temperature, cloud top height, cloud water path and cloud phase are used.

Putting all these working packages together a 24-h-technique for process-related and quantitative estimation of precipitation in mid-latitudes using MSG SEVIRI data is derived. The retrieval's design which incorporates the working packages outlined above is illustrated in figure 2.3. The realisations of the WP's are described in the following chapters.

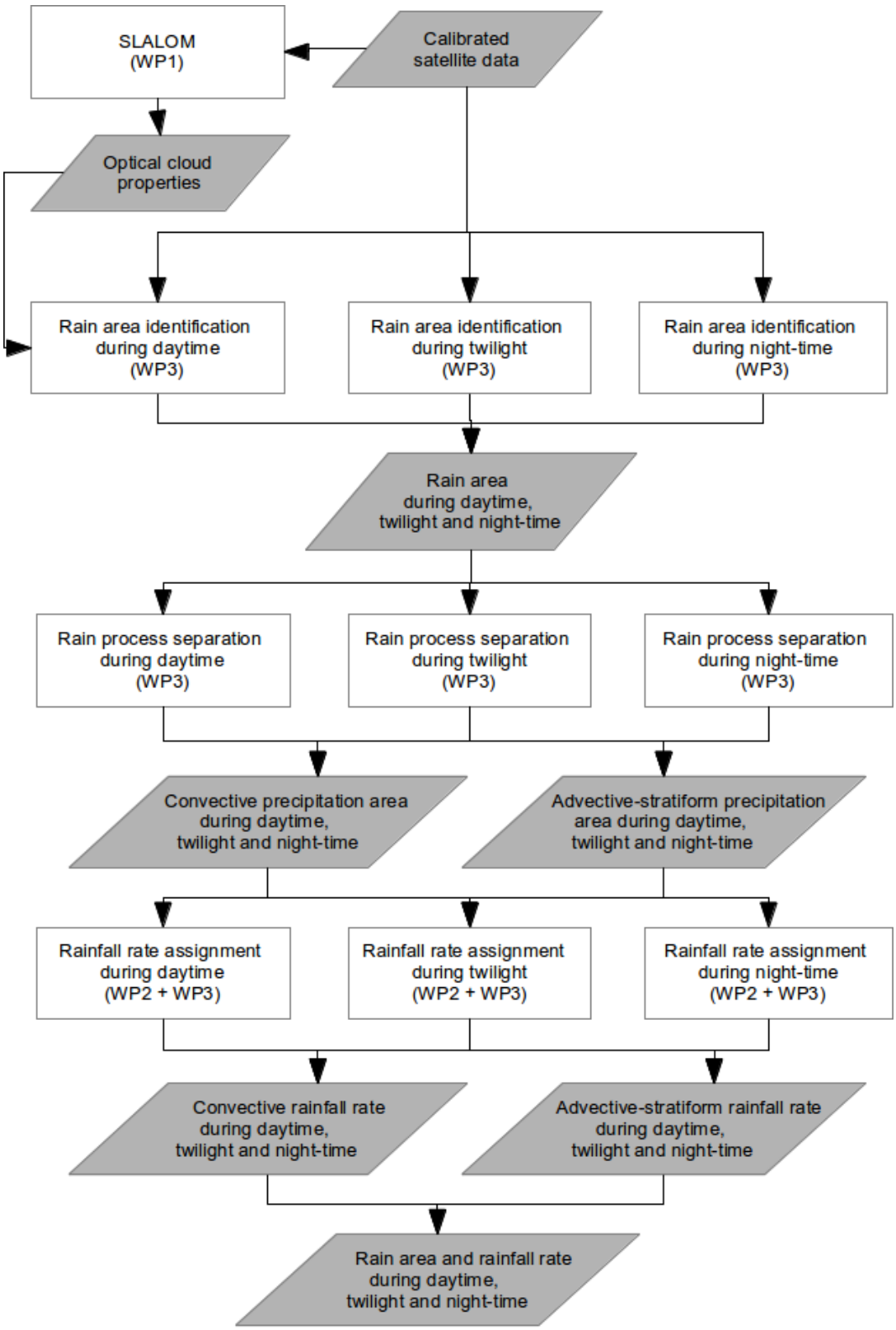

Figure 2.3: *Design of the novel rainfall retrieval.*

Chapter 3

An evaluation of a semi-analytical cloud property retrieval using MSG SEVIRI, MODIS and CloudSat

Meike Kühnlein [(1)], Tim Appelhans [(1)], Boris Thies [(3)], Alexander A. Kokhanovsky [(2)], Thomas Nauß [(1)]

[(1)] Environmental Informatics, Faculty of Geography, Philipps–University Marburg, Marburg, Germany

[(2)] Institute of Environmental Physics, Bremen University, O. Hahn Allee 1, 28334 Bremen, Germany

[(3)] Laboratory for Climatology and Remote Sensing, Faculty of Geography, Philipps-University Marburg, Marburg, Germany

Published in

Atmospheric Research 122 (2013) 111–135

Received 21 April 2012;
Received in revised form 15 October 2012;
Accepted 24 October 2012;

3 An evaluation of a semi–analytical cloud property retrieval using MSG SEVIRI, MODIS and CloudSat

Meike Kühnlein, Tim Appelhans, Boris Thies,
Alexander A. Kokhanovsky, Thomas Nauß

Abstract Knowledge of cloud properties such as cloud effective radius (a_{ef}) and optical thickness (τ) is essential to understand their role in the dynamic radiation budget and climate change. The Spinning Enhanced Visible and Infrared Instrument (SEVIRI) on board Meteosat Second Generation (MSG) with its high temporal resolution (15 minutes), permits a quasi-continuous monitoring of the evolution of cloud properties. This has motivated the adaptation of the SLALOM (SimpLe Approximations for cLOudy Media) algorithm, a semi-analytical cloud property retrieval technique to MSG SEVIRI. The optical properties retrieved by SLALOM are compared against the well known and validated NASA MODIS cloud property product (MODIS 06) as well as the cloud optical depth product (2B-TAU) of CloudSat. The results are shown over the North Atlantic and over the European continent with the intention of determine the relative accuracy between SLALOM and the other retrievals. Over the North Atlantic, SLALOM-based cloud properties retrieved from SEVIRI datasets show a good agreement with the MODIS 06 product with correlation coefficients of 0.93 (τ) and 0.82 (a_{ef}). The largest deviations were found in less homogeneous cloud areas that are characterised by broken clouds and towards the cloud borders. Moreover, SLALOM optical thickness values are well within the range of corresponding CloudSat 2B-TAU optical thickness values which can be found within a SEVIRI pixel, except for $\tau<5$ where SLALOM tends to overestimate τ. Despite the different sensor characteristics and viewing geometries, the retrieved cloud properties compare very well. Over Europe, the evaluation between SLALOM and MODIS 06 showed larger differences. We attribute this to (a) uncertainties related to the surface albedo which is treated differently in the algorithms and is based on different albedo maps and (b) inhomogeneities of clouds which exhibit quite complex structures particularly over land. The latter are detected on different scales by MODIS and SEVIRI because of their different spatial resolutions. Given the demonstrated accuracy of SLALOM using MSG SEVIRI data there is a wide spread of potential applications.

Keywords Cloud properties, Satellite retrieval, SLALOM, MSG SEVIRI, CloudSat, MODIS

3.1 Introduction

Clouds play an important role in the Earth's climate system (Liou, 1992; Platnick and Valero, 1995; Stephens, 2005) and are recognized as a key modifier of the earth-atmosphere radiation budget (Kiehl and Trenberth, 1997) driving atmospheric dynamics, hydrological cycle and global warming. The radiative properties of clouds and hence their influence on the earth-atmosphere radiation budget can largely be explained by a few optical and microphysical cloud parameters, particularly the cloud optical thickness and the effective cloud droplet radius. Despite their great importance, the correct spatio-temporal measurement of clouds is still associated with large uncertainties (Stephans and Kummerow, 2007). This is mainly due to the fact that clouds are characterized by a highly short-lived nature. However, detailed information about clouds are important for a better understanding of the climate system. The possibility to obtain information about cloud properties and their spatial distribution on a global scale can be achieved by satellite remote sensing only.

Over the past decades, various algorithms have been developed to retrieve optical and microphysical cloud properties (e. g. Arking and Childs, 1985; Han et al., 1994; Kawamoto et al., 2001; Kokhanovsky et al., 2003; Kokhanovsky and Nauss, 2005; Liou and Wittman, 1979; Nakajima and King, 1990; Nakajima and Nakajima, 1995; Nauss and Kokhanovsky, 2011; Platnick et al., 2003; Roebeling et al., 2006; Strabala et al., 1994; Twomey and Cocks, 1982). The performance of many cloud retrieval algorithms using satellite data from passive imager instruments is being evaluated by Eumetsat's cloud retrieval evaluation workshop (CREW, see `http://www.icare.univ-lille1.fr/crew/`). Long time series of such and other satellite derived cloud properties are retrieved by the Working Group on Data Management and Analysis (WGDMA) of the International Satellite Cloud Climatology Project (ISCCP) as part of the World Climate Research Program (WCRP) (Schiffer and Rossow, 1983; Rossow, 1989, see `http://isccp.giss.nasa.gov/index.html`) as well as the MODIS cloud product team at NASA (Platnick et al., 2003, see `http://modis-atmos.gsfc.nasa.gov/MOD06_L2/index.html`).

For day-time observations, these techniques rely on the concurrent measurement of the cloud reflectance in absorbing and non-absorbing wavelengths. The reflection of clouds in the visible region strongly depends on cloud optical thickness, whereas the reflection in the near-infrared region is additionally related to the cloud droplet size (i. e. effective cloud droplet radius, see e. g. integrative figure in Nakajima and King (1990)). The majority of the retrieval techniques have been developed for optical sensors aboard polar-orbiting satellites mainly because of the available spectral resolution. Up until now, only a few techniques are applied to geostationary satellite systems (e. g. Feijt et al., 2004; Han et al., 1994; Pandey et al., 2012; Pérez et al., 2011; Roebeling et al., 2006, 2008).

The present European geostationary system Meteosat Second Generation

(MSG), carrying the Spinning Enhanced Visible and Infrared Instrument (SEVIRI, Aminou, 2002), provides sufficient spectral resolution required for cloud retrievals. Furthermore, its relatively high spatial (3 by 3 km at sub satellite point) and very high temporal resolution (15 minutes) permits a quasi-continuous monitoring of the evolution of cloud properties which is of particular importance with respect to the spatio-temporal cloud cover dynamics. Recently, Nauss and Kokhanovsky (2011) developed SLALOM (SimpLe Approximations for cLOudy Media), a cloud property retrieval based on approximations of the asymptotic solutions of the radiative transfer theory using data from the Moderate Resolution Imaging Spectroradiometer (MODIS). The computation speed and accuracy of this retrieval along with the enhanced information content potentially provided by the spatio-temporal high resolution geostationary data motivates the adaptation of the SLALOM algorithm to MSG SEVIRI.

In this study, SEVIRI and MODIS-based results of SLALOM will be compared to the LUT-based approach by Platnick et al. (2003) for the MODIS sensor onboard of the NASA EOS Aqua and Terra satellites (King and Greenstone, 1999) and the cloud optical depth product (2B-TAU) of CloudSat (Polonsky et al., 2008). The validation against CloudSat 2B-Tau product is of great value, since CloudSat serves, based on the detailed information about the vertical structure of the cloud measured by its Cloud Profiling Radar (CPR), a quantitative basis to produce an accurate cloud optical depth profile which forms a very good validation source. Furthermore, CloudSat 2B-Tau provides detailed information on the sub-pixel heterogeneity of the cloud structure within the larger SEVIRI pixel. So far, CloudSat measurements have been used only in a few studies for the validation of cloud retrievals (e. g. Bennartz et al., 2010; Pandey et al., 2012; Pérez et al., 2011). Unfortunately, existing retrieval algorithms applied to geostationary satellite systems are either not yet available to the public or have yet to be adapted to existing processing chains and therefore could not used for the evaluation.

While idealized radiative transfer measurements could be used to evaluate the theoretical accuracy of a retrieval or as an alternative data source for within sensor evaluation set-ups, across-sensor studies show the performance under real conditions and provide an estimate for the uncertainties which must be expected if data from different sources is analysed within the respective studies.

The algorithms presented in this study are capable of retrieving many other cloud parameters like liquid and ice water path and particle absorption length, but only results of effective radius and optical thickness are compared. Furthermore, only the case of liquid clouds is considered since the complexity of light transport through highly inhomogeneous crystalline clouds bears many intrinsic uncertainties and requires a priori knowledge of the shape distributions (e. g. Min et al., 2004). The retrievals are compared over both ocean and land.

The structure of this paper is as follows: The underlying datasets and retrieval

methods used for the evaluation and calibration differences between MODIS and SEVIRI are introduced in section 3.2. In Section 3.3, the across-sensor evaluation design is introduced and the results of the evaluation study are presented and discussed using scenes over both the Atlantic and Central Europe. A summary and final conclusion can be found in section 3.4.

3.2 Data and Methods

3.2.1 Satellite systems used within this study

3.2.1.1 SEVIRI

For this study, data of MSG SEVIRI, MODIS and CloudSat have been used. The Spinning-Enhanced Visible and Infrared Imager (SEVIRI) on board Meteosat Second Generation (MSG) is the present European geostationary satellite system operated by EUMETSAT. The first satellite of the second generation MSG-1 (Meteosat-8) was launched on 28 August 2002 and is currently positioned at an altitude of about 36,000 km above the equator at 3.4° W. In December 2005, MSG-2 (Meteosat-9) was launched and positioned at 0° longitude (see Munro et al., 2002; Schmetz et al., 2002; Schumann et al., 2002). EUMETSAT and ESA are currently working on Meteosat Third Generation as the follow up mission which is intended to be launched in 2015 (Bézy et al., 2005). SEVIRI measures reflected and emitted radiance in 12 channels, three channels at visible and very near infrared wavelengths (between 0.6 and 1.6 μm), eight at mid-infrared to thermal infrared wavelengths (between 3.8 and 14 μm), and one high-resolution visible channel. SEVIRI scans the full disk every 15 minutes and provides a nominal spatial resolution of 1 by 1 km at the sub-satellite point for the high resolution channel, and 3 by 3 km for the other channels (Aminou, 2002; Schmetz et al., 2002). Over our study area in Central Europe, the satellite viewing angles of SEVIRI range from 48° to 61° and increase to a range from 38° to 68° over the North Atlantic. As a consequence the above mentioned spatial resolution is reduced in our study. The MSG SEVIRI data required for this study were downloaded from the EUMETSAT data centre (`www.eumetsat.int`) and were pre-processed by the SCISYS Deutschland GmbH (formerly VCS) software package 2met! (`http://www.scisys.de/`).

3.2.1.2 MODIS

In addition to MSG SEVIRI, data from the MODerate-resolution Imaging Spectroradiometer (MODIS) on board the Terra and Aqua satellites is used. MODIS measures reflected and emitted radiance in 36 spectral bands ranging from 0.42 to 14.24 μm. Depending on the spectral band the spatial resolution for nadir views range from 250 m to 1 km. Beside the MODIS cloud property product (see below), MODIS geolocation dataset (MODIS 03) and MODIS calibrated

radiances with a spatial resolution of 1 km (MODIS 021KM, both available at http://ladsweb.nascom.nasa.gov/) have been used in this study. The latter products provide information about geolocation, solar and satellite angles as well as calibrated radiances for each MODIS tile which is necessary for the application of SLALOM to MODIS.

3.2.1.3 CloudSat

CloudSat is another EOS satellite which was launched on 28 April 2006 and is orbiting in formation within the A-Train, a satellite constellation of 7 satellites in a sun-synchronous polar orbit with a ground-track repeat of 16 days (Stephens et al., 2002). The main instrument aboard the satellite is the Cloud Profiling Radar (CPR), a 94 GHz nadir-looking radar that measures the energy backscattered by precipitation and clouds as a function of distance from the radar (Im et al., 2005).

3.2.2 Algorithms used within the study

3.2.2.1 Semi-analytical approach SLALOM

Nauss and Kokhanovsky (2011) have presented a new cloud property retrieval relying on SimpLe Approximations for cLOudy Media (SLALOM). The forward model is based on approximate solutions of the asymptotic radiative transfer theory (e. g. Germogenova, 1963; King, 1987) and provides increased computation speed since the equations can be efficiently solved during runtime. In order to address the restriction of the technique to weakly absorbing media and small satellite zenith angles, look-up tables for the reflection function of a semi-infinite cloud, the escape function and the spherical albedo are incorporated into the algorithm (see Nauss and Kokhanovsky (2006) for details). The present version of SLALOM is limited to water clouds.

A detailed description of the algorithm can be found in Nauss and Kokhanovsky (2011) and - in contrast to the website mentioned in that publication - the Fortran source code of the techniques as well as all necessary ancillary datasets have been moved to http://code.google.com/p/rtm-cloud-slalom/ for free download (please do not hesitate to contact the author for assistance).

In order to retrieve cloud optical thickness (τ) and cloud effective droplet radius (a_{ef}) from MSG SEVIRI data, a combination of reflectance measurements at visible (0.65 μm or 0.81 μm) and near-infrared (1.64 μm) wavelengths is used. Necessary look-up tables (see above) have been generated for a droplet size of 10 μm using the radiative transfer code from Mishchenko et al. (1999) in analogy to Nauss and Kokhanovsky (2011). For the background albedo, a minimum composite of the reflectances in the visible (0.65 μm and 0.81 μm) and near-infrared (1.64 μm) channel over one month was calculated. Cloud phase as well as cloud masks were derived using the algorithm developed and implemented by Cermak (2006) and Cermak and Bendix (2008). Beside its application to MSG SEVIRI,

SLALOM has also been applied to MODIS data within this study.

3.2.2.2 NASA MODIS 06 cloud product

The first retrieval used for the comparison study was developed in the framework of NASA's Earth Observing System (EOS, King and Greenstone, 1999) and is part of the MODIS product. The LUT-based approach (King et al., 1997; Platnick et al., 2003) also uses a combination of one visible and one near-infrared channel to retrieve the aforementioned cloud properties but in order to minimize the influence of the background reflection, channel 1 (0.65 μm), 2 (0.86 μm) and 3 (1.2 μm) are used over land, ocean, and ice respectively. For the near-infrared, the retrieval is computed for each of channels 6 (1.6 μm), 7 (2.1 μm) and 20 (3.7 μm). The MODIS 06 product (MYD06 for Aqua MODIS and MOD06 for Terra MODIS, both further referred as M06) has a spatial resolution of 1 km and is readily available for download through NASA's Level 1 and Atmosphere Archive and Distribution System (LAADS web, `http://ladsweb.nascom.nasa.gov/`).

3.2.2.3 NASA CloudSat 2B-Tau product

Of interest for this study is the 2B-TAU (Cloud optical depth - off nadir) product (Polonsky et al., 2008). Therefore, the 1B-CPR standard product (radar backscatter profiles) is combined with auxiliary MODIS and ECMWF (European Center for Medium-Range Weather Forecasts) data, as well as the 2B-GEOPROF (Cloud geometric profile), 2B-CLDCLASS (Cloud Classification) and CWC-RO (Radar only combined water content) product to produce the 2B-TAU product within a 2.5 km along-track and 1.4 km cross-track radar footprint with 500 m vertical resolution. The sample rate is 0.16 second per profile (Im et al., 2005; Polonsky et al., 2008).

The product contains the total optical depth for each CloudSat profile. The latter is used in this study, because the radar reflectivity measured by the CPR provides detailed information about the the cloud in the atmospheric column and serves therefore as a quantitative basis to produce an accurate cloud optical depth profile. Caused by the differing horizontal measurement characteristics of CloudSat and SEVIRI, more than one CloudSat profile is generally located within one SEVIRI pixel, providing an insight into the within-pixel horizontal heterogeneity of the respective area sector. To retrieve the cloud parameters of interest for our study, the day time algorithm uses the upwelling reflectivity at the top of the atmosphere (TOA) measured by MODIS channels 2 (0.864 μm) and 7 (2.13 μm), averaged over an associated subset of 3x5 MODIS pixels (Polonsky et al., 2008). For details on the CloudSat mission, CPR and its products, the reader is referred to `http://cloudsat.atmos.colostate.edu` and Stephens et al. (2002). As with M06, the 2B-TAU product is already computed and provided by CloudSat Data Processing Center (`http://www.cloudsat.cira.colostate.edu/`).

3.2.3 Pre-processing of the MODIS and SEVIRI datasets

In order to compare the retrieved cloud properties based on spatially and temporally corresponding MSG SEVIRI and Terra/Aqua MODIS scenes, both datasets have to be re-projected to a common map projection. In a first step, the MODIS datasets were transformed from their initial HDF-EOS swath format to a geographic projection grid using the Reprojection Tool Swath (see https://lpdaac.usgs.gov/tools/modis_reprojection_tool_swath) with a target resolution of 0.01° by 0.01°. This grid forms the basis for the comparison between the NASA MODIS 06 product and the MODIS-based SLALOM results (for details please refer to the evaluation design presented in section 3.3.1). For the evaluation of the SEVIRI-based SLALOM results, the MODIS 06 product as well as the MODIS-based SLALOM results were subsequently averaged to a 0.08° by 0.08° grid. The SEVIRI-reflectances were corrected for sensor ageing (see 3.2.4) and also projected to this geometry. The impact of the different grid resolutions can clearly be seen in figure 3.1. Broken cloud fields which can be recognized in the centre of the MODIS image are displayed as rather homogeneous cloud fields in the SEVIRI image.

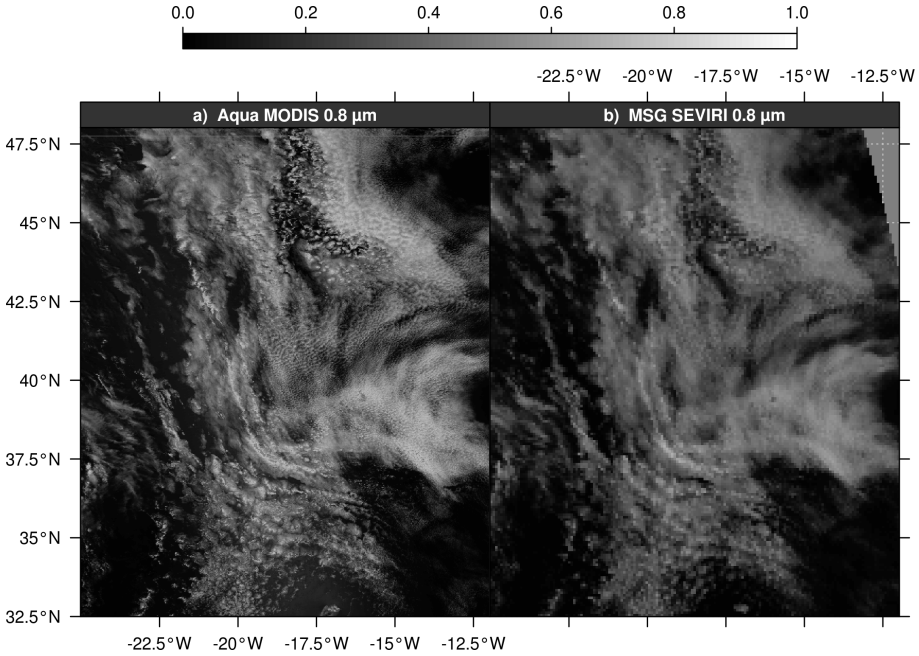

Figure 3.1: *(a) Aqua MODIS and (b) MSG SEVIRI 0.8 µm reflectances over the North Atlantic for 11 June 2008 14:15 and 14:20 UTC (Aqua MODIS: 0.01° x 0.01°, MSG SEVIRI: 0.08° x 0.08°).*

3.2.4 Adjustment of SEVIRI reflectances due to sensor ageing and operational calibration uncertainties

Differences in the reflectance measurements will have a direct impact on the retrieved cloud properties since reflection in the VIS is mainly proportional to the cloud optical thickness and reflection in the NIR is mainly inversely pro-

portional to the cloud effective droplet radius. Hence, cross-calibration between different satellite systems is a crucial aspect for long-term and multi-source observations. The MODIS instrument has onboard calibration for all channels and is well-calibrated with an accuracy about 2% (Guenther et al., 1998). The solar channels are calibrated using an onboard solar diffuser, lunar observations and a solar diffuser stability monitor for tracking the solar diffuser degradation. In contrast, there is no onboard calibration device for the solar channels of the SEVIRI instrument. In order to remedy the problem of degradation, the post-launch calibration is done by vicarious calibration techniques. The accuracy of the calibration of the VIS and NIR channels is expected to be about 5%, when adequate data and calibration targets are used (Govaerts and Clerici, 2004). Additionally it should be noted that the corresponding channels have slightly different response functions which also leads to differences in the measured reflectances of SEVIRI and MODIS. As can be seen in figure 3.2 a, the response functions for the utilized VIS channel are centred at 0.81 μm (SEVIRI) and 0.86 μm (Aqua MODIS) and have different bandwidths and shapes. Both response functions for the NIR are centred around 1.64 μm, but the SEVIRI bandwidth is almost three times wider than the respective Aqua MODIS channel and their shapes also differ (figure 3.2 b).

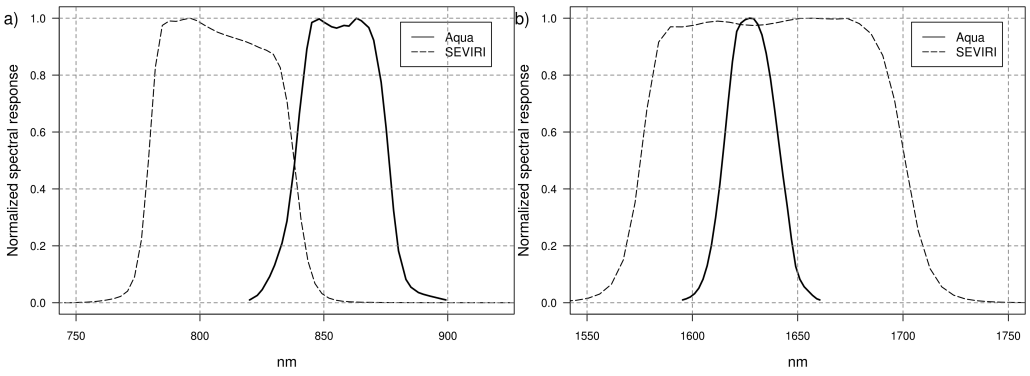

Figure 3.2: *Spectral response functions of the (a) 0.8 μm and (b) 1.6 μm channels from Aqua MODIS and MSG SEVIRI.*

As a consequence of the calibration procedures and the sensor characteristics, radiances registered by the two sensors differ and effects like sensor ageing will additionally deflect the respective measurements.

Using MODIS as the calibration source, Jan Fokke Meirink (KNMI) has already investigated simultaneous nadir overpasses of MODIS and MSG SEVIRI and found that the calibration of SEVIRI seems to be too dark for VIS and slightly too bright for NIR channels (Meirink, personal communication). He compared Aqua- and Terra MODIS granules near 0°N and 0°E with temporally corresponding SEVIRI images with a maximum time-difference of +/-7 minutes. The results show that the slope between MODIS and SEVIRI is 1.08 for the 0.64 μm channel, 1.06 for the 0.81 μm channel and 0.96 for the 1.64 μm channel. The slopes

for Terra MODIS differ by ~1% (Meirink, personal communication). For the present study, these slopes have been applied to the initially received SEVIRI reflectances in order to correct the sensor ageing and calibration deficiencies. However, caused by different viewing conditions (SEVIRI has much larger viewing angles than MODIS) in our study area and the different underlying spatial resolutions and spectral sensor characteristics (e. g. bandwidth, spectral response function) deviations between SEVIRI and MODIS reflectances can be expected, even though the initially received SEVIRI reflectances have been adjusted using the coefficients.

3.3 Validation of SLALOM retrievals for MODIS and SEVIRI

3.3.1 Methodology

Since many different cloud property retrievals are used in national and international research projects, it is important to examine possible differences between them. Therefore, SLALOM is compared to well-known and validated retrieval products. The intention is to get information about the relative accuracy between SLALOM and the other retrievals rather than the absolute accuracy against idealised model data (for this topic see Nauss and Kokhanovsky, 2006, 2011). Therefore, the following retrieval-sensor combinations will be evaluated against each other in a three-step approach:

- First, SLALOM will be applied to Terra/Aqua MODIS data and the results will be compared against the NASA MODIS M06 product on a 0.01° by 0.01° grid (King et al., 1997; Platnick et al., 2003).

- Second, SLALOM will be applied to MSG SEVIRI data and compared to the NASA MODIS M06 product on a 0.08° by 0.08° grid.

- Third, SLALOM-based cloud optical thickness retrieved from MSG SEVIRI and Aqua MODIS data as well as MODIS M06 cloud optical thickness will be compared to the CloudSat 2B-TAU product (Polonsky et al., 2008).

The first step allows a comparison between two algorithms applied to the same sensor. Hence, differences between SLALOM and the MODIS 06 product can be solely attributed to different assumptions within the retrievals and meta-datasets (e. g. background albedo). In contrast, the second step will also imply potential deviations due to quite different sensor characteristics and viewing geometries which is the reason why the MODIS-based SLALOM results are also incorporated in the comparison of the SEVIRI-based SLALOM results. For the evaluation, one scene over the North Atlantic and some over the European continent will be used. While for the North Atlantic scene the influence of auxiliary data is minimized due to a quite homogeneous ocean background albedo, the comparison over Central

Europe exhibits a considerably higher influence on the resulting parameters of both background albedo and larger satellite viewing angles.

Table 3.1 shows the absorbing and non-absorbing channels which are used to retrieve cloud optical thickness and effective radius by the different optical algorithms. Since the 1.6 μm channel of SEVIRI is used to retrieve a_{ef}, the analogous M06 product (see 3.2.2.2) is used for the comparison of a_{ef} to eliminate errors due to different wavelength-dependent penetration depths of the absorbing channels at 1.64 and 2.13 μm into the clouds (see Platnick, 2000).

Retrieval	Non-absorbing channel (μm)	Absorbing channel (μm)	Res. Nadir (km)
SLALOM	0.64 over land and 0.81 over ocean	1.64	3x3
M06	0.65 over land and 0.86 over ocean	1.6	1x1
2B-TAU	0.86	2.13	2.5x1.4

Table 3.1: *Spatial and spectral characteristics of the retrievals.*

Since CloudSat provide an accurate cloud optical depth profile and detailed information on the sub-pixel heterogeneity of the cloud structure within the SEVIRI pixel, the CloudSat 2B-TAU (τ_{2B-TAU}) product was included into the investigation as well. For the comparison of the three types of cloud optical thickness (SLALOM retrieval for MSG SEVIRI ($\tau_{Sslalom}$), SLALOM retrieval for MODIS ($\tau_{Mslalom}$), MODIS M06 (τ_{M06})) to τ_{2B-TAU}, only those SEVIRI/MODIS pixels which are spatially and temporally co-located to the Nadir scan line of CloudSat can be taken into account. Caused by their differing horizontal pixel dimensions (see chapter 3.2.1.1 and 3.2.1.3), more than one CloudSat profile is generally located within one SEVIRI pixel for any given date. Several studies have shown that validation uncertainties can be reduced by using a series of corresponding measurements instead of using only the nearest neighbour (e. g. Greuell and Roebeling, 2009; Schutgens and Roebeling, 2009). Therefore, to remedy the scale differences, a mean over all CloudSat optical depth profiles within each SEVIRI pixel is calculated for the comparison. In order to get information about the sub-pixel heterogeneity of the cloud structure, the value range of the individual profiles is discussed.

Due to the complexity of light transport through highly inhomogeneous crystalline clouds, retrievals for ice clouds are generally less accurate (Min et al., 2004). Therefore, only the case of liquid clouds is considered here and water clouds have been selected using the algorithm developed and implemented by Cermak (2006) and Cermak and Bendix (2008). Furthermore, pixels immediately located along the border of clouds are eliminated since sub-pixel cloudiness and enhanced 3-D effects can lead to significant and algorithm-dependent errors in the retrieved cloud property values. With respect to pixels that are affected

by the rainbow and the glory, SEVIRI pixels with scattering angles between 134° and 140° and close to 180° were excluded. Reflectance at these angles is more influenced by the shape of the size distribution and not only by the effective radius. Finally, since a number of 1.6-μm detectors on the Aqua MODIS instrument are inoperative (Platnick et al., 2009) those pixels are eliminated within this study as well.

3.3.2 Results

3.3.2.1 Results over the North Atlantic

For the comparison over the ocean, the MSG SEVIRI scene from 11 June 2008 at 14:15 UTC (figure 3.1 b) and the spatially and temporally co-located Aqua MODIS scenes from 11 June 2008, 14:15 and 14:20 UTC were chosen (figure 3.1 a). The area covers the Atlantic Ocean adjacent to the west coast of Europe (geographic boundaries are listed in table 3.2) and clouds are located completely over the ocean. Thus, the influence of the background reflection is minimized. The MSG viewing angles range from about 38° to 68°.

Region	Minimum latitude	Maximum latitude	Minimum longitude	Maximum longitude	Minimum satellite zenith	Maximum satellite zenith
North Atlantic	31.00	48.50	-22.50	-12.00	37.70	68.94
Central Europe	40.00	56.00	0.50	18.50	47.30	61.34

Table 3.2: *Geographic coordinates and satellite zenith angles of the study regions.*

To evaluate the performance of SLALOM while minimizing the aforementioned potential error sources that depend on the satellite characteristics, i. e. different viewing geometries, instrument calibration and spectral resolution, a comparison between SLALOM applied to Aqua MODIS data and the NASA MODIS M06 product is carried out first. Hence, differences can be solely attributed to different assumptions within the retrievals and potentially negligible background albedo. Optical thickness (effective radius) of the chosen cloud field can be seen in figures 3.3 a and b (3.4 a and b). Please note that in these figures, the scenes have already been projected to the grid used for the SEVIRI/MODIS comparison but the statistics have been computed using the 0.01° by 0.01° grid. By using only pixels with an optical thickness larger than 5 in both retrievals, a test sample of 369,185 pixels is used for this first comparison. A statistical summary is presented in the first two columns of table 3.3.

Values for the optical thickness of SLALOM and M06 range from 5 to about 40 with very similar mean values of 11.74 (SLALOM) and 11.30 (M06) and standard deviations of 5.42 (SLALOM) and 5.41 (M06) respectively and the two datasets

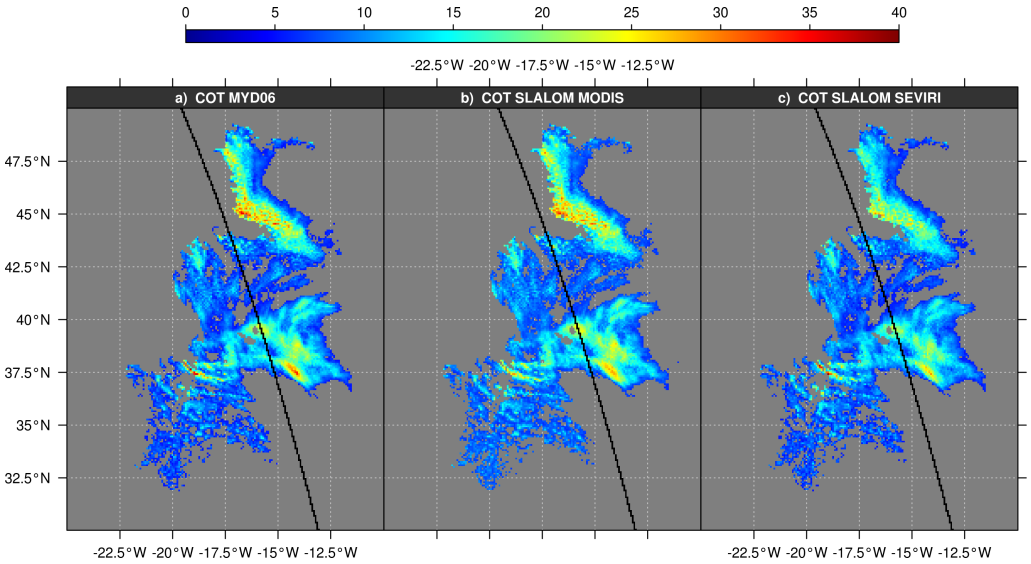

Figure 3.3: *Cloud optical thickness retrieved by (a) M06, (b) SLALOM using Aqua MODIS data and (c) SLALOM using MSG SEVIRI data for MSG SEVIRI scene from 11 June 2008 14:15 UTC and Aqua MODIS scenes from 11 June 2008 14:15 and 14:20 UTC (values shown for τ>5; 0.08° x 0.08°). The CloudSat track is denoted by the black line.*

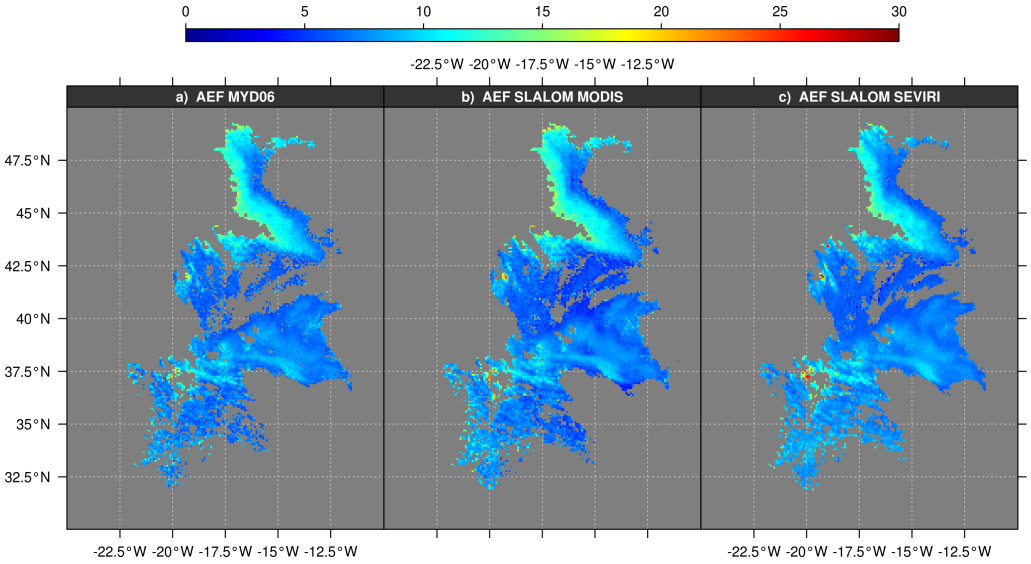

Figure 3.4: *Cloud effective radius retrieved by (a) M06, (b) SLALOM using Aqua MODIS and (c) SLALOM using MSG SEVIRI data for MSG SEVIRI scene from 11 June 2008 14:15 UTC and for Aqua MODIS scenes from 11 June 2008 14:15 and 14:20 UTC (values shown for τ>5; 0.08° x 0.08°).*

Optical thickness	SLALOM MODIS τ>5/τ>10 0.01° x 0.01°	M06 MODIS τ>5/τ>10 0.01° x 0.01°	SLALOM SEVIRI τ>5/τ>10 0.08° x 0.08°	M06 MODIS τ>5/τ>10 0.08° x 0.08°
Min	5.00/10.00	5.00/10.00	5.00/10.00	5.00/10.00
Max	39.68/39.68	40.63/40.63	38.97/38.97	47.00/47.00
Median	10.20/14.70	9.77/14.20	9.43/13.52	9.61/14.59
Mean	11.74/16.04	11.30/15.59	10.45/14.41	11.04/15.72
Std	5.42/4.84	5.41/4.86	4.21/3.56	5.06/4.51
r vs. M06	0.99/0.99		0.95/0.89	
MBE	-0.44/-0.46		0.59/1.31	
Std_Diff	0.51/0.44		5.07/2.08	
Effective radius				
Min	3.00/3.17	1.18/2.49	3.03/5.53	4.95/5.36
Max	37.96/36.74	30.1/29.45	34.51/20.16	19.57/17.75
Median	7.57/8.12	7.32/8.03	7.66/8.24	7.45/7.93
Mean	8.12/8.58	8.07/8.47	8.07/8.68	8.04/8.56
Std	2.88/2.49	2.68/2.36	1.90/1.90	2.00/2.08
r vs. M06	0.94/0.98		0.83/0.91	
MBE	-0.05/-0.10		-0.02/-0.10	
Std_Diff	0.94/0.54		1.17/0.86	

Table 3.3: *Comparison of SLALOM and M06 for scenes from figures 3.5 and 3.6. Columns 1 and 2 are based on the 0.01° x 0.01° projection used for the MODIS only comparison while all other columns represent the 0.08° x 0.08° resolution of the common evaluation grid between MODIS and MSG SEVIRI. Results are presented for values with τ>5/τ>10 (Std = standard deviation, Std_Diff = standard deviation of differences, MBE = mean bias error).*

show a strong linear correlation with a coefficient of correlation r of 0.99 (see figure 3.5 a). The corresponding percentage difference is shown in figure 3.5 c. For τ smaller than 12, the SLALOM values are predominantly larger than the M06 values and the variations range between -20% and +25% with maximum deviations for some pixels as high as 50%. With increasing τ, the spread decreases and the deviation stays mainly between -2% and +10%.

The corresponding effective radius retrievals show similar correlation with r of 0.94 (figure 3.5 b). Again, the deviation decreases towards larger a_{ef}. The rather strong deviations of more than -50% for a_{ef} values around 4 are linked to clouds with small optical thickness (τ<10) and are based on pixels located towards the cloud borders. The coefficient of correlation r increases to 0.98, if just a_{ef} values with τ>10 are considered.

In general, SLALOM and M06 retrieved cloud parameters show only some small differences and the overall agreement is very good. A perfect agreement cannot be expected even between retrievals of the same type (e. g. Nauss et al., 2005). Since both SLALOM and M06 are based on the same MODIS dataset, the results from figure 3.5 and the first two columns of table 3.3 can be regarded as a

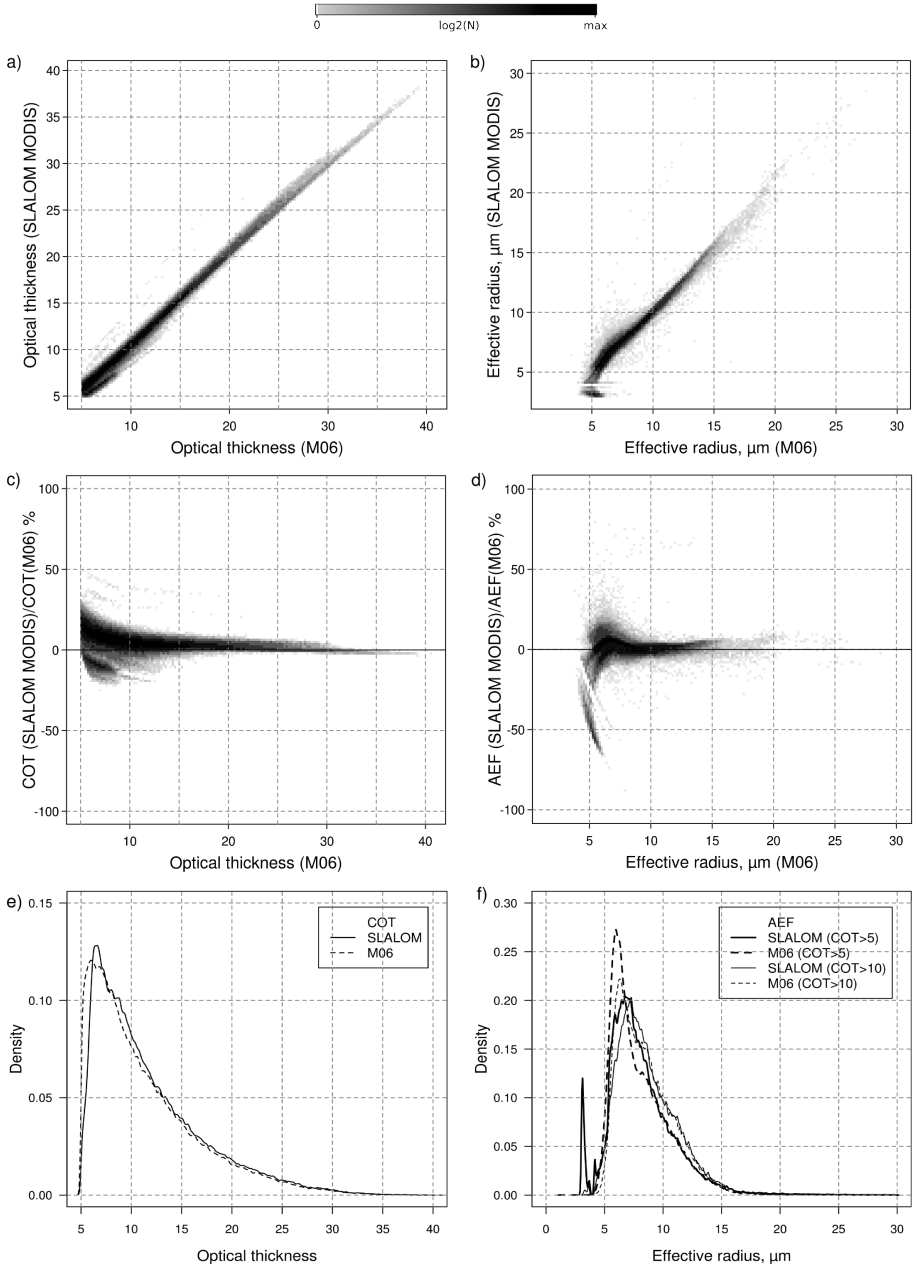

Figure 3.5: *(a) Cloud optical thickness and (b) cloud effective radius retrieved by SLALOM using MODIS data vs. M06, (c,d) corresponding percentage difference and (e,f) distributions for Aqua MODIS scenes from 11 June 2008, 14:15 and 14:20 UTC (values shown for τ>5; 0.01° x 0.01°).*

relative baseline for the following evaluations with respect to the interpretation of the agreement between the SEVIRI-, MODIS-, and CloudSat-based results discussed in the next section.

In the second step, SLALOM is applied to MSG SEVIRI data and compared to the MODIS M06 product on the 0.08° by 0.08° grid as well as the CloudSat 2B-TAU product. In contrast to the previous investigation, this comparison implies potential errors arising from issues related to different satellite sensors characteristics.

First, the differences between SLALOM using MSG SEVIRI data and the MODIS M06 product are evaluated. Optical thickness (effective radius) retrieved by SLALOM and M06 are shown in figures 3.3 a and c (figures 3.4 a and c). Taking into account only values with τ>5 leads to a test sample of 9348. An overview of the dataset is listed in columns three and four of table 3.3. The scatter plots and distributions of $\tau_{Sslalom}$ versus τ_{M06} (figures 3.6a and e) and $a_{efSslalom}$ versus a_{efM06} values (figures 3.6 b and f) reveal that the relative number of smaller values is strongly increased in comparison to the results presented in the preceding study. This can probably be attributed to the different viewing geometries of each satellite.

$\tau_{Sslalom}$ and τ_{M06} show a coefficient of correlation r of 0.95 (figure 3.6 a) and their distributions have a very similar shape and positive skewness (figure 3.6 e). The corresponding percentage difference in figure 3.6 c illustrates that for τ smaller than 10, SLALOM values are predominantly larger than the M06 values and the variations range mainly between -10% and +30%. With increasing τ, the spread decreases. The spatial percentage differences shown in figure 3.7 a show that higher deviations can mainly be found in the southern, less homogeneous cloud area where the differences can mainly be attributed to the different spatial resolutions of the two sensors and the larger SEVIRI pixels may encompass a mixed cloud scenario. Since all retrievals assume plane-parallel and homogeneous clouds throughout each pixel, this type of error generally increases with pixel size because the area of unknown inhomogeneity within the pixel increases (Zinner et al., 2005). To get information about the cloud horizontal homogeneity for each SEVIRI pixel, the standard deviation of MODIS reflection values of the 0.856 μm channel (0.01° to 0.01°) lying within one SEVIRI pixel is calculated. Figure 3.8 shows the distribution of standard deviations considering all SEVIRI pixels within the considered cloud area shown in figure 3.3 with τ>5 (black line). A high number of SEVIRI pixels encompass non-homogeneous clouds with standard deviations of up to 0.19. The horizontal homogeneous index (i. e. the mean over all standard deviations) is 0.05. For comparison, if only pixels lying within the homogeneous area around 38.5 N° and 14° W are taken into consideration, the relative number of smaller standard deviations is increased and values range to 0.085 (dashed line). The corresponding homogeneous index decreases to 0.046.

The scatter plot of $a_{efSslalom}$ against a_{efM06} in figure 3.6 b reveals a correla-

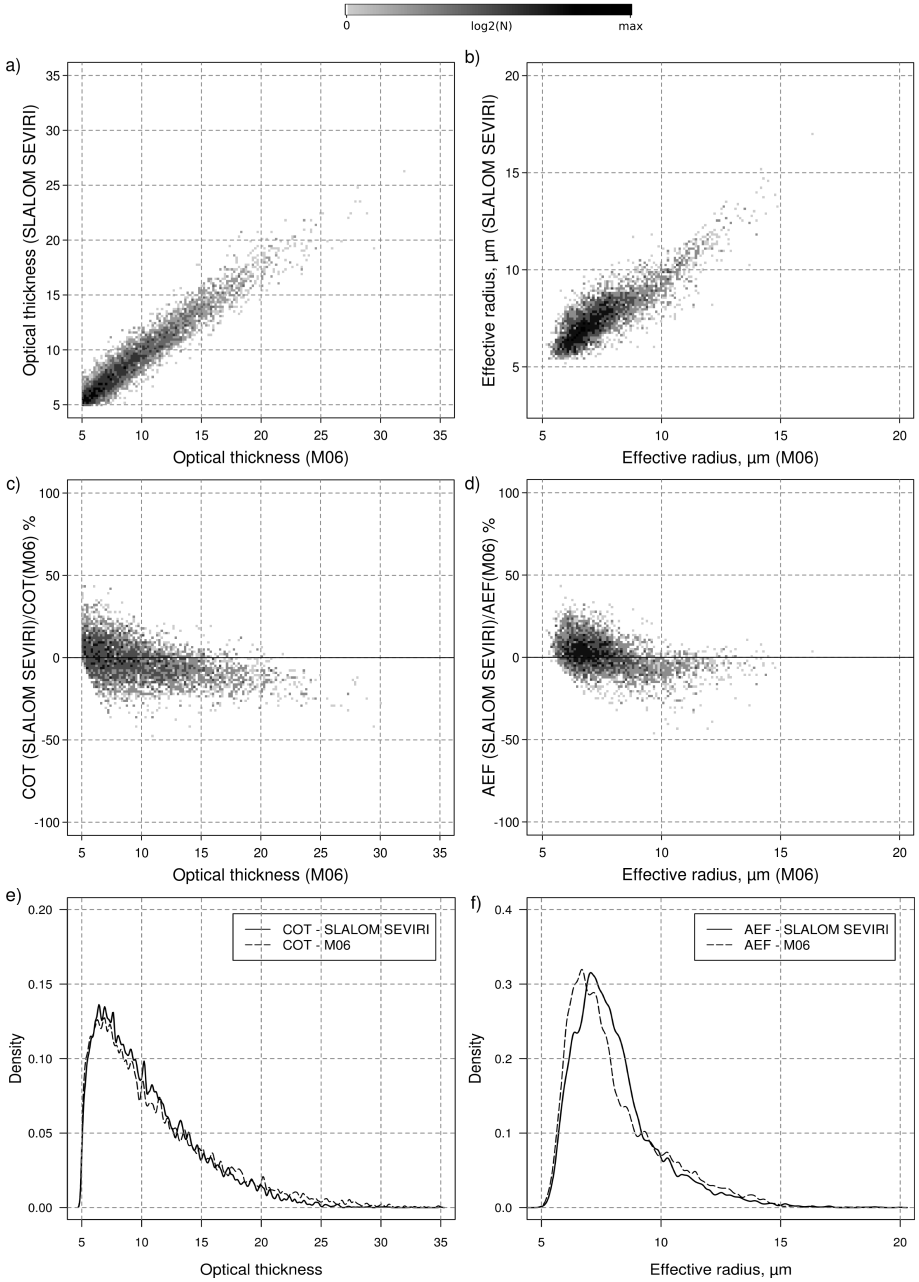

Figure 3.6: *(a) Cloud optical thickness and (b) cloud effective radius retrieved by SLALOM using MSG SEVIRI data vs. M06, (c,d) corresponding percentage difference and (e,f) distributions for scenes from figures 3.3 and 3.4 (values shown for $\tau > 5$; 0.08° x 0.08°).*

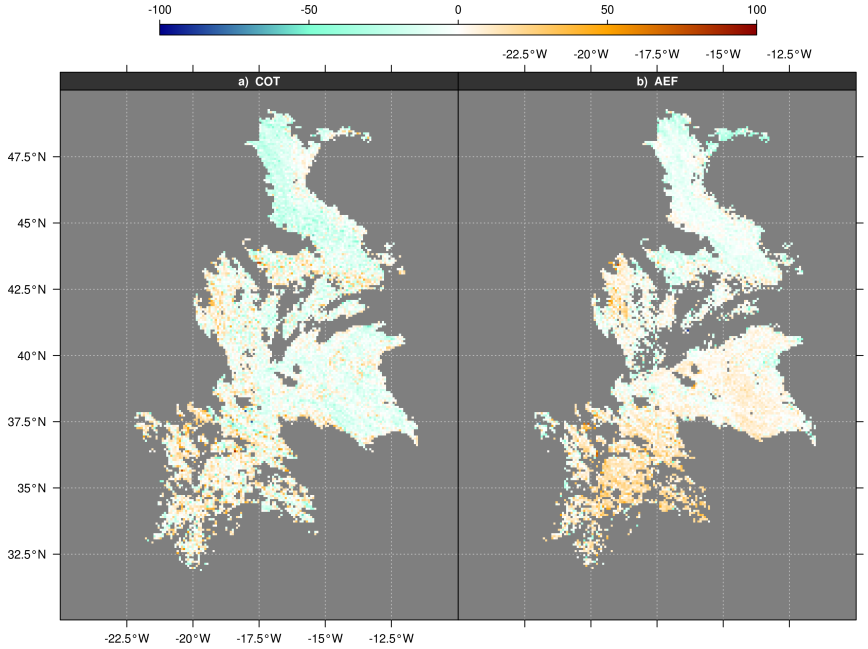

Figure 3.7: *Spatial distribution of the percentage deviation of (a) cloud optical thickness and (b) cloud effective radius retrieved by SLALOM and M06 for scenes from figures 3.3 a and c and 3.4 a and c (values shown for τ>5; 0.08° x 0.08°).*

tion with r of 0.83. As can be seen in figure 3.6 d, the corresponding percentage difference reveals that $a_{efSslalom}$ are generally larger for a_{ef}<9.0 μm and show deviations of up to about +40% for decreasing a_{ef}. For a_{ef}>9.0 μm $a_{efSslalom}$ values are smaller and deviation ranges generally between +5% and -10%. The deviations of some single pixels are around +-50%. The different modes of the $a_{efSslalom}$ and a_{efM06} distributions shown in figure 3.6 f might be caused by the distributions of SLALOM and SEVIRI reflectances at 1.6 μm. A spatial representation of the relative differences between the two datasets is given in figure 3.7 b. The largest deviations with a bias of +-30% and some single pixels up to +-60% can be found mainly in the south-western area and again along the cloud borders. These deviations are likely to be traced back to enhanced 3D effects resulting from horizontally and vertically inhomogeneous clouds and sub-pixel cloudiness often occurring at the cloud edges.

In the third step, the comparison of the three different types of cloud optical thickness $\tau_{Sslalom}$, $\tau_{Mslalom}$ and τ_{M06} against the CloudSat 2B-Tau product is carried out. CloudSat provide an accurate cloud optical depth profile and detailed information on the sub-pixel heterogeneity of the cloud structure within the SEVIRI pixel.

τ_{M06}, $\tau_{Mslalom}$ and $\tau_{Sslalom}$ as well as the CloudSat track are shown in figures 3.3a, b and c, respectively. Datasets of spatially and temporally corresponding pairs of optical thickness derived by M06, $SLALOM_M$, $SLALOM_S$ and CloudSat 2B-TAU are extracted. Taking into account all temporally and spatially cor-

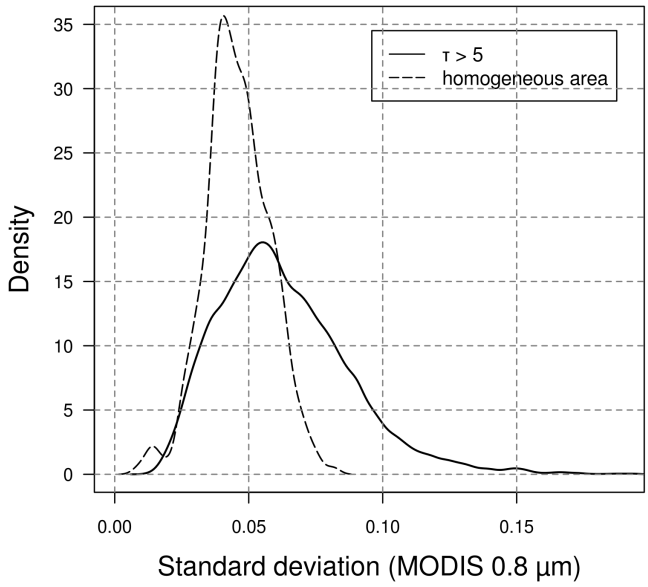

Figure 3.8: *Distributions of standard deviations arose from MODIS 0.8 µm reflectances lying within one MSG SEVIRI pixel.*

responding pixels leads to a test sample of 686 (SLALOM$_S$) and 337 (SLALOM$_M$, M06) value pairs. Note that the test sample of MODIS is reduced in comparison to SEVIRI, since a number of 1.6-μm detectors on the Aqua MODIS instrument are inoperative and therefore those pixels are eliminated within this study. Due to their different viewing characteristics, between one and five CloudSat profiles can be found within one MSG SEVIRI pixel. Therefore a mean over all CloudSat profiles located within one MSG SEVIRI pixel is taken for the present evaluation. As a result, the number of test samples is reduced to 175 (SLALOM$_S$). Tables 3.4 and 3.5 summarize basic descriptive statistics of the mentioned datasets.

Optical thickness	SLALOM MODIS	M06 MODIS	2B-TAU CloudSat
Min	5.28	5.24	1.01
Max	29.27	28.28	31.79
Median	10.89	10.36	10.82
Mean	12.64	12.2	12.92
Std	5.67	5.52	6.28
r vs. 2B-TAU	0.95	0.96	
MBE	0.28	0.72	
Std_Diff	1.98	1.98	

Table 3.4: *Statistical values from the dataset of spatially and temporally corresponding $\tau_{Mslalom}$, τ_{M06} and τ_{2B-TAU} for scenes from figures 3.5 a and b.*

The minimum values of 2B-TAU are lower and the maximum values are higher than those inferred from SLALOM$_M$, M06 and SLALOM$_S$. The latter as well

Optical thickness	SLALOM MSG SEVIRI	2B-TAU CloudSat
Min	3.06	0.55
Max	25.33	30.77
Median	10.01	8.97
Mean	10.75	10.36
Std	5.76	6.76
r vs. 2B-TAU	0.92	
MBE	-0.42	
Std_Diff	2.67	

Table 3.5: *Statistical values from the dataset of spatially and temporally corresponding $\tau_{Sslalom}$ and τ_{2B-TAU} for scene from figure 3.5 c (mean over all CloudSat profiles located within one SEVIRI pixel is taken).*

as the higher standard deviations of the 2B-TAU datasets indicate a typical behaviour when datasets with higher spatial resolution are compared to datasets with lower resolution. The mean optical thickness of 10.75 ($SLALOM_S$) and 10.36 (2B-TAU) as well as the mean values of 12.64 ($SLALOM_M$), 12.2 (M06) and 12.92 (2B-TAU) are very close to each other. As can be seen in the scatter plots shown in figures 3.9 a and c, $\tau_{Mslalom}$ and τ_{2B-TAU} as well as τ_{M06} and τ_{2B-TAU} reveal a clear positive correlation with coefficients of correlation of 0.95 and 0.96. The corresponding percentage differences of both datasets (figures 3.9 b and d) show a very similar spreading. For τ smaller than 10, $\tau_{Sslalom}/\tau_{M06}$ values are larger than the corresponding τ_{2B-TAU} values. With increasing τ, the spread decreases and the deviation ranges mainly between -30% and +25%. $\tau_{Sslalom}$ and τ_{2B-TAU} possess a slightly smaller but still very good correlation with r of 0.92 (figure 3.10 a). The spreading of the corresponding percentage differences (figure 3.10 b) is very similar to those shown in figures 3.9 b and d.

Figures 3.11 and 3.12 show the evolution of τ_{M06}, $\tau_{Mslalom}$ and $\tau_{Sslalom}$ values in comparison to spatially corresponding τ_{2B-TAU} values for each Aqua MODIS/MSG SEVIRI pixel along the tracks depicted in figures 3.3 a, b and c from N-S. Hence the first pixel in the upper part of figures 3.11 and 3.12 corresponds to the northernmost pixel in the satellite images and the following data pairs represent the development along the CloudSat track to the southern part. Figure 3.11 indicates that in most cases τ_{M06} and $\tau_{Mslalom}$ values correspond very well with CloudSat 2B-Tau. In contrast to the MODIS-based comparison where only one CloudSat profile can be located within one MODIS pixel, it is possible to locate one to five CloudSat profiles within one MSG SEVIRI pixel. In order to get also information about sub-pixel cloud heterogeneity within one MSG SEVIRI pixel, instead of the mean over all CloudSat profiles assigned to the accordingly pixel, the evolution of $\tau_{Sslalom}$ values (grey dots) in comparison to the range of spatially corresponding τ_{2B-TAU} values (black lines) within each MSG SEVIRI pixel is shown in figure 3.12. The results indicate that in most cases $\tau_{Sslalom}$

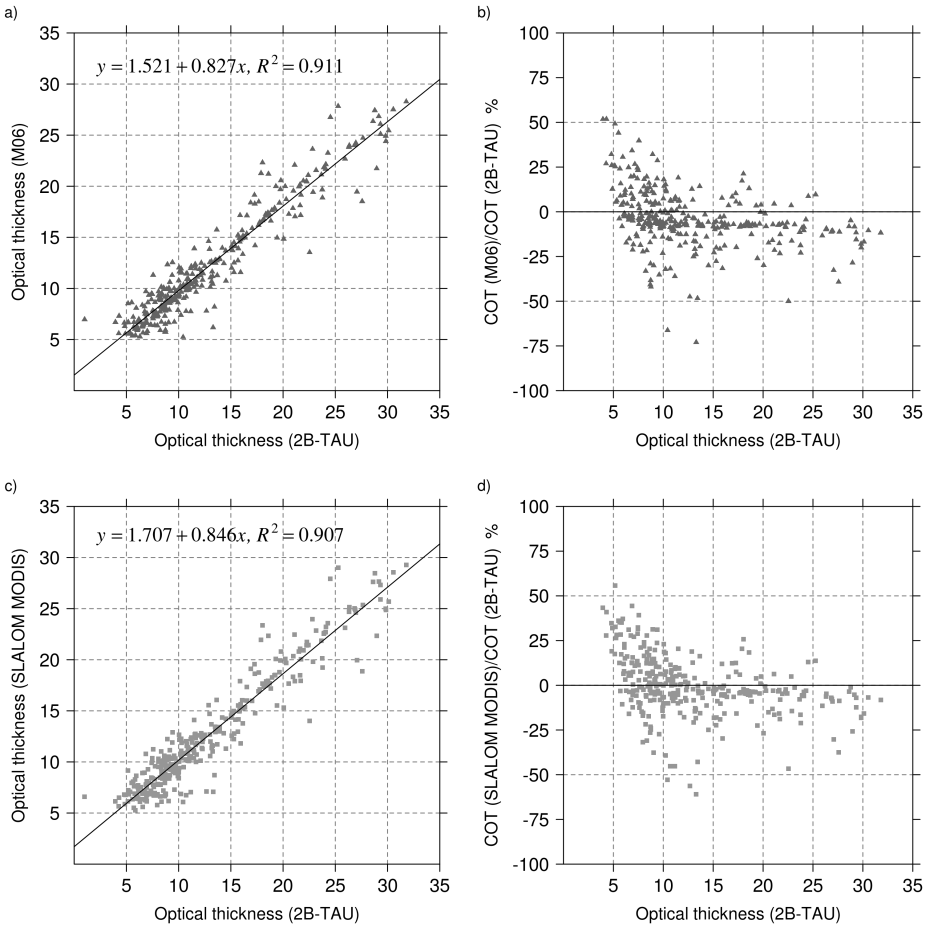

Figure 3.9: *(a) Cloud optical thickness retrieved by SLALOM using Aqua MODIS data and (c) M06 vs. CloudSat 2B-TAU optical thickness and (b,d) corresponding percentage difference for scenes from figures 3.3 a and b.*

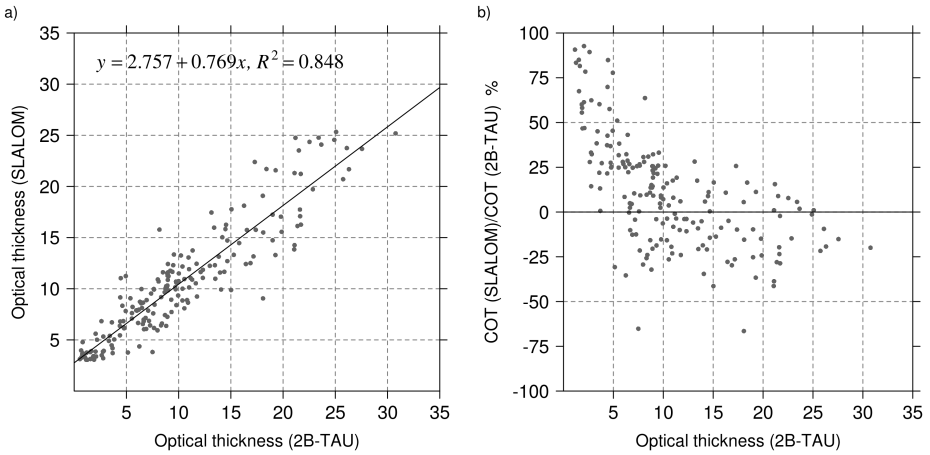

Figure 3.10: *(a) Cloud optical thickness retrieved by SLALOM using MSG SEVIRI data vs. CloudSat 2B-TAU optical thickness and (b) corresponding percentage difference for North Atlantic scene from 11 June 2008, 14:15 UTC (mean over all CloudSat profiles within one SEVIRI pixel is taken).*

values are well within the range of corresponding τ_{2B-TAU} values except for very small optical thickness (smaller than 5) which are overestimated by SLALOM. Cloud inhomogeneities are evident through large ranges of τ_{2B-TAU} values that can be seen within some MSG SEVIRI pixels. Due to the larger MSG SEVIRI pixels, the likelihood of a mixed cloud scenario being located within one pixel is increased, which is likely to result in less accurate estimations of cloud optical thickness in these cloud areas as compared to CloudSat. However, the overall evolution is mapped very well by the $\tau_{Sslalom}$ values.

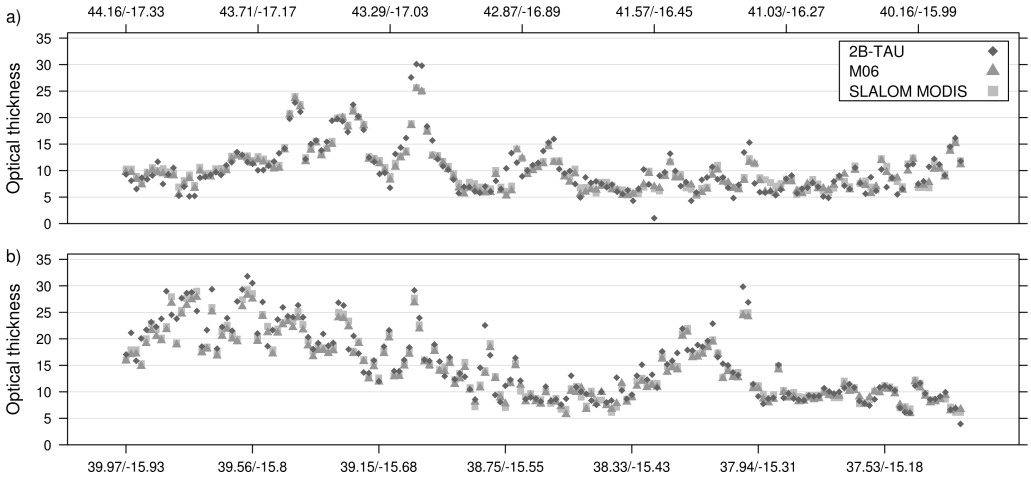

Figure 3.11: *Cloud optical thickness retrieved by SLALOM using Aqua MODIS data, M06 and CloudSat 2B-TAU optical thickness for scenes from figures 3.3 a and b. Retrieved τ values are displayed for each MODIS pixel.*

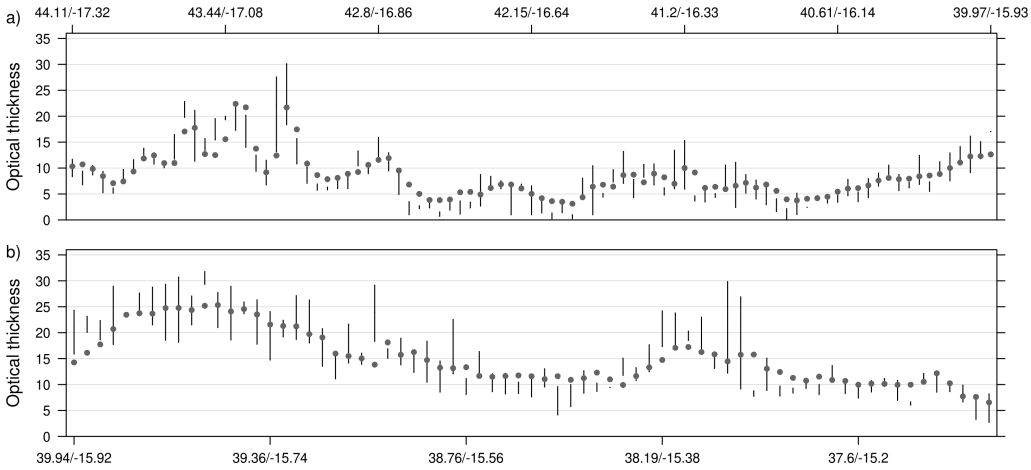

Figure 3.12: *Cloud optical thickness retrieved by SLALOM using MSG SEVIRI data vs. CloudSat 2B-TAU optical thickness for scene from figure 3.3 c. For each SEVIRI pixel the retrieved $\tau_{Sslalom}$ value (grey dot) and the range of spatial corresponding τ_{2B-TAU} values (black line) are displayed.*

While this section has been restricted to an evaluation between the SLALOM, M06 and CloudSat 2B-TAU product over ocean, the situation is complicated over land. There, the influence of the ground surface reflection increases and the resulting consequences are presented in the next section.

3.3.2.2 Results over Central Europe

The MSG SEVIRI scene from 28 June 2008 at 09:45 UTC and the temporally and spatially co-located Terra MODIS scene from 28 June 2008 at 10:00 UTC were chosen for the comparison over land. The area covers Germany according to the geographic boundaries listed in table 3.2.

Just like in the previous comparison, SLALOM will be applied to MODIS data and the results will be compared against the MODIS M06 product first. The results of the analogous comparison of SLALOM against M06 over the ocean is shown in figure 3.5 and the first two columns of table 3.3 reveal differences between the algorithms over the ocean and therefore over an area with quite homogeneous background albedo. These results can be regarded as baseline for the following evaluation with respect to the interpretation of the agreement between SEVIRI and MODIS-based results over land where the influence of background albedo strongly increases.

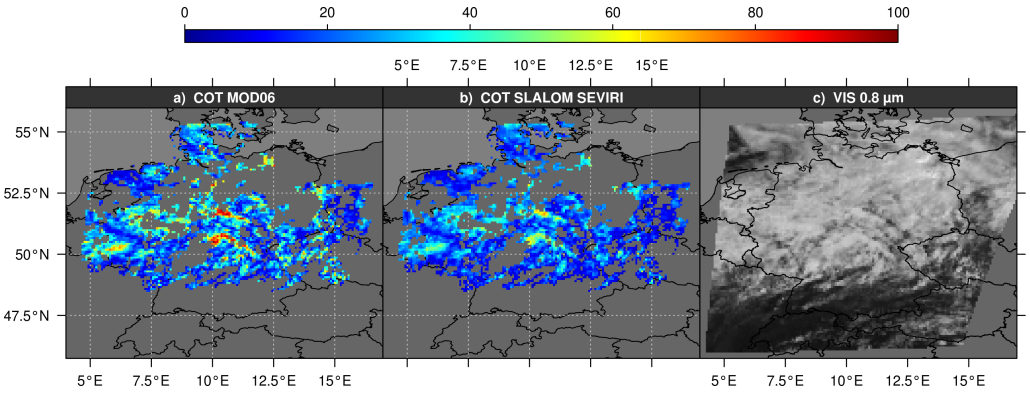

Figure 3.13: *Cloud optical thickness retrieved by (a) M06, (b) SLALOM using MSG SEVIRI data and (c) reflections of MSG SEVIRI 0.84 µm-channel for MSG SEVIRI scene from 28 June 2008 09:45 UTC and Terra MODIS scene from 28 June 2008 10:00 UTC (values shown for τ>5; 0.08° x 0.08°).*

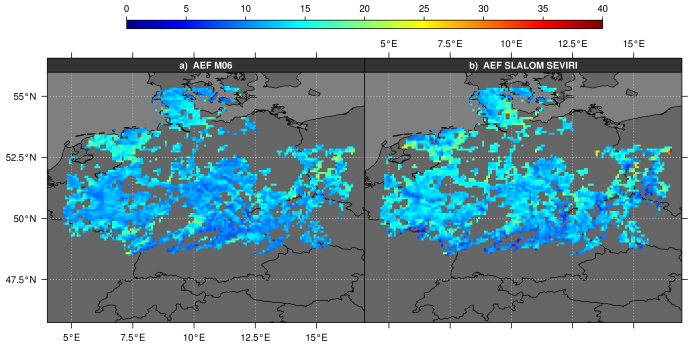

Figure 3.14: *Cloud effective droplet radius retrieved by (a) SLALOM using MSG SEVIRI data and (b) M06 for MSG SEVIRI scene from 28 June 2008 09:45 UTC and Terra MODIS scene from 28 June 2008 10:00 UTC (values shown for τ>5; 0.08° x 0.08°).*

The cloud field which is taken into consideration can be seen in the area of the retrieved optical thickness in 3.13 a (M06) and b (SLALOM) and the re-

trieved effective radius in figures 3.14 a (M06) and b (SLALOM) (please note again, that these images have been reprojected to the 0.08° to 0.08° grid used for the SEVIRI/MODIS comparison but the statistical evaluation has been performed on the initial 0.01° grid). Ice phase clusters within the cloud fields are excluded from the comparison and once again, only clouds with τ>5 are taken into account. A statistical overview about the dataset is listed in the first two columns of table 3.6.

Optical thickness	SLALOM MODIS τ>5/τ>10 0.01° x 0.01°	M06 MODIS τ>5/τ>10 0.01° x 0.01°	SLALOM SEVIRI τ>5/τ>10 0.08° x 0.08°	M06 MODIS τ>5/τ>10 0.08° x 0.08°
Min	5.00/10.00	5.01/10.02	5.33/10.65	5.03/10.02
Max	98.67/98.67	99.98/99.98	72.81/72.81	97.43/97.44
Median	22.84/26.10	22.94/26.42	18.51/21.45	23.31/27.04
Mean	25.23/28.56	26.98/30.57	20.29/23.59	26.68/30.49
Std	13.16/12.54	16.52/15.89	10.62/9.70	15.52/15.03
r vs. M06	0.95/0.92		0.82/0.80	
MBE	2.20/2.55		6.39/6.90	
Std_Diff	7.15/7.75		9.10/9.22	
Effective radius				
Min	3.00/3.10	4.35/4.35	4.55/7.23	6.50/7.30
Max	37.15/37.04	35.05/35.05	36.60/36.60	25.64/25.64
Median	10.53/10.96	11.8/11.7	12.83/13.20	11.85/11.70
Mean	10.84/11.44	12.29/12.13	13.07/13.57	12.30/12.13
Std	3.39/3.10	3.10/2.95	2.93/2.58	2.43/2.27
r vs. M06	0.68/0.89		0.52/0.68	
MBE	1.45/0.68		-0.77/-1.44	
Std_Diff	2.62/1.45		2.68/1.95	

Table 3.6: *Comparison of SLALOM and M06 for scenes from figures 3.13 and 3.14. Columns 1 and 2 are based on the 0.01° x 0.01° projection used for the MODIS only comparison while the other columns represent the 0.08° x 0.08° resolution of the common evaluation grid between MODIS and MSG SEVIRI. Results are presented for values with τ>5/τ>10.*

As can be seen in figure 3.15 a, optical thickness values retrieved by SLALOM and M06 show good agreement with r of 0.95. For τ smaller than 20, the SLALOM deviations range mainly between +-20% and with increasing τ, the spread decreases initially. For τ larger than 50 the spread increases again and the SLALOM values are predominantly smaller than the M06 values (figure 3.15 c). Figure 3.15 e shows that both distributions have a positive skewness and show very good agreement.

The corresponding effective radius shows a correlation coefficient r of 0.64. As can be seen in figure 3.15 d, the percentage difference reveal that $a_{ef Mslalom}$ are generally smaller and for some pixels the deviations increase up to over -100%.

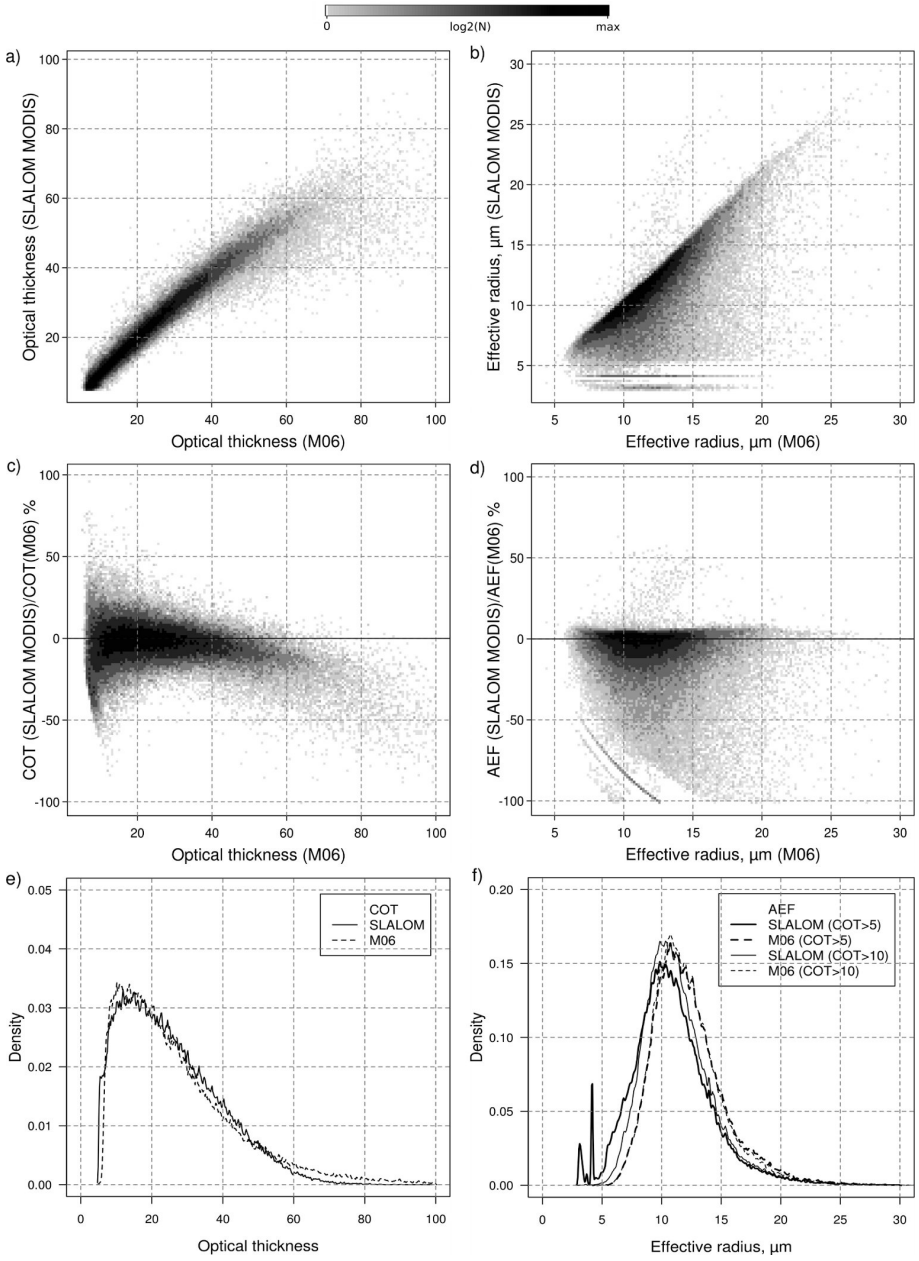

Figure 3.15: *(a) Cloud optical thickness and (b) cloud effective radius retrieved by SLALOM using MODIS data vs. M06, (c,d) corresponding percentage difference and (e,f) distributions for Terra MODIS scene from 28 June 2008 10:00 UTC (values shown for τ>5; 0.01° x 0.01°).*

Like over the North Atlantic, the deviations for $a_{efMslalom}$ values around 4 are linked to clouds with small optical thickness (τ<10) and are based on pixels located towards the cloud borders. If only pixels with τ>10 are considered, the distribution of $a_{efMslalom}$ values shows no deviations around 4 and the coefficient of correlation r increases to 0.89.

In the following section the investigation of differences in cloud properties retrieved by SLALOM using MSG SEVIRI data compared to M06 is carried out. In contrast to the Atlantic scene, the 0.6 μm channel was used this time to retrieve the SLALOM-based cloud optical properties, since M06 uses the 0.65 μm channel over land surfaces. Figures 3.13 a and b show the optical thickness retrieved by SLALOM and M06, respectively. The extracted dataset of spatially corresponding $\tau_{Sslalom}$ and τ_{M06} values consists of a test sample of 4452. τ values range from 5 to 97.44 (M06) and 72.81 (SLALOM) (see table 3.6 for statistical overview of the dataset). A correlation coefficient of 0.80 reveals a good linear correlation between $\tau_{Sslalom}$ and τ_{M06} (figure 3.16 a). The corresponding percentage difference shows that variations range predominantly between -50 and +50% (figure 3.16 c).

Retrieved $a_{efSslalom}$ and a_{efM06} values are presented in figures 3.14 a and b. The corresponding scatter plot in figure 3.16 b and a correlation coefficient r of 0.52 reveal that this correlation is weaker. The main range of a_{ef} values is between 9 and 16 μm and therefore considerably reduced. The values of $a_{efSslalom}$ are mainly larger and show predominantly deviations from -10% to +30% (figure 3.16 d). This is also evident in figure 3.17 b, where the spatial distribution of percentage difference is shown.

The results of $a_{efSslalom}$ and a_{efM06} show substantially larger differences than the corresponding results for the Atlantic scene, but the overall agreement still indicates a correlation. Other than the aforementioned reasons for the uncertainties (e. g. different retrievals, different satellite systems etc.), the major differences over land result from the influence of the surface albedo and small-scale cloud inhomogeneities. In particular, the larger MSG SEVIRI pixels with a resolution of about 5 to 8 km in the study area are likely to encompass mixed cloud scenarios and gaps in the cloud field which can be described much more precisely by the high resolution MODIS sensor (1 km x 1 km).

Just as for the comparison over the Atlantic, the CloudSat 2B-TAU product is included into the investigation. Because of the scattered cloud fields, four scenes from May, August and September 2008 over Central Europe were chosen for the evaluation in order to increase the dataset. Retrieved $\tau_{Sslalom}$ and the corresponding CloudSat tracks are shown for all scenes in figure 3.18. Datasets of spatially and temporally corresponding pairs of optical thickness derived by M06, SLALOMM, SLALOM$_S$ and CloudSat 2B-TAU are extracted. Taking into account all temporally and spatially corresponding pixels leads to a test sample of 1504 (SLALOM$_S$) and 780 (SLALOM$_M$, M06) value pairs. The number of $\tau_{Sslalom}$ test samples is reduced to 409, if a mean over all CloudSat profiles located

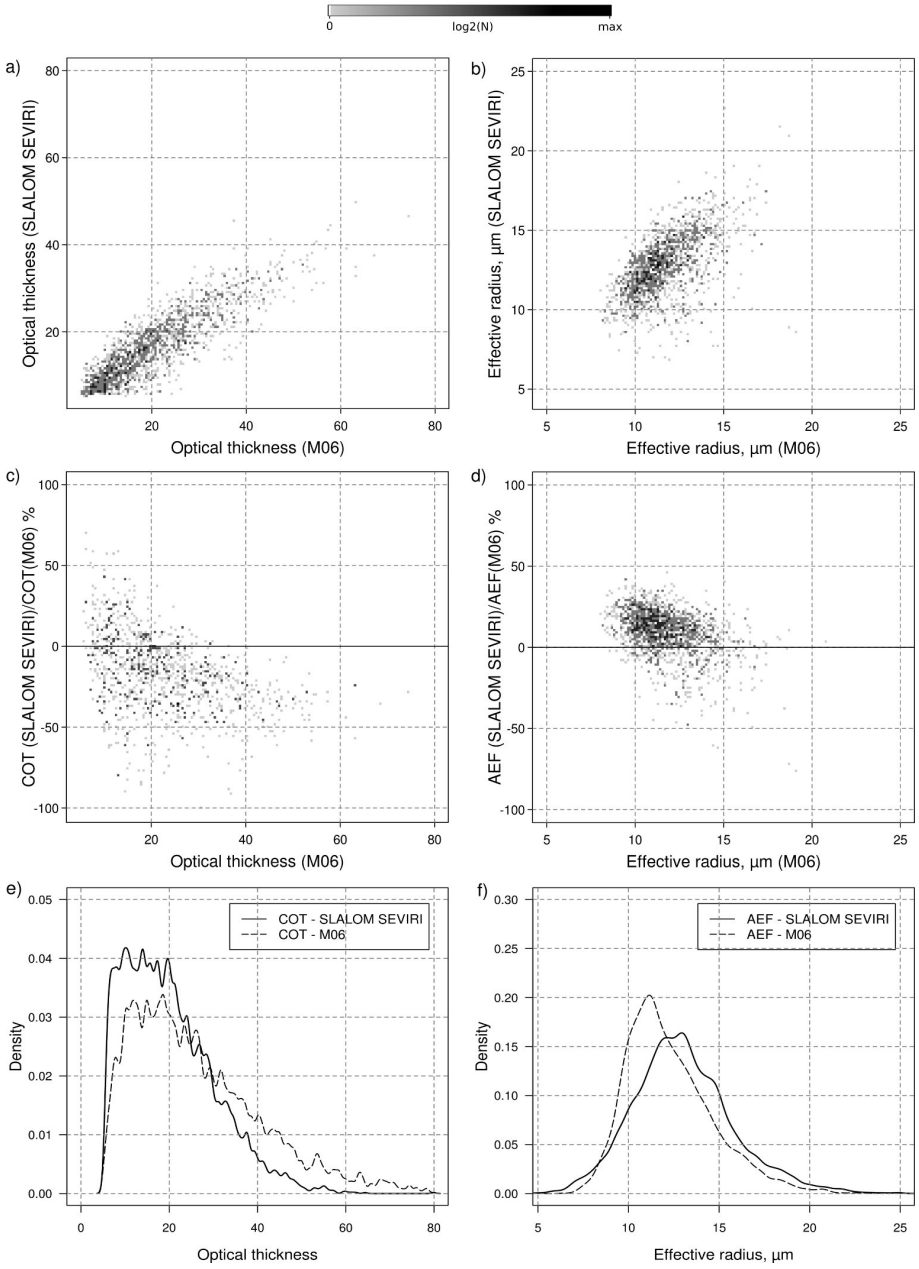

Figure 3.16: *(a) Cloud optical thickness and (b) cloud effective radius retrieved by SLALOM using MSG SEVIRI data vs. M06, (c,d) corresponding percentage difference and (e,f) distributions for MSG SEVIRI scene from 28 June 2008 09:45 UTC and Terra MODIS scene from 28 June 2008 10:00 UTC (values shown for τ>5; 0.08° x 0.08°).*

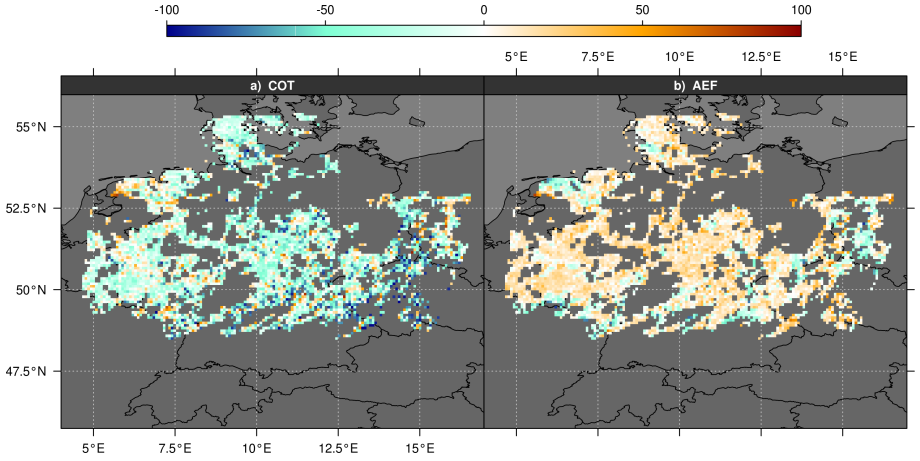

Figure 3.17: *Spatial distribution of the percentage deviation of (a) cloud optical thickness and (b) cloud effective radius retrieved by SLALOM and M06 for scenes from figure 3.13 a and b (optical thickness) and 3.14 a and b (effective radius) (values shown for $\tau > 5$; 0.08° x 0.08°).*

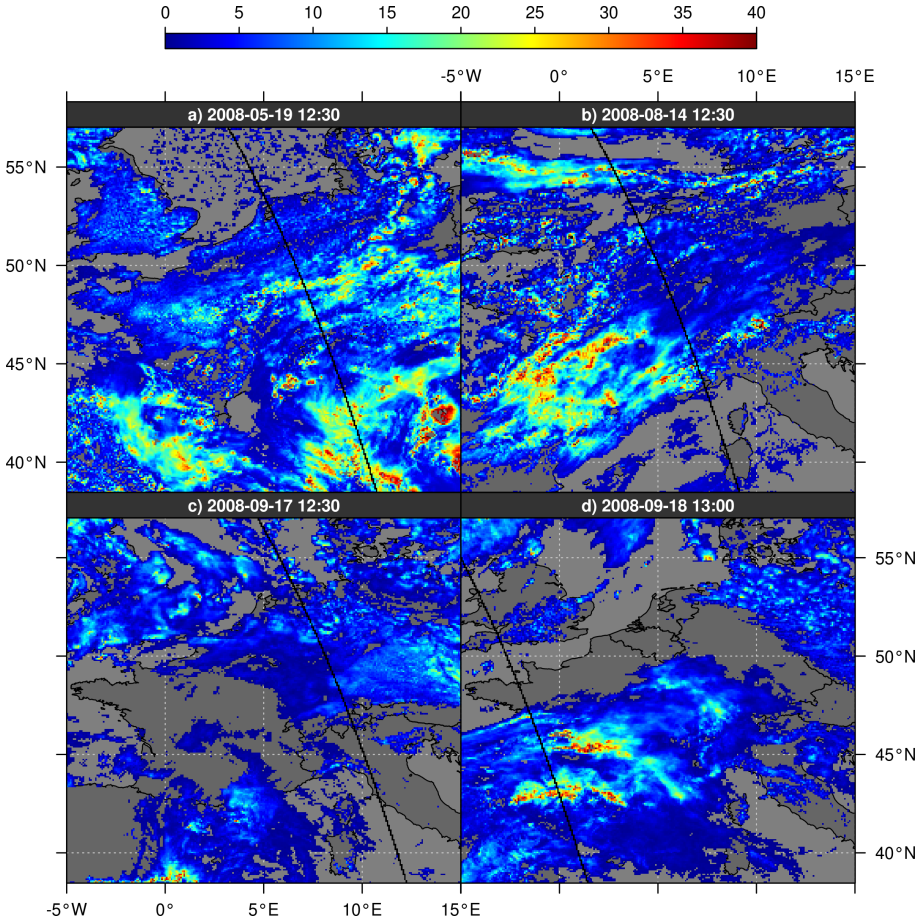

Figure 3.18: *Cloud optical thickness retrieved by SLALOM for scenes over Europe, (a) May 19th 2008 12:30 UTC, (b) August 14th 2008 12:30 UTC, (c) September 17th 2008 12:30 UTC and (d) September 18th 2008 13:00 UTC. The CloudSat tracks are denoted by black lines.*

within one MSG SEVIRI pixel is taken. A summary about the basic descriptive statistics of the mentioned datasets can be found in tables 3.7 and 3.8.

Optical thickness	SLALOM MODIS	M06 MODIS	2B-TAU CloudSat
Min	1.31	1.29	0.10
Max	30.49	72.00	70.21
Median	7.17	9.55	10.15
Mean	8.12	13.08	13.89
Std	5.41	12.51	12.14
r vs.2B-TAU	0.76	0.81	
MBE	5.76	0.81	
Std_Diff	8.78	7.63	

Table 3.7: *Statistical values from the dataset of spatially and temporally corresponding $\tau_{Mslalom}$, τ_{M06} and τ_{2B-TAU} for temporally corresponding Aqua MODIS scenes from figure 3.20.*

Optical thickness	SLALOM MSG SEVIRI	2B-TAU CloudSat
Min	0.03	0.10
Max	32.65	101.33
Median	5.80	10.87
Mean	7.87	15.80
Std	6.79	16.10
r vs.2B-TAU	0.73	
MBE	7.93	
Std_Diff	12.01	

Table 3.8: *Statistical values from the dataset of spatially and temporally corresponding $\tau_{Sslalom}$ and τ_{2B-TAU} for scenes from figure 3.20 (mean over all CloudSat profiles located within one SEVIRI pixel is taken).*

As can be seen in the scatter plot shown in figure 3.19 a, τ_{M06} and τ_{2B-TAU} reveal a clear positive correlation with coefficient of correlation of 0.81. The corresponding percentage difference (figure 3.19 b) illustrates that for τ smaller than 10, τ_{M06} values show the largest deviations up to -120% to +170%. With increasing τ, the spread decreases and the deviation ranges mainly between -50% and +50%. The coefficient correlation between $\tau_{Mslalom}$ and τ_{2B-TAU} possess a slightly smaller but still acceptable correlation with r of 0.76 (figure 3.19 c). However, as can be seen in figure 3.19 d, for τ larger than 10, $\tau_{Mslalom}$ values are always smaller than the corresponding τ_{2B-TAU} values. The scatter plot and percentage difference of $\tau_{Sslalom}$ and τ_{2B-TAU} (figures 3.20 a and b) are very similar to those shown in figures 3.19 c and d ($\tau_{Mslalom}$ vs. τ_{2B-TAU}). $\tau_{Mslalom}$ values are predominantly smaller than the corresponding τ_{2B-TAU} values (r=0.73).

The evolution of τ_{M06}, $\tau_{Mslalom}$ and $\tau_{Sslalom}$ values in comparison to spatially corresponding τ_{2B-TAU} values for each Aqua MODIS/MSG SEVIRI pixel along the tracks depicted in figure 3.18 from N-S is shown in figures 3.21 and 3.22.

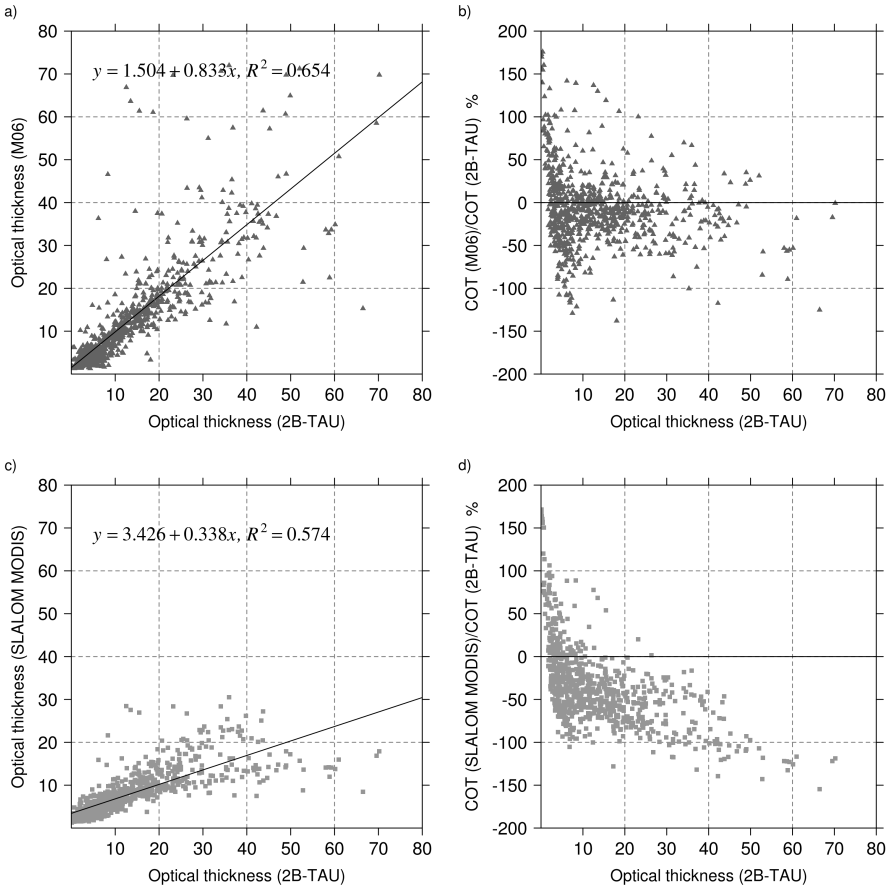

Figure 3.19: *(a) Cloud optical thickness retrieved by SLALOM using Aqua MODIS data and (c) M06 vs. CloudSat 2B-TAU optical thickness and (b,d) corresponding percentage difference for temporally corresponding Aqua MODIS scenes from figure 3.18.*

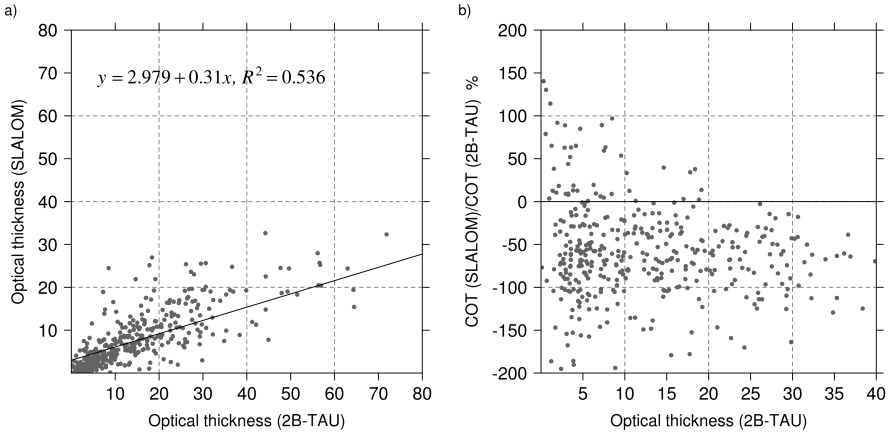

Figure 3.20: *(a) Cloud optical thickness retrieved by SLALOM using MSG SEVIRI data vs. CloudSat 2B-TAU optical thickness and (b) corresponding percentage difference for Europe scenes from figure 3.18 (mean over all CloudSat profiles within one SEVIRI pixel is taken.*

Figure 3.21 indicates that in most cases τ_{M06} values correspond very well with CloudSat 2B-Tau. $\tau_{Mslalom}$ values are generally smaller than the corresponding τ_{2B-TAU} values. This is especially pronounced for large τ_{2B-TAU} values. In figure 3.22 the evolution of $\tau_{Sslalom}$ values (grey dots) in comparison to the range of spatially corresponding τ_{2B-TAU} values (black lines) within each MSG SEVIRI pixel is shown. The large ranges of τ_{2B-TAU} values that can be seen within many SEVIRI pixels indicate the presence of significant cloud inhomogeneities. Due to the larger MSG SEVIRI pixels, the likelihood of a mixed cloud scenario being located within one pixel is increased. The results indicate that the overall evolution is mapped less accurately than the results over the North Atlantic. However, the tendency of the τ_{2B-TAU} cloud thickness evolution is mapped very well.

3.4 Summary and Conclusion

The aim of this study was the evaluation of the semi-analytical cloud retrieval algorithm SLALOM ported to SEVIRI on board MSG. The study is motivated by the enhanced information content potentially provided by a temporal high resolution observation system which is especially important if detailed cloud evolution monitoring is necessary or if the cloud properties are subsequently used for the delineation of instantaneously raining from non-raining cloud areas. The evaluation was realised by using two well-known cloud property retrievals, the LUT-based approach by Platnick et al. (2003) and CloudSat 2B-TAU product Polonsky et al. (2008). The intention was to assess the accuracy of SLALOM using MSG SEVIRI data in real-case situations. To the best of our knowledge it was the first application of SLALOM to geostationary satellite data. For the investigation, two study areas were chosen, one over the North Atlantic where the influence of auxiliary data is minimized (i. e. homogeneous ocean background albedo), and one over Central Europe which exhibits a higher influence of the background albedo and small-scale inhomogeneous clouds. SEVIRI reflectances were recalibrated using calibration coefficients found by Jan Fokke Meirink (KNMI) in order to correct the sensor ageing and calibration deficiencies and reduce calibration differences between MSG SEVIRI and MODIS. For the background albedo over land, a minimum composite of the reflectances in the visible (0.65 μm and 0.81 μm) and near-infrared (1.64 μm) channel over one month was calculated. To investigate the performance of SLALOM, the retrievals were evaluated against each other in three steps. First, SLALOM was applied to MODIS data and compared against the MODIS 06 product. Hence both algorithms were applied to the same sensor and differences can be attributed solely to the different assumptions within the retrievals and meta-datasets. Second, SLALOM was applied to MSG SEVIRI data and compared against the MODIS 06 product on a 0.08° by 0.08° grid. The comparison implied potential errors due to different algorithms and additionally

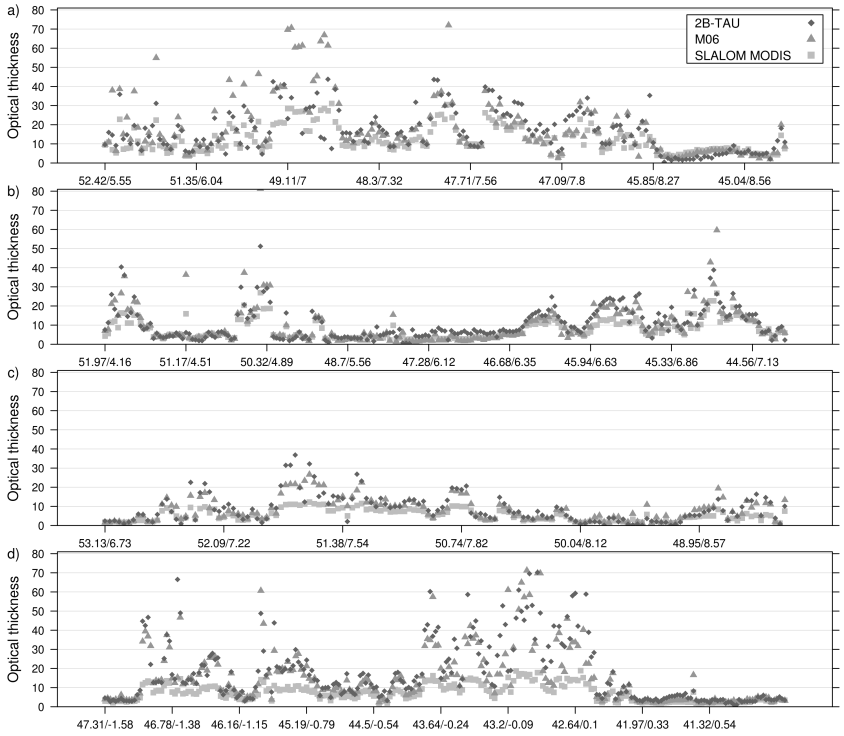

Figure 3.21: *Cloud optical thickness retrieved by SLALOM using Aqua MODIS data, M06 and CloudSat 2B-TAU optical thickness for temporally corresponding Aqua MODIS scenes from figure 3.18. Retrieved τ values are displayed for each MODIS pixel.*

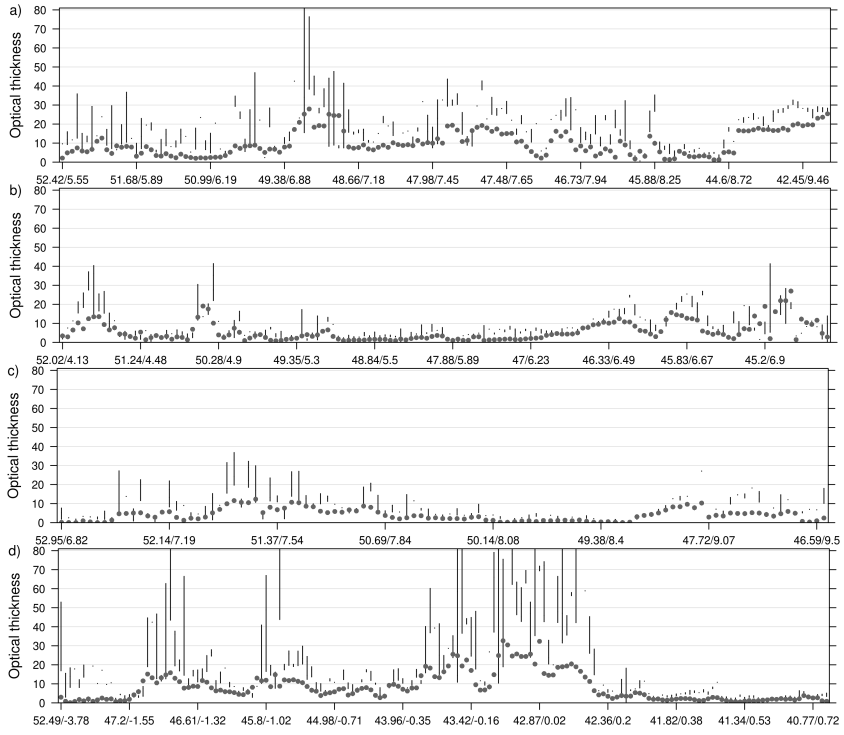

Figure 3.22: *Cloud optical thickness retrieved by SLALOM using MSG SEVIRI data vs. CloudSat 2B-TAU optical thickness for scenes from figure 3.18. For each SEVIRI pixel the retrieved $\tau_{Sslalom}$ value (grey dot) and the range of spatial corresponding τ_{2B-TAU} values (black line) are displayed.*

potential deviations due to different satellite and sensor characteristics. In the third step, SLALOM-based cloud optical thickness retrieved from MSG SEVIRI and MODIS data as well as MODIS M06 cloud optical thickness were compared to the CloudSat 2B-TAU product. CloudSat provide an accurate cloud optical depth profile and detailed information on the sub-pixel heterogeneity of the cloud structure within the SEVIRI pixel.

Over the Atlantic, SLALOMM and M06 retrieved cloud properties showed a high level of agreement with coefficients of correlation of 0.99 (τ) and 0.94 (a_{ef}) if pixels with $\tau > 5$ are considered. The rather strong deviations for $a_{efMslalom}$ values around 4 are linked to clouds with τ<10. The presence of broken clouds and variations between thin and very thin clouds within one pixel may lead to surface visibility at sub-pixel scale resulting in smaller reflectances at 1.6 μm. Based on the cloud reflection function at the absorbing wavelengths, smaller near-infrared reflectances result in larger a_{ef} values. It seems that SLALOM is sensitive to such conditions causing the overestimation. The same effect can be observed in the results over Europe. The correlation coefficient of 0.68 (a_{ef}) increases to 0.89 if values with τ<10 are excluded. The slightly weaker agreement, which is also apparent for the optical thickness (r=0.95), may be caused by the influence of the background albedo. However, the overall analysis showed very close and comparable results.

When comparing SLALOM$_S$ against M06 on the 0.08° by 0.08° grid, it was noticeable that the range of a_{ef} and τ values is significantly smaller in comparison to the results presented in the SLALOM MODIS/M06 comparison. This may be attributed to the less spatial resolution.

Over the ocean, the retrieved cloud properties showed a good agreement with an r of 0.93 (τ) and 0.82 (a_{ef}). The largest deviations were found in less homogeneous cloud areas that are characterised by broken clouds and along the cloud borders. The different spatial resolution of SEVIRI and MODIS may introduce differences since the retrieved optical properties are based on area-averaged reflectance measurements of about 6 by 8 km (SEVIRI) and 1 by 1 km (MODIS) respectively. At sub-pixel resolution, the homogeneity of a cloud remains unknown. It is also indicated that inhomogeneities are present through large ranges of the CloudSat τ_{2B-TAU} values observed within the SEVIRI pixels. Nevertheless, the retrieved $\tau_{Sslalom}$ values are well within the range of the τ_{2B-TAU} values that can be found within one SEVIRI pixel, except for the case of very small optical thickness (smaller than 5) where SLALOM tends to overestimate τ. As $\tau_{Sslalom}$ values, τ_{M06} and $\tau_{Mslalom}$ values also correspond very well with CloudSat 2B-Tau which is evident through clear positive correlations with r of 0.95 ($\tau_{Mslalom}$ vs. τ_{2B-TAU}) and 0.96 (τ_{M06} vs. τ_{2B-TAU}). In addition, the differences found between SLALOM$_S$ and M06 may be attributed to differences in the calibration of both instruments and the different viewing geometries which subsequently affects the retrieved cloud properties. In order to quantify the differences

in the reflectances, a direct comparison of the visible (0.8 μm) and near-infrared (1.6 μm) reflectances of MODIS and SEVIRI was carried out. The variations in SEVIRI and MODIS reflectances showed an acceptable level of agreement with r=0.93 (VIS) and r=0.90 (NIR), but the remaining differences may explain the observed differences between SLALOM$_S$ and M06, since the reflection of clouds in the visible region is strongly related to cloud optical thickness, and the cloud droplet radius influences reflection in the near-infrared region. Moreover, even though the datasets were chosen based on minimum time differences between the observations, slight temporal discrepancies may also contribute to differences between the retrievals, as clouds are highly dynamic and their spatio-temporal location and properties can change quickly. Over Germany, the evaluation between SLALOM$_S$ and M06 showed larger differences and the agreement decreased to correlation coefficients of 0.8 and 0.52 for τ and a_{ef} respectively. This can be explained by inhomogeneities of clouds which exhibit quite complex structures particularly over land which are detected on different scales by MODIS and SEVIRI because of their different spatial resolutions. Moreover uncertainties are caused by using different surface albedo products in both retrievals. The presence of significant cloud inhomogeneities was also confirmed through large ranges of τ_{2B-TAU} values within many SEVIRI pixels. The likelihood of a mixed cloud scenario being located within a larger SEVIRI pixel is increased. However, the tendency of the τ_{2B-TAU} evolution is mapped very well through $\tau_{Sslalom}$. It is conspicuous that for τ larger than 10 values of both SLALOM-based optical thickness products are mainly smaller than the corresponding τ_{2B-TAU} values. It seems that the surface albedo has a significant influence on the SLALOM retrieval. τ_{M06} and τ_{2B-TAU} values show a good agreement with an r of 0.81.

This paper has demonstrated that the cloud retrieval algorithm SLALOM provides robust estimates of cloud optical thickness and cloud effective radius not only for polar orbiting satellites such as MODIS but also for geostationary satellites such as MSG SEVIRI. The rather simple retrieval technique shows comparable results to the well established NASA products. The inverse problem is solved by approximated analytical equations that enables a fast but hardware-undemanding computation speed. Despite the differences between MSG SEVIRI, Terra/Aqua MODIS and CloudSat with respect to their spectral characteristics, spatial resolutions and viewing geometries, the retrieved cloud properties compare well over the Atlantic. The agreement over land is weaker. We attribute this to (a) uncertainties related to the surface albedo which is treated differently in the algorithms and which is based on different albedo maps and (b) the existence of broken clouds.

Given the demonstrated accuracy of SLALOM using MSG SEVIRI data there is a wide spread of potential applications. Beside a stand-alone application it can be used e. g. in the look-up table codes to generate a first guess or to reject unphysical results. SLALOM could also be used to generate first guess scenarios

of physical insights into cloud processes.

Acknowledgements
This work was supported by the Priority Program 1374 "Biodiversity Exploratories" of the German Research Foundation (DFG) Project. The work of A. Kokhanovsky was supported by the ESA Cloud CCI Project. NASA's CloudSat project products are provided online (www.cloudsat.cira.colostate.edu). The MODIS data was obtained from NASA's Level 1 and Atmosphere Archive and Distribution System (LAADS web, http://ladsweb.nascom.nasa.gov/). The authors like to thank Jan Fokke Meirink of KNMI for providing MODIS-SEVIRI calibration coefficients. We acknowledge also the support by Rob Roebeling and Sebastien Wagner of EUMETSAT. Furthermore, the authors would like to thank the anonymous reviewers for valuable remarks and comments which helped to improve the manuscript.

Chapter 4

Improving the accuracy of rainfall rates from optical satellite sensors with machine learning - A random forests-based approach applied to MSG SEVIRI

Meike Kühnlein [(1)], Tim Appelhans [(1)], Boris Thies [(2)], Thomas Nauß [(1)]

[(1)] Environmental Informatics, Faculty of Geography, Philipps–University Marburg, Marburg, Germany

[(2)] Laboratory for Climatology and Remote Sensing, Faculty of Geography, Philipps-University Marburg, Marburg, Germany

Published in

Remote Sensing of Environment 141 (2014) 129–143

Received 25 July 2013;
Received in revised form 26 October 2013;
Accepted 26 October 2013;

4 Improving the accuracy of rainfall rates from optical satellite sensors with machine learning - A random forests-based approach applied to MSG SEVIRI

Meike Kühnlein, Tim Appelhans, Boris Thies, Thomas Nauß

Abstract The present study aims to investigate the potential of the random forests ensemble classification and regression technique to improve rainfall rate assignment during day, night and twilight (resulting in 24-hour precipitation estimates) based on cloud physical properties retrieved from Meteosat Second Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) data.

Random forests (RF) models contain a combination of characteristics that make them well suited for its application in precipitation remote sensing. One of the key advantages is the ability to capture non-linear association of patterns between predictors and response which becomes important when dealing with complex non-linear events like precipitation. Due to the deficiencies of existing optical rainfall retrievals, the focus of this study is on assigning rainfall rates to precipitating cloud areas in connection with extra-tropical cyclones in mid-latitudes including both convective and advective-stratiform precipitating cloud areas. Hence, the rainfall rates are assigned to rain areas previously identified and classified according to the precipitation formation processes. As predictor variables water vapor-IR differences and IR cloud top temperature are used to incorporate information on cloud top height. $\Delta T_{8.7-10.8}$ and $\Delta T_{10.8-12.1}$ are considered to supply information about the cloud phase. Furthermore, spectral SEVIRI channels ($VIS_{0.6}$, $VIS0_{0.8}$, $NIR_{1.6}$) and cloud properties (cloud effective radius, cloud optical thickness) are used to include information about the cloud water path during daytime, while suitable combinations of temperature differences ($\Delta T_{3.9-10.8}$, $\Delta T_{3.9-7.3}$) are considered during night-time.

The development of the rainfall rate retrieval technique is realised in three steps. First, an extensive tuning study is carried out to customise each of the RF models. The daytime, night-time and twilight precipitation events have to be treated separately due to differing information content about the cloud properties between the different times of day. Secondly, the RF models are trained using the optimum values for the number of trees and number of randomly chosen predictor variables found in the tuning study. Finally, the final RF models are used to predict rainfall rates using an independent validation data set and the results are validated against co-located rainfall rates observed by a ground radar network. To train and validate the model, the radar-based RADOLAN RW product from the German Weather Service (DWD) is used which provides area-wide gauge-adjusted hourly precipitation information.

Regarding the overall performance, as indicated by the coefficient of determination (Rsq), hourly rainfall rates show already a good correlation with Rsq=0.5

(day and night) and Rsq=0.48 (twilight) between the satellite and radar based observations. Higher temporal aggregation leads to better agreement. Rsq rises to 0.78 (day), 0.77 (night) and 0.75 (twilight) for 8-h interval. By comparing day, night and twilight performance it becomes evident that daytime precipitation is generally predicted best by the model. Twilight and night-time predictions are generally less accurate but only by a small margin. This may due to the smaller number of predictor variables during twilight and night-time conditions as well as less favourable radiative transfer conditions to obtain the cloud parameters during these periods.

However, the results show that with the newly developed method it is possible to assign rainfall rates with good accuracy even on a hourly basis. Furthermore, the rainfall rates can be assigned during day, night and twilight conditions which enables the estimation of rainfall rates 24 h day.

Keywords Rainfall rate, Rainfall retrieval, Random forests, Machine learning, MSG SEVIRI, Geostationary satellites, Optical sensors

4.1 Introduction

Many ecological and biodiversity-oriented projects require area-wide information on precipitation distribution and quantity at high temporal and spatial resolution. For this purpose satellite-based rainfall retrievals are often the only option. Traditionally, precipitation is observed locally by conventional instruments such as rain (or snow) gauges and where available, weather radar systems. However, the most obvious limitations of station-based precipitation measurements are their spatial incoherencies, uneven global distribution and highly variable density with some regions having a relatively dense network while others have only a few or no gauges. Over oceans, gauges are almost non-existent apart from a few island locations. Ground-based weather radar systems like the ones in Europe, Japan or North America provide spatial measurements of precipitation (100 km from the radar). They are located over land and generally concentrated in regions that are also well covered with rain gauges. In this context, precipitation retrievals from optical sensors aboard geostationary (GEO) weather satellites may be an alternative to fill the information gap by providing area-wide data about rainfall distribution and amount at high spatial and temporal resolutions.

During the last decades, several satellite-based rainfall retrieval techniques for the detection of precipitating clouds and assignment of rainfall rates from optical sensors available on GEO platforms have been developed (see valuable overviews by e. g. Kidd and Levizzani, 2011; Kidd and Huffman, 2011; Prigent, 2010; Thies and Bendix, 2011). Traditionally, GEO system-based retrieval schemes were restricted by the spectral resolution of the sensors, which only allowed concepts that rely on a relationship between cloud top temperatures measured in an infrared

(IR) channel and rainfall probability/intensity (e. g. Adler and Mack, 1984; Arkin and Meisner, 1987). These concepts are based on the assumption that cold cloud tops are associated with high rainfall probabilities/intensities. More advanced IR retrieval methods are able to divide these precipitating cloud areas into different sub-areas, to which rainfall intensities are assigned (e. g. Adler and Negri, 1988; Hong et al., 2004; Hsu et al., 2002; O'Sullivan et al., 1990; Porcu and Levizzani, 1992; Reudenbach et al., 2007; Reudenbach, 2003; Wu et al., 1985). Such IR retrievals show good results in regions with mainly convective clouds, especially in the tropics and sub-tropics, but exhibit considerable drawbacks concerning the detection and quantification of rainfall from stratiform clouds in connection with extra-tropical cyclones (e. g. Adler et al., 2001; Amorati et al., 2000; Ebert et al., 2007; Früh et al., 2007; Levizzani et al., 1990; Negri and Adler, 1993; Pompei et al., 1995). This type of precipitating clouds is characterised by relatively warm and spatially homogeneous cloud top temperatures that do not differ significantly from raining to non-raining regions. Therefore, retrieval techniques based solely on IR cloud top temperature led to uncertainties concerning the assigned rainfall rate (e. g. Ebert et al., 2007; Früh et al., 2007).

With the upcoming of new generation GEO systems, several authors suggested the use of optical and microphysical cloud parameters derived from the now available multispectral data set to improve optical rainfall retrievals (e. g. Ba and Gruber, 2001; Kühnlein et al., 2010; Nauss and Kokhanovsky, 2006; Roebeling and Holleman, 2009; Rosenfeld and Gutman, 1994; Rosenfeld and Lensky, 1998; Thies et al., 2008b,a,d). They were able to show that cloud areas with a high optical thickness and a large effective particle radius possess a high amount of cloud water, and are characterised by a higher rainfall probability and intensity than cloud areas with a low optical thickness and a small effective particle radius.

Thies et al. (2008d) showed the possibility to separate areas of differing precipitation processes and rainfall intensities within the rain area by means of cloud properties retrieved with SEVIRI aboard MSG. The day and night technique for precipitation process separation and rainfall rate differentiation relies on information about the cloud top height, the cloud water path and the cloud phase in the upper parts. It is based on the assumption that areas with higher cloud water path and more ice particles in the upper parts are characterised by higher rainfall intensities. Recently, Kühnlein et al. (2010) used MSG SEVIRI reflection values in the 0.56-0.71 μm ($VIS_{0.6}$) and 1.5-1.78 μm ($NIR_{1.6}$) channels, which provide information about the optical thickness and the effective radius, to estimate rainfall rates over the northern German lowlands. This approach is based on the assumption that high rainfall rates are linked to high optical thickness and large effective particle radius, whereas low rainfall rates are linked to a low optical thickness and a small effective particle radius. The encouraging validation results of both retrievals indicate the high potential for an improved rainfall rate retrieval in the mid-latitudes using optical and microphysical cloud properties

derived from MSG SEVIRI data which provide the enhanced spectral resolution that is needed (Levizzani, 2003; Levizzani et al., 2001b). Furthermore, the relatively high spatial (3 km x 3 km at sub satellite point) and especially temporal resolution (15 min) permits a quasi-continuous observation of rainfall distribution and rainfall rate in near-real time.

In order to relate the retrieved cloud properties to precipitation, parametric statistics are generally used (e. g. Adler and Negri, 1988; Cheng and Brown, 1995; Kühnlein et al., 2010; Levizzani et al., 1990; Thies et al., 2008d). The application is rather straight-forward if only a few input variables are considered. However, cloud-top properties to precipitation relationship is very complex and non-linear, and it is likely beyond the skill of parametric tests and the related conceptual model.

Machine learning algorithms such as support vector machines (Mountrakis et al., 2011), artificial neural networks (Mas and Flores, 2008), decision trees (Breiman et al., 1984) or ensemble classifiers (Breiman, 1996) have been successfully adopted to remote sensing and rainfall applications (Capacci and Conway, 2005; Grimes et al., 2003; Rivolta et al., 2006) and may be suitable to overcome the limitations of the parametric techniques. When faced with high dimensional and complex data, machine learning algorithms provide efficient alternatives and generally show a higher accuracy (Foody, 1995; Friedl and Brodley, 1997; Hansen et al., 1996). In addition, the developments in parallel computing with machine learning offer new possibilities in terms of training and predicting speed resulting in improved real time systems.

In recent years, machine learning techniques which use ensembles of classifications or regressions (e. g. random forests, neural network ensembles, bagging and boosting, see Friedl et al., 1999; Krogh and Vedelsby, 1995; Rodriguez-Galiano et al., 2012; Ruiz-Gazen and Villa, 2007; Steele, 2000) have received increasing interest. They are based on the assumption that a whole set of trees or networks produce a more accurate prediction than a single tree or network (Dietterich, 2002). A new, powerful and promising ensemble classification and regression technique is random forests (Breiman, 2001). It is one of the most accurate learning algorithms available and it offers specific features that make it attractive for remote sensing applications. For example, it runs efficiently on large data sets, it is simple and can easily be applied to parallel computing platforms and it can capture non-linear association patterns between predictors and response. Although widely applied in other disciplines such as bioinformatics (e. g. Cutler and Stevens, 2006), some land-cover classifications using hyper spectral and multispectral satellite data, radar and lidar data (Ghimire et al., 2010; Guo et al., 2011; Pal, 2005; Rodriguez-Galiano et al., 2012) and also in a few ecological studies (e. g. Cutler et al., 2007; Mota et al., 2002; Prasad et al., 2006), the utilisation of random forests in climatology remains rare. This is one of the reasons it has led us to investigate the usefulness of RF approaches for rain rate delineation

from satellite platforms.

In summary, the enhanced information content on cloud properties at high spectral, spatial and temporal resolution offered by current and upcoming GEO systems along with the encouraging results concerning rainfall rate differentiation and rainfall rate assignment shown by Thies et al. (2008d) and Kühnlein et al. (2010) point to a promising potential of optical sensors as basis for reliable rainfall rate retrievals. However, this potential will likely remain unexploited using common parametric approaches. In addition, and not explicitly stated above, all existing optical retrievals that are based on optical and microphysical cloud parameters are restricted to daytime and night-time conditions and do not cover twilight conditions (see e. g. Kidd and Levizzani, 2011; Prigent, 2010).

In this study we show the potential of the random forests method for an improved rainfall rate assignment during day, night and twilight based on MSG SEVIRI data which provide information on cloud properties at high temporal and spatial resolution.

Based on the precipitation processes in connection with extra-tropical cyclones, the retrieval process consists of three steps: (i) identification of precipitating cloud areas, (ii) separation of precipitating areas into predominantly convective and advective-stratiform cloud regions, (iii) individual assignment of rainfall rates to these cloud areas. Since this study focuses on an improved assignment of rainfall rates based on random forests and not on the development of an optimised precipitation retrieval (which includes the delineation of raining from non-raining areas), radar data is used for the first two steps. This means that the derivation of the rain area as well as the rain process is based on observations from the radar network rather than MSG SEVIRI. Rainfall rates are then assigned to the already identified stratiform and convective precipitating areas.

Germany was chosen as study area for the development and validation of the new technique. The region can be regarded as sufficiently representative for mid-latitudes precipitation formation processes since it is dominated by frontally induced precipitation processes in connection with extra-tropical cyclones and shows a prominent maritime to continental gradient from west to east. Moreover, the radar network based and gauge-adjusted, hourly precipitation data set (RADOLAN RW) provided by the German Weather Service (DWD) provides a reliable training and validation basis.

The structure of this paper is as follows: the underlying data sets and methods are introduced in Section 4.2. Section 4.3 gives a presentation of the theoretical background and conceptual design providing the basis for the selection of random forests predictors. In Section 4.4, the adjustment of the random forests models as well as the appraisal of the new rainfall rate assignment technique is introduced. The paper is closed with a summary and some conclusions in Sections 4.5.

4.2 Data and methods

4.2.1 Satellite observations

For this study, MSG SEVIRI data are used. SEVIRI scans the full disk every 15 min and measures reflected and emitted radiance in 12 channels, three channels at visible and very near infrared wavelengths (between 0.6 and 1.6 μm), eight from near-infrared to thermal infrared wavelengths (between 3.9 and 14 μm), and one high-resolution visible channel. The nominal spatial resolution at the sub-satellite point is 1 by 1 km for the high-resolution channel, and 3 by 3 km for the other channels (Aminou, 2002; Schmetz et al., 2002). Over the study area in Germany, the satellite viewing zenith angles of SEVIRI range from 56° to 64°. As a consequence, the above mentioned spatial resolution is reduced in the present study. The follow up mission Meteosat Third Generation is intended to be launched in 2018 (EUMETSAT, 2013). This ensures these data availability and utilisation of applications developed for MSG SEVIRI for the next decades. The MSG SEVIRI data required for this study were downloaded from the EUMETSAT data centre (`www.eumetsat.int`). Processing has been performed based on a newly designed Meteosat processing scheme which has been implemented by Tobias Ebert and Johannes Drönner in co-operation with the working group of Bernhard Seeger from the computer science department at Marburg University. The processing scheme will be available online at `http://umweltinformatik-marburg.de/software/` shortly. Until then please contact the author for a copy of the software.

Cloud properties such as cloud effective radius and cloud optical thickness are retrieved using the semi-analytical approach SLALOM (SimpLe Approximations for cLOudy Media) developed by Nauss and Kokhanovsky (2011). This forward model is based on approximated solutions of the asymptotic radiative transfer theory (e. g. Germogenova, 1963; King, 1987) and provides increased computation speed since the equations can be efficiently solved during runtime. In order to retrieve cloud optical thickness and cloud effective droplet radius from MSG SEVIRI data, a combination of reflectance measurements at visible (0.65 μm) and near-infrared (1.64 μm) wavelengths is used. For the background albedo, a minimum composite of the reflectance in the visible (0.65 μm) and near-infrared (1.64 μm) channel over one month was calculated. A validation of SLALOM over sea and land surface against the well-known NASA MODIS cloud property product (Platnick et al., 2003) as well as the CloudSat 2B-TAU product (Polonsky et al., 2008) showed good agreement and can be found in Kühnlein et al. (2013). The present version of SLALOM is limited to water clouds. Cloud masks and cloud phase were derived using the algorithm by Cermak (2006) and Cermak and Bendix (2008) which have kindly been provided by the authors and are also implemented in the new Meteosat processing scheme.

4.2.2 Weather radar observations

For the development and validation of the new rainfall rate technique, radar-based precipitation data of the German Weather Service is used. The RADOLAN RW product is based on measurements with a C-band Doppler radar. Rain intensity adapted Z-R relationships, statistical clutter filtering and shadowing effects are treated within an online calibration process. Furthermore, precipitation intensities are adapted with ground-based precipitation measurements. The final precipitation product is available at temporal resolution of one hour and is a composite consisting of 16 German radar stations and some from neighbouring countries (e. g. Nancy/France) covering the entire area of Germany at a spatial resolution of 1 by 1 km (Bartels et al., 2004).

4.2.3 Pre-Processing of satellite and weather radar observations

The different temporal and spatial characterises of the satellite data (15 min; 3 by 3 km at sub-satellite point) and weather radar data (more or less continuously over 1 h; 1 by 1 km) must be addressed to ensure the pixel matching between satellite and radar data. An average of the satellite-based products is aggregated over a time interval of one hour. This is done by taking the arithmetic mean of the four scenes available every hour. To assure that only cloudy pixels within the time interval are incorporated, the cloud mask developed and implemented by Cermak (2006) and Cermak and Bendix (2008) is applied. Only those SEVIRI pixels that are classified as cloudy over the entire time interval are taken into account. Because of the differing viewing geometries between both systems, the radar product was projected and spatially aggregated (mean) to the geometry of SEVIRI.

The final data set consists of 1150 scenes of precipitation events between April and September 2010. Scenes with at least 2000 rainy pixels were chosen as precipitating events based on the RADOLAN RW product. Hereby pixels with higher than 0.06 mm/h are considered as rainy. The data set is split into daytime, night-time and twilight data sets. To ensure sufficient solar illumination in the VIS and NIR channels, scenes with a corresponding solar zenith angle less than 70° belong to the daytime data set. Scenes with a solar zenith angle greater than 70° and less than 108° are assigned to the twilight, and those greater than 108° are assigned to the night-time data set. The resulting daytime data set consists of 525 scenes, the night-time data set has 274 and the twilight data set has 351 scenes.

4.2.4 Random forests

The ensemble technique random forests which has been shown to perform very well in a variety of environmental investigations, contains a combination of characteristics that make it well suited for its application in remote sensing. RF runs

efficiently on large data sets, can easily be parallelised and is relatively robust to outliers and noise. Furthermore, it does not require the specification of an underlying data model, it offers the ability to capture non-linear association patterns between predictors and response and it is able to deal with highly correlated predictor variables. It also generates an internal unbiased estimate of the generalisation error (OOB error) and has the ability to determine which variables are important in the regression. Finally, it offers the flexibility to perform several types of statistical data analysis (e. g. regression and classification) and it is computationally lighter than other tree ensemble methods.

Below, a brief overview of the RF procedure is given. For more theoretical details the reader is referred to the literature (Boulesteix et al., 2012; Breiman and Cutler, 2008; Breiman, 2001; Malley et al., 2011; Strobl et al., 2009).

In general, the RF algorithm for regression works as follows:

1. *ntree* bootstrap samples are randomly selected from the data set with replacement. For each bootstrap, a different subset of the data set is used to develop the decision tree model. About one-third of the cases are left out of the sample. This out-of-bag (OOB) is used to get unbiased estimates of the regression error and to get estimates of the importance of the variables used for constructing the tree.

2. A regression tree for each of the bootstrap samples is grown (resulting in *ntree* trees) with the following modification: at each node, a subset of the predictor variables (*mtry*) is selected randomly to create the binary rule. In other words, *mtry* specifies the number of randomly chosen variables upon which the decision for the best split at each node is made. Variable selection is based on the residual sum of squares i. e. the predictor with the lowest residual sum of squares is chosen for the split. *mtry* is held constant during the forest growing.

3. Each of the *ntree* trees is grown to the largest extent possible. There is no pruning.

4. Finally, predictions are calculated by putting each OOB observation or observation of the test data set down each of the *ntree* trees. Then the predictions of all regression trees are averaged to produce the final estimate (Breiman, 2001).

The OOB error is an important feature of RF. As mentioned before, each tree is built on a bootstrap sample that comprises roughly two-third of the training data. The remaining one-third (OOB) of the training data is not included in the learning sample for this tree and can be used to test it. Therefore, the RF model is applied to the OOB data. Then, the deviations between predicted and observed values are used to calculate the OOB error, which is for regression the

mean square error (MSE), and is given by

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(RR_{Predi} - RR_{Obsi})^2 \quad (1)$$

where RR_{Predi} is the ith prediction and RR_{Obsi} is the i^{th} observation. This resulting OOB error provides an unbiased estimate of the generalised error and can be considered as an internal validation. As long as enough trees have been grown, OOB's estimate of error rate is quite accurate (Breiman, 1996, 2001). Otherwise the OOB estimate can bias upward (see Bylander, 2002).

While bagging uses all predictor variables at each node (bagging = *mtry* = number of predictors), RF constructs a tree using different training data subsets created through bagging and bootstrap of the data. By making the tree grow from different bootstrap samples, the diversity of the trees is increased. This increases generality, makes the regression more robust when facing slight variations in the training data and generally increases the overall prediction accuracy (Breiman, 2001). Several studies have shown that methods based on bagging are not sensitive to noise (Briem, 2002; Chan and Paelinckx, 2008). When RF makes a tree grow, it uses the best split based on a number of randomly sampled predictor variables. If all variables were used for each tree, the trees would be very identical and therefore highly correlated (Breiman, 2001). Thus, the randomly chosen subsets of predictor variables at each split of each tree ensure lower correlation between trees that in turn increases model robustness.

Beside the favourable features of RF, some limitations need also be mentioned. One of the most significant drawbacks of RF is that it does lack interpretability. Since the predictions are derived using a forest of trees, it is not possible to easily illustrate how the predictions are made (i. e. no single tree can be drawn to illustrate the decisions upon which the predictions are based). Furthermore, averaging over all trees means that it is neither possible to predict beyond the range of response values in the training data, nor to predict the entire range of response values. As a result, RF tends to overestimate low values and underestimate high values. Furthermore, RF is a truly random statistical method which entails a number of methodological issues related to repeatability and generalisability of the analyses. For in depth descriptions and generally accepted solutions of how to address these issues, the reader is referred to the fundamental statistical literature provided for the method. At the very least, the complete RF procedure needs to be repeated several times to evaluate the general robustness of the obtained predictions.

For our calculations we used an R implementation of the RF library, which was created by Liaw and Wiener (2002) based on the original Fortran code by Leo Breiman and Adele Cutler (for information on the open source software R see R Development CoreTeam (2008)). The package is called "randomForest".

The algorithm falls into the embarrassingly parallel category. This means that the number of trees to grow within a RF model can be divided into independent subsets, since each tree in the forest depends only on the given data set and not on the other trees. The subsets can be built on all available cores or on different machines. Then the resulting RF objects are combined to get the final forest. The parallel execution was realised by using the "foreach" and "doSNOW" packages. The R code used for the analysis at hand is available on our homepage (http://umweltinformatik-marburg.de/software/).

4.3 Selection of predictor variables

Similar to parametric approaches, RF also requires a set of predictor variables to estimate a response variable. While the selection of the response variable i. e. the rainfall rate is obvious for the study objective, the definition of predictor variables should reflect the conceptual framework of the rainfall assignment technique which in turn must consider the mid-latitudinal precipitation processes with a strong focus on extra-tropical cyclones.

Houze (1993) summarised the conceptual model of rainbands dominated by different rainfall processes in connection with extra-tropical cyclones. Following this conceptual model of rainbands, the precipitation field can be decomposed into areas dominated by different rainfall processes: (i) advective-stratiform background and intermediary precipitation which are linked to light precipitation intensities (further referred as advective-stratiform precipitation process) and (ii) narrow cold-frontal, wide cold-frontal and warm-frontal rain bands which are characterised by high rainfall intensities (further referred as convective precipitation process). Hence, rainbands dominated by different rainfall processes lead to differing rainfall rates. For this reason the rainfall rate assignment is realised in a three step approach (identification of precipitating cloud areas, separation according to their process, final rainfall rate assignment).

Regarding the vertical cloud extension, convectively dominated precipitation areas with higher rainfall intensities are characterised by larger cloud depths and cloud tops reaching higher into the troposphere. On the other hand, advective-stratiform precipitation areas are not necessarily connected to cold cloud top temperatures and are also not colder than surrounding non-precipitating cloud areas which explains the limited accuracy of the traditional IR retrievals presented in the introduction. Therefore, several authors have successfully used either optical and microphysical cloud parameters derived from multispectral satellite data or a suitable selection of spectral channels and channel combinations which provide information about cloud parameters, for an improved rain area delineation (Lensky and Rosenfeld, 2003a; Nauss and Kokhanovsky, 2006, 2007; Rosenfeld and Lensky, 1998; Thies et al., 2008b,a), rainfall intensity differentiation (Thies et al., 2008d) and rainfall rate assignment (Kühnlein et al., 2010). They were

able to show that cloud areas with a high cloud water path (i. e. large enough combination of the optical thickness and effective particle radius) possess a higher rainfall probability and higher rainfall rates than cloud areas with a small cloud water path. Since effective rain formation processes are mainly coupled to ice particles in the upper parts of the clouds and the "seeder-feeder" effect (Houze, 1993), advective-stratiform precipitation areas with a higher cloud water path and a higher amount of ice particles in the upper cloud regions are also characterised by higher rainfall intensities (Thies et al., 2008d).

Considering the dominant precipitation processes for convective and advective-stratiform rainfall areas within extra-tropical cyclones, the following cloud physical parameters are chosen for this study:

- Cloud top height (CTH)
- Cloud top temperature (CTT)
- Cloud phase (CP)
- Cloud water path (CWP)

A proper SEVIRI spectral channel selection can be used as surrogates for these cloud physical parameters and therefore as predictor variables for the random forest model.

As a good proxy for the cloud top temperature, the brightness temperature in the 10.8 μm channel ($BT_{10.8}$) can be used. In addition to the cloud top temperature, the brightness temperature difference between the water vapour (WV) and the IR channels are used to gain information about the cloud top height relative to the tropopause level which enables a reliable identification of deep convective clouds (Heinemann et al., 2001; Schmetz et al., 1997; Tjemkes et al., 1997). Thies et al. (2008c) showed that the WV-IR combinations of MSG SEVIRI perform differently for different cloud-top height to tropopause level settings. To include different sensitivities on cloud top height the two channel differences $\Delta T_{WV6.2-IR10.8}$ and $\Delta T_{WV7.3-IR12.1}$ have been chosen.

Since effective rain formation processes are mainly coupled to ice particles in the upper parts of the cloud, the cloud phase is incorporated. The channel differences between 8.7 μm and 10.8 μm ($\Delta T_{8.7-10.8}$) as well as between 10.8 μm and 12.1 μm ($\Delta T_{10.8-12.1}$) can be used to gain information about the cloud phase (Strabala et al., 1994; Ackerman et al., 1998; Thies et al., 2008b). At these two wavelengths, the absorption of ice and water is different (Baum and Platnick, 2006). The increase of water particle absorption is greater between 11 and 12 μm than between 8 and 11 μm. On the other hand the increase of ice particle absorption is greater between 8 and 11 μm than between 11 and 12 μm (Strabala et al., 1994). Therefore the difference $\Delta T_{8.7-10.8}$ of ice clouds are greater than coincident $\Delta T_{10.8-12.1}$ differences and the opposite is true for water clouds.

For day-time observations, the SLALOM retrieval is used to derive optical and microphysical cloud properties such as cloud effective radius (aef) and cloud optical thickness (τ) which are used as predictor variables for advective-stratiform rainfall rates. Because the application of the commonly available cloud property retrievals to ice clouds requires the a-priori definition of ice particle geometries, which in turn heavily influences the retrieved values (see also Kokhanovsky et al., 2005), the authors decided to use the reflectance at 0.6 μm ($VIS_{0.6}$), 0.8 μm ($VIS_{0.8}$) and 1.6 μm ($NIR_{1.6}$) channel directly as input variables for the assignment of rainfall rates to convectively dominated precipitation areas. For night-time, there are MSG SEVIRI retrievals that can compute optical and microphysical cloud properties (e. g. VISST-SIST algorithms developed by NASA Langley Cloud and Radiation Group). Unfortunately, these algorithms derive reliable properties only for non-raining optically thin clouds (Minnis et al., 2011). However, several case studies have shown that implicit information about microphysical and optical cloud properties is available in the emissive channels (Baum et al., 2000; Inoue, 1985; Lensky and Rosenfeld, 2003a; Ou et al., 1993; Stone et al., 1990; Strabala et al., 1994). Based on these studies Thies et al. (2008c) demonstrated that the SEVIRI channel differences $\Delta T_{3.9-10.8}$, $\Delta T_{3.9-7.3}$, $\Delta T_{8.7-10.8}$ and $\Delta T_{10.8-12.1}$ provide information about the CWP, that can be used for rainfall retrieval. Therefore, the respective channel differences are incorporated instead of retrieved cloud properties during night-time. During twilight conditions neither the reflectances of VIS and NIR channels (due to insufficient solar illumination) nor the SEVIRI channel differences $\Delta T_{3.9-10.8}$ and $\Delta T_{3.9-7.3}$ can be used as surrogates for the CWP. The 3.9 μm channel radiance contains both reflected solar radiance and thermal emitted radiance. To use the 3.9 μm channel the solar component must be quantified and eliminated. This itself has been investigated by several studies (Rao et al., 1995; Rosenfeld and Lensky, 1998) and they showed that simplifications and assumptions are necessary which are only acceptable for a certain kind of cloud. To prevent misinterpretation, predictors containing the 3.9 μm channel are not used to gain information about the CWP and therefore there are no predictor variables available representing the CWP during twilight.

In summary, a different set of predictor variables with respect to the degree of solar illumination (day, night, twilight) and the dominant precipitation processes within each cloud region is utilised. An overview is given in table 4.1. Since the RF approach is not limited to a certain number of predictor variables, all channels which are part of the aforementioned channel combinations (e. g. $\Delta T_{10.8-12.1}$), are also incorporated in the test.

	Daytime		Night-time		Twilight	
	Stratiform	Convective	Stratiform	Convective	Stratiform	Convective
CTH	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$	$\Delta T_{6.2-10.8}$ $\Delta T_{7.3-12.1}$
CTH/CTT	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$
CP	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$ $\Delta T_{10.8-12.1}$
CWP	$VIS_{0.6}$ $VIS_{0.8}$ $NIR_{1.6}$ a_{ef} τ CWP	$VIS_{0.6}$ $VIS_{0.8}$ $NIR_{1.6}$	$\Delta T_{3.9-10.8}$ $\Delta T_{3.9-7.3}$	$\Delta T_{3.9-10.8}$ $\Delta T_{3.9-7.3}$		
SEVIRI channels	$WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$	$WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$	$WV_{3.9}$ $WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$	$WV_{3.9}$ $WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$	$WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$	$WV_{6.2}$ $WV_{7.3}$ $IR_{8.7}$ $IR_{12.1}$

Table 4.1: *Overview of RF predictor variables.*
Abbreviations are as follows: CTH, cloud top height; CTT, cloud top temperature; CP, cloud phase; CWP, cloud water path.

4.4 Estimation of rainfall rates

4.4.1 General methodology

On the basis of the theoretical background introduced in section 4.3, the rainfall rate assignment technique is realised in three steps:

(i) precipitating cloud areas are identified,

(ii) the precipitating cloud areas are separated into convective and advective-stratiform dominated precipitation areas and

(iii) rainfall rates are assigned to the convective and advective-stratiform dominated precipitation areas, respectively.

As mentioned in the introduction, the focus of the present study lies upon the development of a new technique for the assignment of rainfall rates. In order to develop and evaluate such a rainfall rate assignment technique, the RADOLAN RW product of the DWD is used to realise the first and second step. This means that both the rain areas, as well as the rain processes, are derived from the radar observations rather than from satellite images. Precipitation areas with more than 1.8 mm/h are considered as convectively dominated and areas between 0.06 and 1.8 mm/h are considered as advective-stratiform dominated (see Thies et al., 2008d). Hence, the proposed technique aims to assign rainfall rates to precipitating cloud areas that are already identified as being convective and advective-stratiform.

For the development and validation of the technique, a data set consisting of 525 daytime, 274 night-time and 351 twilight satellite scenes of precipitation events between April and September 2010 are used. The data set is split into daytime, night-time and twilight data sets which have to be treated separately due to differing illumination conditions. During night-time the channels at visible and very near infrared wavelengths (0.6 to 1.6 μm) are not available. During twilight and daytime the use of the 3.9 μm channel is complicated due to the varying solar component in this channel. This means that depending on the time of the day different random forest models are built and adapted. Each of these three data sets are randomly split into training (1/4 of the scenes) and validation data sets (3/4 of the scenes). The training data sets are used to train the according RF model, whereas the validation data sets are used to validate the predictions afterwards. The precipitation events taken for training are independent from those taken for validation. Since the technique aims to assign rainfall rates to precipitating cloud areas already classified as convective or advective-stratiform, each data set is split into cases of convectively dominated rainfall and cases of advective-stratiform dominated rainfall and treated separately (day = day-C and day-S, night = night-C and night-S, twilight = twilight-C and twilight-S; where C=convective and S=advective-stratiform).

The development of the rainfall rate technique is realised in three steps:

(1) Tuning: the optimal values of *ntree* and *mtry* are assessed for each RF model.

(2) Training: the RF models are trained using the optimum values of model parameters found in the tuning study.

(3) Validation: the RF models are applied to the validation data sets and the predicted rainfall rates are validated against co-located rainfall rates observed by the radar.

All these steps are described in the following sections.

4.4.2 Model tuning

There are basically two parameters to adjust in the R-package "randomForest", the overall number of trees in a forest (*ntree*) and the number of predictor variables randomly sampled for use at each split (*mtry*). The tuning is based on performance of OOB data.

An important consideration is how many trees to grow within the random forest model. Breiman (2001) suggested that the generalisation error converges as the number of trees increases. Adding more and more trees to the model does not result in over-adjustment. The main limitation of increasing *ntree* is the extra computation time. Therefore, the number of trees is not a real parameter in the sense that there is an optimum value, rather the number should be as large as

computationally feasible.

To assess the optimal value of *ntree*, a large number of RF models using randomly selected subsets of each data set (day-C, day-S, night-C, night-S, twilight-C, twilight-S) is created. Each RF model is created using 1000 trees for all possible values of *mtry* (*mtry*=1 to maximum). Then the MSE values of every possible value of *mtry* is averaged. Figure 4.1 shows how the error rate changes with the number of trees for *mtry* equal to the minimum (*mtry*=1), mean (*mtry*=6) and maximum value of *mtry* (*mtry*=12/15, for details see below). From around 500 trees onwards, the MSE of each data set stabilises and the addition of trees neither increases nor decreases the MSE. Therefore, the number of trees in the forest can be regarded as sufficient using *ntree* larger than 500. In order to reduce extra computation time, *ntree* is set to 500 in all models as this value is large enough to produce a stable prediction.

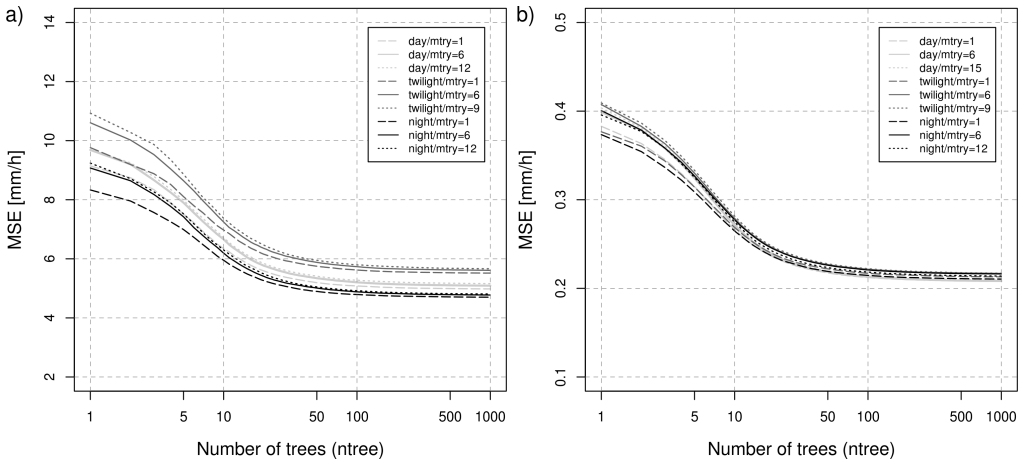

Figure 4.1: *Effect of number of trees (ntree) and random split variables (mtry) on OOB error (MSE). Results shown for (a) convective and (b) stratiform day, night and twilight tuning data sets, respectively.*

In comparison to *ntree*, the number of predictor variables randomly sampled for use at each split (*mtry*) is a real parameter. The suggested default value of *mtry* for regression is p/3 (with p total number of predictors, Breiman, 2001). In practice the best values for *mtry* will depend on the problem at hand and the parameter should be treated as tuning parameter. If the recommended *mtry* is too small in the presence of a large number of noise predictors, then it is more likely to select non-informative predictors. On the other hand, a small value of *mtry* might offer the possibility for strong predictors to be chosen in a scenario with many informative predictors (Boulesteix et al., 2012). Reducing the number of predictor variables reduces the correlation between the trees, but also causes each individual tree to be more biased (Goldstein et al., 2011).

To assess the optimal values of *mtry*, a large number of RF models are created for each data set using different possible values of splitting variables (*mtry*=1 to maximum number of *mtry*) while keeping *ntree*=500 constant. First, the RF model is computed with *mtry*=1. Then, *mtry* is increased by 1 and a new forest

is built with the new *mtry*. This is done until the maximum number of *mtry* is reached. This whole process is repeated several times using randomly selected subsets of the data sets. The effect of *mtry* on the OOB error is exemplarily shown for the day-time models in figure 4.2. Based on these results, the value of *mtry* leading to the smallest OOB error of a forest is selected for the according RF model. Regarding the convective data sets, *mtry*=7 leads to the highest prediction accuracy whereas *mtry*=8 leads to the highest prediction accuracy for the stratiform data sets.

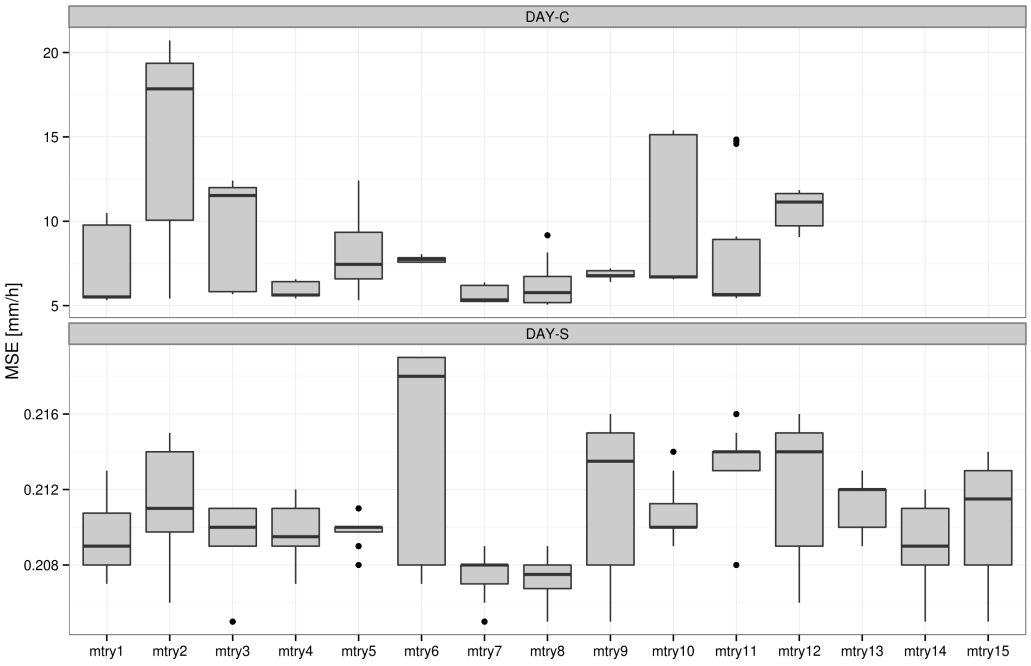

Figure 4.2: *Effect of the number of random split variables (mtry) on OOB error (using ntree = 500) for the daytime scenes (where C=convective and S=advective-stratiform).*
Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars.

4.4.3 Model training

The training of the RF models is done on a pixel and hourly basis by a comparison with the ground-based radar precipitation data using the different training data sets. Hence, various satellite-based products (see section 4.3) are combined with spatially and temporally co-located radar data from which the corresponding rainfall rates are extracted. The RF models are trained with *ntree* and *mtry* found in the tuning study (section 4.4.2). All this is realised for daytime, night-time and twilight scenes as well as for convective and stratiform cases separately.

4.4.4 Model validation

Once the RF models are established, it is possible to assign rainfall rates based on the same predictors used for training. In order to assess the quality of the model,

the established RF models are applied to the day, night and twilight validation data sets, respectively. Then the predicted rainfall rates (RR_{Pred}) are validated against the co-located rainfall rates observed by the radar (RR_{Obs}). For this, the hourly rainfall rates of RR_{Obs} and RR_{Pred} are also summed over 3-h and 8-h intervals where the latter contains all scenes of the day from the daytime, night-time and twilight time intervals respectively. For the appraisal of the retrieval technique, standard continuous verification scores are calculated. As measures of the agreement between observed and predicted values mean error (ME = bias), mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (Rsq) are calculated. These scores are given by:

$$ME = \frac{1}{N}\sum_{i=1}^{N}(RR_{Predi} - RR_{Obsi}) \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|RR_{Predi} - RR_{Obsi}| \tag{3}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(RR_{Predi} - RR_{Obsi})^2} \tag{4}$$

$$Rsq = \left(\frac{\sum_{i=1}^{N}(RR_{Predi} - \overline{RR_{Pred}})(RR_{Obsi} - \overline{RR_{Obs}})}{\sqrt{\sum_{i=1}^{N}(RR_{Predi} - \overline{RR_{Pred}})^2}\sqrt{\sum_{i=1}^{N}(RR_{Obsi} - \overline{RR_{Obs}})^2}}\right)^2 \tag{5}$$

where RR_{Predi} is the ith prediction and RR_{Obsi} is the i^{th} observation.

In order to assess the model performances four different evaluation strategies are employed:

(a) The overall performance of the rainfall rate assignment technique during daytime, night-time and twilight conditions is investigated.

(b) The performance of the rainfall rate assignment technique is investigated on a scene-by-scene basis.

(c) The diurnal performance of the RF models is investigated.

(d) The influence of different training data sets on the model performances is investigated.

First, the overall performance of the rainfall rate retrieval during daytime, night-time and twilight conditions is investigated. By extracting the data pairs of the validation data sets on a pixel basis, a total of 1309721 pairs of RR_{Obs} and RR_{Pred} are made available on an hourly basis, containing 678600 daytime, 280776 night-time and 350345 twilight data pairs. The hourly rainfall rates have also been aggregated for 3- and 8-h intervals. With increasing temporal aggregation, the number of data pairs decreases accordingly. The performance of the models across different aggregation times is summarised in table 4.2. As table 4.2 shows,

	Time period (tp)	Predicted 1-h	Observed 1-h	Predicted 3-h	Observed 3-h	Predicted 8-h	Observed 8-h
day	Min [mm/tp]	0.16	0.06	0.18	0.06	0.18	0.07
	Max [mm/tp]	26.51	56.10	32.91	49.43	48.05	80.00
	Median [mm/tp]	0.78	1.00	1.01	1.26	1.70	1.88
	Mean [mm/tp]	1.61	1.48	2.38	2.18	2.58	3.28
	Std [mm/tp]	1.57	1.68	2.54	2.67	4.10	4.28
	Rsq	0.50		0.68		0.78	
	RMSE [mm/tp]	1.26		1.58		2.07	
	MAE [mm/tp]	0.72		0.88		1.13	
	ME [mm/tp]	0.13		0.20		0.30	
	N	678600		422205		305871	
night	Min [mm/tp]	0.14	0.06	0.14	0.06	0.14	0.06
	Max [mm/tp]	9.02	40.13	18.53	49.00	27.55	62.12
	Median [mm/tp]	0.88	1.09	1.66	1.79	1.86	1.96
	Mean [mm/tp]	1.76	1.83	2.87	3.01	3.38	3.51
	Std [mm/tp]	1.43	2.07	2.65	3.55	3.39	4.34
	Rsq	0.50		0.69		0.77	
	RMSE [mm/tp]	1.46		2.00		2.13	
	MAE [mm/tp]	0.80		1.06		1.13	
	ME [mm/tp]	-0.07		-0.14		-0.13	
	N	280776		149608		146136	
twilight	Min [mm/tp]	0.18	0.06	0.22	0.07	0.19	0.07
	Max [mm/tp]	21.02	47.58	26.03	52.62	32.79	58.24
	Median [mm/tp]	0.87	1.09	1.54	1.65	1.66	1.76
	Mean [mm/tp]	1.79	1.69	2.87	2.71	3.30	3.12
	Std [mm/tp]	1.58	1.80	2.83	3.09	3.49	3.73
	Rsq	0.48		0.68		0.75	
	RMSE [mm/tp]	1.35		1.78		1.91	
	MAE [mm/tp]	0.79		1.03		1.10	
	ME [mm/tp]	0.10		0.16		0.18	
	N	350345		190758		145857	

Table 4.2: *Statistical values from the validation data sets.*
The scores are based on data pairs of 394 daytime, 205 night-time and 264 twilight precipitation scenes from April to September 2010. The values are calculated from the RR_{Obs} and RR_{Sat} data pairs extracted from the whole day, night and twilight data set, respectively. "Min" and "Max" denote the minimum and maximum value. "Mean" and "Median" signify the average and median value. "Std" signifies the standard deviation. Abbreviations are as follows: Rsq, coefficient of determination; RMSE, root mean square error; MAE, mean absolute error; ME, mean error; N, number of data pairs considered.

the rainfall rates observed by the radar at a 1-h interval range between 0.06 and 56.10. However, the spread of rainfall rates predicted by RF is much smaller and ranges between 0.16 and 26.51 for the same time interval. This highlights that very high and low observations cannot be captured by RF as the response variable is calculated by putting each object down the decision trees and then averaging the predictions of all trees. As a consequence, the averaging reduces the very high rainfall rates and increase the very small rainfall rates, which leads to an underestimation of the high rainfall rates and overestimation of the small rainfall rates. A consistent under-estimation (over-estimation) of the maximum (minimum) is apparent for all sets of models, regardless of aggregation or time of day,

especially maximum rainfall which can be severely under-estimated. However, mean, median and standard deviations generally show good agreement (please note, that it is also not possible to predict beyond the range of the response values in the training data).

As indicated by the Rsq, hourly RR_{Obs} and RR_{Pred} already show a good correlation with Rsq=0.5 (day and night) and Rsq=0.48 (twilight). Higher temporal aggregation leads to better agreement (Rsq as high as 0.78 (day), 0.77 (night) and 0.75 (twilight) for 8-h scenes). However, this is paralleled by an increased spread in the error scores. Small positive ME are seen for all daytime and twilight time intervals (ranging from 0.10 to 0.30), indicating a slight overestimation of the rainfall rate. During night-time, a slight under-estimation of rainfall rates are indicated (ME ranging from -0.07 to -0.13 for the different aggregation intervals).

In the next step the performance of the rainfall rate assignment technique on a scene-by-scene basis is investigated with results are summarised in figure 4.3. The standard verification scores calculated from RR_{Obs} and RR_{Pred} on a scene basis are plotted as box plots to facilitate a visual analysis of the performance.

There is a general pattern of daytime precipitation being best estimated by the model, regardless of aggregation times. Twilight and night-time predictions are generally less accurate but only by a small margin. Overall, there are no significant differences in central tendency arising from the time of day and this is remarkable particularly for twilight conditions. The most obvious difference between different times of day lies in the consistency of the prediction performance indicated by the spread of the Rsq. Here, twilight and night-time scenes clearly show greater variability, as indicated by the elevated inter quartile ranges (i.e. height of the boxes). As mentioned earlier, elevated aggregation exhibits better agreement between predictions and observations, which, is paralleled by an increase in the error scores in both magnitude and range. A look at the ME distribution reveals a general tendency of the model to overestimate precipitation, especially during day and twilight hours. Night-time predictions are more balanced in this respect. In general, it becomes apparent that night-time errors are smaller than during other times of the day.

Figures 4.4 a and 4.5 a show examples of hourly rainfall rates observed by the radar and predicted by RF model. Both figures illustrate that for higher RR_{Obs} (>10 mm/h), the corresponding RR_{Pred} are predominantly smaller. At the same time, the areas in the surrounding of high rainfall areas are slightly overestimated by RF. As a result of the aggregation process, it is possible to better reproduce the observed rainfall rate (figure 4.4 b and 4.5 b) but again, very high rainfall rates (>20 mm/3h) cannot be captured by the model. As already mentioned, the observed mismatch between predicted and observed extremes is characteristic of random forests.

General diurnal performance of the RF model is shown in figure 4.6. Here, all model verification scores are shown for each hour of the day across all scenes.

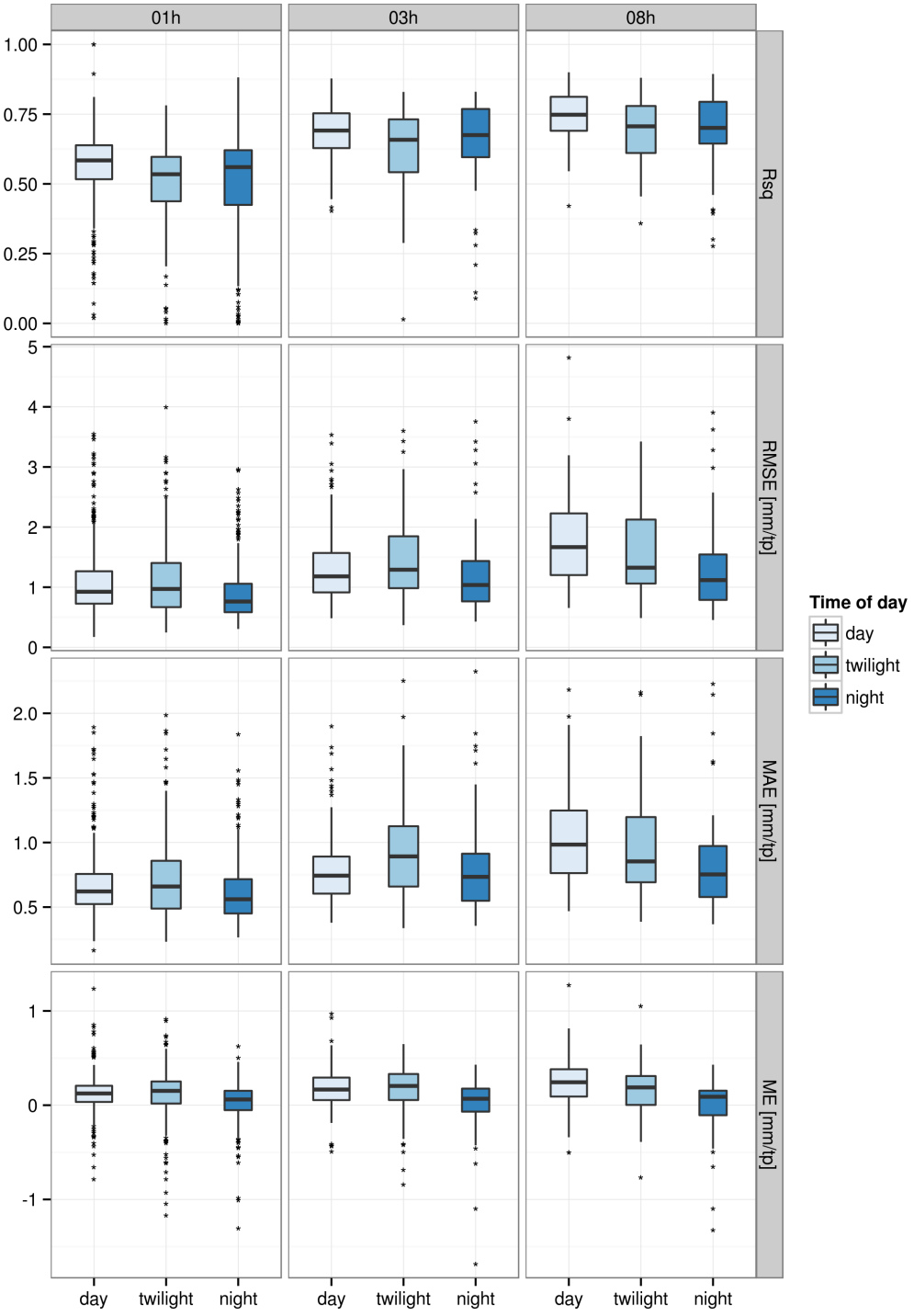

Figure 4.3: *Performance of rainfall rate assignment on a scene by scene basis.*
Box and whisker plots showing distribution of standard verification scores (rows) of RR_{Obs} vs. RR_{Pred} for different aggregation times (columns) according to time of day (colours). Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars.
Abbreviations are as follows: Rsq, coefficient of determination; RMSE, root mean square error; MAE, mean absolute error; ME, mean error; tp, time period.

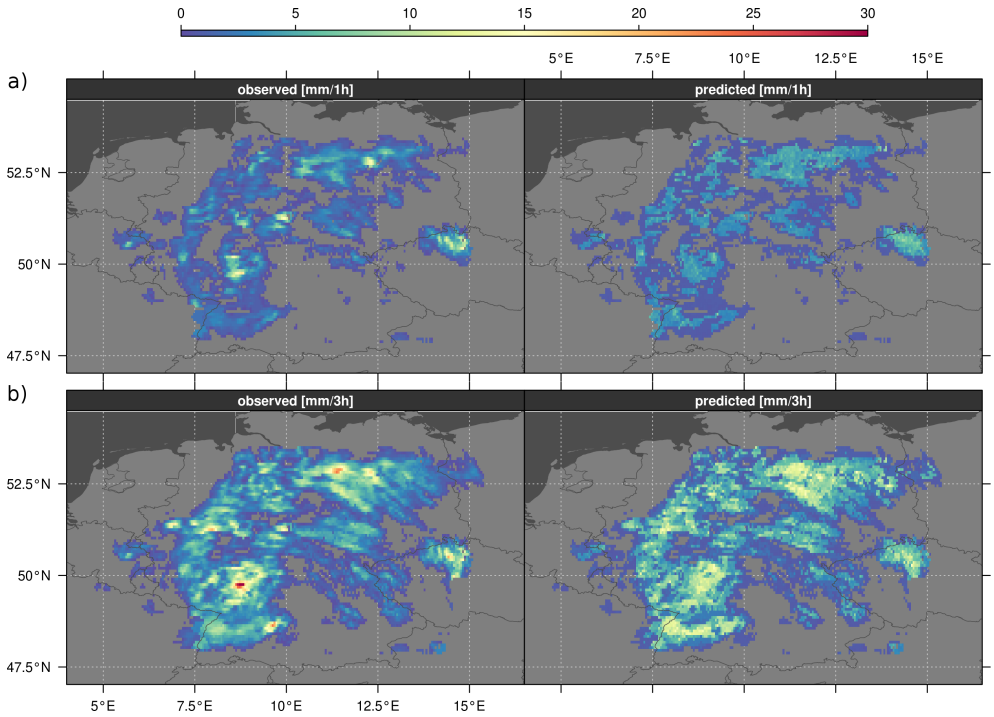

Figure 4.4: *Rainfall rates observed by radar (left) and predicted by RF model (right) during daytime. (a) rainfall rates for scene from 6 May 2010 15:00 UTC, (b) aggregated rainfall rates for scenes from 6 May 2010 14:00 to 16:00 UTC.*

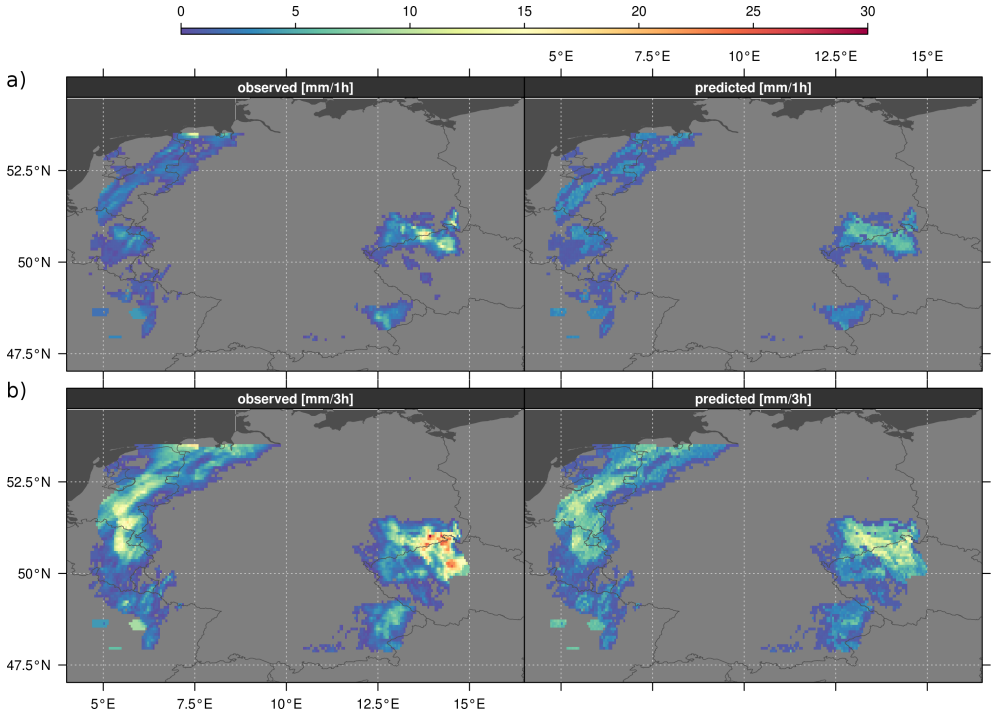

Figure 4.5: *Rainfall rates observed by radar (left) and predicted by RF model (right) during nighttime. (a) rainfall rates for scene from 15 August 2010 23:00 UTC, (b) aggregated rainfall rates for scenes from 15 August 2010 23:00 UTC to 16 August 2010 1:00 UTC.*

The comparison reveals reasonable agreement between RR_{Obs} and RR_{Pred} for all hours of the day. Best performance is seen during late morning and midday (increased Rsq and reduced spread in the error measures) most likely because of the favourable solar observation geometries. Twilight hours reveal rather variable prediction performance, especially toward the end of the day, whereas night-time performance is rather consistent. The maximum divergences occur during twilight conditions. It is to be expected that performance is different. Depending on solar illumination in twilight conditions the retrieval of cloud properties is more difficult and there are fewer predictor variables that can be used due to solar illumination. Furthermore, there are a reduced number of scenes which is indicated by the width of the boxes.

As already mentioned, the data used for the training of the model may influence the prediction. In order to assess how different input data sets may influence the RF results we have repeated the complete analysis 10 times, each with different randomly split sets of independent training and validation sets. For each of these iterations, the same standard verification scores as above are calculated on a pixel-basis and presented in figure 4.7 according to the time of the day. The daytime and night-time validation data sets show the highest median values for all aggregation levels with the night-time validation data set having a bigger range of Rsq values with increased temporal aggregation. By comparison to day and night, the twilight results are slightly below. The Rsq values are less and the range of Rsq values is bigger. As can be seen in column 2 and 3, the correlations increase considerably at lower temporal resolutions of 3- and 8-h. These results indicate that the selection of randomly chosen training and validation data sets influences the overall prediction accuracy, but in a small margin. This could be expected since the observations of the training sets are used to build the RF model and therefore influence the prediction. However, even if the results show that the different training sets influence the overall prediction, the standard verification scores reveal the same patterns which were shown in the foregoing investigations. This means that they reflect the same behaviour according to temporal aggregation and time of day.

4.5 Summary and conclusions

The aim of the present study was to investigate the potential of MSG SEVIRI for improved rainfall rate assignment using the ensemble classification and regression technique random forests as a fundamental algorithm. The novel approach differs from the most state-of-the-art satellite-based rainfall retrievals since it is not using a conventional parametric approach but a machine learning algorithm. RF is one of most accurate learning algorithm available and offers specific features that make it attractive for remote sensing applications, e. g. it runs efficiently on large data sets, it is simple and easy to parallelise. One of the key advan-

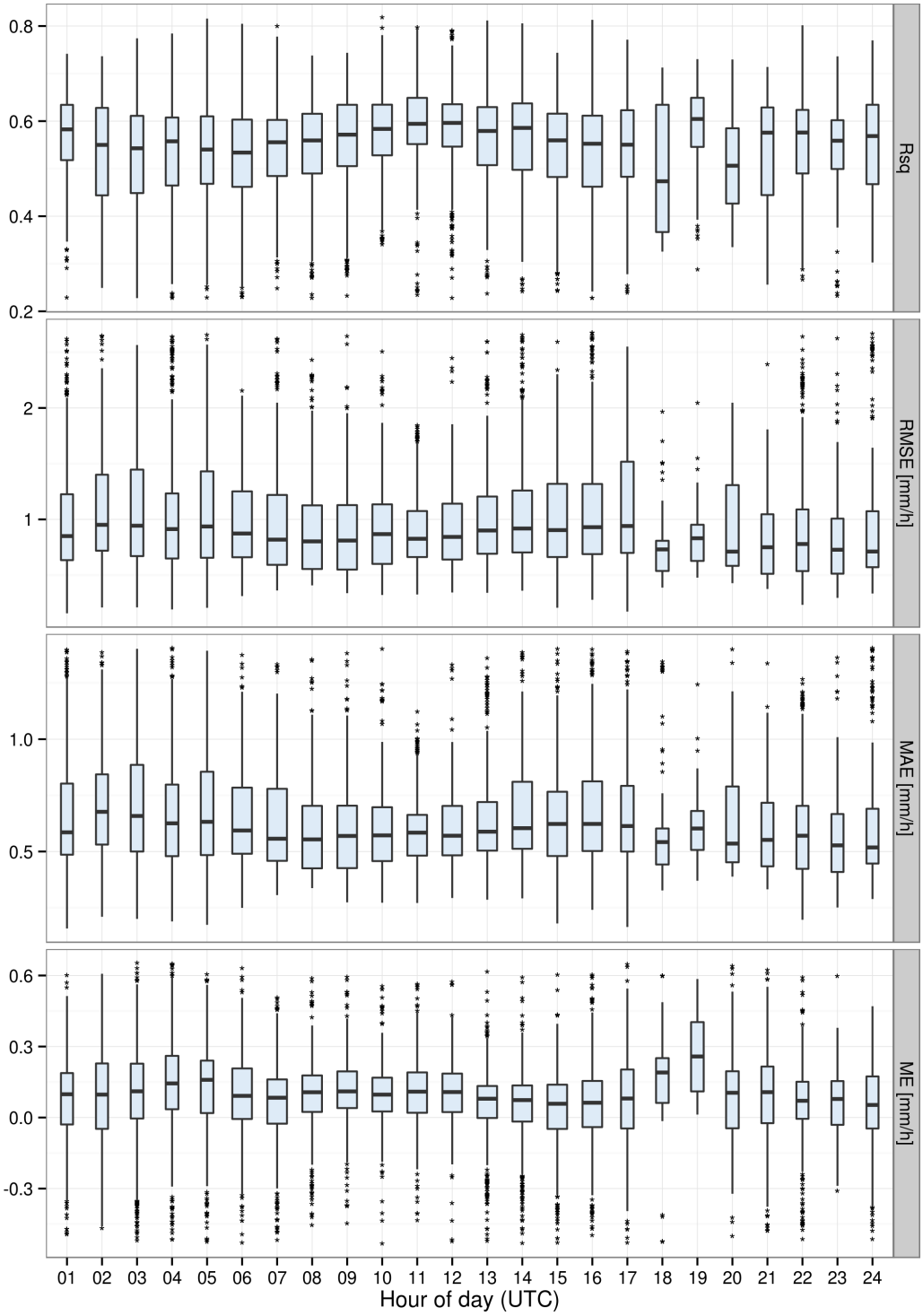

Figure 4.6: *Diurnal performance of the rainfall rate assignment technique.*
Box and whisker plots showing distribution of standard verification scores (rows) of RR_{Obs} vs. RR_{Pred} for each hour of the day. Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars. Box widths are relative to number of observations. Extreme outliers (beyond mean +/- 2 times standard deviation) have been removed.
Abbreviations are as follows: Rsq, coefficient of determination; RMSE, root mean square error; MAE, mean absolute error; ME, mean error.

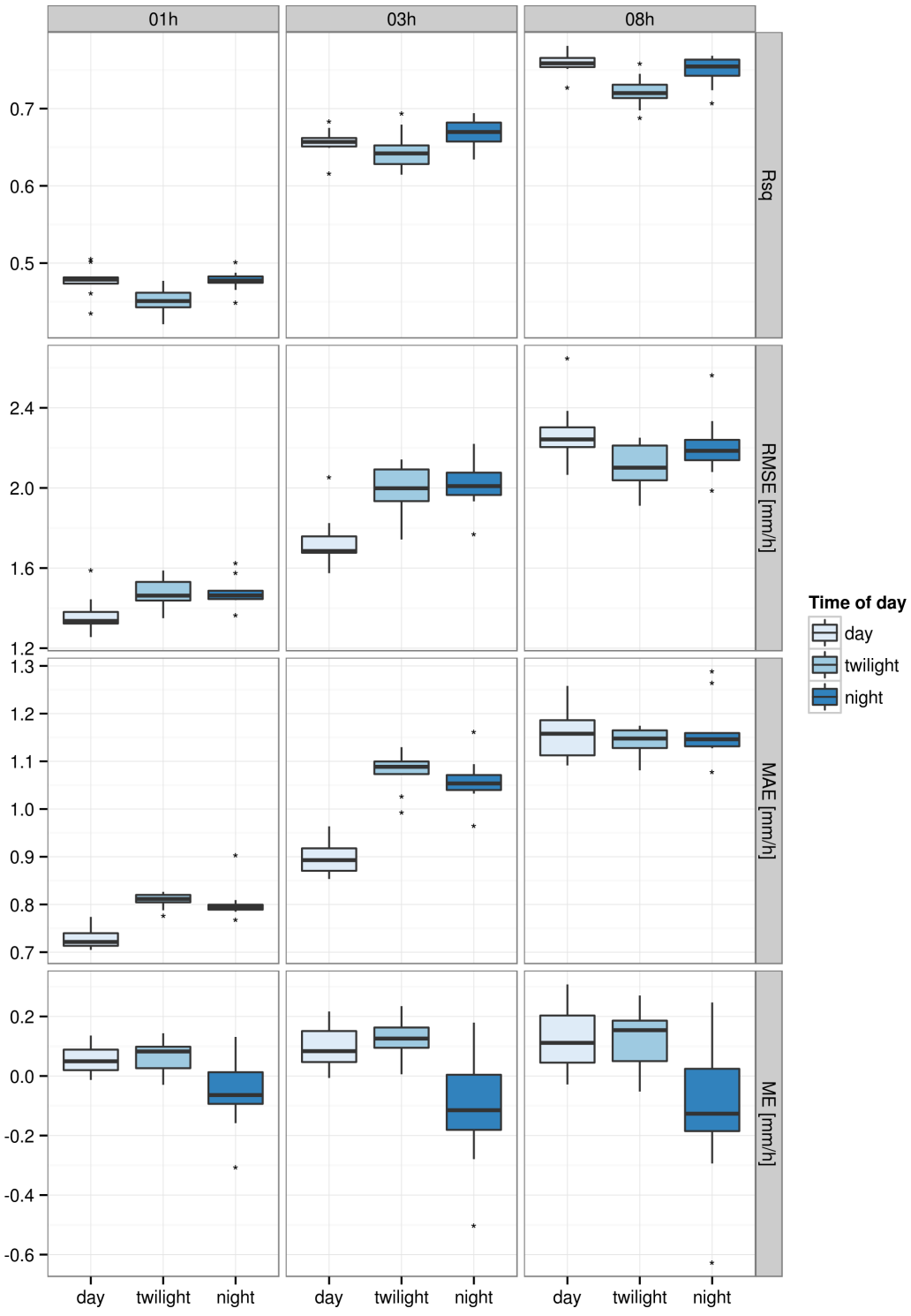

Figure 4.7: *Influence of different input data sets to the performance of the rainfall rate assignment technique.*
Box and whisker plots showing distribution of standard verification scores (rows) of RR_{Obs} vs. RR_{Pred} for different aggregation times (columns) according to time of day (colours). Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars. The complete analysis was rerun 10 times, each with different randomly split training and validation sets.
Abbreviations are as follows: Rsq, coefficient of determination; RMSE, root mean square error; MAE, mean absolute error; ME, mean error.

tages is the ability to capture non-linear association patterns between predictors and response, which becomes important when dealing with a very complex non-linear event like precipitation. Due to the deficiencies of existing rainfall retrieval techniques based on the IR cloud-top temperature concerning the detection and quantification of rainfall from stratiform clouds, the aim of the present study was to capture rainfall rates from both, advective-stratiform and convective precipitating cloud areas. Furthermore, the satellite-based estimates of rainfall were realised during day, night and twilight conditions resulting in a 24-hour prediction.

The final rainfall rate assignment technique was realised in three steps:

(i) Precipitating cloud areas are identified.

(ii) The precipitating cloud areas were separated into convective and advective-stratiform dominated precipitation areas.

(iii) Rainfall rates were assigned to the convective and advective-stratiform dominated precipitation areas, respectively.

Since the purpose of this study was to explicitly evaluate the potential of random forests for an improved rainfall rate assignment the rain area and rain process detected by the radar network was taken as basis for the investigation. Considering the dominant precipitation processes of convective and stratiform precipitation areas within extra-tropical cyclones, satellite-based information on the cloud top height, cloud top temperature, cloud phase and cloud water path were chosen as predictor variables. Precipitation events between April and September 2010 were chosen. Because of differing information content about the cloud properties during daytime, night-time and twilight conditions, the data set of precipitating events were split accordingly and treated separately. For the training and validation, the radar-based RADOLAN RW product from the DWD which provide area-wide gauge-adjusted hourly precipitation information, was used.

The development of the rainfall rate retrieval technique was realised in three steps. First, an extensive tuning study was carried out to optimise each of the RF models. Second, the RF models were trained using the optimum values for *ntree* and *mtry* found in the tuning study. Third, the RF models were applied to the validation data sets, respectively. Then the predicted rainfall rates are validated against co-located rainfall rates observed by the radar. The training, as well as prediction, were completed on an hourly basis. The rainfall rates were aggregated and also evaluated against each other for 3-h and 8-h interval.

In order to assess the model performances four different evaluation strategies were employed. First, the overall performance of the rainfall rate assignment technique during daytime, night-time and twilight conditions was investigated. Then, the performance on a scene-by-scene basis was considered closely before the diurnal performance of the RF models was shown. Finally, the influence of

different training data sets on the model performances was investigated.

Regarding the overall performance, as indicated by Rsq, hourly RR_{Obs} and RR_{Pred} show already a good correlation with Rsq=0.5 (day and night) and Rsq=0.48 (twilight). Higher temporal aggregation leads to better agreement. Rsq rises to 0.78 (day), 0.77 (night) and 0.75 (twilight) for 8-h interval. However, a consistent under-estimation (over-estimation) of the maximum (minimum) is apparent for all sets of models, regardless of aggregation or time of day. This shows that very high and low observed rainfall rates can not be captured by RF because of the averaging of the individual predictions over all trees. In addition, it is not possible to predict beyond the range of response values in the training data.

Comparing day, night and twilight performance show that daytime precipitation is generally predicted best by the model. Twilight and night-time predictions are generally less accurate but only by a small margin. The most obvious difference between the time of day lies in the spread of Rsq values. In comparison to daytime scenes, twilight and night-time scenes clearly show greater variability. The smaller number of predictor variables during twilight and night-time conditions as well as more difficult conditions to get cloud parameters during twilight might be a reason. Nevertheless, concerning the considerable problems of existing optical retrievals particularly during twilight, these results reveal a clear and unprecedented improvement and offer the possibility for 24 hour rainfall rate estimation.

The investigation on the influence of different training data sets on the model performances shows that different training data sets do influence the overall prediction accuracy of the according validation data set. Since the observations of the training data sets are used to build the RF model, these results are not surprising. However, the ranges of the standard verification scores are in a small margin. It seems that the biggest influence is on twilight data. The daytime data sets show the highest median and smallest range of Rsq values for all aggregation levels.

The results show that with the newly developed technique it is possible to assign rainfall rates with good accuracy even on an hourly basis. Furthermore, the rainfall rates can be assigned during day, night and twilight conditions which enables the estimation of rainfall rates for 24 h of a day. This shows great potential for upcoming optical rainfall retrievals. In this context, the spectral resolution provided by MSG SEVIRI offers the possibility for area-wide rainfall rate retrieval in near-real time and in quasi-continuous manner. Furthermore, the potential of rainfall rate assignment based on random forests is confirmed. However, further investigations are necessary to develop a final operational retrieval technique. In the next step, the rainfall rate assignment technique will be combined with a rain area detection and process separation technique. A combined evaluation scheme of precipitation detection, process separation and rainfall rate assignment is in-

tended and absolutely necessary. The complete rainfall retrieval technique will be extensively validated against radar-based RADOLAN RW data. The latter allows a quantification of the total error of the final operational rainfall retrieval technique.

Acknowledgements
The authors are grateful to the German Weather Service (DWD) for providing the RADOLAN and the EUMETSAT Earth Observation Portal (https://eoportal.eumetsat.int/) for providing the Meteosat data set. We would like to thank Jan Cermak from the University of Bochum for providing his cloud mask algorithm, Bernhad Seeger from the University of Marburg for the fruitful co-operation in the context of the Meteosat processing scheme and Tobias Ebert and Johannes Drönner for an excellent implementation. Finally, we thank the members of the Ecosystem Informatics PhD program funded by the University of Marburg for the valuable discussions. The work of M. Kühnlein has partly been supported by the Priority Program 1374 "Biodiversity Exploratories" and the work of T. Appelhans by the Research Unit 1246 "Kilimanjaro ecosystems under global change", both funded by the German Research Foundation (DFG), see www.environmentalinformatics-marburg.de.

Chapter 5

Precipitation estimates from MSG SEVIRI daytime, night-time and twilight data with random forests

Meike Kühnlein [(1)], Tim Appelhans [(1)], Boris Thies [(2)], Thomas Nauß [(1)]

[(1)] Environmental Informatics, Faculty of Geography, Philipps–University Marburg, Marburg, Germany

[(2)] Laboratory for Climatology and Remote Sensing, Faculty of Geography, Philipps-University Marburg, Marburg, Germany

Published in

Journal of Applied Meteorology and Climatology 53 (2014) 2457–2480.

Received 27 February 2014;
Received in revised form 4 August 2014;
Accepted 6 August 2014;

5 Precipitation estimates from MSG SEVIRI daytime, night-time and twilight data with random forests

Meike Kühnlein, Tim Appelhans, Boris Thies, Thomas Nauß

Abstract A new rainfall retrieval technique to determine rainfall rates in a continuous manner (day, twilight and night) resulting in a 24-hour estimation applicable to mid-latitudes is presented. The approach is based on satellite-derived information on cloud top height, cloud top temperature, cloud phase and cloud water path retrieved from Meteosat Second Generation (MSG) Spinning Enhanced Visible and InfraRed Imager (SEVIRI) data and uses the machine learning algorithm random forests. The technique is realized in three steps: (i) Precipitating cloud areas are identified. (ii) The areas are separated into convective and advective-stratiform precipitating areas. (iii) Rainfall rates are assigned separately to the convective and advective-stratiform precipitating areas. Validation studies were carried out for each individual step as well as for the overall procedure using co-located ground-based radar data. Regarding each individual step, the models for rain area and convective precipitation detection produce good results. Both retrieval steps show a general tendency towards elevated prediction skill during summer months and daytime. The RF models for rainfall rate assignment exhibit similar performance patterns, yet it is noteworthy how well the model is able to predict rainfall rates during night-time and twilight. The performance of the overall procedure shows very promising potential to estimate rainfall rates at high temporal and spatial resolutions in an automated manner. The near-real-time continuous applicability of the technique with acceptable prediction performances at 3 to 8 hourly intervals is particularly remarkable. This provides a very promising basis for future investigations into precipitation estimation based on machine learning approaches and MSG SEVIRI data.

Keywords Rainfall rate, Rainfall area, Precipitation process, Rainfall retrieval, Random forests, Machine learning, MSG SEVIRI, Geostationary satellites, Optical sensors

5.1 Introduction

Various investigations within biodiversity and ecological-oriented projects require area-wide precipitation information at high temporal and spatial resolution. However, despite its great importance, the high variability in space, time and intensity of this parameter still impedes its correct spatio-temporal detection and quantification. Moreover, rain gauges or ground-based radar networks which are generally used to observe precipitation, are sparse or unavailable in many regions. In this context, precipitation retrievals from optical sensors aboard geostationary (GEO) weather satellites may fill the gap by providing area-wide information about rainfall distribution and amount at high spatial and temporal resolution.

Traditionally, the spectral resolution of optical sensors available on GEO platforms was rather poor. This restriction only allowed schemes that rely on a relationship between cloud top temperatures measured in an infrared (IR) channel and rainfall probability/intensity (e. g. Adler and Mack, 1984; Arkin and Meisner, 1987). Such IR retrievals are most applicable for convective clouds, that can be easily identified in the IR and/or water vapour channels, and thus work best in the tropics (e. g. Levizzani et al., 2001b; Levizzani, 2003). However, they show considerable drawbacks concerning the detection and quantification of rainfall from advective-stratiform clouds in connection with extra-tropical cyclones (e. g. Ebert et al., 2007; Früh et al., 2007). This type of clouds is characterised by relatively warm and spatially homogeneous cloud top temperatures that do not differ significantly from raining to non-raining regions. Stratiform rain is usually not associated with high rainfall rates, but it often covers large areas and therefore contributes to a significant portion of the rainfall in a region. As a result, the use of retrieval techniques based solely on IR cloud top temperature leads to an underestimation of the detected precipitation area and to uncertainties concerning the assigned rainfall rate in such cases.

The enhanced spectral resolution of more recent GEO systems along with enhanced information content on cloud properties led several authors to suggest the utilisation of optical and microphysical cloud parameters derived from now available multispectral data set to improve optical rainfall retrievals (e. g. Ba and Gruber, 2001; Feidas and Giannakos, 2010; Kühnlein et al., 2010; Nauss and Kokhanovsky, 2006; Roebeling and Holleman, 2009; Rosenfeld and Gutman, 1994; Rosenfeld and Lensky, 1998; Thies et al., 2008b,a,d). They were able to show that cloud areas with a high optical thickness and a large effective particle radius possess a high amount of cloud water and are characterised by a higher rainfall probability and intensity than cloud areas with a low optical thickness and a small effective particle radius. Beside the use of optical and microphysical cloud parameters, many of the retrieval techniques make use of convective and stratiform precipitation area classification schemes to improve the accuracy of the satellite rainfall estimation (e. g. Adler and Negri, 1988; Arkin and Meisner,

1987; Anagnostou and Kummerow, 1997). In this context, Thies et al. (2008d) used cloud properties retrieved from MSG SEVIRI data to separate areas of differing precipitation processes within the rain area as part of a satellite-based retrieval scheme during day and night. Recently, Feidas and Giannakos (2012) and Giannakos and Feidas (2013) introduced techniques that classify convective and stratiform precipitation areas based on spectral and textural features of MSG SEVIRI data.

So far, most of the retrieval techniques use parametric approaches to relate cloud properties to precipitation (e. g. Adler and Negri, 1988; Cheng and Brown, 1995; Kühnlein et al., 2010; Levizzani et al., 1990; Thies et al., 2008d). These approaches require the specification of the underlying parametric tests and the related conceptual models. Moreover, these methods are only convenient for use with a few input variables. To address the consideration of multiple input variables as well as to avoid the assumption of parametric tests and underlying conceptual models it is necessary to consider other techniques. In this context, machine learning algorithms have been successfully adopted to remote sensing and rainfall applications (Capacci and Conway, 2005; Giannakos and Feidas, 2013; Grimes et al., 2003; Hsu et al., 1997; Islam et al., 2012a,b, 2014b; Rivolta et al., 2006) which show that machine learning algorithms might be suitable to overcome these limitations. Especially, the recent developments in parallel computing with machine learning offer new possibilities in terms of training and predicting speed and therefore make the usage and improvement of real time systems feasible.

In recent years, the machine learning technique random forest (RF, Breiman, 2001) has received increasing interest. This ensemble classification and regression technique is based on the assumption that a whole set of trees or networks produce a more accurate prediction than a single tree or network (Dietterich, 2002). It is one of the most accurate learning algorithms available and although RF has been shown to perform very well in a variety of environmental investigations (e. g. Cutler et al., 2007; Ghimire et al., 2010; Guo et al., 2011; Islam et al., 2014a; Mota et al., 2002; Pal, 2005; Prasad et al., 2006; Rodriguez-Galiano et al., 2012), the utilisation of RF in atmospheric sciences is rare. Yet, it offers a number of features that makes it well suited for use in remote sensing applications (e. g. it runs efficiently on large data sets, it can capture non-linear association patterns between predictors and response).

In summary, the encouraging results concerning rain area detection, rain process separation and rainfall rate assignment along with the enhanced information content on cloud properties at high spectral, spatial and temporal resolution point to a quite promising potential of current and upcoming GEO systems as basis for reliable rainfall retrievals. However, there is still a great deficit regarding the detection of rain areas and assignment of rainfall rates in the mid-latitudes, especially in connection with extra-tropical cyclones in a continuous manner (day, night, twilight) resulting in a 24-hour estimation at high temporal resolution. In

particular, all existing optical retrievals that are based on optical and microphysical cloud parameters are restricted to daytime and night-time conditions and do not cover twilight conditions (see e. g. Kidd and Levizzani, 2011). Hereby the potential offered by machine learning approaches will likely help to benefit from the potential offered by GEO systems. The usefulness of RF for satellite-based rainfall rate prediction has been successfully shown by Kühnlein et al. (2014a).

Building upon these results, we propose a new coherent daytime, twilight and night-time rainfall retrieval based on MSG SEVIRI data which provide information on cloud properties at high temporal and spatial resolution. The technique aims to retrieve rainfall rates for precipitation events in connection with extratropical cyclones in the mid-latitudes in a continuous manner at hight temporal resolution. Based on the dominant precipitation processes, the proposed rainfall retrieval consists of three steps which are applied consecutively to get the final product: (i) Identification of precipitating cloud areas. (ii) Separation of precipitating areas into predominately convective and advective-stratiform cloud regions. (iii) Individual process-oriented assignment of rainfall rates to these cloud areas. Hereby, the technique uses the relationship between cloud top temperature, cloud top height, cloud water path and cloud phase to retrieve information about precipitation. The ensemble method RF is used to develop the prediction algorithms. Beside the overall performance of the technique, the performance of each step (rain area detection, rain process separation, rainfall rate assignment) is investigated as well. The technique either can be used as general retrieval scheme or each retrieval step as standalone algorithm to detect rain area, discriminate convective and advective-stratiform precipitating cloud areas or rainfall rate assignment for nowcasting or climate purpose. The land area of Germany was chosen as study area for the development and validation of the new technique.

The structure of this paper is as follows. Section 5.2 introduces the underlying data sets and methods. The general methodology of the technique including theoretical background and conceptual design is presented in Section 5.3. The adjustment of the RF models as well as the appraisal of the new developed rainfall retrieval is outlined in Section 5.4. Finally, Section 5.5 provides conclusions.

5.2 Data and Methods

5.2.1 Data

For this study, data of the present European geostationary satellite system SEVIRI on board Meteosat Second Generation together with corresponding ground-based radar data are required. The MSG SEVIRI data were downloaded from the EUMETSAT data centre (`www.eumetsat.int`) and were pre-processed based on a newly designed Meteosat processing scheme implemented in co-operation with the computer science department at Marburg University. SEVIRI measures re-

flected and emitted radiance in 12 channels (between 0.6 and 14 μm) every 15 minutes and the data have a nominal spatial resolution of 3 by 3 km at sub satellite point (Aminou, 2002). In order to retrieve cloud properties such as cloud effective radius and cloud optical thickness the semi-analytical approach SLALOM (SimpLe Approximations for cLOudy Media) is used (Nauss and Kokhanovsky, 2011; Kühnlein et al., 2013). The radar-based precipitation data are provided by the German Weather Service (DWD). The so-called RADOLAN RW product is based on measurements with a C-band Doppler radar and the precipitation intensities are adapted with ground-based precipitation measurements (Bartels et al., 2004). The scan interval is one hour and the final product is a composite consisting of 16 German radar stations and some from neighbouring countries (e. g. Nancy/France) covering the whole area of Germany at a spatial resolution of 1 km x 1 km. The RADOLAN RW product provides a reliable training and validation basis. Moreover, Germany can be regarded as sufficiently representative for mid-latitudes precipitation formation processes since it is dominated by frontally induced precipitation processes in connection with extra-tropical cyclones. Because of the different temporal characteristics of both data sets, the satellite-based products are aggregated over the time interval of one hour. This is done by taking the arithmetic mean of the four scenes available every hour (e. g. 11:45 UTC, 12:00 UTC, 12:15 UTC and 12:30 UTC according to the aggregation period of the ground-based radar product). Hereby only those pixels which are classified as cloudy over the entire time interval are considered. In order to retrieve the cloud mask the algorithm developed and implemented by Cermak (2006) and Cermak and Bendix (2008) is used. Because of the differing viewing geometries between both systems, the radar product was spatially projected and aggregated (mean) to the geometry of SEVIRI. This is done by taking a mean over all RADOLAN RW pixels that are located within one MSG SEVIRI pixel.

The final data set consists of 967 scenes of precipitation events between January and December 2010. Based on the RADOLAN RW product, scenes with at least 2000 rainy pixels (higher than 0.10 mm/h) were chosen as precipitating events. The data set of precipitating events is split because of the differing information content about the cloud properties between the different times of day. Based on solar zenith angle the data is divided into day (<70°), twilight (70°-108°) and night (>108°) data sets.

5.2.2 Random forests

Breiman (2001) proposed the ensemble technique RF. Apart from its proven performance in a variety of environmental studies (e.g. Cutler and Stevens, 2006; Ghimire et al., 2010; Guo et al., 2011; Mota et al., 2002; Pal, 2005; Prasad et al., 2006; Rodriguez-Galiano et al., 2012; Ruiz-Gazen and Villa, 2007) we have opted for RF for a number of (largely practical) inherent characteristics of the method (e. g. RF runs efficiently on large data sets, RF does not require the specification

of an underlying data model).

Here, we only can give a brief general overview of the algorithm. For more theoretical details the reader is referred to the literature (Boulesteix et al., 2012; Breiman and Cutler, 2008; Breiman, 2001; Malley et al., 2011; Strobl et al., 2009).

In general, RF fits many classification and regression trees to a data set, and then combines the predictions from all the trees. At first n bootstrap samples are randomly selected with replacement from the data set. For each bootstrap, a different subset of roughly two-third of the bootstrap sample is used to grow a decision tree. The remaining one-third of the training data is not included in the learning sample for this tree and can be used to test it as an out of bag sample (OOB). Therefore, the RF model is applied to the OOB data. Then, the deviations between predicted and observed values are used to calculate the OOB error which is for classification the error rate defined as the proportion between misclassifications and the total number of OOB cases. For regression, the mean square error (MSE) is used. These resulting OOB errors provide an unbiased estimate of the generalised error and can be considered as an internal validation. After the selection of n bootstrap samples, one tree for each of the bootstrap samples is grown (resulting in n trees) with the following characteristic: at each node, a subset of m predictor variables is selected randomly to create a binary rule. In other words, m specifies the number of randomly chosen variables upon which the decision for the best split at each node is made. m is held constant during the forest growing. Each of the n trees is grown to the largest extent possible. Finally, predictions are calculated by putting each OOB observation or observation of the test data set down each of the trees and evaluate the predictions of the n trees. This means for classification that for each observation, each individual tree votes for one class and the forest predicts the class that has the majority of votes (Breiman and Cutler, 2008). For example, if 500 trees are grown and 400 of them predict that this particular pixel is “rain” and 100 as no rain, the predicted output will be rain. For regression, the predictions of all regression trees are averaged to produce the final estimate (Breiman, 2001).

Beside the favourable features of RF, some limitations need also be mentioned. One of the severest drawbacks of RF is that it does lack interpretability. Since the predictions are derived using a forest of trees, it is not possible to easily illustrate how the predictions are made (i.e. no single tree can be drawn to illustrate the decisions upon which the predictions are based). Furthermore, RF classification may perform poorly when learning from an extremely unbalanced data set. Hence, a strategy has to be applied to overcome this limitation. In addition, averaging over all regression trees means that it is neither possible to predict beyond the range of response values in the training data nor to predict the entire range of response values. Furthermore, RF is a truly random statistical method which entails a number of methodological issues related to repeatability and generalisability of the analyses. For in depth descriptions and generally

accepted solutions of how to address these issues, the reader is referred to the fundamental statistical literature provided for the method.

The R (R Development Core Team, 2014) implementation of the RF library (package "randomForest"), which was created by Liaw and Wiener (2002) based on the original Fortran code by Leo Breiman and Adele Cutler, was used to implement the RF algorithm. This algorithm falls into the embarrassingly parallel category. This means that the number of trees to grow within a RF model can be divided into independent subsets, since each tree in the forest depends only on the given data set and not on the other trees. The subsets can be built on all available cores or on different machines. Then the resulting RF objects are combined to get the final forest. While the parallel execution of RF classification was realised by using the "parallel" package (R Development Core Team, 2014), we used the "foreach" (Revolution Analytics, 2012b) and "doSNOW" (Revolution Analytics, 2012a) packages for the parallel execution of RF regression in order to reduce computation time.

5.2.3 Verification scores

For the development and appraisal of the technique, standard categorical and continuous verification scores following the suggestions of the International Precipitation Working Group (IPWG, Turk and Bauer, 2006) are calculated on a pixel basis for each scene in comparison with corresponding ground-based radar precipitation measurements from the DWD. As categorical verification scores the bias, probability of detection (POD), probability of false detection (PFD), false alarm ratio (FAR), critical success index (CSI), equitable threat score (ETS), Hansen-Kuipers discriminante (HKD) and Heidke skill score (HSS) are used. The basis of these categorical verification scores is a contingency table showing agreement and disagreement in the prediction and observation data set (see table 5.1). "Hits" indicate the number of correctly predicted "rain" or "convective rain" pixels. "False alarms" denote estimates of "rain" or "convective rain" pixels that do not correspond to observed events, while "misses" indicates the number of "rain" or "convective rain" pixels that were not predicted by RF. "Correct negatives" gives the number of correctly predicted pixels of "no rain" or "advective-stratiform". Based on the elements of this table the categorical verification scores are calculated as described in table 5.2. For the rainfall rates, continuous verification scores are calculated on a pixel basis for each scene. As measures of the agreement between observed and predicted values mean error (ME = bias), mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (Rsq) are chosen. Further information about the verification scores can be found in the WMO report of Stanski et al. (1989) or Ebert (2002).

	Observation yes	**Observation no**	
Prediction yes	hits (H)	false alarms (F)	H+F
Prediction no	misses (M)	correct negatives (C)	M+C
	H+M	F+C	Total(T=H+M+F+C)

Table 5.1: *Contingency table.*

Name	Equation	Range	Optimum
Bias	$bias = \frac{H+F}{H+M}$	0...∞	1
Probability of detection	$POD = \frac{H}{H+M}$	0...1	1
Probability of false detection	$PFD = \frac{F}{F+C}$	0...1	0
False Alarm Ratio	$FAR = \frac{F}{H+F}$	0...1	0
Critical Success Index	$CSI = \frac{H}{H+F+M}$	0...1	1
Equitable Threat Score	$ETS = \frac{H-p_H}{((H+F+M)-p_H)}$ where $p_H = \frac{(H+M)(H+F)}{T}$	-1/3...1	1
Heidke Skill Score	$HSS = \frac{(H+F)(H+M)+(C+F)(C+M)}{T}$	-∞....1	1
Hansen-Kuipers discriminant	$HKD = \frac{H}{H+M} - \frac{F}{F+C}$	-1...1	1

Table 5.2: *Categorical verification scores with computation, theoretical range of values, and optimum value.*

5.3 General Methodology

Following the conceptual model of rainbands dominated by different rainfall processes in connection with extra-tropical cyclones (Houze, 1993), the precipitation field can be decomposed into areas dominated by different rainfall processes. Advective-stratiform background and intermediary precipitation which are linked to light precipitation intensities (further referred as advective-stratiform precipitation process) and narrow cold-frontal, wide cold-frontal and warm-frontal rain bands which are characterised by high rainfall intensities (further referred as convective precipitation process).
This is the reason why the final rainfall retrieval is realised as a three step process:

i. Identification of precipitating cloud areas.

ii. Separation of precipitating cloud areas according to their process into convective and advective-stratiform cloud regions.

iii. Individual assignment of rainfall rates to these cloud areas.

These three parts are realized subsequently by means of individual RF models, binary classification models for steps (i) and (ii), and regression models for step (iii).

While the choice of the response variables is obvious for steps (i) and (iii), the response variable of step (ii) has to depict the delineation between advective-stratiform and convective precipitating clouds. There is not a standard threshold to distinguish between these two types of precipitating clouds. We followed the threshold used by (Thies et al., 2008d). Precipitation areas with higher than 1.8 mm/hr are flagged as convective, with 0.1-1.8 mm/hr are flagged as advective-stratiform. All response variables needed for steps (i), (ii) and (iii) can be identified in the RADOLAN RW product.

Regarding the predictor variables, careful consideration of the conceptual framework of the rainfall retrieval is necessary. This means that the dominant precipitation processes with a strong focus on extra-tropical cyclones must be considered. The choice of predictor variables leans on the work of Thies et al. (2008d).

Convectively dominated precipitation areas with higher rainfall intensities are characterised by a larger cloud depth and a cloud top reaching higher into the troposphere which result in colder cloud tops. The brightness temperature in the 10.8 μm channel ($BT_{10.8}$) serves information about the cloud top temperature. Using the brightness temperature difference between the water vapour (WV) and IR channels information about the cloud top height relative to the tropopause level can be obtained which enables a reliable identification of deep convective clouds (Tjemkes et al., 1997; Schmetz et al., 1997; Heinemann et al., 2001). The differences $\Delta T_{WV6.2-IR10.8}$ and $\Delta T_{WV7.3-IR12.1}$ have been used to include different sensitivities on cloud top height.

While convectively dominated precipitation areas are characterised by cold cloud top temperatures, advective-stratiform precipitation areas are not necessarily connected to the latter. Frequently, they are not even colder than surrounding non-precipitating cloud areas. Therefore, several authors decided to use optical and microphysical cloud parameters derived from multispectral satellite data or a suitable selection of spectral channels and channel combinations which provide information about cloud parameters, for rain area delineation (Rosenfeld and Lensky, 1998; Lensky and Rosenfeld, 2003a; Ba and Gruber, 2001; Nauss and Kokhanovsky, 2006, 2007; Thies et al., 2008a,b) rain process differentiation (Thies et al., 2008d; Feidas and Giannakos, 2012), rainfall intensity differentiation (Thies et al., 2008d) and rainfall rate assignment (Kühnlein et al., 2010, 2014a). They showed that cloud areas with a high cloud water path (i.e. large enough

combination of the optical thickness and effective particle radius) possess a higher rainfall probability and higher rainfall rates than cloud areas with small cloud water path. For daytime, the cloud retrieval SLALOM is used to derive optical and microphysical cloud properties (e.g. cloud effective radius (a_{ef}), cloud optical thickness (τ)). Because the application of the commonly available cloud property retrievals to ice clouds requires the a-priori definition of ice particle geometries which in turn heavily influences the retrieved values (see also Kokhanovsky et al. (2005)), we decided to use additionally the reflectance at 0.6 μm ($VIS_{0.6}$), 0.8 μm ($VIS_{0.8}$) and 1.6 μm ($NIR_{1.6}$) channels directly as input variables for the rain area delineation, rain process separation and rainfall rate assignment.

For night-time, there are MSG SEVIRI retrievals that can compute optical and microphysical cloud properties (e.g. VISST-SIST algorithms developed by NASA Langley Cloud and Radiation Group). Unfortunately, these algorithms derive reliable properties only for non-raining optically thin clouds (Minnis et al., 2011). However, several case studies have shown that implicit information about microphysical and optical cloud properties is available in the emissive channels (Baum et al., 2000; Inoue, 1985; Lensky and Rosenfeld, 2003a; Ou et al., 1993; Stone et al., 1990; Strabala et al., 1994). Based on these studies Thies et al. (2008a) demonstrated that the SEVIRI channel differences $\Delta T_{3.9-10.8}$, $\Delta T_{3.9-7.3}$, $\Delta T_{8.7-10.8}$ and $\Delta T_{10.8-12.1}$ provide information about the CWP. This is due to the differing sensitivities of the respective channels on microphysical and optical cloud properties. During twilight conditions neither the reflectances of VIS and NIR channels (due to insufficient solar illumination) nor the SEVIRI channel differences $\Delta T_{3.9-10.8}$ and $\Delta T_{3.9-7.3}$ can be used as surrogates for the CWP. The 3.9 μm channel radiance contains both reflected solar radiance and thermal emitted radiance. To use the 3.9 μm channel the solar component must be quantified and eliminated. This itself has been investigated by several studies (Rao et al., 1995; Rosenfeld and Lensky, 1998) and they showed that simplifications and assumptions are necessary which are only acceptable for a certain kind of cloud. To prevent misinterpretation, predictors containing the 3.9 μm channel are not used to gain information about the CWP and therefore there are no predictor variables available representing the CWP during twilight.

Since it is well-known that effective rain formation processes are mainly coupled to ice particles in the upper part of the cloud and the "seeder-feeder" effect (Houze, 1993), the cloud phase, which can be obtained through the channel differences between 8.7 μm and 10.8 μm channel ($\Delta T_{8.7-10.8}$) as well as between 10.8 μm and 12.1 μm ($\Delta T_{10.8-12.1}$), is incorporated (Strabala et al., 1994; Ackerman et al., 1998; Thies et al., 2008b).

To consider the dominant precipitation processes with a strong focus on extratropical cyclones as cloud physical parameters the cloud top height (CTH), cloud top temperature (CTT), cloud phase (CP) and cloud water path (CWP) are chosen. As shown above, a proper SEVIRI channel selection can be used as

surrogates for these cloud physical parameters and therefore as predictor variables for the RF models. An overview about the predictor variables with regard to day, night and twilight as well as precipitation processes is given in tables 5.3 and 5.4. In addition, all MSG SEVIRI channels are also incorporated.

	Daytime	Twilight	Night-time
CTH	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$
	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$
CTH/CTT	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$
CP	$\Delta T_{8.7-10.8}$	$\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$
	$\Delta T_{10.8-12.1}$	$\Delta T_{8.7-10.8}$	$\Delta T_{10.8-12.1}$
CWP	$VIS_{0.6}$		$\Delta T_{3.9-10.8}$
	$VIS_{0.8}$		$\Delta T_{3.9-7.3}$
	$NIR_{1.6}$		
SEVIRI channels	$WV_{6.2}$	$WV_{6.2}$	$IR_{3.9}$
	$WV_{7.3}$	$WV_{7.3}$	$WV_{6.2}$
	$IR_{8.7}$	$IR_{8.7}$	$WV_{7.3}$
	$IR_{10.8}$	$IR_{10.8}$	$IR_{8.7}$
	$IR_{12.1}$	$IR_{12.1}$	$IR_{10.8}$
			$IR_{12.1}$

Table 5.3: *Overview of RF predictor variables used within the RF models for rain area detection and rain process separation.*

	Daytime		Twilight		Night-time	
	Stratiform	Convective	Stratiform	Convective	Stratiform	Convective
CTH	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$	$\Delta T_{6.2-10.8}$
	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$	$\Delta T_{7.3-12.1}$
CTH/CTT	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$
CP	$\Delta T_{8.7-10.8}$	$\Delta T_{8.7-10.8}$	$\Delta T_{8.7-10.8}$	$\Delta T_{8.7-10.8}$	$\Delta T_{8.7-10.8}$	$\Delta T_{8.7-10.8}$
	$\Delta T_{10.8-12.1}$	$\Delta T_{10.8-12.1}$	$\Delta T_{10.8-12.1}$	$\Delta T_{10.8-12.1}$	$\Delta T_{10.8-12.1}$	$\Delta T_{10.8-12.1}$
CWP	$VIS_{0.6}$	$VIS_{0.6}$			$\Delta T_{3.9-10.8}$	$\Delta T_{3.9-10.8}$
	$VIS_{0.8}$	$VIS_{0.8}$			$\Delta T_{3.9-7.3}$	$\Delta T_{3.9-7.3}$
	$NIR_{1.6}$	$NIR_{1.6}$				
	a_{ef}					
	τ					
	CWP					
SEVIRI channels	$WV_{6.2}$	$WV_{6.2}$	$WV_{6.2}$	$WV_{6.2}$	$WV_{3.9}$	$WV_{3.9}$
	$WV_{7.3}$	$WV_{7.3}$	$WV_{7.3}$	$WV_{7.3}$	$WV_{6.2}$	$WV_{6.2}$
	$IR_{8.7}$	$IR_{8.7}$	$IR_{8.7}$	$IR_{8.7}$	$WV_{7.3}$	$WV_{7.3}$
	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{10.8}$	$IR_{8.7}$	$IR_{8.7}$
	$IR_{12.1}$	$IR_{12.1}$	$IR_{12.1}$	$IR_{12.1}$	$IR_{10.8}$	$IR_{10.8}$
					$IR_{12.1}$	$IR_{12.1}$

Table 5.4: *Overview of RF predictor variables used within the RF models for rainfall rate assignment.*

Because of different sets of predictor variables during daytime, night-time and twilight conditions, the respective precipitation events must be treated separately.

For example, during night-time the channels at visible and very near infrared wavelengths (0.6 to 1.6 μm) are not available. During twilight and daytime the use of the 3.9 μm channel is complicated due to the varying solar component in this channel. This means that depending on the time of the day different RF models are built and adapted for each retrieval step. Furthermore, to capture the peculiarities of convective and advective-stratiform dominated rainfall events in the different seasons our study period January to December 2010 is divided into two sub periods for which two different models are developed independently. The first period is from April to September (AS) and the second includes January to March and October to December (OM). Thus, for the complete retrieval, 24 separate RF models need to be formulated (incl. tuning, training and validation). An overview of the stepwise scheme for rainfall rate assignment is shown in figure 5.1.

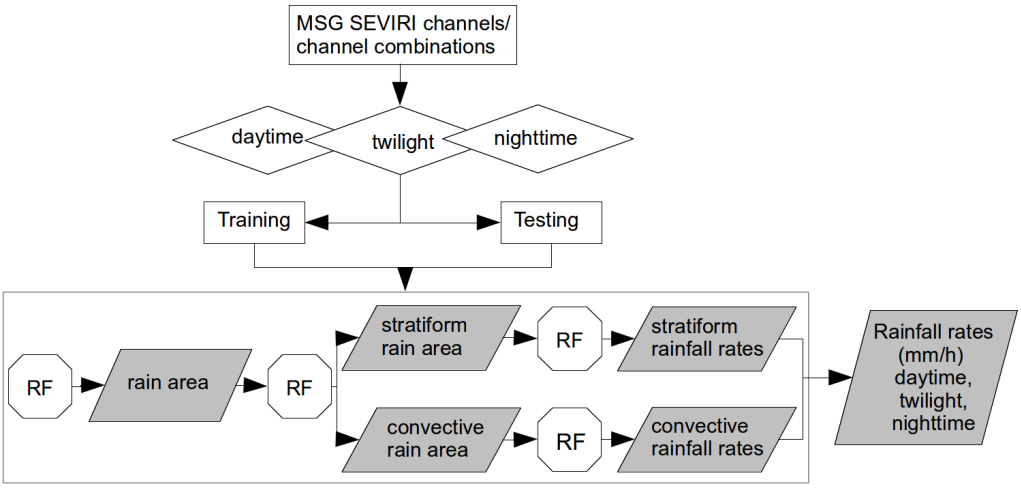

Figure 5.1: *Overview of the stepwise scheme for rainfall rate assignment.*

For the development and validation of the technique, a data set consisting of 311 (27) daytime, 305 (95) twilight and 117 (112) night-time satellite scenes of precipitation events for AS (OM) are used. Each of these three data sets are randomly split into training (1/10 of the scenes) and validation data sets (9/10 of the scenes). The training data sets are used to train the respective RF model whereas the validation data sets are used to validate the predictions afterwards. The precipitation events taken for training are independent from those taken for validation.

The development and validation of the new rainfall retrieval is realized as follows: First, a tuning study is presented to customise each of the RF models. Then, the RF models are trained using the optimum values of model parameters found in the tuning study. Once the RF models are established, it is possible to predict the appropriate response variable based on the same predictors used for training. In order to assess the quality of the technique, the RF models are applied to the day-AS, day-OM, night-AS, night-OM, twilight-AS and twilight-

OM validation data sets and the predictions are validated against co-located radar measurements. Details are described in the following sections.

5.4 A new technique for satellite-based rainfall rate assignment

5.4.1 Tuning of the RF models

The implementation of the RF classification and regression models is realised using the R-package randomForest. This package requires the adjustment of two parameters, the overall number of trees to grow within a forest (*ntree*) and the number of predictor variables randomly sampled for use at each split (*mtry*). Beside the assessment of optimal values of *ntree* and *mtry* it has to be considered that the classification training data sets are extremely unbalanced: for example the class rain represents only about 20% of the whole database of day-AS, only 30% correspond to the class convective. Since RF classification may perform poorly when learning from an extremely unbalanced data set (see e. g. Liu et al., 2006) a strategy has to be found to overcome this restriction. A possibility for handling the class distribution imbalance happens at the data level. This means that balancing the class distribution is done by either down-sampling the majority class or/and over- sampling the minority class. Liu et al. (2006) explore several re-balancing schemes based on both down-sampling and over-sampling. The results show that down-sampling seems to be the better strategy. Moreover, Dupret and Koda (2001) showed that the optimal re-balancing strategy is not necessarily to down-sample the majority class at the same level as the minority class.

To address the unbalance of class distribution two techniques are investigated. In the first option the training data sets are re-balanced by keeping all the observations belonging to the minority class (rain or convective) and by randomly sampling (without replacement) a given (smaller than the original) number of observations from the majority class (no rain or advective-stratiform). The other approach is to maintain a fixed class distribution used for all trees, each class is sampled separately. Hence, a given number of samples from the minority and majority class is used to build the trees. To assess which sampling method is the best, RF models using these two different sampling techniques with different numbers of observations from the minority and majority class while keeping *ntree* and *mtry* constant (both are set to default values, $ntree = 500$ and $mtry = \sqrt{p}$, with p total number of predictors) are created for each training data set (day-AS, day-OM, night-AS, night-OM, twilight-AS, twilight-OM). The RF-models are applied to the validation data sets and the predictions are compared to corresponding ground based radar data. Verification scores are calculated on a scene-by-scene basis and a mean is taken over all scenes. Table 5.5 summarises how the performance of the rain area detection technique changes using the afore-

	Number of samples minority class	Number of samples majority class	Area (%), Satellite	Area (%), Radar	Bias	POD	PFD	FAR	CSI	ETS	HSS	HKD
Rain area re-balanced training data sets	R	1.0*R	32.86	18.17	1.82	0.82	0.22	0.54	0.40	0.27	0.42	0.56
	R	1.2*R	27.72	18.17	1.55	0.76	0.18	0.51	0.41	0.29	0.44	0.55
	R	1.4*R	27.53	18.17	1.55	0.75	0.17	0.51	0.41	0.29	0.44	0.55
	R	1.6*R	25.86	18.17	1.45	0.73	0.16	0.49	0.41	0.29	0.45	0.54
	R	1.8*R	24.15	18.17	1.37	0.70	0.14	0.48	0.41	0.30	0.45	0.53
	R	2.0*R	23.33	18.17	1.31	0.68	0.14	0.47	0.41	0.29	0.45	0.51
	R	2.2*R	21.55	18.17	1.23	0.66	0.12	0.46	0.40	0.29	0.45	0.51
	R	2.4*R	20.47	18.17	1.16	0.64	0.11	0.44	0.40	0.30	0.46	0.50
	R	2.6*R	19.56	18.17	1.10	0.62	0.11	0.44	0.40	0.29	0.45	0.48
	R	2.8*R	18.79	18.17	1.06	0.59	0.10	0.43	0.39	0.29	0.45	0.47
	R	3.0*R	18.02	18.17	1.01	0.58	0.09	0.42	0.39	0.29	0.45	0.46
Rain area fixed sampsize	0.1*R	1.0*0.1*R	34.09	18.17	1.87	0.84	0.23	0.54	0.40	0.27	0.42	0.57
	0.1*R	1.2*0.1*R	30.60	18.17	1.72	0.80	0.20	0.53	0.41	0.28	0.44	0.57
	0.1*R	1.4*0.1*R	28.25	18.17	1.59	0.78	0.18	0.51	0.42	0.29	0.45	0.56
	0.1*R	1.6*0.1*R	26.49	18.17	1.48	0.75	0.16	0.49	0.42	0.30	0.46	0.56
	0.1*R	1.8*0.1*R	24.78	18.17	1.39	0.72	0.15	0.48	0.42	0.31	0.46	0.55
	0.1*R	2.0*0.1*R	23.33	18.17	1.31	0.70	0.13	0.46	0.42	0.31	0.47	0.54
	0.1*R	2.2*0.1*R	22.04	18.17	1.24	0.68	0.12	0.45	0.41	0.31	0.47	0.52
	0.1*R	2.4*0.1*R	20.84	18.17	1.16	0.65	0.11	0.44	0.41	0.31	0.46	0.51
	0.1*R	2.6*0.1*R	19.73	18.17	1.10	0.63	0.10	0.43	0.41	0.31	0.46	0.50
	0.1*R	2.8*0.1*R	18.77	18.17	1.04	0.61	0.10	0.42	0.40	0.30	0.46	0.49
	0.1*R	3.0*0.1*R	18.00	18.17	1.00	0.60	0.09	0.42	0.40	0.30	0.46	0.48

Table 5.5: *Performance of the rain area and convective precipitation detection technique using different sample techniques.*
The scores are based on data pairs of 311 daytime precipitation scenes from April to September (day-AS data set). The data set is randomly split into training (1/10 of the scenes) and validation data set (9/10 of the scenes). Scores are calculated on a scene-by-basis and a mean is taken over all scenes. R denotes the number of observations of raining cases. Re-balanced training data sets are training data sets which are re-balanced by keeping all the observations belonging to the minority class (rain) and by randomly sampling (without replacement) a given (smaller than the original) number of observations from the majority class (no rain). Fixed sampsize means that each class is sampled separately by the RF models.

mentioned techniques to address the unbalance of class distribution with given numbers of samples from the minority and majority class. Shown is the performance for day-AS, however, the other time periods exhibit very similar results. Overall, there are no big differences arising from the two sample techniques. Using re-balanced data sets, the performance generally drops by a small margin. The bigger influence is caused by a different number of samples from the minority and majority class. Using the number of samples from the minority rain class (0.1*R) and majority no-rain class (1.0*0.1*R) leads to overestimation of the rain area predicted by RF (area (%) satellite = 34.09, area (%) radar = 18.17, bias=1.82). As indicated by the POD, 84% of the radar observed raining pixels are also identified by RF. However, the PFD of 0.23 and FAR of 0.54 show also a false identification. Increasing the number of samples from the majority no-rain class decreases the percentage rain area predicted by RF which results in a smaller

bias. A consistent overestimation is apparent till the number of samples from the majority class reaches three times the number of observations from the minority class. With increasing the number of samples from the majority no-rain class the POD values decrease. This is paralleled by decreasing the false identification shown by smaller PFD and FAR values. The CSI, ETS and HSS values vary only by a small margin, regardless of the number of samples. Obviously, it is generally desirable to maximise the ratio between recognition and false alarms. However, the final decision, as to which strategy to adopt also need to consider the overall goal of the investigation. To address the unbalance of class distribution within this study, we decided to use a fixed sample size. This means a given number of samples from the minority and majority class is used to build the RF models.

To assess the optimal value of *ntree*, a large number of RF models using randomly selected subsets of each data set (day-AS, day-OM, twilight-AS, twilight-OM, night-AS, night-OM) are created using 1000 trees for all possible values of *mtry* (*mtry*=1 to maximum). Then the OOB mean error rates (classification) and OOB mean square errors (regression) of every possible value of *mtry* are averaged. The change of the error rates for rain area detection and rain process separation as a result of the number of trees is shown for different *mtry* values (minimum *mtry*=1, mean *mtry*=5/6, maximum *mtry*=10/13) in figure 5.2. Breiman (2001)

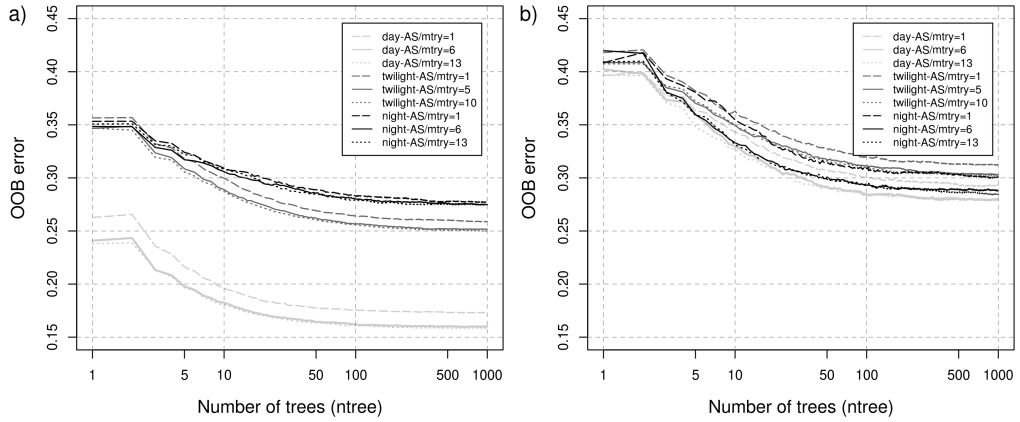

Figure 5.2: *Effect of number of trees (ntree) and random split variables (mtry) on OOB error (mean error rate). Results shown for (a) rain area and (b) rain process day-AS, night-AS and twilight-AS tuning data sets.*

suggested that the generalisation error converges as the number of trees increases. Adding more and more trees to the model does not result in over-adjustment. The main limitation of increasing *ntree* is the extra computation time. As can be seen in figure 5.2, from around 500 trees onwards, the mean error rate of each data set stabilises and the addition of trees neither increases nor decreases the mean error rate. As a result, *ntree* is set to 500 in all classification models. This is also done in all regression models as Kühnlein et al. (2014a) has shown that *ntree* of 500 is large enough to produce a stable rainfall rate prediction.

The suggested default value of *mtry* for classification is $\sqrt{p}$ and for regression

p/3 (Breiman, 2001). However, this should only be regarded as a recommendation, since in practice the best values for *mtry* will depend on the problem at hand. In a similar approach to the *ntree* tuning, the optimal values of *mtry* are assessed with a large number of RF models using different possible values of splitting variables (*mtry*=1 to maximum number of *mtry*) while keeping *ntree*=500 constant. After the RF model is computed with *mtry*=1, *mtry* is increased by 1 till the maximum number of *mtry* is reached, and a new forest is built with each of the new *mtry*. This whole process is repeated several times using randomly selected subsets of the data sets. Then, the value of *mtry* leading to the smallest OOB error is selected for the according RF model.

An overview of the final model parameters chosen for each model is given in table 5.6.

		Day-AS	Day-OM	Twilight-AS	Twilight-OM	Night-AS	Night-OM
Rain area	*ntree*	500	500	500	500	500	500
	mtry	5	4	4	4	4	4
	sampsize rain	0.1*R	0.1*R	0.1*R	0.1*R	0.1*R	0.1*R
	sampsize no rain	2.0*(0.1*R)	0.5*(0.1*R)	1.6*(0.1*R)	1.4*(0.1*R)	1.6*(0.1*R)	1.25*(0.1*R)
Convective precipitation	*ntree*	500	500	500	500	500	500
	mtry	5	4	4	5	6	4
	sampsize c	0.1*C	0.1*C	0.1*C	0.1*C	0.1*C	0.1*C
	sampsize a	0.9*(0.1*C)	0.5*(0.1*C)	0.85*(0.1*C)	1.3*(0.1*C)	0.85*(0.1*C)	1.2*(0.1*C)
Stratiform rainfall rate	*ntree*	500	500	500	500	500	500
	mtry	7	7	6	6	6	6
Convective rainfall rate	*ntree*	500	500	500	500	500	500
	mtry	7	7	6	6	6	6

Table 5.6: *Overview about model parameters used within the RF models for rain area and convective precipitation process detection as well as rainfall rate assignment (convective and stratiform).*
To address the imbalance in the class distribution, a fixed sample size is used. This means that each class is sampled separately by the RF models. R denotes the number of observations of raining cases and C indicates the number of observations of convective cases. Abbreviations are as follows: a, advective-stratiform; c, convective.

5.4.2 Validation of the RF models

After the successful adjustment of the 24 RF models for rain area detection, rain process separation and rainfall rate assignment, each model is trained using the respective values of re-balancing, *ntree* and *mtry* found in the tuning study (section 5.4.1).

In order to assess both the model performances at each step as well as the overall performance, two separate evaluation studies are carried out. First, each step is evaluated individually to eliminate the influence of error propagation. This means that at each step, the evaluation reference is derived from the radar data. For example, when assessing the accuracy of the process delineation, the rain area as identified by the radar data is used. Second, to assess the overall performance of the retrieval technique, all RF steps are applied and only the final predicted

rainfall rates (RR_{Pred}) are validated against the observed rainfall rates (RR_{Obs}).

5.4.3 Performance of each rainfall retrieval step

5.4.3.1 Rain area detection

The verification scores for rain area detection presented in figure 5.3 are calculated on a pixel basis for each scene. Given that the evaluation is done on a number of scenes depending on the time period considered, results are summarised as box and whisker plots and shown separately for day-AS, day-OM, twilight-AS, twilight-OM, night-AS and night-OM.

As a general pattern it can be seen that predictions perform best during daytime and summer. Overall, the central tendencies of the verification scores are very promising, however, the spread is regularly rather wide indicating that the ability to capture the observed rain area is extremely variable.

In particular, the bias indicates very good agreement between the rain area identified by the radar and RF. For all periods the bias is around 1.2 indicating a slight overestimation. As indicated by POD, there is a general tendency of enhanced ability to detect precipitating areas during AS in comparison to OM. In general, about 70% of the radar observed raining pixels are also identified by RF during the AS daytime period, even though a rather wide spread of the POD is apparent. Twilight-AS and night-AS values are generally slightly smaller with a median of 0.6 and 0.58, respectively. The encouraging POD values are paralleled by small PFD values across all periods. In contrast, the FAR exhibits a clear seasonal dependency with less false alarms during summer months regardless of time of day. The overall good performance is further supported by the CSI (e. g. 0.41 (day-AS), 0.34 (twilight-AS), 0.35 (night-AS)) and ETS (e. g. 0.32 (day-AS), 0.18 (twilight-AS), 0.20 (night-AS)).

In summary, RF shows reasonable performance in rain area detection for all validation data sets (e. g. HSS between 0.23 and 0.49). This is particularly remarkable for twilight conditions, especially considering the high temporal and high spatial resolution. The enhanced performance during daytime is probably due to the higher number of predictor variables and the increased information content about the CWP inherent in the VIS and NIR channels (see section 5.3). The fact that summer months are generally predicted slightly better, might simply be a consequence of the larger precipitation area during this season.

To provide an impression of the spatial prediction distribution, figure 5.4 b shows an example of the rain area detected by RF during daytime conditions in comparison to the ground-based radar data. It illustrates the good agreement of the observed and predicted rain area. However, it can also be seen that RF tends to overestimate the area, though, this is mainly restricted to the direct vicinity of the observed raining areas.

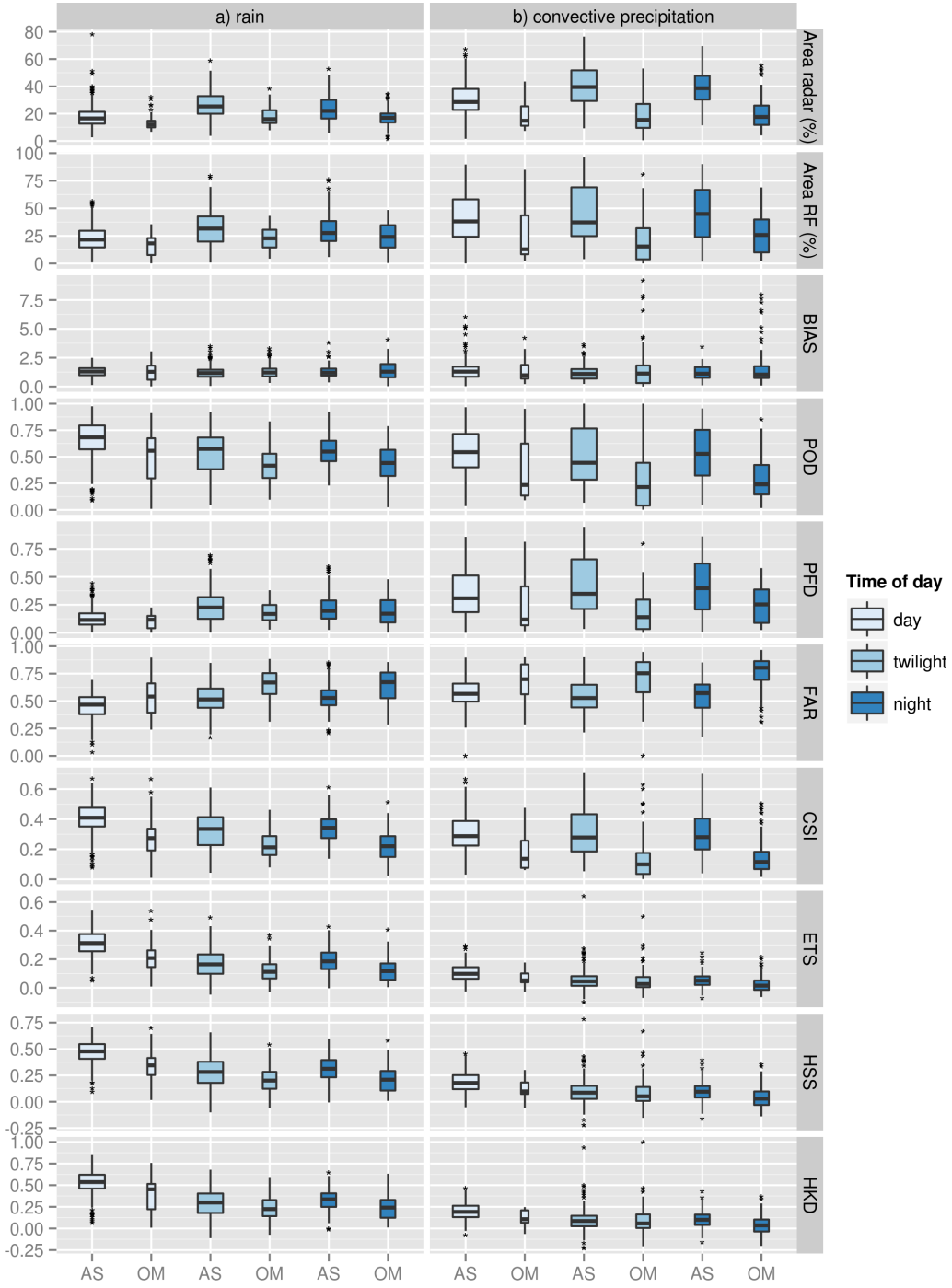

Figure 5.3: *Standard verification scores for (a) rain area and (b) convective precipitation process area across all classified scenes by RF.*
The scores are calculated on a scene-by-scene basis and are based on 311 (27) daytime, 305 (95) twilight and 117 (112) night-time precipitation scenes from 2010 during AS (OM). Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars. Box widths are relative to number of observations.

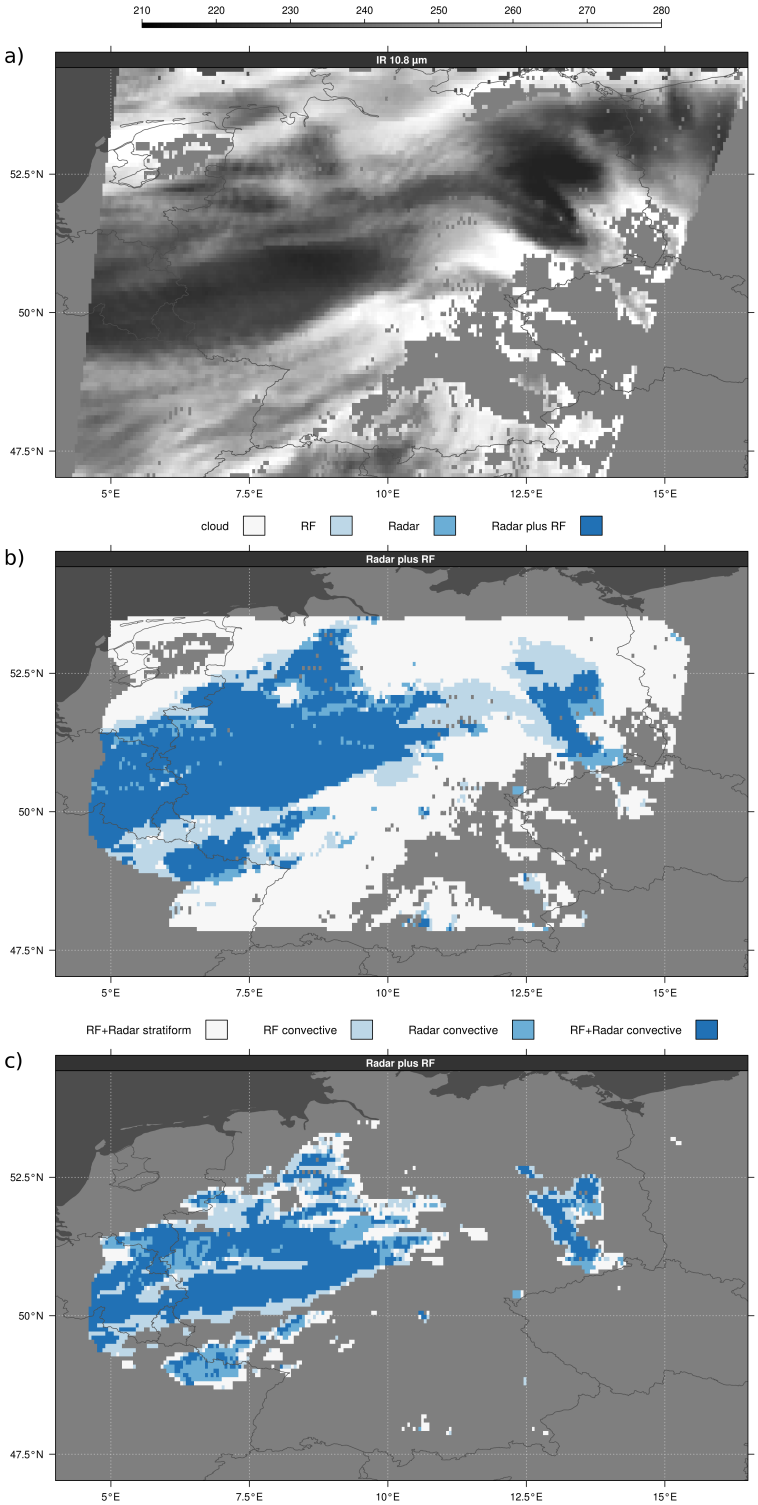

Figure 5.4: *Example for delineation of raining cloud areas and separation between convective and advective-stratiform precipitation process areas for scene from 15 August 2010 14:00 UTC. (a) IR image, (b) rain area and (c) precipitation processes delineated by RF in comparison to the radar data.*

5.4.3.2 Precipitation process separation

The equivalent performance evaluation for the precipitation process separation is shown in the right column of figure 5.3. Similar to the rain area detection, daytime performance is again best and there is an evident difference between summer and winter seasons, with summer being predicted better than winter. In comparison to area detection, this seasonal discrepancy is more pronounced. In general, the performance for process separation is slightly worse than that for area identification.

A look at the AS-period results shows that the bias again indicates very good agreement between the predictions and observations. In comparison to rain area detection, the POD is generally slightly decreased, yet shows acceptable values for all times of the day. As a consequence, PFD is increased showing values of 0.37 (day-AS), 0.42 (twilight-AS) and 0.44 (night-AS). The slightly decreased prediction skill for process separation is consistently reflected in all performance skills with FAR showing slightly elevated values, while CSI, ETS, HSS and HKD are slightly reduced.

One of the main reasons for the reduced ability to properly separate processes during winter is the reduced convective activity during this season. This leads to a significantly reduced amount of potentially detectable pixels, which in turn distorts the distribution of potential hits and misses underlying the statistics. Furthermore, the reduced spatial extent of convective areas increases potential misalignments between satellite and radar pixels due to their differing viewing geometries. Other potential sources of pixel displacements stem from the pre-processing, namely the spatial aggregation and reprojection of the radar data. Finally, wind drift of rain drops is also an ever prominent issue for satellite-based rain retrievals.

Figure 5.4 c shows an example of the convective precipitation area detected by RF in comparison to the ground-based radar data. As can be seen, the majority of the convective precipitation area detected by the radar is also predicted by RF. But it is also noticeable that RF tends to overestimate the convective precipitation area by quite some margin.

5.4.3.3 Rainfall rate assignment

Evaluation statistics for rainfall rate predictions are shown in table 5.7 and figure 5.5. Comparisons of RR_{Obs} and RR_{Pred} are made on a pixel basis for 311 (27) daytime, 305 (95) twilight and 117 (112) night-time scenes during AS (OM). The exact amount of pixel pairs available for the respective periods is shown in table 5.7. In general, the RF model is able to predict rainfall rates rather accurately. Especially, prediction of the central tendency is very accurate, whereas upper (lower) extremes tend to be under-estimated (over-estimated). This is due to the inherent RF characteristic of model averaging which results in an underestimation of extremes. Similar to the evaluation studies presented before,

seasonal performance varies with AS periods being generally predicted more accurately than OM periods. In contrast to the area and process delineation, rainfall rate assignment shows less diurnal variability. Performance during twilight is slightly worse, whereas predictions during day and night tend to be equally accurate. In general, a slightly positive mean error (ME) during twilight and night-time indicates over-estimation, whereas daytime precipitation seems to be slightly under-estimated, regardless of season. Temporal aggregation of rainfall rates leads to significantly better results with a significant increase in mean correlation between predictions and observations with a concurrent reduction in the observed spread. Aggregation also leads to increased ability to predict maximum rainfall more accurately. On the other hand, error scores tend to increase with higher aggregation times. A detailed description of the temporal aggregation strategies can be found in Kühnlein et al. (2014a).

An example of the spatial distribution of RR_{Obs} and RR_{Pred} is shown in figure 5.6. As can be seen in figure 5.6 a, for higher RR_{Obs} (>10 mm/h), the corresponding RR_{Pred} are predominantly smaller. On the other hand, the direct surroundings of high rainfall areas are slightly overestimated. With decreased temporal resolution, it is possible to better reproduce the observed rainfall rate (figure 5.6 b and 5.6 c). However, it is still not possible to capture very high rainfall rates (>20 mm /3hr or /8hr).

5.4.4 Overall performance of the rainfall retrieval technique

As has been shown in the above sections, RF is able to produce reliable results when each step of the rainfall retrieval is evaluated separately. Hereby, the results presented in section 5.4.3.3 can be considered as the upper limit of rainfall rate predictability of the RF model in its current form.

To assess the overall performance of the RF-based rainfall rate retrieval the identified RF models for each retrieval step need to be combined and the final product be evaluated. In this approach, all uncertainties/errors arising from each prediction step will propagate through the models and accumulate in the final rainfall rate. Validation of the rainfall rates in this integrated approach will be only possible for areas that were classified as raining by both, RF and radar. Similar to section 5.4.3.3 the results are presented both in tabular form (for all pixels) and as box and whisker plots (for a scene-by-scene comparison) in table 5.8 and figure 5.7. Note, that the scenes used for evaluation of the overall performance are not equal to those used for the isolated rainfall rate validation in the previous section.

In comparison to section 5.4.3.3, the performance of the integrated model is reduced to a high degree. Most obviously, the correlation between predictions and observations is decreased (e. g. Rsq (day-AS, 8-h scenes) is reduced from 0.72 to 0.38). The RF now shows a tendency to over-estimate rainfall rates as indicated by generally large positive mean errors ranging from 0.0 to 1.5.

	Time period (tp)	Predicted 1-h	Observed 1-h	Predicted 3-h	Observed 3-h	Predicted 8-h	Observed 8-h
day-AS	Min [mm/tp]	0.25	0.10	0.25	0.10	0.25	0.10
	Max [mm/tp]	25.51	56.80	32.41	68.58	49.58	79.58
	Median [mm/tp]	0.90	1.10	1.59	1.65	2.70	2.35
	Mean [mm/tp]	1.69	1.81	2.48	2.63	3.57	3.80
	Std [mm/tp]	1.30	2.27	2.22	3.19	3.35	4.46
	Rsq	0.49		0.67		0.76	
	RMSE [mm/tp]	1.80		2.09		2.57	
	MAE [mm/tp]	0.77		0.93		1.18	
	ME [mm/tp]	-0.1		-0.14		-0.23	
	N	807441		454586		374901	
day-OM	Min [mm/tp]	0.23	0.10	0.23	0.10	0.24	0.10
	Max [mm/tp]	6.22	26.95	10.03	38.38	15.18	38.38
	Median [mm/tp]	0.76	1.00	1.08	1.26	0.97	1.09
	Mean [mm/tp]	1.15	1.28	2.00	2.30	1.82	2.05
	Std [mm/tp]	0.88	1.35	1.77	2.56	1.85	2.49
	Rsq	0.52		0.68		0.75	
	RMSE [mm/tp]	0.95		1.52		1.30	
	MAE [mm/tp]	0.52		0.75		0.67	
	ME [mm/tp]	-0.13		-0.30		-0.23	
	N	65720		37824		22954	
twilight-AS	Min [mm/tp]	0.20	0.10	0.20	0.10	0.20	0.10
	Max [mm/tp]	27.16	67.35	42.59	90.58	46.41	117.29
	Median [mm/tp]	0.97	1.29	2.45	2.08	2.98	2.41
	Mean [mm/tp]	2.19	2.15	3.55	3.50	4.19	4.12
	Std [mm/tp]	1.99	2.53	3.43	4.18	4.31	5.02
	Rsq	0.43		0.59		0.69	
	RMSE [mm/tp]	1.94		2.68		2.82	
	MAE [mm/tp]	0.99		1.35		1.45	
	ME [mm/tp]	0.04		0.05		0.07	
	N	752479		324961		387560	
twilight-OM	Min [mm/tp]	0.27	0.10	0.31	0.10	0.31	0.10
	Max [mm/tp]	11.39	38.29	22.31	45.48	23.99	46.79
	Median [mm/tp]	0.77	1.00	0.96	1.19	1.20	1.25
	Mean [mm/tp]	1.21	1.25	1.87	1.98	2.03	2.10
	Std [mm/tp]	1.11	1.38	1.78	2.44	2.11	2.68
	Rsq	0.36		0.57		0.59	
	RMSE [mm/tp]	1.14		1.61		1.72	
	MAE [mm/tp]	0.58		0.76		0.78	
	ME [mm/tp]	-0.05		-0.11		-0.08	
	N	197429		111350		80162	
night-AS	Min [mm/tp]	0.28	0.10	0.31	0.10	0.28	0.10
	Max [mm/tp]	18.04	40.67	20.36	45.15	25.06	49.62
	Median [mm/tp]	0.96	1.21	1.85	1.95	2.70	2.25
	Mean [mm/tp]	2.04	2.01	3.15	3.11	3.79	3.72
	Std [mm/tp]	1.69	2.21	2.85	3.57	3.57	4.30
	Rsq	0.47		0.64		0.72	
	RMSE [mm/tp]	1.62		2.16		2.26	
	MAE [mm/tp]	0.89		1.17		1.27	
	ME [mm/tp]	0.03		0.04		0.07	
	N	264794		110027		134247	
night-OM	Min [mm/tp]	0.29	0.10	0.29	0.10	0.31	0.10
	Max [mm/tp]	13.36	37.13	23.38	43.71	29.63	44.71
	Median [mm/tp]	0.80	1.00	1.25	1.29	1.52	1.62
	Mean [mm/tp]	1.28	1.25	2.04	1.99	2.62	2.56
	Std [mm/tp]	1.15	1.12	2.01	1.99	2.71	2.74
	Rsq	0.48		0.68		0.77	
	RMSE [mm/tp]	0.89		1.19		1.34	
	MAE [mm/tp]	0.53		0.68		0.77	
	ME [mm/tp]	0.03		0.06		0.06	
	N	244460		127531		118510	

Table 5.7: *Results of the standard verification scores applied to the rainfall rates predicted by RF.*
The scores are based on data pairs of 311 (27) daytime, 305 (95) twilight and 117 (112) night-time precipitation scenes from 2010 during AS (OM). The values are calculated from the RR_{Obs} and RR_{Pred} data pairs extracted from the whole day-AS, day-OM, twilight-AS, twilight-OM, night-AS and night-OM data sets, respectively.

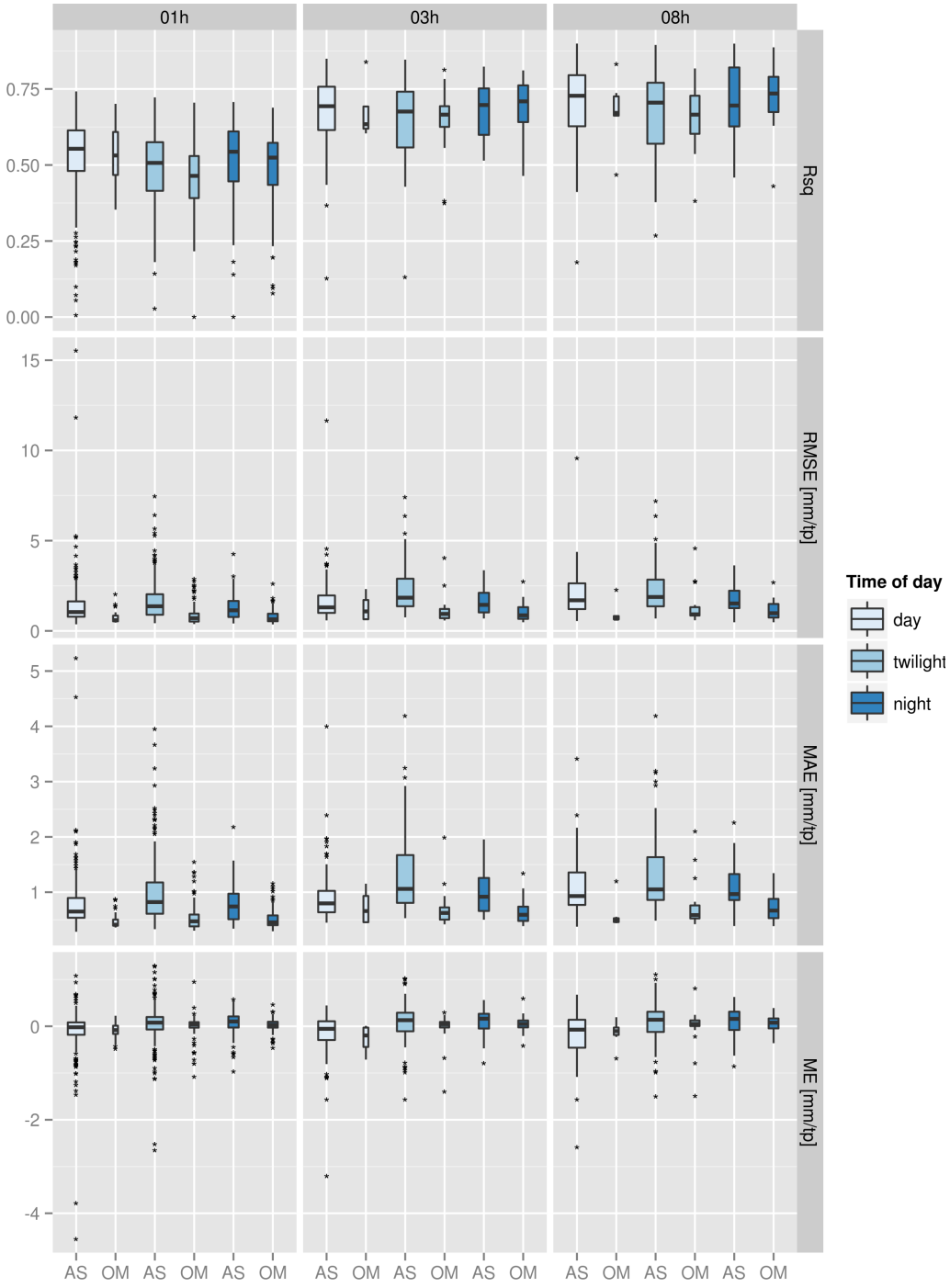

Figure 5.5: *Results of rainfall rate assignment on a scene-by-scene basis.*
The scores are calculated on a scene-by-scene basis and are based on 311 (27) daytime, 305 (95) twilight and 117 (112) night-time precipitation scenes from 2010 during AS (OM). Box and whisker plots showing distribution of standard verification scores (rows) of RR_{Obs} vs. RR_{Pred} for different aggregation times (columns) according to time of day (colours). Boxes indicate 25th, 50th and 75th percentiles. Whiskers extend to the most extreme data point within 1.5 times the interquartile range (75th – 25th percentiles). Outliers shown as stars. The abbreviation tp stands for time period.

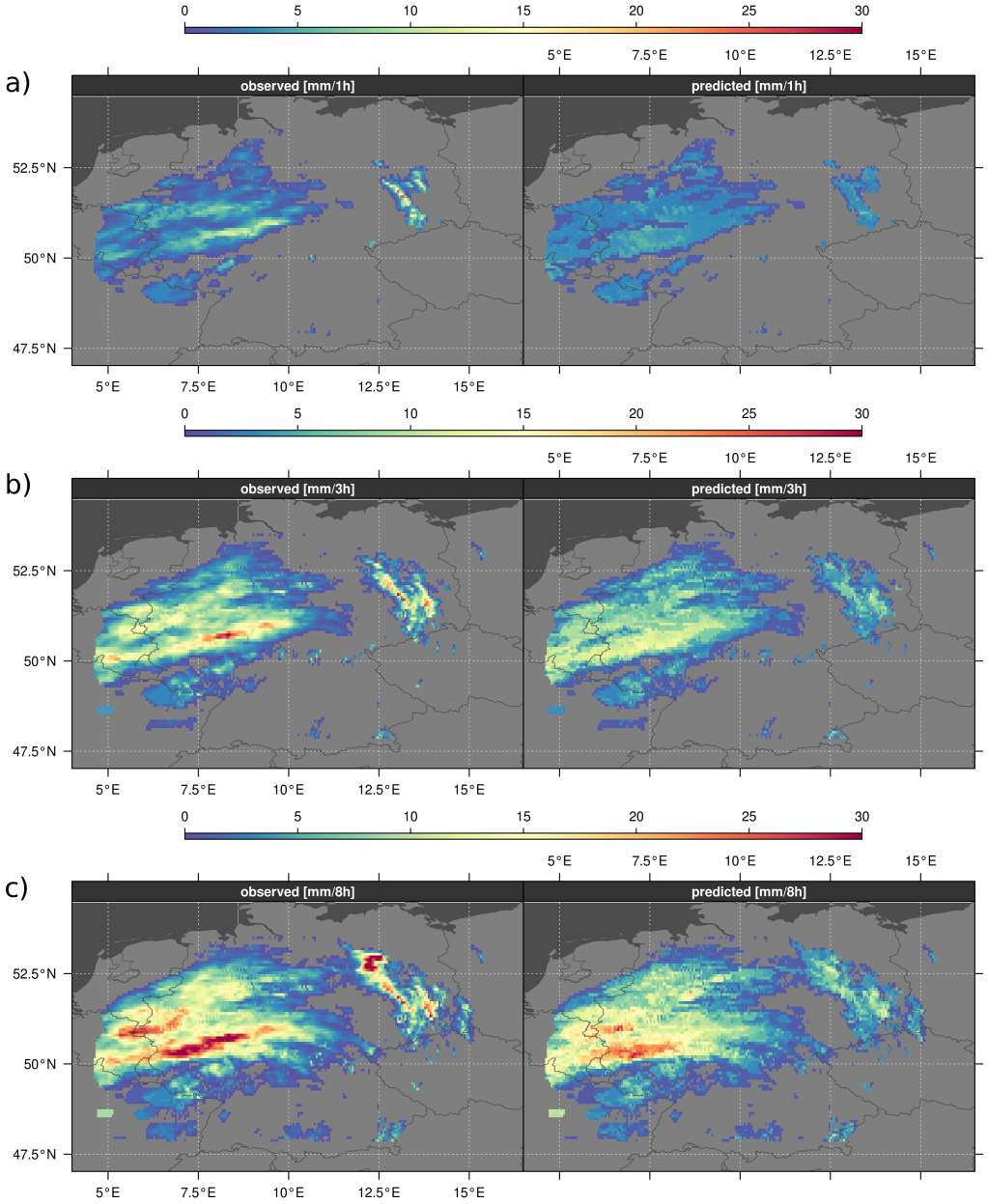

Figure 5.6: *Rainfall rates observed by radar (left) and predicted by RF model (right) during daytime. (a) rainfall rates for scene from 15 August 2010 14:00 UTC, (b) aggregated rainfall rates for scenes from 15 August 2010 13:00 UTC to 15:00 UTC and (c) all scenes of the 15 August 2010 from the daytime interval.*

This is primarily due to an increased over-estimation of the central tendencies which show much higher discrepancies between predictions and observations than before. With regard to seasonal performance, it is apparent that the overall model tends to perform better during the AS period as indicated by the decreased MAE values during this period. This is likely due to the process separation step which tends to over-estimate convective areas which in turn results in the assignment of higher rainfall rates. On the other hand, incorrectly assigned advective-stratiform process areas lead to under-estimation of precipitation rates which, all together, results in the low correlations seen in figure 5.7. Temporal aggregation is able to alleviate this problem to some degree with significantly increased coefficients of determination (Rsq as high as 0.38 (day-AS), 0.21 (day-OM), 0.24 (twilight-AS), 0.2 (twilight-OM), 0.24 (night-AS) and 0.3 (night-OM) for 8-h scenes).

The most significant problem with the integrated retrieval model seems to be the process separation. It needs to be kept in mind that for the distinction between convective and advective-stratiform processes a rainfall rate of 1.8 mm/hr is used while training the process separation model. This means, that by definition the RF model cannot assign more than this threshold rainfall rate to areas classified as advective-stratiform. On the other hand, areas classified as convective must be assigned values>1.8 mm/hr. Hence, process misclassification greatly increases rainfall rate assignment errors. The observed over-estimation of rain rates thus suggests a clear tendency to misclassify rain processes in favour of convective rainfall areas. This is especially critical during winter months as convective activity is reduced during these times resulting in lower amounts of potentially detectable pixels. A direct consequence of this is a relative distortion of the distribution of potential hits and misses underlying the statistics (evident in the huge increase of FAR). In light of this, the results at 3 hours and 8 hours temporal resolutions are very encouraging. As before, spatial distribution examples of the overall performance are also given (figure 5.8).

5.5 Summary and conclusions

In this paper, we introduced a new rainfall retrieval technique based on MSG SEVIRI data which determines rainfall rates of precipitating events in connection with extra-tropical cyclones in the mid-latitudes in a continuous manner (day, twilight and night) resulting in a 24-hour estimation at high temporal resolution. The approach utilises satellite-derived information on cloud top height, cloud top temperature, cloud phase and cloud water path and uses a machine learning algorithm (RF) to develop the predictions. The technique is realized in three steps: (i) Precipitating cloud areas are identified. (ii) The areas are separated into convective and advective-stratiform precipitating areas. (iii) Rainfall rates are assigned separately to the convective and advective-stratiform precipitating areas. Additionally accounting for seasonal differences in precipitation processes

	Time period (tp)	Predicted 1-h	Observed 1-h	Predicted 3-h	Observed 3-h	Predicted 8-h	Observed 8-h
day-AS	Min [mm/tp]	0.32	0.10	0.32	0.10	0.32	0.10
	Max [mm/tp]	19.07	66.80	36.63	68.58	42.77	128.58
	Median [mm/tp]	3.08	1.13	3.60	1.71	4.10	2.26
	Mean [mm/tp]	2.67	1.81	4.08	2.66	5.48	3.68
	Std [mm/tp]	1.72	2.29	3.19	3.22	4.79	4.33
	Rsq	0.12		0.35		0.51	
	RMSE [mm/tp]	2.68		3.54		4.39	
	MAE [mm/tp]	1.83		2.47		3.03	
	ME [mm/tp]	0.86		1.42		1.80	
	N	659499		367212		318138	
day-OM	Min [mm/tp]	0.30	0.10	0.30	0.10	0.30	0.10
	Max [mm/tp]	19.34	19.25	19.51	19.25	19.51	21.54
	Median [mm/tp]	2.36	1.00	2.96	1.29	2.65	1.14
	Mean [mm/tp]	2.14	1.31	3.47	2.10	3.33	2.05
	Std [mm/tp]	1.64	1.31	2.72	2.13	3.15	2.32
	Rsq	0.11		0.34		0.45	
	RMSE [mm/tp]	1.91		2.65		2.68	
	MAE [mm/tp]	1.36		1.91		1.85	
	ME [mm/tp]	0.82		1.37		1.28	
	N	53765		32091		25225	
twilight-AS	Min [mm/tp]	0.33	0.10	0.33	0.10	0.33	0.10
	Max [mm/tp]	19.59	72.20	39.16	78.55	54.44	95.43
	Median [mm/tp]	3.59	1.50	4.28	2.14	4.50	2.38
	Mean [mm/tp]	3.42	2.36	5.15	3.60	5.88	4.06
	Std [mm/tp]	2.11	2.73	3.96	4.30	5.04	4.92
	Rsq	0.07		0.29		0.42	
	RMSE [mm/tp]	3.15		4.26		4.57	
	MAE [mm/tp]	2.31		3.05		3.27	
	ME [mm/tp]	1.06		1.55		1.82	
	N	478816		224309		274863	
twilight-OM	Min [mm/tp]	0.33	0.10	0.35	0.10	0.33	0.10
	Max [mm/tp]	16.44	28.29	29.90	40.26	34.74	44.83
	Median [mm/tp]	0.87	1.00	0.96	1.05	1.04	1.05
	Mean [mm/tp]	1.77	1.40	2.29	1.91	2.40	1.93
	Std [mm/tp]	1.79	1.67	3.03	2.64	3.02	2.76
	Rsq	0.21		0.50		0.49	
	RMSE [mm/tp]	1.84		2.26		2.33	
	MAE [mm/tp]	1.16		1.35		1.40	
	ME [mm/tp]	0.37		0.37		0.47	
	N	86294		59982		43738	
night-AS	Min [mm/tp]	0.34	0.10	0.38	0.10	0.34	0.10
	Max [mm/tp]	15.01	40.67	27.28	67.91	37.65	67.91
	Median [mm/tp]	2.86	1.38	3.07	1.90	3.63	2.18
	Mean [mm/tp]	2.91	2.17	3.89	3.01	4.67	3.49
	Std [mm/tp]	2.33	2.34	3.66	3.44	4.33	3.82
	Rsq	0.07		0.25		0.39	
	RMSE [mm/tp]	2.94		3.65		4.17	
	MAE [mm/tp]	2.12		2.50		2.94	
	ME [mm/tp]	0.74		0.87		1.19	
	N	176960		104654		78688	
night-OM	Min [mm/tp]	0.34	0.10	0.34	0.10	0.34	0.10
	Max [mm/tp]	10.02	37.13	14.82	38.38	26.99	38.67
	Median [mm/tp]	0.99	1.00	2.30	1.21	2.83	1.52
	Mean [mm/tp]	1.85	1.29	2.73	1.93	3.47	2.42
	Std [mm/tp]	1.30	1.23	2.15	1.96	3.01	2.58
	Rsq	0.02		0.19		0.33	
	RMSE [mm/tp]	1.74		2.33		2.81	
	MAE [mm/tp]	1.24		1.62		1.93	
	ME [mm/tp]	0.55		0.81		1.05	
	N	154844		87763		81803	

Table 5.8: *Overall performance of the proposed rainfall retrieval. Results of the standard verification scores applied to the rainfall rates predicted by RF.*
The scores are based on data pairs of 311 (27) daytime, 305 (95) twilight and 117 (112) night-time precipitation scenes from 2010 during AS (OM). The values are calculated from the RR_{Obs} and RR_{Pred} data pairs extracted from the whole day-AS, day-OM, twilight-AS, twilight-OM, night-AS and night-OM data sets, respectively.

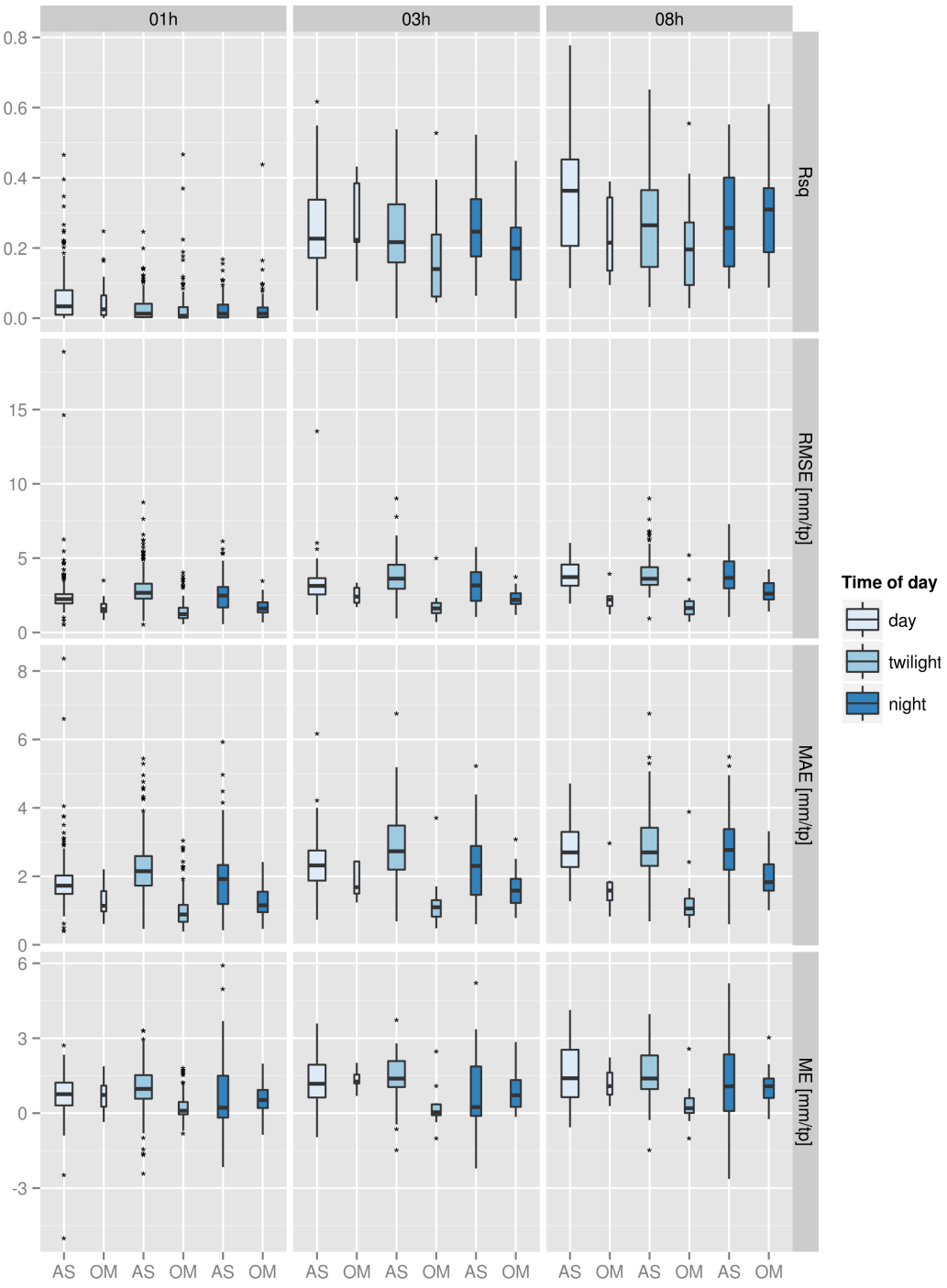

Figure 5.7: *As in figure 5.5, but for the rainfall retrieval technique.*

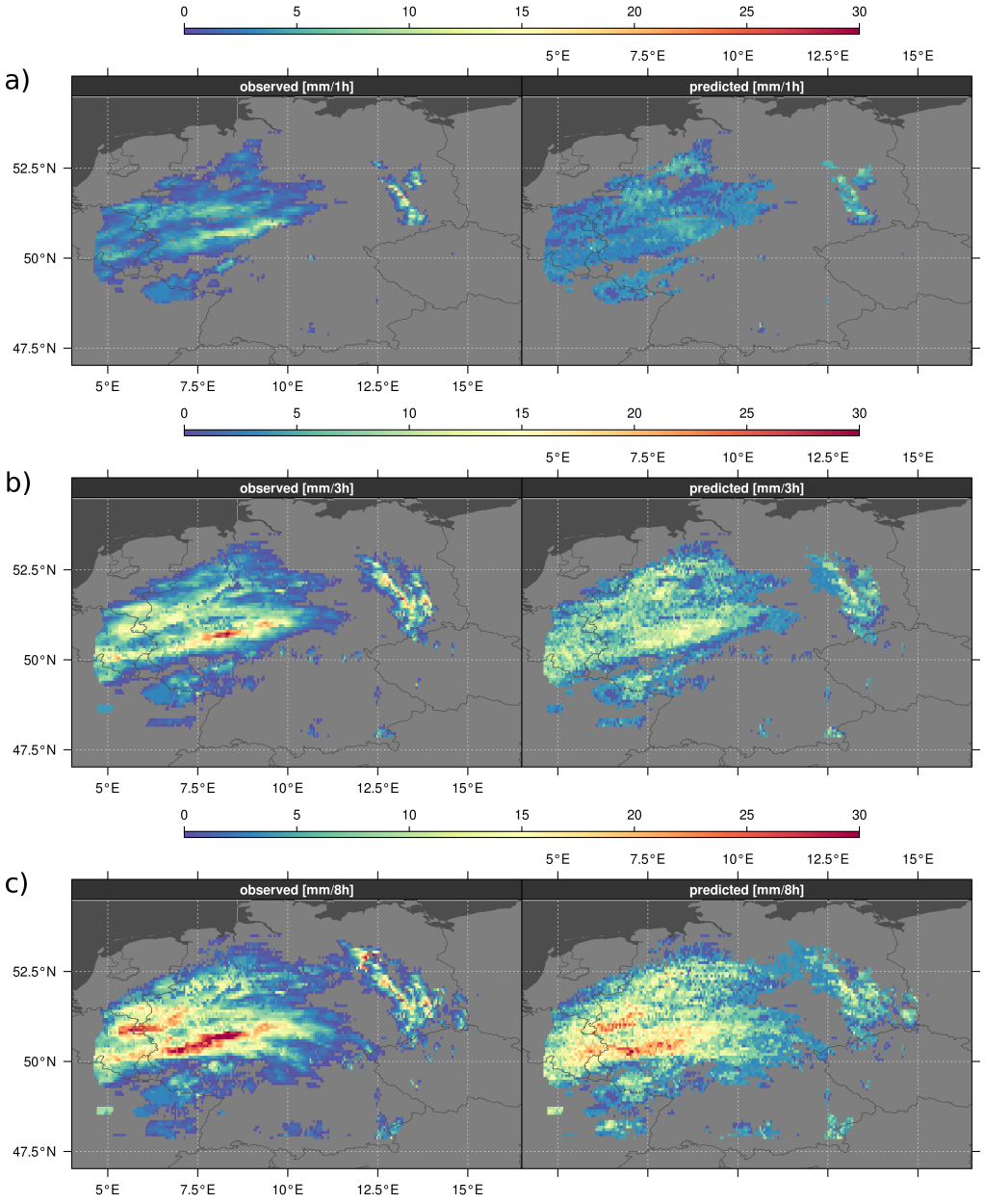

Figure 5.8: *Rainfall rates observed by radar (left) and predicted by RF model (right) during daytime. (a) rainfall rates for scene from 15 August 2010 14:00 UTC, (b) aggregated rainfall rates for scenes from 15 August 2010 13:00 UTC to 15:00 UTC and (c) all scenes of the 15 August 2010 from the daytime interval. Rainfall rates are the result of the rainfall retrieval which combines rain area detection, process separation and rainfall rate assignment.*

in the mid-latitudes, the retrieval is composed of 24 separate RF models.

As the proposed retrieval is an iterative approach, it is possible to apply each of the 24 models separately or to use the complete set together as a stand-alone rainfall retrieval. Here, validation studies were carried out for each individual step (rain area, rain process and rainfall rate) as well as for the overall procedure. For the development and validation of the technique, precipitation events from 2010 were chosen. The radar-based RADOLAN RW product which provide area-wide gauge-adjusted hourly precipitation information was used for the tuning, training and validation of the different models.

Regarding the rain area detection, the model produces very good results. A general tendency towards elevated prediction skill during summer months is apparent. This is likely a consequence of the generally higher precipitation area percentage during this season. With respect to diurnal performance, there is a general pattern of daytime rain area being predicted best by the model. The higher number of predictor variables as well as the higher information content about the CWP inherent in the VIS and NIR channels is most likely the reason for this.

The results of the detection of convective precipitation areas are also quite promising and reveal the same seasonal and diurnal patterns. Again, the area percentage of the respective area seems to have an influence on the prediction results such that the better performance during summer is likely to be a consequence of the higher area percentage of convective precipitation areas. In general, the higher area fraction of continuous convective precipitation areas makes potential spatial misalignments between radar and satellite data, which is a well known issue for convective clouds, less likely. As before, the higher information content about the CWP leads to a better performance during the day.

With regard to rainfall rate assignment, the RF models exhibit similar performance patterns as before with the model for summer being slightly superior. Also, daytime predictions are generally more accurate, yet it is noteworthy how well the model is able to predict rainfall rates during night-time and especially twilight periods, given the reduced amount of spectral information. Additionally, the achieved prediction accuracy at temporal resolutions of one hour to several hours is also very encouraging.

When comparing the rainfall rate prediction performance of the overall procedure with that of the stand-alone rainfall rate assignment, it is expectable that the former is less accurate than the latter, even though the general patterns, such as enhanced predictability of AS-day periods still hold. The less accuracy is particularly true at high temporal resolutions of one hour. Temporal aggregation to 3/8 hours intervals significantly increases prediction performance during all times of the day. Concerning the the considerable problems of existing optical retrievals particularly during twilight, these results reveal a clear improvement and offer the possibility for 24 hours rainfall rate estimation. The most crucial

problem of the integrated retrieval model seems to be the process separation. By definition a RF model cannot predict beyond the identified range during model training. This can lead to difference between observations and predictions due to misclassified process areas. As the classification threshold for process separation is based upon the rainfall rate itself (1.8 mm/hr), it is impossible to assign realistic rainfall rates for misclassified areas. Hence, process misclassification greatly increases rainfall rate assignment errors. In light of this, the results at 3 hours and 8 hours temporal resolutions are very encouraging.

In summary, the presented retrieval technique shows very promising potential to predict rainfall rates at high temporal and spatial resolutions in an automated manner. This becomes especially apparent, if the final results are compared against the performance of the IR-based GOES precipitation index (GPI) as shown in Kidd et al. (2012). Amongst others, the performance of the GPI is analysed against ground-based radar data at temporal resolution of 3 hours and spatial resolution of 0.25° x 0.25°. In terms of correlation, the GPI shows a seasonal cycle with correlation coefficients ranging mainly from 0.1 to 0.4 during winter and 0.3 to 0.5 during summer. Summarising the results of the overall procedure (shown in figure 5.8) according to summer/winter season, a better agreement is indicated by correlation coefficients which lie predominantly between 0.4 and 0.6 (summer) and 0.3 and 0.6 (winter) at the same temporal resolution, but even at a higher spatial resolution (3 by 3 km at sub-satellite point). Assessing the performance of rain area detection, CSI values from 0.2 to 0.3 during the cold season indicate a better performance of the GPI during winter than summer (0.1 - 0.2). CSI values of smaller than 0.2 are only reached during twilight/night-time in winter. Especially during summer months, CSI values of 0.25 to 0.5 indicate the ability of RF to differentiate between raining and non-raining areas. This provides a very promising basis for future investigations into rainfall rate assignments based on machine learning approaches and MSG SEVIRI data which provides the sufficient spectral resolution.

Because of the deficiencies regarding the assignment of instantaneous rainfall rates at the ground, the proposed rainfall retrieval is of valuable benefit especially for now-casting purposes. In this context, the high spatial and temporal resolution provided by MSG SEVIRI is of valuable benefit. It allows a quasi-continuous detection of the spatio-temporal rainfall distribution in near-real time, which also accounts for short-time precipitation events, such as intensive convective precipitation.

Acknowledgements

The authors are grateful to the German Weather Service (DWD) for providing the RADOLAN and the EUMETSAT Earth Observation Portal (https://eoportal.eumetsat.int/) for providing the Meteosat data set. We would like to thank Jan Cermak from the University of Bochum for providing his cloud

mask algorithm, Bernhad Seeger from the University of Marburg for the fruitful co-operation in the context of the Meteosat processing scheme and Tobias Ebert and Johannes Drönner for an excellent implementation. Finally, we thank the members of the Ecosystem Informatics PhD program funded by the University of Marburg for the valuable discussions. The work of M. Kühnlein has partly been supported by the Priority Program 1374 "Biodiversity Exploratories" and the work of T. Appelhans by the Research Unit 1246 "Kilimanjaro ecosystems under global change", both funded by the German Research Foundation (DFG), see www.environmentalinformatics-marburg.de.

Chapter 6

Summary and Outlook

6 Summary and Outlook

From a scientific point of view, reliable precipitation estimates are of great value in climatology, hydrology and ecology. Moreover, knowledge of precipitation has an economic value since it plays a key role in the fields of agriculture, water engineering and risk management. Because of its great importance, the correct detection and quantification, especially at high temporal and spatial resolution, is a crucial but to a certain degree still unresolved task. Satellite imagery are the obvious data source to retrieve area-wide precipitation measurements in areas where ground-based radar networks are not available.

Caused by the poor spectral resolution of the first satellite sensor systems, many of the existing optical rainfall retrievals rely on a relationship between the cloud top temperature measured in an infrared channel and rainfall probability/intensity. The application of these methods is limited to deep convective and therefore cold clouds which is not suitable for precipitation detection in the mid-latitudes. The ongoing developments of satellite sensor systems accompanied by an improved spectral resolution enabled the development of improved rainfall retrieval techniques using optical and microphysical cloud parameters. For the first time, techniques of this kind are allowed to consider all different precipitation areas of an extra-tropical cyclone. However, these methods show still a great deficit regarding the identification of precipitating cloud areas in a continuous manner (daytime, twilight, night-time), especially during twilight. Moreover, the quantification of rainfall is still largely unsolved during all times of the day.

When faced with a very complex and non-linear cloud top to precipitation relationship and a variety of precipitation processes, machine learning algorithms might be suitable to overcome these deficits and will likely help to benefit from the potential offered by current GEO systems. With the advent of Meteosat Second Generation (MSG) Spinning-Enhanced Visible and InfraRed Imager (SEVIRI) in 2004, a geostationary system with sufficient spectral and temporal resolution has become available. The follow up mission Meteosat Third Generation is intended to be launched in 2015 (Bézy et al., 2005) which ensures the data availability and utilisation of applications developed for MSG SEVIRI for the next decades.

Hence, the aim of the present study was to develop a 24-h-technique for the process-related and quantitative estimation of precipitation in connection with extra-tropical cyclones in the mid-latitudes based on MSG SEVIRI data using the machine learning algorithm random forest.

This aim was based on the following hypothesis:

- Based on spectrally and temporally adequate MSG SEVIRI data, it is possible to devise a process-related and quantitative precipitation scheme for daytime, twilight and night-time conditions suitable for precipitating clouds in connection with extra-tropical cyclones using the machine learning algo-

rithm random forest.

The verification of the hypothesis required the development of an entirely new methodology. The algorithms and approaches needed were successfully developed and implemented within three working packages:

(WP1) The cloud property retrieval SLALOM, first developed for Terra MODIS, was successfully transferred and adapted to the specific requirements of the SEVIRI system and an extensive validation study was carried out. The cloud optical properties retrieved by SLALOM, namely cloud effective radius (a_{ef}) and cloud optical thickness (τ) that were needed for satellite-based rainfall estimation in WP2 and WP3, were compared against the well known and validated NASA MODIS cloud property product (MODIS 06) as well as the cloud optical depth product (2B-TAU) of CloudSat. The suitability of SLALOM has been shown over the North Atlantic and over the European continent (chapter 3, Kühnlein et al., 2013).

(WP2) A new 24-h-technique for rainfall rate assignment was developed for MSG SEVIRI using the machine learning algorithm random forest as fundamental prediction algorithm. Based on the precipitation processes in connection with extra-tropical cyclones, rainfall rates were assigned to advective-stratiform and convective precipitating areas by means of individual RF models. As predictor variables for the RF models satellite-based information on cloud top height, cloud top temperature, cloud phase and cloud water path were chosen. The different illumination conditions (daytime, twilight and night-time) were taken into account with a proper SEVIRI spectral channel selection as surrogates for theses cloud physical parameters. The development was realised in three steps: First, an extensive tuning study was carried out to customise each of the RF models. Secondly, the RF models were trained using the optimum model parameter values found in the tuning study. Finally, the final RF models were used to predict rainfall rates using an independent validation data set and the results were validated against co-located rainfall rates observed by the RADOLAN RW product of the DWD. The outstanding validation results during all times of the day confirmed the ability of RF as tool for the rainfall rate assignment technique from MSG SEVIRI data (chapter 4, Kühnlein et al., 2014a).

(WP3) A new coherent daytime, twilight and night-time rainfall retrieval was developed for MSG SEVIRI. The technique aims to retrieve rainfall rates for precipitation events in connection with extra-tropical cyclones in the mid-latitudes in a continuous manner resulting in a 24 hour prediction. Based on the dominant precipitation processes, the proposed rainfall retrieval consists of three steps which are applied consecutively by means of individual RF models to get the final product: (i) Identification of precipitating cloud

areas. (ii) Separation of precipitating areas into predominately convective and advective-stratiform cloud regions. (iii) Individual process-oriented assignment of rainfall rates to these cloud areas. Again, the relationship between cloud top temperature, cloud top height, cloud water path and cloud phase was used to retrieve information about precipitation and according to the illumination conditions, a suitable selection of the predictor variables were taken into account as input to the RF models (chapter 5, Kühnlein et al., 2014b).

The newly developed rainfall retrieval technique was tested in an extensive validation study over Germany using the radar-based RADOLAN RW product as reference data. This region can be regarded as sufficiently representative for mid-latitude precipitation processes. Beside the overall performance, the performance of each retrieval step was assessed separately. The validation results show reliable performance of the new technique concerning rain area detection, rain process separation as well as rainfall rate assignment during all times of the day which enables the estimation of precipitation for 24 hours of a day. Hereby, the twilight applicability of the technique as well as good rainfall rate prediction performances even on an hourly basis are particularly remarkable and set this study apart from other rainfall retrievals.

The validation results demonstrated that the hypothesis can be verified with unprecedented accuracy: The estimation of precipitation is possible for precipitating clouds in connection with extra-tropical cyclones during daytime, twilight and night-time conditions using MSG SEVIRI data and the machine learning algorithm random forest.

With the successful implementation of the working packages (see section 2.4), a solid, objective and reliable approach for the quantitative estimation of precipitation in a quasi-continuous manner applicable during all times of the day was developed which overcomes the limitations of previously existing optical rainfall retrievals. For the first time, a 24-h precipitation monitoring becomes possible for precipitating clouds of not only convective but also of advective-stratiform character, opening many areas of application. For example, these data can be used as input for data assimilation for the initialisation of numerical weather and climate models. Furthermore, quantitative information on precipitation is of great benefit in climatological studies e. g. concerning spatio-temporal distribution of precipitation events. Because of the high temporal and spatial resolution provided by MSG SEVIRI, which allows to consider short-term dynamics, the new scheme can improve the reliability of very-short change and now-casting applications, e. g. risk management, flood prediction and monitoring. Thus, the proposed rainfall retrieval is of valuable benefit for precipitation estimation in space in near-real-time.

Climatological and ecological studies were cited as motivation for the work in the introduction, as area-wide precipitation data sets at high spatio-temporal re-

solution are requested by a variety of climatological and ecological investigations. This is the reason why the newly developed rainfall retrieval will be adapted shortly to the Kilimanjaro region in Tanzania where the DFG research unit 1246 (Phase II) "Kilimanjaro ecosystems under global change: Linking biodiversity, biotic interactions and biogeochemical ecosystem processes" is located. As the retrieval does not depend on any region-specific assumptions, the transferability to other areas may be safely assumed. Due to the use of a machine learning approach, the algorithm can be applied automatically. For the training of the prediction models, precipitation measurements from the in-situ rain gauge network which has been established in the research area, will be used. The provided data sets can be considered as baseline information for the ecosystem functioning – biodiversity analysis, e. g. water budget, carbon cycle as well as remotely-sensed prediction of biodiversity. Such high-resolution observations also allow insights into locally induced rainfall dynamics (i.e. local convection) which are essential for tackling the question on the extent of existing feedbacks between land-cover change and rainfall in the area. To characterise the rainfall dynamics of the study area in a longer perspective, the rainfall retrieval technique will be applied to Meteosat-8 and 9 data for the time periods from 2002 to today. These results still cannot be interpreted as precipitation climatology, but typical characteristics of the spatial precipitation distribution should be identifiable (e. g. mountain ranges, larger valleys, basins, as well as seasonal and diurnal dynamics).

Beside the adaptation to the Kilimanjaro region, the new rainfall retrieval will be used within the BMBF Research Program SPACES (Science Partnerships for the Assessment of Complex Earth System Processes) in South Africa for the development of a hybrid method intending to merge in-situ rain gauges, the passive microwave (PMW) based retrieval technique CMORPH (Joyce et al., 2004) and data from GEO satellite systems in order to make benefit of the assets of the different techniques (e. g. PMW sensors are directly sensitive to precipitation, MSG SEVIRI is characterised by high temporal (15 minutes) and high spatial (3 by 3 km) resolution, gauges are accurate on the plot scale). The CMORPH product will be used as basis for the hybrid technique. To account for CMORPH's intensity bias, a correction function for each scene (3 hourly) will be computed using co-located in-situ rainfall measurements. In order to increase the spatial resolution of the final rainfall product, the MSG SEVIRI data will be used. As a result the new developed retrieval technique for MSG SEVIRI offers the potential to enhance the quality and reliability of the merged rainfall product. Beside grazing (feeding and grass), browsing (feeding on woody plants) and fire, rainfall is one of the potential drivers of savanna dynamics. Hence, a precipitation product of this kind is of great value for this project in order to improve the understanding of the interplay between management and savanna ecology.

Chapter 7

Zusammenfassung

7 Zusammenfassung

Aus wissenschaftlicher Sicht sind Informationen über die raum-zeitliche Verteilung von Niederschlag von großem Nutzen in der Klimatologie, Hydrologie und Ökologie. Darüber hinaus besitzt die Kenntnis über diesen Parameter einen ökonomischen Wert, da Niederschlag ein wesentliches Element in den Bereichen Landwirtschaft, Wasserbau und Risikomanagement darstellt. Aufgrund der hohen Bedeutung ist die korrekte Erfassung und Quantifizierung, besonders in hoher räumlicher und zeitlicher Auflösung, eine äußerst wichtige wenn auch immer noch zum großen Teil ungelöste Aufgabe. Für die Bereitstellung flächendeckender Datensätze stellen in Gebieten ohne bodengebundene Radarnetzwerke satellitengestütze Ansätze die einzige Möglichkeit dar.

Aufgrund der unzureichenden spektralen Auflösung der ersten Satellitensensorsysteme basiert der Großteil existierender optischer Niederschlagsverfahren ausschließlich auf einem Zusammenhang zwischen der Wolkenoberflächentemperatur in einem Infrarot-Kanal und der Regenwahrscheinlichkeit/-intensität. Die Anwendung dieser Verfahren ist auf konvektive und somit kalte Wolkenoberflächentemperaturen limitiert, was den Einsatz in den Mittelbreiten nur begrenzt zulässt. Die Weiterentwicklung der Satellitensensorsysteme und der damit verbundenen erhöhten spektralen Auflösung ermöglichte die Entwicklung verbesserter Techniken. Diese Verfahren beruhen auf der Verwendung optischer und mikrophysikalischer Wolkenparameter und ermöglichen zum ersten Mal die Betrachtung aller Niederschlagsgebiete innerhalb eines außertropischen Zyklons. Doch auch diese Methoden weisen immer noch große Defizite bei der kontinuierlichen Identifikation der Niederschlagsfläche (Tag, Dämmerung, Nacht), besonders während der Dämmerung auf. Darüber hinaus ist die Quantifizierung des Niederschlags zu allen Tageszeiten größtenteils ungelöst.

Angesichts des sehr komplexen und nicht-linearen Zusammenhangs zwischen Wolkenoberflächeneigenschaften und Niederschlag sowie der Vielfalt an Niederschlagsprozessen die abgebildet werden müssen, können maschinelle Lernverfahren helfen die genannten Defizite zu überwinden und das Potential derzeitiger geostationärer Satellitensysteme auszuschöpfen. Mit dem Meteosat Second Generation (MSG) Spinning-Enhanced Visible and InfraRed Imager (SEVIRI) steht seit 2004 ein geostationäres Satellitensystem mit ausreichend spektraler und zeitlicher Auflösung zur Verfügung. Der geplante Start der Folgemission Meteosat Third Generation 2015 stellt sicher, dass auch über die Laufzeit von MSG hinaus Daten für die entwickelten Anwendungen vorliegen.

Daraus ergab sich das Ziel der vorliegenden Arbeit: Die Entwicklung einer neuen 24-Stunden-Methode zur prozessbasierten und quantitativen Niederschlagserfassung, basierend auf MSG SEVIRI Daten und dem maschinellen Lernverfahren Random Forest. Der Fokus lag dabei auf Niederschlagsprozessen im Zu-

sammenhang mit außertropischen Zyklonen in den Mittelbreiten.

Als Grundlage für die Untersuchungen wurde folgende These formuliert:

- Unter Verwendung von spektral und zeitlich adäquaten MSG SEVIRI Daten ist es mit Hilfe des maschinellen Lernverfahrens Random Forest möglich, ein prozessbasiertes und quantitatives Niederschlagsverfahren für Tag, Dämmerung und Nacht zu entwickeln, das alle regnenden Wolkenbereiche in Verbindung mit außertropischen Zyklonen berücksichtigen kann.

Die Untersuchung der Hypothese erforderte die Entwicklung einer neuen Methodik. Die dafür notwendigen Algorithmen und Ansätze wurden in den folgenden Arbeitspaketen erfolgreich umgesetzt:

(WP1) Das ursprünglich für Terra MODIS implementierte Verfahren zur Ableitung von Wolkeneigenschaften SLALOM wurde an die spezifischen Anforderungen des SEVIRI Systems angepasst und erfolgreich übertragen. In einer umfangreichen Validierungsstudie wurden die mit Hilfe von SLALOM abgeleiteten Wolkeneigenschaften effektiver Wolkenradius und optische Wolkendicke mit dem NASA MODIS Wolkenprodukt (MODIS 06) sowie dem optische Wolkendicke Produkt (2B-Tau) von CloudSat verglichen. Die Eignung von SLALOM konnte für eine Region im Nordatlantik sowie über dem europäischen Kontinent gezeigt werden (Kapitel 3, Kühnlein et al., 2013). Die abgeleiteten Wolkenparameter dienten als Basis für eine verbesserte Niederschlagserfassung in den Arbeitspaketen WP2 und WP3.

(WP2) Ein neues Verfahren für die Zuweisung von Niederschlagsraten wurde für MSG SEVIRI entwickelt. Hierbei diente das maschinelle Lernverfahren Random Forest als grundlegender Vorhersagealgorithmus. Basierend auf den dominierenden Niederschlagsprozessen in Verbindung mit außertropischen Zyklonen erfolgte die Zuweisung der Niederschlagsraten durch einzelne RF Modelle für bereits identifizierte advektiv-stratiforme und konvektive Regenflächen. Die Identifikation der Flächen erfolgte mit Hilfe von Radardaten. Als Prediktorvariablen für die RF Modelle wurden satellitenbasierte Informationen über die Wolkenhöhe, Wolkenoberflächentemperatur, Wolkenphase und Wolkenwasserweg herangezogen. Entsprechend der unterschiedlichen Beleuchtungsbedingungen (Tag, Dämmerung und Nacht) wurde eine passende Auswahl an SEVIRI Spektralkanälen gewählt. Die Entwicklung erfolgte in drei Schritten: Als erstes wurde eine intensive Studie zur Anpassung der einzelnen RF Modelle durchgeführt um die Modellparameter individuell anzupassen. Unter der Verwendung dieser Modellparamter erfolgte im nächsten Schritt das Training der RF Modelle. Diese RF Modelle wurden dann für die Zuweisung der Niederschlagsraten für unabhängige Validierungsdatensätze verwendet und die Ergebnisse gegen Niederschlagsraten des radarbasierten RADOLAN RW Produkts des Deut-

schen Wetterdienstes verglichen. Die hervorragenden Validierungsergebnisse zu allen Tageszeiten bestätigten die Eignung von Random Forest als Tool für die Ableitung der Niederschlagsraten von MSG SEVIRI Daten (Kapitel 4, Kühnlein et al., 2014a).

(WP3) Ein neues Niederschlagsverfahren das gleichermaßen am Tage, zur Dämmerung und nachts anwendbar ist, wurde für MSG SEVIRI entwickelt. Ziel des Verfahrens ist es Niederschlagsraten für Niederschlagsereignisse in Verbindung mit außertropischen Zyklonen in den Mittelbreiten kontinuierlich zu bestimmen, so dass eine 24 Stunden Vorhersage erzielt wird. Basierend auf den dominanten Niederschlagsprozessen besteht das vorgestellte Verfahren aus drei Schritten, die mit Hilfe individueller Random Forest Modelle nacheinander ausgeführt werden: (i) Identifikation regnender Wolkenbereiche. (ii) Aufteilung dieser Wolkenbereiche in konvektiv und advektiv-stratiform dominierte Wolkenbereiche. (iii) Prozessbasierte Zuweisung der Niederschlagsraten. Wie bereits in WP2 wird der Zusammenhang zwischen satellitenbasierten Informationen über Wolkenhöhe, Wolkenoberflächentemperatur, Wolkenphase und Wolkenwasserweg zur Ableitung der Niederschlagsinformationen verwendet und entsprechend der Beleuchtungsbedingungen eine passende Auswahl an Prediktorvariablen als Inputvariablen für die Random Forest Modelle gewählt (Kapitel 5, Kühnlein et al., 2014b).

Das neu entwickelte Niederschlagsverfahren wurde mit Hilfe des radargestützten RADOLAN RW Produkts im Rahmen einer umfangreichen Studie über Deutschland validiert. Dieses Gebiet kann als ausreichend repräsentativ für Niederschlagsprozesse in den Mittelbreiten angesehen werden. Darüber hinaus modifiziert die naturräumliche Gliederung die idealtypischen Niederschlagsprozesse (vgl. Kapitel 2.1), so dass die Übertragbarkeit eines funktionsfähigen Verfahrens auf Regionen, die ebenfalls im Einzugsgebiet außertropischer Zyklonen liegen, angenommen werden kann. Neben dem Gesamtverfahren wurden die Ergebnisse der einzelnen Verfahrensschritte untersucht. Die Ergebnisse der Validierungsstudie zeigten eine überzeugende Performanz der neuen Technik sowohl hinsichtlich der identifizierten Niederschlagsfläche, der differenzierten Niederschlagsprozesse sowie der abgeleiteten Niederschlagsraten zu allen Tageszeiten. Dies ermöglicht die Beobachtung und Erfassung des Niederschlags 24 Stunden am Tag. Hervorzuheben sind an dieser Stelle die Anwendbarkeit während der Dämmerung sowie die guten Niederschlagsvorhersagen auf stündlicher Basis.

Auf Basis der Validierungsergebnisse kann die anfangs formulierte Hypothese mit noch nie dagewesener Genauigkeit bestätigt werden: Eine prozessbasierte und quantitative Erfassung des Niederschlags regnender Wolken in Verbindung mit außertropischen Zyklonen ist gleichermaßen während dem Tag, der Dämmerung und der Nacht unter Verwendung von MSG SEVIRI Daten und dem maschinellen

Lernverfahren Random Forest möglich.

Mit der erfolgreichen Umsetzung der genannten Arbeitspakete (vgl. Kapitel 2.4) wurde ein solides, objektives und verlässliches Verfahren zur quantitativen quasi-kontinuierlichen Beobachtung und Erfassung des Niederschlags in den Mittelbreiten entwickelt. Damit wurde ein Ansatz gefunden, der die Defizite vorheriger optischer Niderschlagsverfahren überwindet und erstmals die Niederschlagserfassung nicht nur konvektiver sondern auch advektiv-stratiform regnender Wolkenbereiche 24 Stunden am Tag ermöglicht. Hierdurch ergeben sich viele neue Anwendungsgebiete. Eine Verwendungsmöglichkeit ist im Rahmen der Datenassimilierung zur Initialisierung numerischer Wetter- und Klimamodelle zu sehen. Des Weiteren bilden quantifizierbare Niederschlagsinformationen eine wichtige Grundlage für klimatologische und hydrologische Studien wie z. B. die Untersuchung raum-zeitlicher Verteilungen von Niederschlagsereignissen. Die Verwendung des quasi-kontinuierlichen Beobachtungssystems MSG SEVIRI ermöglicht darüber hinaus die Berücksichtigung kurzzeitiger Niederschalgsdynamiken. Dies ist von großem Nutzen bei Kurzfristvorhersagen und Echtzeitanwendungen unter anderem im Risikomanagement, bei Hochwasservorhersagen und -beobachtung. Dadurch stellt die vorgestellte Technik einen wichtigen Beitrag für die flächenhafte Niederschlagserfassung in Nah- bis Echtzeit dar.

In Kürze wird das neue Niederschlagsverfahren im Rahmen der DFG Forschergruppe 1246 (Phase II) “Kilimanjaro ecosystems under global change: Linking biodiversity, biotic interactions and biogeochemical ecosystem processes” in der Region Kilimanjaro in Tansania sowie im Rahmen des BMBF Forschungsprogramms SPACES (Science Partnerships for the Assessment of Complex Earth System Processes) in Südafrika zum Einsatz kommen.

References

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, and L. E. Gumley, 1998: Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research*, **103 (D24)**, 32 141–32 157.

Adler, R. F., C. Kidd, G. Petty, M. Morissey, and H. M. Goodman, 2001: Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3). *Bulletin of the American Meteorological Society*, **82 (7)**, 1377–1396.

Adler, R. F. and R. A. Mack, 1984: Thunderstorm cloud height-rainfall rate relations for use with satellite rainfall estimation techniques. *Journal of Climate and Applied Meteorology*, **23**, 280–296.

Adler, R. F. and A. J. Negri, 1988: A Satellite Infrared Technique to Estimate Tropical Convective and Stratiform Rainfall. *Journal of Applied Meteorology*, **27**, 30–51.

Alpaydin, E., 2010: *Introduction to machine learning*. The MIT Press, Cambridge, Massachusetts, London, England, 537 pp.

Aminou, D. M. A., 2002: MSG's SEVIRI Instrument. *ESA Bulletin*, **111**, 15–17.

Amorati, R., P. P. Alberoni, V. Levizzani, and S. Nanni, 2000: IR-based satellite and radar rainfall estimates of convective storms over northern Italy. *Meteorological Applications*, **7 (1)**, 1–18.

Anagnostou, E. N., 2004: Overview of overland satellite rainfall estimation for hydro-meteorological applications. *Surveys in Geophysics*, **25**, 511–537.

Anagnostou, E. N. and W. F. Krajewski, 1999a: Real-Time Radar Rainfall Estimation. Part I: Algorithm Formulation. *Journal of Atmospheric and Oceanic Technology*, **16**, 189–197.

Anagnostou, E. N. and W. F. Krajewski, 1999b: Real-time radar rainfall estimation. Part II: Case study. *Journal of Atmospheric and Oceanic Technology*, **16**, 198–205.

Anagnostou, E. N. and C. Kummerow, 1997: Stratiform and convective classification of rainfall using SSM/I 85-GHz brightness temperature observations. *Journal of Atmospheric and Oceanic Technology*, **14**, 570–575.

Arkin, P. A., 1979: The relationship between the fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array. *Monthly Weather Review*, **107**, 1382–1387.

Arkin, P. A. and P. E. Ardanuy, 1989: Estimating climatic-scale precipitation from space: A review. *Journal of Climate*, **2**, 1229–1238.

Arkin, P. A., R. Joyce, and J. E. Janowiak, 1994: The estimation of global monthly mean rainfall using infrared satellite data: The GOES Precipitation Index (GPI). *Remote Sensing Reviews*, **11**, 107–124.

Arkin, P. A. and B. N. Meisner, 1987: The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Monthly Weather Review*, **115**, 51–74.

Arking, A. and J. D. Childs, 1985: Retrieval of cloud cover parameters from multispectral satellite images. *Journal of Applied Meteorology*, **24**, 322–333.

Ba, M. B. and A. Gruber, 2001: GOES multispectral rainfall algorithm (GMSRA). *Journal of Applied Meteorology*, **40**, 1500–1514.

Ba, M. B. and S. E. Nicholson, 1998: Analysis of Convective Activity and Its Relationship to the Rainfall over the Rift Valley Lakes of East Africa during 1983–90 Using the Meteosat Infrared Channel. *Journal of Applied Meteorology*, **37 (10)**, 1250–1264.

Barry, R. G. and R. J. Chorley, 2010: *Atmosphere, Weather and Climate*. Routledge, Oxon, 516 pp.

Bartels, H., E. Weigl, T. Reich, P. Lang, A. Wagner, O. Kohler, N. Gerlach, and MeteoSolutions GmbH, 2004: Projekt RADOLAN - Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer). Tech. rep., Deutscher Wetterdienst, 111 pp., Offenbach.

Barth, M. C. and D. B. Parsons, 1996: Microphysical processes associated with intense frontal rainbands and the effect of evaporation and melting on frontal dynamics. *Journal of the Atmospheric Sciences*, **53 (11)**, 1569–1586.

Baum, B. A. and S. Platnick, 2006: Introduction to MODIS cloud products. *Earth Science Satellite Remote Sensing. Vol.1: Science and Instruments*, J. J. Qu, W. Gao, M. Kafatos, R. E. Murphy, and V. V. Salomonson, Eds., Springer, Berlin, vol. 1 ed., chap. 5, 74–91.

Baum, B. A., P. Soulen, K. I. Strabala, M. D. King, S. A. Ackerman, W. P. Menzel, P. Yang, and A. Steven, 2000: Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS 2. Cloud thermodynamic phase. *Journal of Geophysical Research*, **105 (D9)**, 11 781–11 792.

Behrangi, A., K.-L. Hsu, B. Imam, S. Sorooshian, G. J. Huffman, and R. J. Kuligowski, 2009: PERSIANN-MSA: A Precipitation Estimation Method from Satellite-Based Multispectral Analysis. *Journal of Hydrometeorology*, **10 (6)**, 1414–1429.

Bellerby, T. J., 2004: A feature-based approach to satellite precipitation monitoring using geostationary IR imagery. *Journal of Hydrometeorology*, **5**, 910–921.

Bendix, J., 1997: Adjustment of the Convective-Stratiform Technique (CST) to estimate 1991/93 El Niño rainfall distribution in Ecuador and Peru by means of Meteosat-3 IR data. *International Journal of Remote Sensing*, **18 (6)**, 1387–1394.

Bendix, J., 2000: Precipitation dynamics in Ecuador and northern Peru during the 1991/92 El Nino: a remote sensing perspective. *International Journal of Remote Sensing*, **21 (3)**, 533–548.

Bennartz, R., P. Watts, J. F. Meirink, and R. Roebeling, 2010: Rainwater path in warm clouds derived from combined visible/near-infrared and microwave satellite observations. *Journal of Geophysical Research*, **115 (D19120)**, 1–16.

Bézy, J.-L., D. Aminou, P. Bensi, R. Stuhlman, S. Tjemkes, and A. Rodriguez, 2005: Meteosat Third Generation: The future European Geostationary Meteorological Satellite. *Esa Bulletin*, **123**, 29–32.

Boulesteix, A.-L., S. Janitza, J. Kruppa, and I. R. König, 2012: Overview of Random Forest Methodology and Practical Guidance with Emphasis on Computational Biology and Bioinformatics. Tech. Rep. 129, Department of Statistics, University of Munich, 31 pp., Munich.

Breiman, L., 1996: Bagging predictors. *Machine Learning*, **24**, 123–140.

Breiman, L., 2001: Random Forest. *Machine Learning*, **45**, 5–32.

Breiman, L. and A. Cutler, 2008: Random forests – Classification manual. URL http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone, 1984: *Classification and regression trees*. Wadsworth and Brooks, Monterey, CA.

Briem, G., 2002: Multiple classifiers applied to multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, **40 (10)**, 2291–2299.

Browning, K. A. and N. M. Roberts, 1996: Variation of frontal and precipitation structure along a cold front. *Quarterly Journal of the Royal Meteorological Society*, **122 (536)**, 1845–1872.

Bylander, T., 2002: Estimating generalization error on two-class datasets using out-of-bag estimates. *Machine Learning*, **48**, 287–297.

Capacci, D. and B. J. Conway, 2005: Delineation of precipitation areas from MODIS visible and infrared imagery with artificial neural networks. *Meteorological Applications*, **12**, 291–305.

Capacci, D. and F. Porcù, 2009: Evaluation of a Satellite Multispectral VIS–IR Daytime Statistical Rain-Rate Classifier and Comparison with Passive Microwave Rainfall Estimates. *Journal of Applied Meteorology and Climatology*, **48 (2)**, 284–300.

Cermak, J., 2006: SOFOS – A new satellite-based operational fog observation scheme. Ph.D. thesis, Philipps-University Marburg, 165 pp.

Cermak, J. and J. Bendix, 2008: A novel approach to fog/low stratus detection using Meteosat 8 data. *Atmospheric Research*, **87**, 279–292.

Chan, J. C.-W. and D. Paelinckx, 2008: Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment*, **112 (6)**, 2999–3011.

Cheng, M. and R. Brown, 1995: Delineation of precipitation areas by correlation of METEOSAT visible and infrared data with radar data. *Monthly Weather Review*, **123 (9)**, 2743–2757.

Coppola, E., D. I. F. Grimes, M. Verdecchia, and G. Visconti, 2006: Validation of improved TAMANN neural network for operational satellite-derived rainfall estimation in Africa. *American Meteorological Society*, **45 (11)**, 1557–1572.

Cutler, A. and J. R. Stevens, 2006: Random Forests for Microarrays. *Methods in Enzymology*, A. Kimmel and O. Brian, Eds., Academic Press, San Diego, Vol. 411, 422–432.

Cutler, D. R., T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, 2007: Random forests for classification in ecology. *Ecology*, **88 (11)**, 2783–2792.

Dietterich, T., 2002: Ensemble learning. *The handbook of brain theory and neural networks, Second edition*, M. Arbib, Ed., The MIT Press, Cambridge.

Dupret, G. and M. Koda, 2001: Bootstrap re-sampling for unbalanced data in supervised learning. *European Journal of Operational Research*, **134 (1)**, 141–156.

Ebert, E. E., 2002: Verifying satellite precipitation estimates for weather and hydrological applications. *Proceedings of the 1st intl. Precipitation Working Group (IPWG) Workshop*, IPWG, Ed., Madrid, Spain, 23-27 September 2002, IPWG Workshop Report, Vol. 1, 10.

Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bulletin of the American Meteorological Society*, **88 (1)**, 47–64.

EUMETSAT, 2013: Meteosat Third Generation (MTG) will see the launch of four new satellites from 2018. URL http://www.eumetsat.int/website/home/Satellites/FutureSatellites/MeteosatThirdGeneration/index.html.

Feidas, H. and A. Giannakos, 2010: Identifying precipitating clouds in Greece using multispectral infrared Meteosat Second Generation satellite data. *Theoretical and Applied Climatology*, **104 (1-2)**, 25–42.

Feidas, H. and A. Giannakos, 2012: Classifying convective and stratiform rain using multispectral infrared Meteosat Second Generation satellite data. *Theoretical and Applied Climatology*, **108 (3-4)**, 613–630.

Feijt, A., D. Jolivet, R. Koelemeijer, and H. Deneke, 2004: Recent improvements to LWP retrievals from AVHRR. *Atmospheric Research*, **72**, 3–15.

Ferraro, R. R., 2007: Past, present and future of microwave operational rainfall algorithms. *Measuring precipitation from space*, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Springer Netherlands, Dordrecht, chap. 15, 189–198.

Ferreira, F., P. Amayenc, S. Oury, and J. Testud, 2001: Study and test of improved rain estimates from the TRMM precipitation radar. *Journal of Applied Meteorology*, **40**, 1878–1899.

Foody, G., 1995: Land cover classification by an artificial neural network with ancillary information. *International Journal of Geographical Information Systems*, **9 (5)**, 527–542.

Friedl, M. A. and C. E. Brodley, 1997: Decision tree classification of land cover from remotely sensed data. *Remote Sensing of Environment*, **61 (3)**, 399–409.

Friedl, M. A., C. E. Brodley, and A. H. Strahler, 1999: Maximizing land cover classification accuracies produced by decision trees at continental to global scales. *IEEE Transactions on Geoscience and Remote Sensing*, **37 (2)**, 969–977.

Früh, B., J. Bendix, T. Nauss, M. Paulat, A. Pfeiffer, J. W. Schipper, B. Thies, and H. Wernli, 2007: Verification of precipitation from regional climate simulations and remote-sensing observations with respect to ground-based observations in the upper Danube catchment. *Meteorologische Zeitschrift*, **16 (3)**, 275–293.

Germogenova, T. A., 1963: Some formulas to solve the transfer equation in the plane layer problem. *Spectroscopy of Scattering Media. Academy of Sciences of BSSR*, 36–41.

Ghimire, B., J. Rogan, and J. Miller, 2010: Contextual land-cover classification: incorporating spatial dependence in land-cover classification models using random forests and the Getis statistic. *Remote Sensing Letters*, **1 (1)**, 45–54.

Giannakos, A. and H. Feidas, 2013: Classification of convective and stratiform rain based on the spectral and textural features of Meteosat Second Generation infrared data. *Theoretical and Applied Climatology*, **113 (3-4)**, 495–510.

Goldstein, B. A., E. C. Polley, and F. B. S. Briggs, 2011: Random forests for genetic association studies. *Statistical applications in genetics and molecular biology*, **10 (1)**, 32.

Govaerts, Y. and M. Clerici, 2004: Evaluation of radiative transfer simulations over bright desert calibration sites. *IEEE Transactions on Geoscience and Remote Sensing*, **42**, 176–187.

Greuell, W. and R. A. Roebeling, 2009: Toward a standard procedure for validation of satellite-derived cloud liquid water path: A study with SEVIRI data. *Journal of Applied Meteorology*, **48**, 1575–1590.

Grimes, D., E. Coppola, M. Verdecchia, and G. Visconti, 2003: A neural network approach to real-time rainfall estimation for Africa using satellite data. *Journal of Hydrometeorology*, **4**, 1119–1133.

Grimes, D., E. Pardo-Igúzquiza, and R. Bonifacio, 1999: Optimal areal rainfall estimation using raingauges and satellite data. *Journal of Hydrology*, **222 (1-4)**, 93–108.

Guenther, B., G. D. Godden, X. Xiong, E. J. Knight, S.-Y. Qiu, H. Montgomery, M. M. Hopkins, M. G. Khayat, and Z. Hao, 1998: Prelaunch algorithm and data format for the Level 1 calibration products for the EOS-AM1 Moderate Resolution Imaging Spectroradiometer (MODIS). *IEEE Transactions on Geoscience and Remote Sensing*, **36 (4)**, 1142–1151.

Guo, L., N. Chehata, C. Mallet, and S. Boukir, 2011: Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66 (1)**, 56–66.

Han, Q., W. B. Rossow, and A. A. Lacis, 1994: Near-global survey of effective droplet radii in liquid water clouds using ISCCP data. *Journal of Climate*, **7**, 465–497.

Hansen, M., R. Dubayah, and R. Defries, 1996: Classification trees: an alternative to traditional land cover classifiers. *International Journal of Remote Sensing*, **17 (5)**, 1075–1081.

Heinemann, G., C. Reudenbach, E. Heuel, J. Bendix, and M. Winiger, 2001: Investigation of summertime convective rainfall in Western Europe based on a synergy of remote sensing data and numerical models. *Meteorology and Atmospheric Physics*, **76 (1-4)**, 23–41.

Heymsfield, A. J., 1977: Precipitation development in stratiform ice clouds: A microphysical and dynamical study. *Journal of the Atmospheric Sciences*, **34**, 367–381.

Hollinger, J., J. Peirce, and G. Poe, 1990: SSM/I instrument evaluation. *IEEE Transactions on Geoscience and Remote Sensing*, **28 (5)**, 781–790.

Hong, Y., K.-L. Hsu, S. Sorooshian, and G. Hiaogang, 2004: Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, **43**, 1834–1852.

Houze, R. A., 1993: *Cloud dynamics*. Academic Press, San Diego.

Hsu, K., X. Gao, S. Soorshian, and H.-V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, **36**, 1176–1190.

Hsu, K. K.-L., X. Gao, and S. Sorooshian, 2002: Rainfall estimation using cloud texture classification mapping. *Proceedings of the 1st intl. Precipitation Working Group (IPWG) Workshop*, IPWG, Ed., Madrid, Spain, 23-27 September 2002, 6, September.

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. E. Janowiak, A. McNab, B. Rudolf, U. Schneider, and J. Janoviak, 1997: The global precipitation climatology project (GPCP) combined precipitation dataset. *Bulletin of the American Meteorological Society*, **78 (1)**, 5–20.

Iguchi, T., T. Kou, R. Meneghini, J. Awaka, and K. Okamoto, 2000: Rain-profiling algorithm for the TRMM precipitation radar. *Journal of Applied Meteorology*, **39**, 2038–2052.

Im, E., W. Chialin, and S. L. Durden, 2005: Cloud profiling radar for the cloudsat mission. *IEEE International Radar Conference*, Arlington, United States, 9-12 May 2005, Vol. 20, 483–486.

Inoue, T., 1985: On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the 10-μm window region. *Journal of the Meteorological Society of Japan*, **63**, 88–99.

Islam, T., M. a. Rico-Ramirez, and D. Han, 2012a: Tree-based genetic programming approach to infer microphysical parameters of the DSDs from the polarization diversity measurements. *Computers & Geosciences*, **48**, 20–30.

Islam, T., M. a. Rico-Ramirez, D. Han, and P. K. Srivastava, 2012b: Artificial intelligence techniques for clutter identification with polarimetric radar signatures. *Atmospheric Research*, **109-110**, 95–113.

Islam, T., M. a. Rico-Ramirez, P. K. Srivastava, and Q. Dai, 2014a: Non-parametric rain/no rain screening method for satellite-borne passive microwave radiometers at 19–85 GHz channels with the Random Forests algorithm. *International Journal of Remote Sensing*, **35 (9)**, 3254–3267.

Islam, T., P. K. Srivastava, M. a. Rico-Ramirez, Q. Dai, D. Han, and M. Gupta, 2014b: An exploratory investigation of an adaptive neuro fuzzy inference system (ANFIS) for estimating hydrometeors from TRMM/TMI in synergy with TRMM/PR. *Atmospheric Research*, **145-146**, 57–68.

Joyce, R. J., J. E. Janowiak, P. Arkin, P. A. Arking, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, **5**, 487–503.

Kawamoto, K., T. Nakajima, and T. Y. Nakajima, 2001: A global determination of cloud microphysics with AVHRR remote sensing. *Journal of Climate*, **14**, 2054–2068.

Kerrache, M. and J. Schmetz, 1988: A precipitation index from the ESOC climatological data set. *ESA Journal*, **12**, 379–383.

Kidd, C., 2001: Satellite rainfall climatology: a review. *International Journal of Climatology*, **21 (9)**, 1041–1066.

Kidd, C., P. Bauer, J. Turk, G. J. Huffman, R. Joyce, K.-L. Hsu, and D. Braithwaite, 2012: Intercomparison of High-Resolution Precipitation Products over Northwest Europe. *Journal of Hydrometeorology*, **13 (1)**, 67–83.

Kidd, C. and G. Huffman, 2011: Review Global precipitation measurement. *Meteorological Applications*, **18**, 334–353.

Kidd, C. and V. Levizzani, 2011: Status of satellite precipitation retrievals. *Hydrology and Earth System Sciences*, **15 (4)**, 1109–1116.

Kidd, C., V. Levizzani, J. Turk, R. Ferraro, and V. Turk, 2009: Satellite precipitation measurements for water resource monitoring. *Journal of the American Water Resource Association*, **45 (3)**, 567–579.

Kidder, S. Q. and T. H. Vonder Haar, 1995: Satellite meteorology: An introduction. *Academic Press, London*, 466 pp.

Kiehl, J. and K. Trenberth, 1997: Earth's annual global mean energy budget. *Bulletin American Meteorological Society*, **78 (2)**, 197–208.

King, M. D., 1987: Determination of the scaled optical thickness of clouds from reflected solar radiation measurements. *Journal of the Atmospheric Sciences*, **44 (13)**, 1734–1751.

King, M. D. and R. Greenstone, 1999: 1999 EOS reference handbook: a guide to NASA's Earth Science Enterprise and the Earth Observing System. Tech. Rep. NASA NP-1999-08-134-GSFC, NASA/Goddard Space Flight Center, Greenbelt, Md.

King, M. D., S.-C. Tsay, S. E. Platnick, M. Wang, and K.-N. Liou, 1997: Cloud retrieval algorithms for MODIS: Optical thickness, effective particle radius, and thermodynamic phase. Algorithm Theoretical Basis Document ATBD-MOD-05, NASA.

Kokhanovsky, A. A. and T. Nauss, 2005: Satellite based retrieval of ice cloud properties using a semi-analytical algorithm. *Journal of Geophysical Research*, **110 (D19206)**.

Kokhanovsky, A. A., V. Rozanov, P. Zege, H. Bovensmann, and J. P. Burrows, 2003: A semianalytical cloud retrieval algorithm using backscattered radiation in 0.4-2.4 μm spectral region. *Journal of Geophysical Research*, **108 (D1)**, 1–19.

Kokhanovsky, A. A., V. V. Rozanov, T. Nauss, C. Reudenbach, J. S. Daniel, H. L. Miller, and J. P. Burrows, 2005: The semianalytical cloud retrieval algorithm for SCIAMACHY – I. The validation. *Atmospheric Chemistry and Physics Discussions*, **5**, 1995–2015.

Kraus, H., 1995: Das neue Bild von den atmosphärischen Fronten. *Erdkunde*, **49 (2)**, 81–105.

Krogh, A. and J. Vedelsby, 1995: Neural network ensembles, cross validation, and active learning. *Advances in Neural Information Processing Systems*, **7**, 231–238.

Kühnlein, M., T. Appelhans, B. Thies, A. A. Kokhanovsky, and T. Nauss, 2013: An evaluation of a semi-analytical cloud property retrieval using MSG SEVIRI, MODIS and CloudSat. *Atmospheric Research*, **122**, 111–135.

Kühnlein, M., T. Appelhans, B. Thies, and T. Nauss, 2014a: Improving the accuracy of rainfall rates from optical satellite sensors with machine learning – A random forests-based approach applied to MSG SEVIRI. *Remote Sensing of Environment*, **141**, 129–143.

Kühnlein, M., T. Appelhans, B. Thies, and T. Nauss, 2014b: Precipitation estimates from MSG SEVIRI daytime, night-time and twilight data with random forests. *Journal of Applied Meteorology and Climatology*, **53**, 2457–2480.

Kühnlein, M., B. Thies, T. Nauss, and J. Bendix, 2010: Rainfall rate assignment using MSG SEVIRI data – a promising approach to spaceborne rainfall rate retrieval for midlatitudes. *Journal of Applied Meteorology and Climatology*, **49 (7)**, 1477–1495.

Kuligowski, R., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *Journal of Hydrometeorology*, **3**, 112–130.

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

Lensky, I. M. and D. Rosenfeld, 2003a: A night-time delineation algorithm for infrared satellite data based on microphysical considerations. *Journal of Applied Meteorology*, **42 (9)**, 1218–1226.

Lensky, I. M. and D. Rosenfeld, 2003b: Satellite-based insights into precipitation formation processes in continental and maritime convective clouds at nighttime. *Journal of Applied Meteorology*, **42**, 1227–1233.

Levizzani, V., 2000: Satellite rainfall estimates: A look back and a perspective. *EUMETSAT Meteorological Satellite Data Users' Conference*, EUMETSAT, Bologna, Italy, 29 May - 2 June 2000, EUMETSAT, 344–353, No. 2000 in EUMETSAT Proceedings.

Levizzani, V., 2003: Satellite rainfall estimates: new perspectives for meteorology and climate from the EURAINSAT project. *Annals of Geophysics*, **46 (2)**, 363–372.

Levizzani, V., P. Bauer, A. Buzzi, D. E. Hinsman, A. Khain, C. Kidd, F. S. Marzano, F. Meneguzzo, A. Mugnai, J. P. V. Poiares Baptista, F. Prodi, J. F. W. Purdom, D. Rosenfeld, J. Schmetz, E. A. Smith, F. Tampieri, F. J. Turk, and G. A. Vicente, 2001a: EURAINSAT – European satellite rainfall analysis and monitoring at the geostationary scale. *11th Conf. Satellite Meteorol. Oceanography*, AMS, Madison, 650–653, October.

Levizzani, V., F. Porcu, and F. Prodi, 1990: Operational rainfall estimation using Meteosat infrared imagery. An application in Italy's Arno river basin. Its potential and drawbacks. *ESA Journal*, **14**, 313–323.

Levizzani, V., J. Schmetz, H. J. Lutz, J. Kerkmann, P. P. Alberoni, and M. Cervino, 2001b: Precipitation estimations from geostationary orbit and prospects for Meteosat Second Generation. *Meteorological Applications*, **8 (1)**, 23–42.

Liaw, A. and M. Wiener, 2002: Classification and Regression by randomForest. *R News*, **2 (3)**, 18–22.

Liou, K. N., 1992: *Radiation and cloud processes in the atmosphere. Theory, Observation and Modeling*. Oxford University Press, New York, 504 pp.

Liou, K.-N. and G. D. Wittman, 1979: Parameterization of the radiative properties of clouds. *Journal of the Atmospheric Sciences*, **36 (7)**, 1261–1273.

Liu, Y., N. V. Chawla, M. P. Harper, E. Shriberg, and A. Stolcke, 2006: A study in machine learning from imbalanced data for sentence boundary detection in speech. *Computer Speech and Language*, **20 (4)**, 468–494.

Malley, J. D., K. G. Malley, and S. Pajevic, 2011: *Statistical Learning for Biomedical Data*. Cambridge University Press.

Marrocu, M. A., A. Pompei, G. Dalu, G. L. Liberti, and A. J. Negri, 1993: Precipitation estimation over Sardinia from satellite infrared data. *International Journal of Remote Sensing*, **14**, 115–134.

Mas, J. F. and J. J. Flores, 2008: The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, **29 (3)**, 617–663.

Matejka, T. J., R. A. Houze, J. R. A. Houze, and P. V. Hobbs, 1980: Microphysics and dynamics of clouds associated with mesoscale rainbands in extratropical cyclones. *Quarterly Journal of the Royal Meteorological Society*, **106**, 29–56.

Menz, G. and A. Zock, 1997: Regionalisation of precipitation models in east Africa using Meteosat data. *Journal of Climate*, **17**, 1011–1027.

Min, Q., P. Minnis, and M. M. Khaiyer, 2004: Comparison of cirrus optical thickness depths derived from GOES 8 and surface measurements. *Journal of Geophysical Research*, **109**, D15 207.

Minnis, P., S. Sun-Mack, D. F. Young, P. W. Heck, D. P. Garber, Y. Chen, D. a. Spangenberg, R. F. Arduini, Q. Z. Trepte, W. L. Smith, J. K. Ayers, S. C. Gibson, W. F. Miller, G. Hong, V. Chakrapani, Y. Takano, K.-N. Liou, Y. Xie, and P. Yang, 2011: CERES Edition-2 Cloud Property Retrievals Using TRMM VIRS and Terra and Aqua MODIS Data – Part I: Algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, **49 (11)**, 4374–4400.

Mishchenko, M. I., J. M. Dlugach, E. G. Yanovitskij, and N. T. Zakharova, 1999: Bidirectional reflectance of flat, optically thick particulate layers: an efficient radiative transfer solution and applications to snow and soil surfaces. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **63**, 409–432.

Mohri, M., A. Rostamizadeh, and A. Talwalkar, 2012: *Foundations of Machine Learning*. The MIT Press, 412 pp.

Mota, J. F., F. J. Perez-Garcia, M. L. Jimenez, J. J. Amate, and J. Penas, 2002: Phytogeographical relationships among high mountain areas in the Baetic Ranges (South Spain). *Global Ecology and Biogeography*, **11**, 497–504.

Mountrakis, G., J. Im, and C. Ogole, 2011: Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66 (3)**, 247–259.

Munro, R., A. Ratier, J. Schmetz, and D. Klaes, 2002: Atmospheric measurements from the MSG and EPS systems. *Advances in Space Research*, **29 (11)**, 1609–1618.

Nakajima, T. and M. D. King, 1990: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory. *Journal of the Atmospheric Sciences*, **47 (15)**, 1878–1893.

Nakajima, T. Y. and T. Nakajima, 1995: Wide-Area Determination of Cloud Microphysical Properties from NOAA AVHRR Measurements for FIRE and ASTEX Regions. *American Meteorological Society*, **52**, 4043–4059.

Nauss, T. and A. A. Kokhanovsky, 2006: Discriminating raining from non-raining clouds at mid-latitudes using multispectral satellite data. *Atmospheric Chemistry and Physics*, **6 (1)**, 5031–5036.

Nauss, T. and A. A. Kokhanovsky, 2007: Assignment of rainfall confidence values using multispectral satellite data at mid-latitude: First results. *Advances in Geosciences*, **10**, 99–102.

Nauss, T. and A. A. Kokhanovsky, 2011: Retrieval of warm cloud optical properties using simple approximations. *Remote Sensing of Environment*, **115 (6)**, 1317–1325.

Nauss, T., A. A. Kokhanovsky, T. Y. Nakajima, C. Reudenbach, and J. Bendix, 2005: The intercomparison of selected cloud retrieval algorithms. *Atmospheric Research*, **78 (1-2)**, 46–78.

Negri, A. J. and R. F. Adler, 1993: An intercomparison of three satellite infrared rainfall techniques over Japan and surrounding waters. *Journal of Applied Meteorology*, **32**, 357–373.

New, M., M. Todd, M. Hulme, and P. Jones, 2001: Precipitation measurements and trends in the twentieth century. *International Journal of Climatology*, **21 (15)**, 1889–1922.

O'Sullivan, F., C. H. Wash, M. Stewart, and C. E. Motell, 1990: Rain estimation from infrared and visible GOES satellite data. *Journal of Applied Meteorology*, **29**, 209–223.

Ou, S. C., K. N. Liou, W. M. Gooch, and Y. Takano, 1993: Remote sensing of cirrus cloud parameters using advanced very-high-resolution radiometer 3.7- and 10.9-μm channels. *Applied Optics*, **32 (12)**, 2171–2180.

Pal, M., 2005: Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, **26 (1)**, 217–222.

Pandey, P., K. De Ridder, D. Gillotay, and N. P. M. van Lipzig, 2012: Estimating cloud optical thickness and associated surface UV irradiance from SEVIRI by implementing a semi-analytical cloud retrieval algorithm. *Atmospheric Chemistry and Physics Discussions*, **12 (1)**, 691–721.

Pérez, J., A. González, and M. Armas-Padilla, 2011: Remote sensing of water cloud properties from MSG/SEVIRI nighttime imagery. *Remote Sensing of Environment*, **115 (2)**, 738–746.

Petty, G. W., 1995: The Status of Satellite-Based Rainfall Estimation over Land. *Remote Sensing of Environment*, **51 (1)**, 125–137.

Platnick, S., 2000: Vertical photon transport in cloud remote sensing problems. *Journal of Geophysical Research*, **105**, 22 919–22 935.

Platnick, S., P. A. Hubanks, G. Wind, M. D. King, S. A. Ackerman, B. Maddux, T. Zinner, and A. Ackerman, 2009: The MODIS Cloud Optical and Microphysical Product: An Evaluation of Effective Radius Retrieval Statistics and Model Simulations, Hyperspectral Imaging and Sensing of the Environment, OSA Technical Digest (CD). *Advances in Imaging*, Optical Society of America, HWB1.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, **41 (2)**, 459–473.

Platnick, S. and F. P. J. Valero, 1995: A Validation of a Satellite Cloud Retrieval during ASTEX. *Journal of Atmospheric Sciences*, **52**, 2985–3001.

Polonsky, I. N., L. C. Labonnote, and S. Cooper, 2008: Level 2 cloud optical depth product process description and interface control document, version 5.0. CloudSat Project, CIRA. Tech. rep., Colorado State University, Fort Collins, 21 pp.

Pompei, A., M. Marrocu, P. Boi, and G. Dalu, 1995: Validation of retrieval algorithms for the infrared remote sensing of precipitation with the Sardinian gauge network data. *II Nuovo Cimento*, **18 C**, 483–496.

Porcu, F. and V. Levizzani, 1992: Cloud Classification using METEOSAT VIS-IR imagery. *International Journal of Remote Sensing*, **13 (5)**, 893–909.

Prasad, A. M., L. R. Iverson, and A. Liaw, 2006: Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems*, **9 (2)**, 181–199.

Prigent, C., 2010: Precipitation retrieval from space: An overview. *Comptes Rendus Geoscience*, **342 (4-5)**, 380–389.

Pruppacher, H. R. and J. D. Klett, 1997: *Microphysics of clouds and precipitation*. Kluwer Academic, Doedrecht.

R Development Core Team, 2014: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.r-project.org/.

R Development CoreTeam, 2008: *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.r-project.org.

Rao, N. X., S. C. Ou, and K. N. Liou, 1995: Removal of the Solar Component in AVHRR 3.7-μm Radiances for the Retrieval of Cirrus Cloud Parameters. *Journal of Applied Meteorology*, **34 (2)**, 482–499.

Reudenbach, C., 2003: Konvektive Sommerniederschläge in Mitteleuropa Eine Kombination aus Satellitenfernerkundung und numerischer Modellierung zur automatischen Erfassung mesoskaliger Niederschlagsfelder. *Bonner Geographische Abhandlungen*, **109**, 152.

Reudenbach, C., T. Nauss, and J. Bendix, 2007: Retrieving precipitation with GOES, Meteosat and Terra/MSG at the tropics and mid-latitudes. *Measuring precipitation from space*, V. Levizzani, P. Bauer, and F. J. Turk, Eds., Springer Netherlands, Vol. 88, 509–519.

Revolution Analytics, 2012a: doSNOW: Foreach parallel adaptor for the snow package. R package version 1.0.6, URL http://cran.r-project.org/package=doSNOW.

Revolution Analytics, 2012b: foreach: Foreach looping construct for R. R package version 1.4.0, URL http://cran.r-project.org/package=foreach.

Rivolta, G., F. S. Marzano, E. Coppola, and M. Verdecchia, 2006: Artificial neural-network technique for precipitation nowcasting from satellite imagery. *Advances in Geosciences*, **7**, 97–103.

Rodriguez-Galiano, V. F., B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, 2012: An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, **67**, 93–104.

Roebeling, R. A., H. M. Deneke, and A. J. Feijt, 2008: Validation of cloud liquid water path retrievals from SEVIRI using one year of CloudNET observations. *Journal of Applied Meteorology and Climatology*, **47 (1)**, 206–222.

Roebeling, R. A., A. J. Feijt, and P. Stammes, 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. *Journal of Geophysical Research*, **111 (D20210)**, 1–16.

Roebeling, R. A. and I. Holleman, 2009: SEVIRI rainfall retrieval and validation using weather radar observations. *Journal of Geophysical Research*, **114 (D21)**, 1–13.

Rosenfeld, D. and G. Gutman, 1994: Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data. *Atmospheric Research*, **34**, 259–283.

Rosenfeld, D. and I. M. Lensky, 1998: Satellite-based insights into precipitation formation processes in continental and maritime convective clouds. *Bulletin American Meteorological Society*, **79**, 2457–2476.

Rossow, W. B., 1989: Measuring cloud properties from space: A review. *Journal of Climate*, **2**, 201–213.

Ruiz-Gazen, A. and N. Villa, 2007: Storms prediction: Logistic regression vs random forest for unbalanced data. *Case Studies in Business, Industry and Government Statistics*, **1 (2)**, 91–101.

Rutledge, S. A. and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. VIII: a model for the "seeder-feeder" process in warm-frontal rainbands. *Journal of the Atmospheric Sciences*, **40**, 1185–1206.

Schiffer, R. A. and W. B. Rossow, 1983: The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Programme. *Bulletin of the American Meteorological Society*, **64 (7)**, 779–784.

Schmetz, J., P. Pili, S. Tjemkes, D. Just, J. Kerkmann, S. Rota, and A. Ratier, 2002: An introduction to Meteosat Second Generation (MSG). *American Meterological Society*, **83 (7)**, 977–992.

Schmetz, J., S. A. Tjemkes, M. Gube, and L. van de Berg, 1997: Monitoring deep convection and convective overshooting with Meteosat. *Advances in Space Research*, **19 (3)**, 433–441.

Schumann, W., H. Stark, K. McMullan, D. Aminou, and H.-J. Luhmann, 2002: The MSG System. *ESA Bulletin*, **111**, 11–14.

Schutgens, N. A. J. and R. A. Roebeling, 2009: Validating the validation: the influence of liquid water distribution in clouds on the intercomparison of satellite and surface observations. *Journal of Atmospheric and Ocean Technology*, **26 (8)**, 1457–1474.

Scofield, R. A. and R. J. Kuligowski, 2003: Status and Outlook of Operational Satellite Precipitation Algorithms for Extreme-Precipitation Events. *Weather and Forecasting*, **18**, 1037–1051.

Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. *Bulletin of the American Meteorological Society*, **81 (9)**, 2035–2046.

Stanski, H. R., L. J. Wilson, and W. R. Burrows, 1989: Survey of Common Verfication Methods in Meteorology. Tech. rep., WMO World Weather Watch No.8, WMO/TD No. 358, Geneva.

Steele, B., 2000: Combining multiple classifiers: an application using spatial and remotely sensed information for land cover type mapping. *Remote Sensing of Environment*, **74 (3)**, 545–556.

Stephans, G. L. and C. D. Kummerow, 2007: The Remote Sensing of Clouds and Precipitation from Space: A Review. *Journal of the Atmospheric Sciences*, **64**, 3742–3765.

Stephens, G., 2005: Cloud feedbacks in the climate system: A critical review. *Journal of Climate*, **18**, 237– 273.

Stephens, G. L., D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, and C. S. Team, 2002: The Cloudsat Mission and the A-Train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, **83 (12)**, 1771–1790.

Stone, R. S., G. L. Stephens, C. M. R. Platt, and S. Banks, 1990: The remote sensing of thin cirrus cloud using satellites, lidar and radiative transfer theory. *Journal of Applied Meteorology*, **29 (5)**, 353–366.

Strabala, K. I., S. A. Ackerman, and W. P. Menzel, 1994: Cloud properties Inferred from 8-12-μm data. *Journal of Applied Meteorology*, **33**, 212–229.

Strobl, C., J. Malley, and G. Tutz, 2009: An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. *Psychological Methods*, **14 (4)**, 323–348.

Tapiador, F. J., C. Kidd, V. Levizzani, and F. S. Marzano, 2004: A neural networks-based fusion technique to estimate half- hourly rainfall estimates at 0.1 degrees resolution from satellite passive microwave and infrared data. *Journal of Applied Meteorology*, **43 (4)**, 576–594.

Thies, B. and J. Bendix, 2011: Review Satellite based remote sensing of weather and climate: recent achievements and future perspectives. *Meteorological Applications*, **295**, 262–295.

Thies, B., T. Nauss, and J. Bendix, 2008a: Discriminating raining from non-raining cloud areas at mid-latitudes using Meteosat Second Generation SEVIRI daytime data. *Atmospheric Chemistry and Physics*, **8**, 2341–2349.

Thies, B., T. Nauss, and J. Bendix, 2008b: Discriminating raining from non-raining cloud areas at mid-latitudes using Meteosat Second Generation SEVIRI nighttime data. *Meteorological Applications*, **15**, 219–230.

Thies, B., T. Nauss, and J. Bendix, 2008c: First results on a process-oriented rain area classification technique using Meteosat Second Generation SEVIRI night-time data. *Advances in Geosciences*, **8**, 1–9.

Thies, B., T. Nauss, and J. Bendix, 2008d: Precipitation process and rainfall intensity differentiation using Meteosat Second Generation SEVIRI data. *Journal of Geophysical Research*, **113 (D23206)**.

Thornes, J., W. Bloss, S. Bouzarovski, X. Cai, L. Chapman, J. Clark, S. Dessai, S. Du, D. van der Horst, M. Kendall, C. Kidd, and S. Randalls, 2010: Communicating the value of atmospheric services. *Meteorological Applications*, **17 (2)**, 243–250.

Tjemkes, S. A., L. van de Berg, and J. Schmetz, 1997: Warm water vapour pixels over high clouds as observed by Meteosat. *Beiträge zur Physik der Atmosphäre*, **70 (1)**, 15–21.

Todd, M. C., E. C. Barrett, M. J. Beaumont, and T. J. Bellerby, 1999: Estimation of daily rainfall over the upper Nile river basin using a continuously calibrated satellite infrared technique. *Meteorological Applications*, **6 (3)**, 201–210.

Todd, M. C., E. C. E. Barrett, M. J. Beaumont, and J. L. Green, 1995: Satellite identification of rain days over the upper Nile river basin using an optimum infrared rain no-rain threshold temperature model. *Journal of Applied Meteorology*, **34**, 2600–2611.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The Changing Character of Precipitation. *Bulletin of the American Meteorological Society*, **84 (9)**, 1205–1217.

Turk, J. and P. Bauer, 2006: The International Precipitation Working Group and Its Role in the Improvement of Quantitative Precipitation Measurements. *Bulletin of the American Meteorological Society*, **87 (5)**, 643–647.

Twomey, S. and T. Cocks, 1982: Spectral reflectance of clouds in the near-infrared: Comparison of measurements and calculations. *Journal of the Meteorological Society of Japan*, **60**, 583–592.

Vicente, G. A., J. C. Davenport, and R. A. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *International Journal of Remote Sensing*, **23 (2)**, 221–230.

Vicente, G. A., R. A. Scofield, and P. W. Menzel, 1998: The Operational GOES Infrared Rainfall Estimation Technique. *Bulletin of the American Meteorological Society*, **79 (9)**, 1883–1898.

Weng, F. W., L. Zhao, R. Ferraro, G. Poe, X. Li, and N. C. Grody, 2003: Advanced Microwave Sounding Unit (AMSU) cloud and precipitation algorithms. *Radio Science*, **38**, 8068–8083.

Wilheit, T. T., R. F. Adler, S. Avery, E. Barrett, P. Bauer, W. Berg, A. Chang, J. Ferriday, N. Grody, S. Goodman, C. Kidd, D. Kniveton, C. Kummerow, A. Mugnai, W. Olsen, G. Petty, A. Shibata, E. A. Smith, and R. Spencer, 1994: Algorithms for the retrieval of rainfall from passive microwave measurements. *Remote Sensing Reviews*, **11**, 163–194.

Wu, R., J. A. Weinman, and R. T. Chin, 1985: Determination of rainfall rates from GOES satellite images by a pattern recognition technique. *Journal of Atmospheric and Oceanic Technology*, **2**, 314–330.

Zinner, T., L. Bugliaro, and B. Mayer, 2005: Remote sensing of inhomogeneous clouds with MSG/SEVIRI. *Proceedings of the EUMETSAT Meteorological Satellite Conference 2005*, Dubrovnik, Croatia, 46.