Evaluation of ERA5, ERA5-Land, and IMERG-F precipitation with a particular focus on elevation-dependent variations

A comparative analysis using observations from Germany and Brazil

Dissertation
zur Erlangung des Doktorgrades der Naturwissenschaften
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Geowissenschaften
der Universität Hamburg

vorgelegt von Tobias Olliver Kawohl
aus Hamburg, Deutschland

Hamburg, 2019

Als Dissertation angenommen am Fachbereich Geowissenschaften	
Tag des Vollzugs der Promotion:	17.01.2020
Gutachter/Gutachterinnen:	Prof. Dr. Jürgen Böhner Dr. Frank Lunkeit
Vorsitzender des Fachpromotionsausschusses Geowissenschaften:	Prof. Dr. Dirk Gajewski
Dekan der Fakultät MIN:	Prof. Dr. Heinrich Graener

Abstract

Gridded precipitation datasets from ERA5 and ERA5-Land reanalysis, as well as Integrated Multi-satellite Retrievals for GPM (IMERG) satellite-based precipitation estimates were compared with station observations in Brazil and Germany. Because there are notable differences between grid cell and station elevations in areas with complex terrain, there is an elevation-dependent bias. To adjust the gridded precipitation to the station elevation, two models were built from ERA5 and ERA5-Land using the elevation-precipitation relationship found in the model data, in the study area Germany. One model is based on a linear regression, the other model takes into account the model cloud base height. Because values averaged above space and time are used it is desirable to modify the approach to a more physically based one. In the Brazilian study area, it was not possible to find a relationship between model precipitation and model elevation. The correction of altitude differences led to a reduced bias overall, but the effect depends strongly on the individual station, therefore altitude correction can also increase the bias. It is therefore difficult to make general statements about the bias. The altitude adjustment influences performance measures for binary events not notable. However, for higher rainfall intensities and stations with a pronounced elevation difference the adjustment affects single measure. Generally, the reanalyses performed better in the German study area than in the Brazilian study area, and IMERG-F showed better results in the Brazilian study area than in the German study area. Moreover, IMERG-F produced better results during summer than during winter, which is opposite to the reanalyses.

Zusammenfassung

Tägliche Niederschläge aus den Reanalysen ERA5- und ERA-Land sowie IMERG Niederschläge, die auf Satellitenmessungen beruhenden wurden in dieser Arbeit mit Stationsmessungen in Deutschland und einem Untersuchungsgebiet in Brasilien verglichen. Da Rasterdaten im Allgemeinen eher räumliche Mittel als Punktinformation wiedergeben und in Gebieten mit komplexer Topographie zudem deutliche Höhenunterschiede zwischen Modellorographie and Stationshöhe auftreten können wurden die Niederschlagswerte höhenkorrigiert. Hierzu wurden zwei Korrekturansätze aus langjährig gemittelten Niederschlagswerten und der ERA5 bzw. ERA5-Land Orographie erstellt. Eine Höhenkorrektur der IMERG Niederschläge war auf Grund der Messmethode nicht möglich. Außerdem wurde keine Höhenabhängigkeit bzw. keine Zunahme des Niederschlags mit der Höhe in dem Brasilianischen Untersuchungsgebiet gefunden. Die Berücksichtigung und Korrektur von Höhenunterschieden zwischen Modell und Station führt generell zu einem verringerten Bias. Allerdings sind allgemeingültige Aussagen schwierig, da es stark von der einzelnen Station abhängt ob die Höhenkorrektur einen positiven oder negativen Effekt hat. Aus diesem Grund besteht auch Bedarf den Korrekturansatz so anzupassen, dass unter Verwendung weiterer Variablen räumlich und zeitlich begrenzte Verhältnisse berücksichtigt werden. Ereignisbezogene Messgrößen wie der Heidke Skill Score werden nur gering durch die Höhenkorrektur beeinflusst, Hit Rate und False Alarm Rate werden hingegen an Stationen mit deutlichem Höhenunterschied und bei Niederschlagsereignisse mit höheren Intensitäten merkbar beeinflusst. Generell hat sich gezeigt, dass die Reanalysen in dem Deutschen Untersuchungsgebiet bessere Ergebnisse liefern als die IMERG Niederschläge. In dem Brasilianischen Untersuchungsgebiet waren die IMERG Niederschläge näher an den Stationsmessungen als die beiden Reanalyseprodukte. Unabhängig vom Untersuchungsgebiet waren die Ergebnisse der Reanalysen im Winter besser als im Sommer und andersherum waren die Ergebnisse der IMERG Niederschläge im Sommer besser als im Winter.

Contents

List of Figures

List of Tables

List of Abbreviations

ANA Agencia Nacional De Aguas

ANN artificial neural network

a.s.l above sea level

CarBioCial Carbon sequestration, biodiversity and social structures in Southern Amazonia

CBH cloud base height

CHELSA Climatologies at High resolution for the Earth's Land Surface Areas

CHOMPS CICS High-Resolution Optimally Interpolated Microwave Precipitation from Satellites

CMAP CPC Merged Analysis of Precipitation

CMORPH CPC MORPHing technique

CP convective precipitation

CPTEC Centro de Previsão do Tempo e Estudos Climáticos

CSI critical success index

DEM digital elevation model

DPR dual-frequency phased precipitation radar

DWD Deutscher Wetterdienst

ECMWF European Centre for Medium-Range Weather Forecasts

ENSO El Niño Southern Oscillation

F false alarm rate

FAR false alarm ratio

FB frequency bias

FGGE First Global Atmospheric Research Program Global Experiment

GMI GPM microwave imager

GPCC Global Precipitation Climatology Centre

GPCP Global Precipitation Climatology Project

GPM Global Precipitation Measurement

GPROF Goddard Profiling Algorithm

GSMaP Global Satellite Mapping of Precipitation

GSOD Global Summary of the Day

H hit rate

HSS Heidke Skill Score

IFS Integrated Forecast System

IMERG Integrated Multi-satellite Retrievals for GPM

INMET Instituto Nacional de Meteorologia

INPE Instituto Nacional de Pesquisas Espaciais

IR infrared

ITCZ intertropical convergence zone

JAXA Japan Aerospace Exploration Agency

JRA-55 Japanese 55-year Reanalysis
LM linear model
LTM long-term mean
MAE mean absolute error
MERRA Modern-Era Retrospective analysis for Research and Applications
met. meteorological
MSWEP Multi-Source Weighted-Ensemble Precipitation
MW microwave
NASA National Aeronautics and Space Administration
NCAR National Center for Atmospheric Research
NCEP National Centers for Environmental Prediction
NLM non-linear model
NRT near real time
NWP numerical weather prediction
P precipitation
PERSIANN Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
POD probability of detection
PR precipitation radar
Q-Q quantile-quantile
r correlation coefficient
rH relative humidity
RMSE root mean square error
SD standard deviation
SRTM Shuttle Radar Topography Mission
Stn. station
TMPA Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis
TP total precipitation
TRMM Tropical Rainfall Measuring Mission
UTC coordinated universal time
WP-KS-KW Waldproduktivität-Kohlenstoffspeicherung-Klimawandel
3D three-dimensional

1 Introduction

Information about the spatial distribution of precipitation is of great interest for many sectors because the presence or absence of precipitation and rain influences humanity both directly and indirectly. Precipitation can affect the potential withdrawal of cooling water for power plants or the production of electricity from reservoir dams (Silva et al., 2007; Coelho et al., 2016b). Agriculture is especially affected by rainfall since both droughts and heavy rainfall can lead to crop shortfall and crop damage (Coelho et al., 2016a; Marengo et al., 2008; McGregor and Nieuwolt, 1998). Thus, it is not surprising that the World Meteorological Organization has classified precipitation as the most important climate variable because of its impacts on humanity (WMO, 2019). However, precipitation is highly variable in space and time and is therefore difficult to model (Arvor et al., 2014; Roy and Avissar, 2002; Seth et al., 2004). The spatial distribution of precipitation is influenced by several factors. Some of these factors such as topography are static, while others such as land use may vary over a certain period of time. The effect of land use change on precipitation has been extensively investigated, especially in the context of climate scenarios (Werth and Avissar, 2002; Nobre et al., 2009; Khanna et al., 2017; Huntingford et al., 2013). Studies have often focused on areas where the potential of land use change is high, such as in the Amazon. However, in remote areas there is often a lack of high-density precipitation observation networks (Huffman et al., 2010), which makes rainfall analysis difficult. Today, there is gridded precipitation data from different sources for the past decades, often with global or virtually global coverage. These gridded datasets may originate from point observations, from spaceborne estimates, or as climate reanalysis from numerical weather prediction (NWP) models. The spatial resolution of these datasets has increased continuously, reaching a horizontal resolution with the dimension of $10^1\,\mathrm{km}$.

Recently, the European Centre for Medium-Range Weather Forecasts (ECMWF) released two new climate reanalysis products, namely the ERA5 and ERA5-Land climate reanalysis datasets. Likewise, a backward extension of IMERG has been released, along with a version update. These three datasets are the most recent gridded datasets that provide precipitation data with virtually global coverage. Gridded data is often evaluated against station observation, because there is no dataset with a spatial truth. This results in some difficulty in comparing model output, which should be interpreted as area average rather than as point information with point measurements. It is obvious that the modelled precipitation has a bias which comes from many sources including sub-grid scale processes and insufficient representation of the topography, which affect processes such as topographically induced precipitation and windward and leeward effects. But the elevation likewise affects the amount of precipitation due to evaporation of precipitation, and it is the issue this thesis is concerned with.

The consideration of evaporation as one influencing factor seems appropriate since it is known that ERA5 has an insufficient representation of evaporation, resulting in too much light rain reaching the surface (cf. Sec. 2.2). Such an altitude adjustment by a lapse rate is commonly known and applied to the temperature, often using the environmental lapse rate

of $0.65\,°/100\mathrm{m}$, even though substantial spatial and seasonal differences exist (Minder et al., 2010; Whiteman, 2000). Nevertheless, any kind of altitude adjustment for precipitation is unusual, probably because of its small-scale variability and rather low physically robust relationship. This thesis demonstrates that such a concept does not work in all environments and would need some improvements to be ubiquitously applicable.

Thus, followed by an evaluation of the named datasets, a concept is presented that takes into account differences in altitude between the model orography and a target orography. The latter can either be the orography from another grid or station altitude. The general idea of the concept is to take into account the presumable evaporation that would occur averaged over space and time for a certain elevation difference in ERA5 or ERA5-Land. The supposed evaporation is not calculated using the evaporation parameterisation in the model because it is difficult or even impossible to meet all the assumptions and values that would be necessary. Instead, the evaporation is estimated using either a linear or non-linear statistical model. Moreover, this theoretically allows application to models with other precipitation parameterisation schemes. However, this approach does not necessarily reduce the model bias compared to stations observations; it aims to correct the bias by the amount resulting from the altitude differences. The non-linear model was developed during two research projects – 'Waldproduktivität-Kohlenstoffspeicherung-Klimawandel' WP-KS-KW and 'Carbon sequestration, biodiversity and social structures in Southern Amazonia' CarBioCial – which are briefly introduced at the end of this section.

The evaluation of the precipitation products is done in Germany and a large part of Brazil to obtain results from different environmental conditions. In Germany the weather is influenced by the westerlies. In contrast the Brazilian area considered is dominated by a tropical climate, with the southern part of the area demonstrating a strong seasonality in rainfall. This thesis aims to answer the following questions:

a) Are the differences between the performance of satellite-based precipitation and climate reanalysis found in the literature still persistent in the most recent precipitation datasets, and does the quality dependent on the elevation (as often stated) or on elevation differences, when using station observation?
b) How does the concept of adjusting elevation differences between model orography and target orography influence the bias and performance measures of modelled P?

Broadly speaking this thesis is structured as follows: first, a short section about the two aforementioned research projects is presented. In Sec. 2, the most recent gridded precipitation datasets are introduced, and their benefits and disadvantages are addressed. This is followed by descriptions of the two study areas in Section 3. Section 4 describes the difficulties with the used datasets and the necessary preparations. Likewise, the method for the altitude adjustment of precipitation is presented. The results are given in Section 5, followed by a discussion of the same in Section 6.

Project Overview

The Carbon sequestration, biodiversity and social structures in Southern Amazonia (CarBioCial) project was a joint research project between several universities in the context of sustainable land use funded by the German Federal Ministry of Education and Research. More specifically, it was about carbon-optimised land management strategies for southern Amazonia, mainly focusing on the area around the Highway BR-163. Along this highway there is a distinct land use gradient characterised by agricultural expansion proceeding northward (Gerold et al., 2018). Data from different climate models were used for crop and land use modelling.

The project Waldproduktivität-Kohlenstoffspeicherung-Klimawandel (WP-KS-KW) was a cooperative project between different forest research units and the University of Hamburg and was funded by the German Federal Ministry of Food and Agriculture. In this project, the focus was on how the forest would be affected by climate change. Therefore, simulations from different climate models were used and processed.

2 State of the Research

Today, there are many gridded datasets for different variables that often provide virtually global coverage. Additionally, various datasets are connected with each other, or one incorporates another. The drawbacks and benefits resulting from the various types of gridded precipitation data are discussed in this section with a focus on ERA5, ERA5-Land, and IMERG.

In general, global gridded datasets based on observation from meteorological (met.) stations are of coarse spatial or temporal resolution (Schneider et al., 2014; Schamm et al., 2014; Harris et al., 2014), but high-resolution precipitation data of this type only exists for certain regions. While such datasets exist for Germany (Rauthe et al., 2013) and South America (Xavier et al., 2015), the quality differs considerably because of the input stations. Problems with Brazilian or South American stations are discussed in Section 4. Both types – met. station observations and modelled precipitation – benefit from each other. Either station observations are used to correct modelled data, or modelled precipitation is used to close the gaps of observation. In this respect, WorldClim version 2, a global dataset primary based on station observation that provides among other variables gridded precipitation climatologies, incorporates satellite-based precipitation estimates to handle the problem of partly sparse observations (Fick and Hijmans, 2017). Likewise, Global Precipitation Climatology Project (GPCP) data combines met. station observations and precipitation estimates from satellites (Adler et al., 2003). Depending on the version, CPC Merged Analysis of Precipitation (CMAP) also combines met. stations, satellite estimates and forecasts from NWP models (Xie and Arkin, 1997). On the other hand, Climatologies at High resolution for the Earth's Land Surface Areas (CHELSA) uses climate reanalyses and Global Precipitation Climatology Centre (GPCC) data for its monthly precipitation data (Karger et al., 2017). Before observations from space became available as they are today, the need for a consistent global climate dataset arose and was addressed in the 1980s by Bengtsson and Shukla (1988).

2.1 Climate Reanalysis

The reason that climate reanalyses (hereafter reanalyses) are quite beneficial for climate analysis is that they provide many variables in a physically consistent manner and can be used for multiple purposes, such as to investigate the current climate and its variability including monsoon systems or the El Niño Southern Oscillation (ENSO) (Trenberth et al., 2008). Trenberth and Olson (1988) have described the benefits of an internally consistent dataset, because at the time of their research, reanalysis was only available as a by-product, of the initial conditions for NWP. This by-product suffered inconsistencies due to model changes. Moreover a significant number of meteorological bulletins were not included in the analysis because of shortcomings in the global telecommunication system (Bengtsson and Shukla, 1988). Generally, a reanalysis is a system that includes available observations and uses these with a NWP model, usually in a hindcast mode. Nevertheless, virtually all reanalyses have some

inconsistencies. These inconsistencies can emerge due to new observations or observation systems that are included to produce the initial condition. Because of this, reanalysis has never been intended to be a long-term homogeneous time series but rather the best possible analysis for each time step (Thorne and Vose, 2010). A comprehensive summary of the physical processes, parameterisation, and difficulties of NWP can be found in Bauer et al. (2015), for instance.

Today, several meteorological institutions provide climate reanalysis as, either global or regional datasets. Reanalyses with global coverage include the Modern-Era Retrospective analysis for Research and Applications (MERRA)-2 from the National Aeronautics and Space Administration (NASA), which begins in 1980 with a latitudinal resolution of roughly $50\,\mathrm{km}$ and replaced the first MERRA because of an increasing number of available satellite observations (Rienecker et al., 2011; Gelaro et al., 2017). Likewise, Version 2 of the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis project begins with the availability of satellite observation (Kanamitsu et al., 2002; National Centers for Environmental Information (NCEI), n.d.). Both version 1 and 2 of the NCEP/NCAR reanalysis are available in $2.5°$ (horizontal resolution). The Japanese 55-year Reanalysis (JRA-55) begins with the global availability of radiosonde observation and covers the time period from 1958 to present in roughly $55\,\mathrm{km}$ (Ebita et al., 2011; Kobayashi et al., 2015). The ECMWF provided several climate reanalyses, beginning with First Global Atmospheric Research Program Global Experiment (FGGE) and followed by ERA-15, ERA-40, and ERA-Interim; recently the fifth generation ERA5 was released, providing 1-hourly data in $0.25°$ followed by ERA5-Land which provides surface and near-surface variables in $0.1°$. Next, a ERA5 near real time (NRT) product is announced, with only a one-week delay. Both the ERA5 and ERA5-Land are addressed below. Because this thesis focuses on datasets with virtually global coverage, regional reanalysis products are not mentioned here, though they can provide a higher horizontal resolution.

2.2 The ERA5 Climate Reanalyses

As previously mentioned, ERA5 is the most recent climate reanalysis and the fifth generation of the ECMWF reanalyses. Two versions of ERA5 exist, namely ERA5-HRES, the high-resolution realisation ($0.25°$), and ERA5-EDA, a 10-member ensemble with reduced horizontal resolution ($0.5625°$), where the latter intents to account for uncertainties (ECMWF, n.d.a). Hereafter, if not otherwise stated, ERA5 refers to ERA5-HRES.

The ERA5 model has 137 model levels that are interpolated to 37 pressure levels; the top model level reaches up to $0.1\,\mathrm{hPa}$ (ECMWF, n.d.a). The ERA5 time series is not processed sequentially but by several parallel experiments. As a result, discontinuities can exist at the transition periods. In the troposphere, these discontinuities are smaller than the ensemble spread, and they are also smaller than the discontinuities in ERA-Interim (Hersbach and Dee, 2016; ECMWF, 2018). Major changes from the ERA5 predecessor ERA-Interim to ERA5 are

summarised in Table 1. Additionally, the assimilation scheme of ERA5 incorporates more observations than that of ERA-Interim (Hersbach and Dee, 2016; ECMWF, 2019b).

ERA5 is based on the ECMWF Integrated Forecast System (IFS), which is coupled with a soil model, an ocean wave model, and the four-dimensional variational data assimilation system data assimilation system (CY41r2). Data assimilation is the key component and is used to estimate the state of the atmosphere as accurately as possible, incorporating observations and a previous forecast, which is a background (Kleist et al., 2009). Although precipitation is assimilated in the assimilation system, it is not directly connected to the modelled precipitation values; instead, it is used to gain information on backscatter signals that influence satellite wind speed measurements (ECMWF, 2016a).

The precipitation consists of short forecasts, initialised twice per day at 06 and 18 coordinated universal time (UTC) ECMWF (n.d.a); moreover, for precipitation the variables total precipitation (TP) and convective precipitation (CP) are available and are later used to distinguish the proportion of convective and large-scale precipitation. However, precipitation in the IFS model also underlies evaporation below the cloud. For convective precipitation, the parameterisation scheme follows Kessler (1969). The default evaporation parameterisation for large-scale precipitation follows Abel and Boutle (2012) nonetheless there is likewise a scheme that follows Kessler (1969). In the convective scheme precipitation below the cloud base e_{subcld} begins evaporating when the relative humidity (rH) is below 90% over water bodies and below 70% over land. The evaporation rate is assumed to proportionally on the saturation deficit ($rH_{cr}q_{sat}^{-} - \bar{q}$) and on the rain density ρ_{rain}.

$$e_{subcld} = \alpha_1(rH\bar{q}_{sat} - \bar{q})\rho_{rain}^{13/20} \tag{1}$$

with $\alpha_1 = 0$ if $\bar{q} > rH\bar{q}_{sat}$, else $\alpha_1 = 5.44 \times 10^{-4}s^{-1}$.

The density of rain is not part of the convective scheme but can be indirectly calculated from the precipitation flux (ECMWF, 2016b, p. 85). The evaporation at a certain level can be expressed as follows:

$$e_{subcld} = C_{conv}\alpha_1(RH\bar{q}_{sat} - q)\left[\frac{\sqrt{p/p_{surf}}}{\alpha_2}\frac{P}{C_{conv}}\right]^{\alpha_3} \tag{2}$$

with $\alpha_2 = 5.09 \times 10^{-3}$, $\alpha_3 = 0.5777$, and constant $C_{conv} = 0.05$. C_{conv} describes the fraction of the cell with convective precipitation (ECMWF, 2016b, p. 85).

Looking at the non-default large-scale precipitation scheme based on Kessler (1969), the micro-physical constants α_1 to α_3 remain unchanged. Additionally, here evaporation begins in the clear air and the fraction of precipitation in a grid cell, and the sub-grid heterogeneity

decreases in proportion to the evaporation.

$$S_{evap} = a_p^{clr} \alpha_1 (q_{sat} - q_v^{env}) \left[\frac{\sqrt{p/p_0}}{\alpha_2} \frac{p^{clr}}{a_p^{clr}} \right]^{\alpha_3} \tag{3}$$

In both schemes, the altitude is considered in terms of p/p_0. The calculation of the critical rH value for evaporation in the large-scale scheme depends on the fraction covered by clouds (ECMWF, 2016b). Evaporation of raindrops is generally higher for light rain, which is often formed from drizzling clouds (ECMWF, 2016b).

Topographic effects such as topographic barriers and rain shadow effects are underestimated in ERA5 due to the coarse horizontal resolution. Nevertheless, because of the higher resolution of ERA5-HRES compared to ERA5-EDA, the HRES dataset represents these effects better than the ensemble data. Another issue is the over-production of light rain and drizzle in stratocumulus together with insufficient evaporation, which results in too much light rain reaching the surface. The over-production of light rain and drizzle can become problematic when, due to incorrect modelling of the boundary layer, low clouds are over-persistent. Finally, although the effect should be rather small because of the coarse resolution, there is potential for miss-assignment of rain reaching the surfaces due to incorrect downwind drift of the rain (ECMWF, n.d.b).

2.3 The ERA5-Land Climate Reanalysis

Currently, ERA5-Land is a single simulation producing data at a 9 km horizontal resolution, forced by ERA5's low atmospheric fields, but it is coupled with neither the atmospheric module of the IFS nor the ocean module. Additionally, no data assimilation takes place, which makes it computationally affordable for updates, such as at the land surface model. ERA5-Land uses the tiled ECMWF scheme for surface exchanges over land incorporating land surface hydrology, which uses version CY45R1 of the IFS. At present, uncertainty information from ERA5 are used; in the future an ERA5-Land ensemble run might also be possible (Muñoz Sabater, 2019; ECMWF, 2019a).

2.4 Satellite-based Precipitation

Satellite-based estimates of precipitation became more common in the 1980s (Huffman et al., 2010). Today, there are several satellite precipitation products that can help to close the gaps of measurements. To estimate precipitation from space, infrared (IR) as well as active and passive microwave (MW) sensors are used. Because IR sensors measure the cloud-top temperature, they are particularly useful for tropical convective regimes given that cold cloud-top temperatures are usually connected with rain. However, IR rainfall estimates fail on cold convective clouds and on low-reaching ones. Additionally, IR sensors have the advantage of a high sampling frequency of approximately 15 to 30 minutes. However, so far they have only

Table 1: Major differences between ERA-Interim, ERA5 and ERA5-Land. Source: Muñoz Sabater (2019); ECMWF (2019b).

	ERA-Interim	ERA5	ERA5-Land
Time Period	1979 to present	(1950)/1979 to present	(1950)/1979 - present
Spatial Resolution	79 km (0.75°) 60 levels	31 km (0.25°) 137 levels	9 km (0.1°) single level (near) surface
Temporal Resolution	3 or 6 hourly	1 hourly	1 hourly
Model version	IFS (+TESSEL)	IFS (+HTESSEL)	HTESSEL IFS Cycle 45r1
Assimilation System	IFS Cycle 31r2	IFS Cycle 41r2	

been installed on geostationary satellites. Because MW measurements are more directly connected to the size of hydrometeors, they are better at distinguishing different hydrometeors. Though, due to the revolving orbit of satellites with MW sensors they have a lower sampling frequency. Therefore, the common aim is to combine estimates of these two sensor types (Joyce et al., 2004; Ashouri et al., 2015; Hou et al., 2014).

As previously mentioned, some datasets incorporate satellite-based rainfall estimates. Those estimates in turn merge different sources and sensors. For example, CICS High-Resolution Optimally Interpolated Microwave Precipitation from Satellites (CHOMPS) uses data from different satellites and different passive MW sensors (Joseph et al., 2009). The CPC MORPHing technique (CMORPH) combines the propagation of motion from geosynchronous IR estimates with passive MW measurements to estimate precipitation (Joyce et al., 2004), and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) utilises IR and MW measurements with artificial neural networks (ANNs) to estimate daily precipitation. Additionally, PERSIANN incorporates GPCP data (Ashouri et al., 2015). The most recent dataset, that uses some of the aforementioned techniques is generated from the Global Precipitation Measurement (GPM) mission. Like the other products, it combines different sources, sensors, and algorithms into one dataset (Huffman et al., 2010). The following section describes the GPM IMERG product in more detail.

2.5 Integrated Multi-Satellite Retrievals for GPM

The predecessor of IMERG is the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis (TMPA) of the Tropical Rainfall Measuring Mission (TRMM) which began in the spring of 1998; GPM was then launched in 2014 . Because TRMM sensors were specially designed to measure moderate to heavy rain in the tropics and sub-tropics, only the area between 37°N/S was covered. The Global Precipitation Measurement extends the spatial coverage to 68°N/S and extends the measurement range to light rain and snow, which contributes for a relevant proportion of precipitation in middle and high latitudes. The mission is a

joint project between NASA and the Japan Aerospace Exploration Agency (JAXA), which furthermore involves a multinational cooperation with other countries and agencies. The aim of the mission is to more accurately measure precipitation distribution and frequency on a global scale and to provide information about the vertical structure of clouds. It is also intended to provide a calibration standard to enable the unification of measurements.

The core satellite of the mission has a non-sun-synchronous circular orbit at a $65°$ inclination approximately $407\,\mathrm{km}$ above the earth. It is equipped with a dual-frequency phased precipitation radar (DPR) and a conical scanning multi-channel microwave imager named GPM microwave imager (GMI). The measurements of the core satellite are used to build a database with the microphysical properties of precipitation particles over different environmental conditions and climates. This database is then used as a common reference to unify measurements before, during, and after the lifetime of the GPM core satellite. A lifetime of three years with sufficient fuel for five years was planned for the core satellite. In addition to the core observatory, several satellites from other countries and institutions with either a DPR or microwave imager aboard are affiliated. The orbit of the core satellite enables coincident measurements with these affiliated constellation satellites, which allows inter-sensor calibration over $90\,\%$ of the surface of the earth JAXA (2016); Hou et al. (2014); Huffman et al. (2010, 2007). A schematic view of the core satellite is shown in Figure 1.

The DPR operates a Ka-band and a Ku-band of $35.5\,\mathrm{GHz}$ and $13.6\,\mathrm{GHz}$, respectively. Both bands provide a $5\,\mathrm{km}$ co-aligned footprint on the surface with swath widths of $125\,\mathrm{km}$ and $245\,\mathrm{km}$ for the Ka-band and Ku-band, respectively. In the inner swath of $120\,\mathrm{km}$, data from both bands are acquired almost simultaneously with a vertical resolution of $250\,\mathrm{m}$. The GPM Ku-band has higher precision than the TRMM precipitation radar (PR) and Ka-band because the Ku-band extends the sensitivity down to approximately $0.2\,\mathrm{mm/hr}$. Because of the two different frequencies, the DPR is able to measure the three-dimensional (3D) structure of precipitation. Moreover, because the attenuation of the echoes from both bands are dependent on frequency and raindrop size, the simultaneous measurements enable the calculation of the raindrop size distribution. However, while the Ka-band aims to detect weak rainfall as well as snowfall, this cannot be measured by the Ku-band, which detects heavier precipitation. The time lag between DPR and the GMI is approximately 67 seconds due to geometry and spacecraft motion (Hou et al., 2014).

The GMI has 13 channels with frequencies optimised for different rain frequencies. The following list provides a short overview of the channels taken directly taken from Hou et al. (2014, p.708):

- 10-GHz channel optimal for sensing of liquid precipitation
- 19- and 37-GHz channels for sensing moderate to light precipitation over ocean
- 21-GHz channel for correcting emission by water vapour
- 89-GHz channel for the detection ice particles for precipitation over ocean and land
- 166-GHz channel for sensing light precipitation (typical outside the tropics)

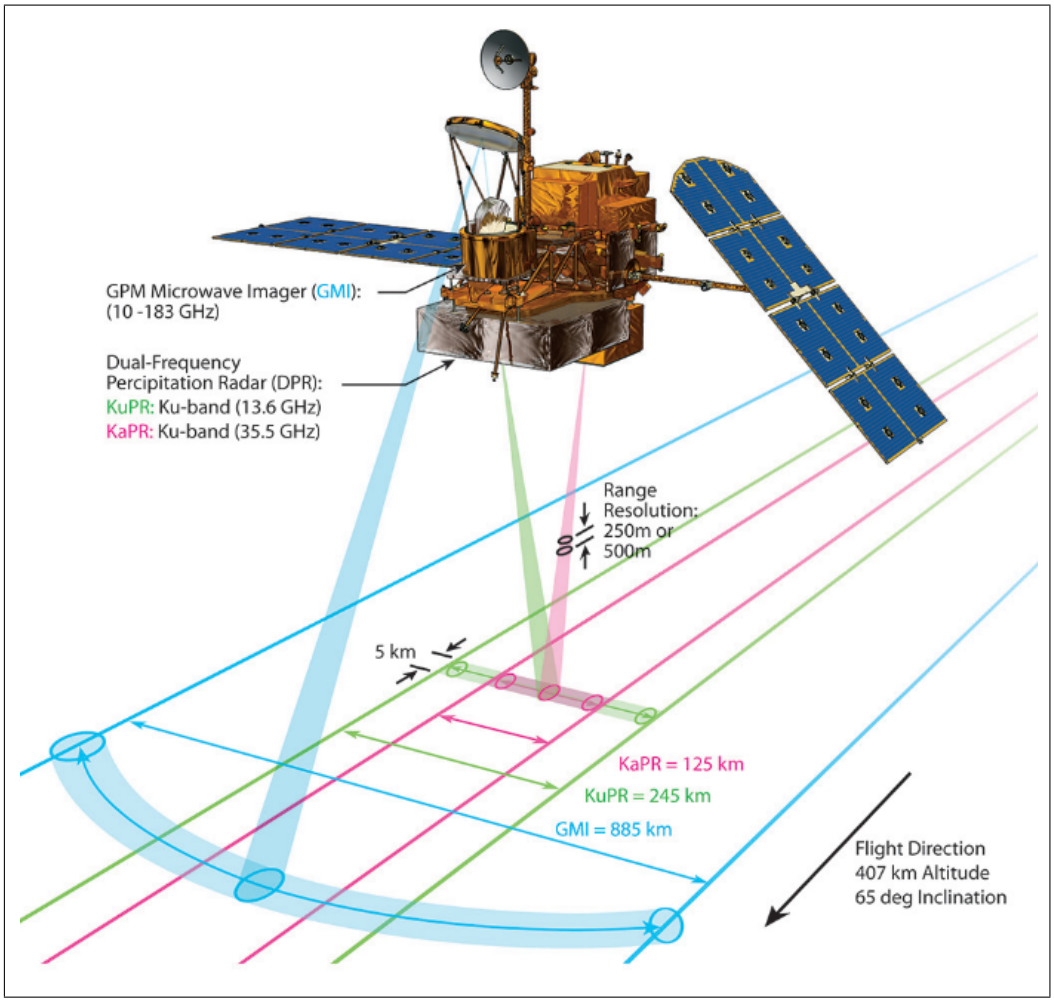

Figure 1: Schematic illustration of the GPM core observatory with swath widths and resolution of installed sensors. Source:Hou et al. (2014, p. 707).

- 183-GHz channel for detecting scattering signals due to small ice particles and estimating light rainfall and snowfall over snow-covered land

The IMERG combines the different measurements, and depending on the processing level, three products, so-called runs with different time delays and data quality are available. The early run or NRT is processed within 4 hours, the late run is available within 12 hours, and the final run or research product can be used following a 2.5-month delay. The different measurements are processed to different levels. Level 0 data consists of unprocessed sensor data at full resolution, calibrated DPR power, GMI brightness temperature, and inter-calibrated brightness temperature from other radiometers. Level 2 data are geolocated, geophysical data and DPR reflectivity. Level 3 data are statistically, spatially, and temporally processed gridded data with time and space coordinates from core and partner satellites. Additionally, a Level 4 product is planed to be composed from remotely sensed values and global NWP model output (Hou et al., 2014; Huffman et al., 2018a).

Level 3 IMERG data combines different algorithms, and at the beginning brightness temperatures from different passive MW sensors are intercalibrated with the core instruments. In-

frared measurements are used to generate motion vectors with the methods from the CMORPH scheme (Joyce et al., 2004) and to the PERSIANN - Cloud Classification Scheme (Hong et al., 2004). The merging of these data generates the *Uncal* field. After bias correction with monthly GPCC data, the *PrecipitationCal* field is obtained. The precipitation phase separation is based on pressure, temperature, and humidity. The systems that process the final run support reprocessing; thus, updated versions are possible for the final run. To improve precipitation estimates, the inclusion of daily-observed precipitation fields is planned for the future (Huffman et al., 2018a; Kummerow et al., 2001). The described process is illustrated in greater detail in Figure 2. At latitudes higher than 60°N/S the spatial coverage of IMERG is incomplete, because no IR estimates form geosynchronous satellites are available and MW estimates are only available when satellites overpassed. Nevertheless, the field *precipitationCal* exists up to 90°N/S latitude. The different fields provided by IMERG are listed below (Huffman et al., 2019a).

- *HQprecipitation* – daily accumulated high-quality precipitation from all available MW sources
- *HQprecipitation cnt* – count of valid half-hourly *HQprecipitation* retrievals for the day
- *precipitationCal* – daily accumulated precipitation (combined microwave-IR) estimates
- *precipitationCal cnt* – count of valid half-hourly *precipitationCal* retrievals for the day
- *randomError* – daily total error of precipitation estimates
- *randomError cnt* – count of valid half-hourly random error retrievals for the day

The highest temporal resolution for IMERG data is half-hourly. For half-hourly and monthly data, improvements on the weights of gauge analysis are included in addition to the data (Huffman et al., 2018b). One of the major changes between IMERG v05 and v06 is the switch in input data to calculate the motion or displacement vectors. In v05, IR measurements were used, while in v06, MERRA-2 and 'Goddard Earth Observing System Forward Processing' data are used (Huffman et al., 2019b). Subsequently -E, -L, and -F indicate the early, late, and final (or research) runs of the IMERG data.

2.6 Quality of Gridded Precipitation

Since the methods of satellite-based precipitation estimates (hereafter satellite precipitation) and climate reanalyses differ fundamentally, it is interesting to see if one method or the other produces better results, and under which circumstances. Because all three datasets used here are relatively new, so far there are only a few studies regarding these products. This is especially true for ERA5-Land and IMERG-F v06 data, because these datasets were released in the summer of 2019. Additionally, the backward extension of ERA5 was recently released. Nevertheless, to answer the question of whether the general differences found in predecessor datasets still prevail in the actual datasets, it is necessary to review the general findings from older datasets.

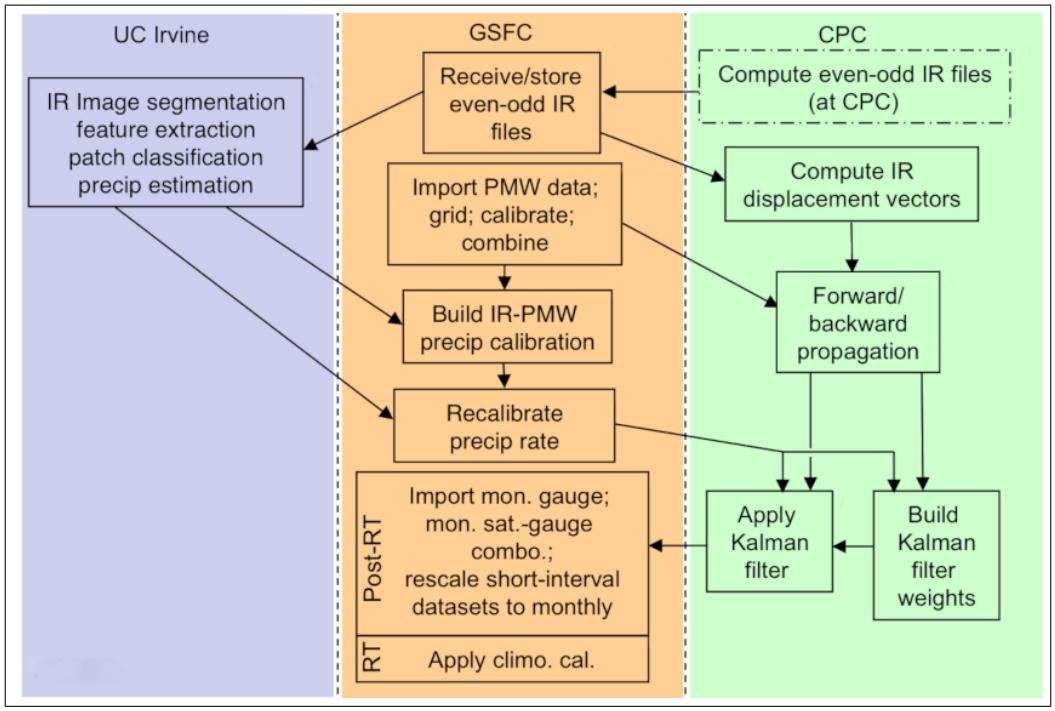

Figure 2: Data flow of major modules in IMERG. The colours illustrate the different institutions from which IMERG uses algorithms to build the integrated system. Source: modified from Huffman et al. (2018a, p. 9).

Since many sources of precipitation data exist and some sources have already been combined, Beck et al. (2017) have argued that global precipitation datasets could be better, if more of the available resources would be combined. Thus, Beck et al. developed a dataset called Multi-Source Weighted-Ensemble Precipitation (MSWEP), which incorporates seven different precipitation products from gridded gauge-only and, over satellite-only, to reanalysis datasets. Beck et al. (2018) then compared 26 precipitation datasets against radar precipitation in the continental USA. Uncorrected precipitation from reanalyses provided better results than uncorrected satellite estimates in terms of correlation coefficients and event identification. The correlation coefficient of MSWEP is higher than 0.8, IMERG and ERA5-HRES have r values between 0.6 and 0.8, and those of ERA5-EDA are slightly higher. Regarding the gauge-corrected IMERG-F and TMPA 3B42 products, IMERG-F showed improvements in the correlation coefficient and in the variability ratio. From the uncorrected datasets, ERA5-HRES and IMERG-E half-hourly V05 followed by ERA5-EDA had the best results. ERA5-EDA demonstrated a lower variability than ERA5-HRES, probably due to ensemble averaging. IMERG-E with half-hourly precipitation performed better than the JAXA algorithm Global Satellite Mapping of Precipitation (GSMaP) Version 6. It was also found that climate reanalysis underestimates the variability of rainfall with a tendency to overestimate rainfall frequency. Additionally, the more complex the terrain, the lower the performance of all uncorrected datasets. Concerning ERA5-HRES and IMERG half-hourly V05, ERA5-HRES provided better results in complex terrain, and IMERG did better in convective regimes. Over all the analysed datasets, MSWEP V2.2 represented precipitation best, followed by daily IMERG-F with a small negative bias.

Beck et al. have noted that the results emphasise the importance of daily gauge-based correction and knowledge about the gauge retrieval times, which is addressed again in Section 4.

Beginning with the few studies that deal with ERA5, Albergel et al. (2018) coupled ERA5 with a land surface model and found the hydrological cycle much better represented than when using ERA-Interim for coupling. Additionally, two other studies, albeit not particularly focused on hydrological variables record a better performance for ERA5 than for older reanalyses. Urraca et al. (2018) have found the radiation in ERA5 to be improved compared to ERA-Interim, especially in areas with low cloudiness, and focusing on wind power modelling, Olauson (2018) have found ERA5 to perform better than MERRA-2. Betts et al. (2019) have examined the precipitation and other variables of ERA5 in Canada. For the cold season with predominantly large-scale precipitation, ERA5 produced approximately the same monthly values as ERA-Interim, but during the warm season, ERA5 produced approximately +14% precipitation compared to ERA-Interim. During the warm season, the bias of ERA5 at the five stations in the province of Saskatchewan is between $11\,\%$ and $18\,\%$ depending on whether the original observations or a corrected observation dataset are used. During winter, the bias of ERA5 was $-22\,\%$, which was probably due to a too-high snow correction in the corrected observation data. However, because of the spatial variability of rainfall and the uncertainty of station observations, Betts et al. (2019) were unable to declare whether or not ERA5 is biased.

However, looking to older reanalyses, namely ERA15, ERA40, and the NCEP reanalysis, a general underestimation of heavy rainfall was found across Europe, and higher correlation coefficients during winter (Zolina et al., 2004). Likewise, Kidd et al. (2012) found that numerical models performed better during winter, because the convective representation during summer was poor. Regarding satellite-based precipitation, Kidd et al. detected a poorer performance during winter with respect to correlation, bias and false alarm rate, across Northwest Europe. This was partly attributed to the inability of satellite precipitation to detect low rain rates. Sun et al. (2018) have compared different reanalyses and satellite precipitation products, including IMERG, from a global perspective. They have found that all products underestimate rainfall almost year-round in Northwest Europe. Only during summer precipitation is overestimated in Germany. On the other hand, satellite precipitation was found to overestimate heavy precipitation events in southern Brazil as well as in Iran, where Sharifi et al. (2016) have compared IMERG, ERA-Interim, and TMPA precipitation and found that the three datasets tend to underestimate the amount of rainfall. Nonetheless, IMERGF-F had the highest correlation and the best performance in terms of probability of detection (POD), false alarm ratio (FAR) and critical success index (CSI). Though, ERA-Interim had a lower RMSE compared to the other datasets.

Focusing on the improvements from TMPA to IMERG, Tang et al. (2016) have stated that especially on sub-daily values, the improvement of IMERG-F over TMPA becomes clear. Moreover, IMERG performs better than TMPA in high latitudes and regions characterised

by a dry climate. However, the overall performance of IMERG-F is better in low to middle latitudes. In contrast to most other studies, Yuan et al. (2017) found no significant improvement of IMERG-F over TMPA 3B42 V7. Instead, they found the opposite and stated that TMPA provides better results than IMERG-F on daily and monthly bases. The correlations found by Yuan et al. at five stations in Myanmar for rain >10 mm/day lie between -0.09 and 0.21. For what they classified as light rain, namely 0.1 to 10.0 mm/day, the correlation coefficients were slightly negative at all five stations. An underestimation of heavy rainfall was likewise prevalent in this study and more distinct in IMERG-F than in TMPA. Using gridded rainfall data for India, Beria et al. (2017) found IMERG-F to perform better than TMPA 3B42, but the overlapping time period of one year was rather short. In terms of hit rate (H) and false alarm rate (F), IMERG-F produced better results for low to high rainfall intensities. In many catchments, correlation coefficients above 0.8 and sometimes even around 0.9 were found in 2019. The most distinct improvements occurred at low rain rates. This is probably due to the limitation of a minimum detectable rain rate of $\leq 0.5\,\mathrm{mm/h}$ by TRMM that is lower in GPM sensors, as already mentioned in Section 2.5. Prakash et al. (2018) and Khodadoust Siuki et al. (2017) have also found this better representation of low rain rates. Even though Chen et al. (2018) have also stated a higher performance for IMERG-F v05 compared to TMPA 3B42 v7 in the Hauihe River basin, the Pearson correlation coefficients they found of 0.41 for IMERG-F daily data and 0.36 for TMPA were much lower than the high correlation between IMERG-F and gridded rainfall in India found by Beria et al. (2017). While the relative bias was approximately -0.17, the POD depended strongly on the rainfall threshold. For rain rates up to 2 mm/day, the POD lies between 0.6 and 0.7, and for a rainfall threshold of 100 mm/day, the POD decreased to 0.2. Consistent with this light rain between 0 mm and 0.5 mm as well as heavy rainfall (>25 mm) was underestimated by IMERG-F, while rainfall between 0.5 mm and 25 mm was overestimated. Moreover, Beria et al. (2017) found that the performance decreased with elevation.

Similarly, Chen et al. (2016) found a lower quality of IMERG-F and TMPA 3B42 data in mountainous regions. In addition, frozen precipitation was found to be unreliable in IMERG-F. Even though the general skill depended strongly on the region, they also found better skills for IMERG-F than for TMPA. Asong et al. (2017) also found a performance decrease in mountainous areas for half-hourly IMERG-F v03 data in southern Canada, where the quality of IMERG-F was generally better during summer than during winter. Regarding heavy precipitation, a tendency towards overestimation was found. Additionally, others studies have discovered an elevation-depended bias. A decreasing performance in mountainous regions has been found by various other authors e.g. (Kim et al., 2017; Prakash et al., 2018; Hirpa et al., 2010; Guo et al., 2017). The often stated lower performance of satellite precipitation in complex terrain can be attributed to warm orographic rain, which is problematic for IR sensors as well as for passive MW, since scattering from ice is reduced, which results in an underestimation of surface rain (Dinku et al., 2007). Therefore, Shige et al. (2013) have developed an orographic/non-orographic precipitation classification for the Japanese Kii peninsula to im-

prove GSMaP, which is the JAXA counterpart of NASA's IMERG product. Depending on the vertical wind component and surface moisture convergence, the algorithm distinguishes between different look-up tables of vertical orographic/non-orographic precipitation profiles. During summertime, the classification system improved the GSMaP rainfall, but during wintertime it was not able to achieve any improvements. Yamamoto and Shige (2015) have noted, that the approach of Shige et al. (2013) was developed in warm maritime air and cannot work under exceptionally conditions such as many solid hydrometeors in the atmosphere. Therefore, Yamamoto et al. (2017) have further improved the approach, taking into account low-level wind speed and the accompanying delay until rain reaches the surface.

The results of the studies cited above can also be found when focusing more on the two study areas in this thesis. Ramsauer et al. (2018) have also discovered the repeatedly mentioned overestimation of precipitation by IMERG-F for Germany, especially during the winter season. Spatially averaged IMERG-F precipitation was virtually always above the radar based precipitation estimates. During winter, IMERG-F produced a plus of $76\,\%$ compared to the radar precipitation. In keeping with this, during winter the number of wet days was higher in IMERG-F than in the radar based precipitation data, and the spatial correlation was lowest during wintertime. Moreover, IMERG-F missed some topographically induced rainfall events in mountainous regions.

Rozante et al. (2018) have examined TMPA, IMERG-F, and GSMaP on a daily basis remapped to the TMPA grid ($0.25°$). They divided Brazil into five regions, with Region 2 being approximately the Brazilian study area of this thesis. While TMPA and IMERG-F tend to mainly overestimate rainfall during summer, the gauge-corrected GSMaP product is closer to the observations, probably because the correction based on daily gauge values in the GSMaP works better than the monthly correction with GPCC data in IMERG-F. The mean error and the RMSE were $0.59\,\mathrm{mm/day}$ and $0.73\,\mathrm{mm/day}$ for IMERGF-F, $0.52\,\mathrm{mm/day}$ and $0.68\,\mathrm{mm/day}$ for TMPA and $-0.06\,\mathrm{mm/day}$ and $0.26\,\mathrm{mm/day}$ forGSMaP, which therefore had the lowest mean error and RMSE. Additionally, in northern Brazil, IMERG-F had a mean error of $1.01\,\mathrm{mm/day}$ and a RMSE of $1.31\,\mathrm{mm/day}$.

In the region of Manaus, Oliveira et al. (2016) validated sub-daily IMERG data and the Goddard Profiling Algorithm (GPROF), which is part of IMERG with ground-based radar precipitation. The question of interest was if GPROF and IMERG are able to reproduce the diurnal cycle of rainfall in the examined region. A periodical overestimation of the frequency of heavy rainfall by GPROF was found. It was argued that this is probably due to a poor calibration over water bodies such as the Amazon River. Rainfall above $20\,\mathrm{mm/h}$ was classified as heavy rainfall and predominantly occurs during the drier months, while rainfall between 5 and $10\,\mathrm{mm/h}$ occurs mostly during the wetter months. During the wetter months the volume frequency of moderate and heavy precipitation is overestimated, while it is underestimated during the dry season. However, during the dry season the underestimation is compensated by an overestimation of light rain in the range of 1 to $10\,\mathrm{mm/h}$, though IMERG-F does not detect light rain between 0.2 and $0.4\,\mathrm{mm/h}$ as good as the radar does. Finally, Oliveira et al.

have noted a positive (negative) bias during wetter (drier) months, for IMERG-F.

The findings of Betts et al. (2019) regarding the influence of observation on the evaluation of gridded precipitation has already been mentioned above. To assess how much station network density affects the measured performance of gridded rainfall data, Tang et al. (2018) generated 500 artificially sparse station networks from a high-density network in a flat basin in southern China. Examining different sub-daily temporal aggregations, Tang et al. found that sparse networks generally underestimated the performance of IMERG-L precipitation and that the results for most evaluation metrics improved with increasing spatial and temporal aggregation. Similarly, O et al. (2017) used a very high station network to analyse two grid cells of IMERG-E, -L, and -F in Austria. For the two cells, 39 and 40 gauge stations were accessible. Over a two-year period between April and October of 2014 and 2015, all IMERG estimates produced higher mean, maximum, and SD values compared to the gridded gauge measurement. Additionally, the percentage of no rain is slightly higher in IMERG. On a daily basis, IMERG tends to overestimate heavy precipitation. Interestingly, also tested the ability of IMERG to model the correct time of a precipitation event on a sub-daily scale. The final run met the time of peak rainfall during a precipitation event better than the other runs; nevertheless, there is a temporal shift between IMERG-F and the observation network. Bases on the IR estimates, this seems to be caused by the backward/forward propagation of clouds (O et al., 2017).

3 Study Areas

In this thesis, two study areas – Germany and a large part of Brazil – are used for two reasons. First, these are the study areas of the two aforementioned research projects, CarBioCial and WP-KS-KW, and second, using two study areas allows different environmental conditions to be considered. Both study areas are described below, and Figure 3 shows the orography of both study areas with the boundaries that were used for the gridded data, which are:

- 5 °E to 16 °E and 47 °N to 55.5 °N for Germany and
- 60 °W to 40 °W and 0 ° to 20 °S for the Brazilian study area.

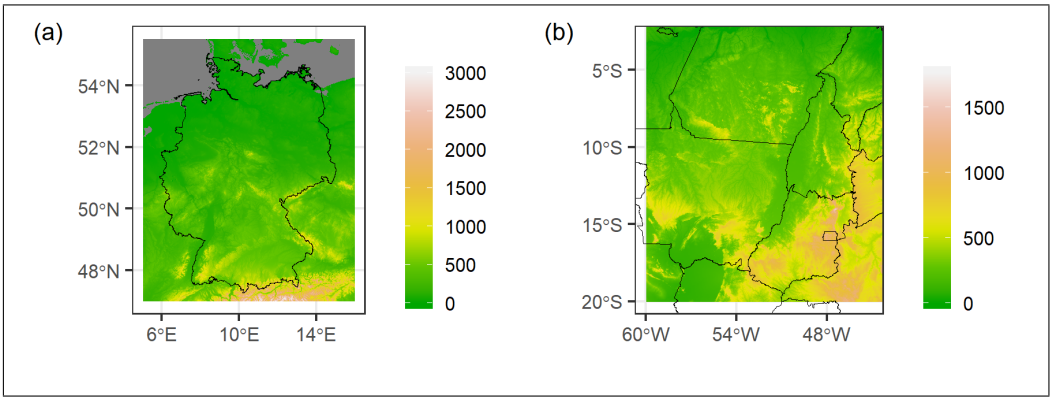

Figure 3: Orography of the German (a) and Brazilian (b) study area in 0.01 ° based on SRTM data (Jarvis et al., 2008). For visibility, altitudes in (a) are limited to 3000 m.

3.1 Study Area – Germany

Germany is characterised in the North by the North German Plain and southward the Central German Uplands have elevations of up to 1500 m. The highest point is the Zugspitze in the German Alps, which is almost 3000 m above sea level (a.s.l). In general, more than half of German land is used by agriculture while forestry uses approximately one third of the land area. Human settlement, transportation infrastructure, industrial land, recreational areas, and cemeteries take up approximately 14% of the area (BfN, n.d.).

The location of Germany between 47 °N and 55 °N on the western part of Europe results in a temperate climate. Basically, the climate in Europe is characterised by air masses from different origins, whereby the North Atlantic Subtropical High (Azores High) and the Icelandic Low, representing the major pressure gradient, mainly influence the origin of the incoming air. The Gulf Stream contributes substantially to the relatively warm climate of Western Europe; thus, in more continental areas, where the effect of the Gulf Stream decreases, the intra-annual temperature variability increases. There are four seasons in Western Europe, namely a moderately warm summer, mild winters, and two broad transition periods. During the winter season, a cold high-pressure zone over Scandinavia and Eastern Europe together with the Azores High and the Icelandic Low influence the weather over Central Europe. The

pressure differences are low or even reversed when the polar jet stream meanders strongly and the often rain-bringing cyclones are routed by the jet stream and its location (Weischet and Endlicher, 2000; Müller-Westermeier, 2006).

Because of the westerlies, predominantly maritime air reaches Germany, but the distance to the ocean and mountains located in the main wind direction increase the continental effects. Precipitation in Europe is subject to seasonality: during winter, the entire eastern part of Central Europe is dryer compared to the western part, which is due to the frequent influence of high-pressure systems. On the other hand, during summer, rain-bringing thunderstorms diminish the west-east gradient, making summer the wetter season in Southern and South-eastern Europe. South of the Danube River, there is twice as much summer rainfall as there is winter precipitation. Focusing on Germany, the uplands are often characterised by many days of fog when located in moisture-bringing winds, which indicates that the uplands are located around the condensation level. A windward-leeward effect can be observed with more clouds on the windward side and cloud dissolution on the leeward side. Additionally, there is an increase in precipitation with elevations, up to 3000 or $3500\,\mathrm{m}$ observed (Müller-Westermeier, 2006; Weischet and Endlicher, 2000). The north face of the Alps, the Black Forest, and Vosges Mountains receive approximately $2000\,\mathrm{mm/yr}$ of precipitation, while other uplands receive approximately 1000 to $1500\,\mathrm{mm/year}$. In the northwest German lowlands, precipitation from 600 to $800\,\mathrm{mm/year}$ occurs, and northeast Germany reeives 500 to $600\,\mathrm{mm/year}$ (Müller-Westermeier, 2006; Weischet and Endlicher, 2000). Figure 4 shows annual and seasonal average precipitation values for the three datasets used in this thesis, the differences and similarities of which are discussed in Section 5.

3.2 Study Area – Brazil

The environmental conditions in Brazil are quite different from those in Germany. The Brazilian study area covers the area between $0°$ and $20°$S and $40°$W and $60°$W. While the northern part is rather flat, in the south-east the Brazilian uplands reach up to $1500\,\mathrm{m}$. The area is characterised by a historical land use gradient and northward expansion of agriculture (Gerold et al., 2018). Nevertheless, the Amazon rainforest holds one third of the earth's tropical forest and is the largest river basin. Hence, it is one of the most important components of the global carbon cycle (Yoon and Zeng, 2010; Nepstad et al., 2014), and the climate in Amazonian also affects the climate of more distant regions (McGregor and Nieuwolt, 1998).

The general circulation in the area east of the Andes is controlled by the intertropical convergence zone (ITCZ). It is characterised by relatively low surface pressure, high temperatures, high rainfall, trade wind convergence and confluence, and thus rising air masses. The zone of low pressure, maximum surface temperature, and surface wind confluence is separated from the zone of maximum cloudiness and rainfall (McGregor and Nieuwolt, 1998). Because of the course of the ITCZ, it is contentiously wet near the equator, but there is a dry season further south. In the continuously wet area, rainfall is favoured by the vegetation,

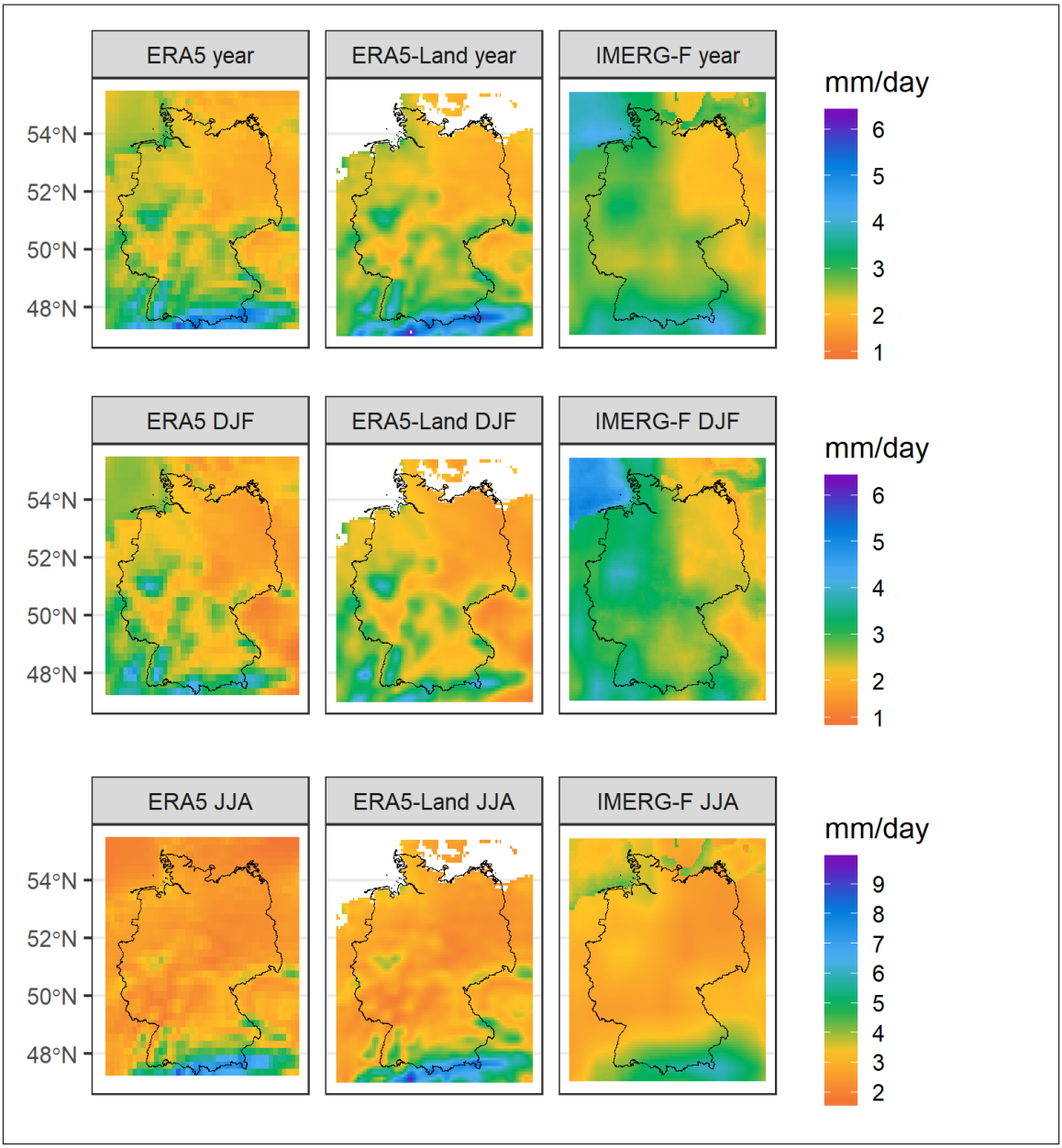

Figure 4: Average precipitation in mm/day for ERA5 (1981-2010), ERA5-Land (2001-2010) and IMERG-F (2001-2010), given in the raw model resolution. Annual average values are listed in the the upper row, winter in the middle row, and summer in the lower row. The scales for annual average and winter are the same, while the scale for summer is different.

which generates huge amounts of water vapour. Likewise, local convection at the coastline and orographic lifting near the Andes foothills produces large amounts of precipitation. The season of maximum rainfall corresponds to the sun position, usually with a delay of one or two months. South of the equator, the dry season is roughly between March and October when southeasterly winds prevail that have lost their moisture along the southeast coast of Brazil and have not yet passed the Amazon Basin. During the rest of the year, northerly winds prevail that have already passed the Amazonia when they reach the south. Although the Amazon Basin, southern Brazil, and parts of northeast Brazil are controlled by the same

driver of rainfall, patterns of intra- and inter-annual rainfall anomalies can be different, meaning that wet anomalies in southern Brazil can correspond to dry anomalies in the Amazon Basin (McGregor and Nieuwolt, 1998). Furthermore, the precipitation in this region is influenced by several mechanisms, such as ENSO, Oscillation, the Madden-Julian Oscillation and the South American monsoon system (Marengo et al., 2016; Shimizu et al., 2017; Alvarez et al., 2016; Liebmann and Mechoso, 2011). The precipitation distribution in the study area for the three datasets used in this thesis is shown in Figure 5. A northwest to southeast gradient exists with $2000\,\mathrm{mm/yr}$ in the northwestern part of the study area and, occasionally more than $2500\,\mathrm{mm/yr}$. In the southeastern part of the study area, annual rainfall is approximately $1300\,\mathrm{mm}$, and in the western part it is roughly $1000\,\mathrm{mm/yr}$.

Land use change is an important issue, especially in Amazonia and the adjacent regions, because the change, mainly in forms of deforestation, changes the local circulation. Results from different studies indicate that the impact of deforestation depends on various effects. Using numerical models, Werth and Avissar (2002) have demonstrated that deforestation reduces evaporation, cloudiness, and precipitation in the corresponding areas. Nobre et al. (2009) have utilised numerical simulation and found that fractional and spatial continuity of deforested areas are important for modulating the local circulation. To evaluate the impact of the size of the deforested patches, Khanna et al. (2017) have used satellite images and numerical simulations. Deforested areas have been found to increase cloudiness and rainfall under limited conditions, because small-scale deforestation triggers thermal meso-scale circulation between pasture and forest. When the deforested patches increase beyond approximately $20\,\mathrm{km}$, the thermally triggered circulation weakens and a redistribution of clouds and precipitation takes place.

Kilian (2017) has implemented land use changes in a regional climate model and found a high spatial variability in the temperature and precipitation response to the land use change. Generally, temperature increased and precipitation decreased thus, the hydrological cycle was weakened. Moreover, Kilian (2017) has found that the trend from induced by land use change is contrary to the trend from climate change and that the effect of the latter is considerably higher. One beneficial characteristic of both satellite precipitation and climate reanalyses is the indirect consideration of such land use change effects because satellite precipitation is based on measured cloud properties and hydrometeors, and reanalyses utilise observations in the data assimilation.

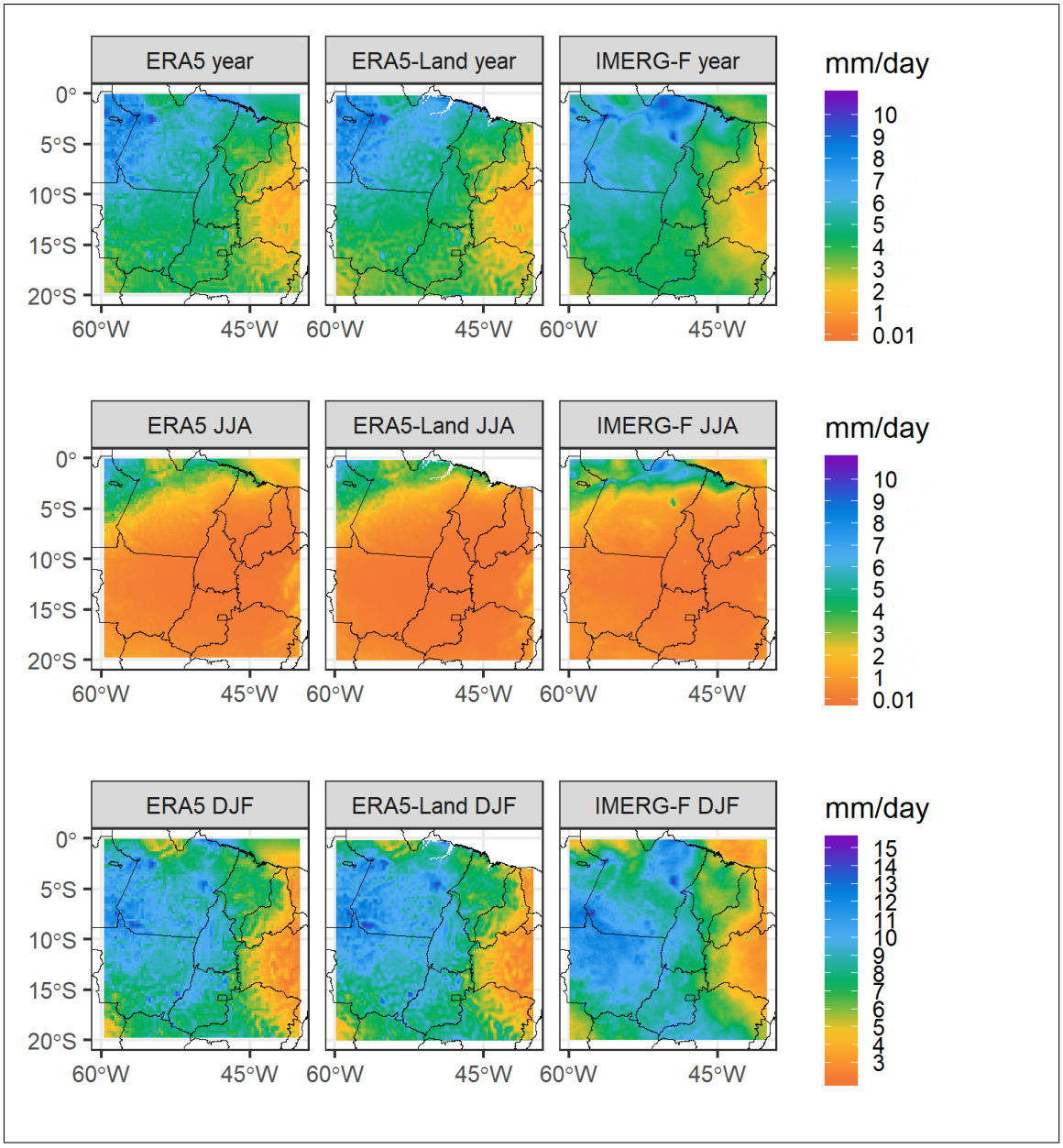

Figure 5: Average precipitation in mm/day for ERA5 (1981-2010), ERA5-Land (2001-2010) and IMERG-F (2001-2010), given in the raw model resolution. Annual average values are listed in the the upper row, austral winter in the middle row, and austral summer in the lower row. The scales for the annual average and austral winter season are the same, while the scale is different for austral summer.

4 Data and Methods

The following section deals with the data used in this thesis, which includes a short remark about the gridded data sets and a more detailed section about station observations. As aforementioned this thesis used ERA5 (Copernicus CCS, 2017), ERA5-Land (Copernicus Climate Change Service (C3S), 2019), and IMERG data (Huffman et al., 2019a). Concerning IMERG-F a data gap of 22 days (2002-01-10 to 2002-01-31) exists in both study areas because of an insufficient data control during data processing for this study.

4.1 Preparation of Station Observations

For the evaluation and model building, daily as well as long-term mean (LTM) daily precipitation values were used. Therefore, two subsets were necessary; one with daily observations that does not necessarily require continuous time series, and LTM values. Finally, those stations were used that were part of both the daily and the LTM subset. The spatial distribution of the 1504 stations used in in Germany and the 407 stations used in the Brazilian study area is shown in Figure A.1 in the appendix. One can see that the number of stations in Mato Grosso and Pará is rather low, but moving eastward in the direction of more populated areas, the number of stations increases.

The use of German station observations does not require much effort. The national German weather service Deutscher Wetterdienst (DWD) provides free access to numerous stations after a law amendment in 2017. In particular, the precipitation network has a high density and measurements are already quality controlled. The level of quality control depends on the classification as recent or historical. Data classified as historical has undergone more intense quality control than the recent observations; thus, this thesis only uses historical observations. Moreover, precipitation data is contained in two different collectives – the KL and the more_precip collective –. The DWD indicates that if stations are contained in both collectives, the quality of the KL collective is better. However, because the more_precip collective contains distinctly more stations and because a formal quality control is applied in any case, the more_precip collective was used here. The DWD automatically checks the data for consistency and for gross errors, but it does not apply systematic correction or homogenisation procedures (DWD, 2018). More information on the quality control can be found in Spengler (2002) and Kaspar et al. (2013). Detailed station metadata and information on relocation are likewise available. The LTM data was obtained from DWD (n.d.).

In Brazil, different agencies theoretically provide met. station data, including the Brazilian national meteorological service Instituto Nacional de Meteorologia (INMET), the national water agency Agencia Nacional De Aguas (ANA), and Instituto Nacional de Pesquisas Espaciais (INPE)/Centro de Previsão do Tempo e Estudos Climáticos (CPTEC). The number of actually usable sources and thus the number of stations is certainly smaller. Considering the access to data, the Lei de acesso à informação (Access to Information Act) has made access

to more stations possible, but unrestricted and easy access is still not possible. For instance, the Brazilian institutions provide no bulk download. However, the R package *inmetr* Tatsch (2019) allows easy download of INMET data. Nonetheless, an authorisation is still necessary. To download station observation from the ANA hidroweb interface does not require an authorisation, but a bulk download is neither possible. However, ANA provides observations from different sources including stations from the INMET network as well as stations from different energy companies, municipalities, and public and local authorities (Agência Nacional de Águas, 2019).

Even though a high number of stations can be obtained from the aforementioned sources, most of the data is unusable if longer time series are needed, because measurement periods are often just a few days or weeks. Furthermore, there are some major problems with the Brazilian stations. As Liebmann and Allured (2005) have stated, a test of quality can be difficult for South American precipitation observations because it is impossible to distinguish if a reported $0\,\mathrm{mm}$ event represents a missing value or if a missing value is actually a $0\,\mathrm{mm}$ rainfall value. Another issue is the occurrence of unexpected and implausibly high values, which is the reason why data from INPE/CPTEC was not used in this study. After a long series of reports of $0\,\mathrm{mm}$ precipitation, very high values appeared at irregular intervals, which are presumably hidden accumulated values. The often used Global Summary of the Day was not used because of long series of $0\,\mathrm{mm}$ rainfall. The accumulation to annual sums resulted in values that do not nearly meet value that could be considered correct. This is a known issue (Funk et al., 2015). Another problem is that there is hardly any information about the time of measurement (e.g., UTC or local time). Moreover, a rainfall event reported for 12:00 UTC might mean that the measured amount was either reported on the day the data was read or on the previous day, suggesting that the majority of rain occurred on the previous day. All these issues have caused difficulties in former studies (Liebmann and Allured, 2005; Arvor et al., 2014; Xavier et al., 2015; Alvares et al., 2013).

Nevertheless, some constraints were applied to the observations in this study. Generally, from the ANA collective, only those stations are considered that the ANA has marked as consistent. Nonetheless the sum of $0\,\mathrm{mm}$ reports and missing values is often almost 100%, which can cause errors in the later event-based analysis. The application of a threshold to this problem seems difficult because of seasonality and spatial variability of precipitation in the Brazilian study area. Additionally, following Liebmann and Allured (2005) and Xavier et al. (2015), rainfall values above $450\,\mathrm{mm/day}$ were removed. Information about station relocation and other metadata were not available. Furthermore, the precision of coordinates varies from two to four decimals, which would make a test for duplicated stations difficult, as is also discussed in Liebmann and Allured and Xavier et al.. For some stations from the ANA altitude information were missing and were thus added from a $0.0083°$ SRTM-based digital elevation model (DEM) (Jarvis et al., 2008).

In addition to the aforementioned constraints, the following restrictions were used to build a selection of stations that can provide daily values as well as LTM daily precipitation for

the period 1981 to 2010. As previously mentioned, the first constraint was the consistency; additionally only those months, and the days in the respective month were considered where information about the number of rainy days were already provided in the raw data, because otherwise it was assumed that some information in the respective entries had to be wrong. A further restriction was the requirement of at least 10 years of measurements. These daily ANA stations were also used to build the LTM subset. Instead of using daily INMET stations for the LTM subset, monthly values directly obtained from INMET could be used. Again, months without information regarding the number of rainy days were removed, and only those stations that had at least 10 entries for each of the 12 calender months were considered. For the subset of daily values, the same constraints as for the ANA stations were applied.

4.2 Altitude Adjustment for Precipitation

Several studies named in Section 2 have demonstrated that the quality of gridded precipitation data in complex terrain is regularly lower than in flat terrain, which is often accompanied by an elevation-dependent bias. This dependency affects the results when datasets from different grid systems are compared or when met. stations are used for the validation. The effects of topography and elevation have been known for a long time, so Sawyer (1956) found increased rainfall on elevated areas and windward sides. Brunsdon et al. (2001) have stated that many studies have found that linear models fit well for middle latitudes and relatively small areas. In consulting older studies, they found different slopes between $1.5\,\mathrm{mm}/100\mathrm{m}$ and $4.5\,\mathrm{mm}/100\mathrm{m}$ in Great Britain based on met. stations. Additionally, Brunsdon et al. have noted that their findings are highly influenced by the distribution of stations and local effects. Moreover, they have suggested that it is advantageous to use area average elevation and precipitation instead of actual stations´ elevation to build models. Additionally, Ceccherini et al. (2015); Goovaerts (2000) and Berndt and Haberlandt (2018) have used elevation as covariate to predict precipitation. Daly et al. (1994) have found that a linear relationship between precipitation and elevation is often assumed, because it is easy to use and an acceptable approximation even though under certain conditions the relationship is better described by a log-normal or exponential relationship. Moreover, Daly et al. have argued that orographic effects estimated from relatively coarse DEMs (2 to $15\,\mathrm{km}$) are often more highly correlated to local precipitation than point-based orographic features are.

Here it is argued that precipitation evaporates to some extent while it is falling; therefore, elevation differences are also differences in precipitation evaporation. Additionally, as mentioned in Section 2.2, it is known that the evaporation in ERA5 is insufficiently represented (ECMWF, n.d.b). Thus, considering the elevation-dependent bias in gridded precipitation data and the elevation dependency of precipitation, this thesis attempts to correct precipitation in terms of the effect of elevation. The direction of the correction or the adjustment towards any given elevation is possible.

As illustrated in Figure 7, the classification of elevation data and the average precipitation

in these classes, for both the gridded datasets and the met. stations in Germany have quite similar pattern. Moreover, the differences between annual, summer, and winter data are small. Beginning from the altitude class of approximately 1000 to 1100 m the relationship ceases and there is no longer any characterisable change of precipitation with altitude. Additionally, the seasonal differences become more pronounced. The altitude where the relationship ceases roughly corresponds to the average cloud base height (CBH) of ERA5. On the one hand there are two things that relativise this match. First, the CBH is given in metres above the model orography and thus the average CBH above sea level (a.s.l) is roughly 1400 m, and second, above the 1000 m class, the number of grid cells per class decreases below 10. On the other hand observations and ERA5-Land show the same elevation-precipitation pattern and ERA5-Land has more than 30 grid cell per class up to 1500 m and still 23 grid cells in the class 1500 to 1600 m. The average CBH values are only averaged from wet-days.

Regarding the Brazilian study area, as shown in Figure 8, no increase of rainfall with elevation was found. Furthermore, stations located lower recorded higher rainfall values. This might be due to grid cells in the northern part of the study area where elevation is low and annual precipitation is high, while elevated areas further south receive less rainfall. Moreover, rainfall in the tropics is often convective with limited spatial extent. Thus, one could argue that the examined area in Brazil is too large but the pattern from Figure 8 does not change notably when only a limited area between 13 °S and 20 °S and 40 °W and 50 °W is considered which is roughly the area of Goías and Minas Gerais. In this area, there are elevation differences in a region of more or less the same seasonality. However, Ragette and Wotawa (1998) have found low evaporation in tropical regions and the IFS documentation mentions that that there is less evaporation in high-intensity rain events compared to events of low intensity (ECMWF, 2016b). Nevertheless, in the Brazil study area, the modelled and observed values are close together. Seasonal differences are more pronounced than in Germany and become clearly visible at approximately 600 m. In agreement with this, areas with such elevation are in the southeast of the study area where a pronounced seasonality exist.

The proportion of CP in ERA5 in the Brazilian study area is almost always high, roughly 2% of the precipitation events in 2001 and 2002 had a CP proportion of less than 25%, and only around 15% of the events had a CP proportion of less than 75%. In Germany, the variability is higher, roughly 30% of the precipitation events had a CP proportion of less than 25%, and roughly 65% of the events had a CP proportion of less than 75%. Thus, a differentiation of the convection fraction was only done for the German study area, resulting in a convective proportion of more than 75% or less than 25%. Somewhat unexpectedly, there were no distinct differences in the altitude precipitation patterns of the two subsets, as shown in Figure 6 (b), which also demonstrated the cumulative percentage of CP proportion for wet days in the Germany and Brazilian study area. Because of the similar pattern in Figure 6 (b), no further differentiation was made between the precipitation types. However, the differentiation of precipitation depends on the ERA5 parameterisation.

The patterns found for Germany and Brazil are also prevalent in the IMERG-F data (cf.

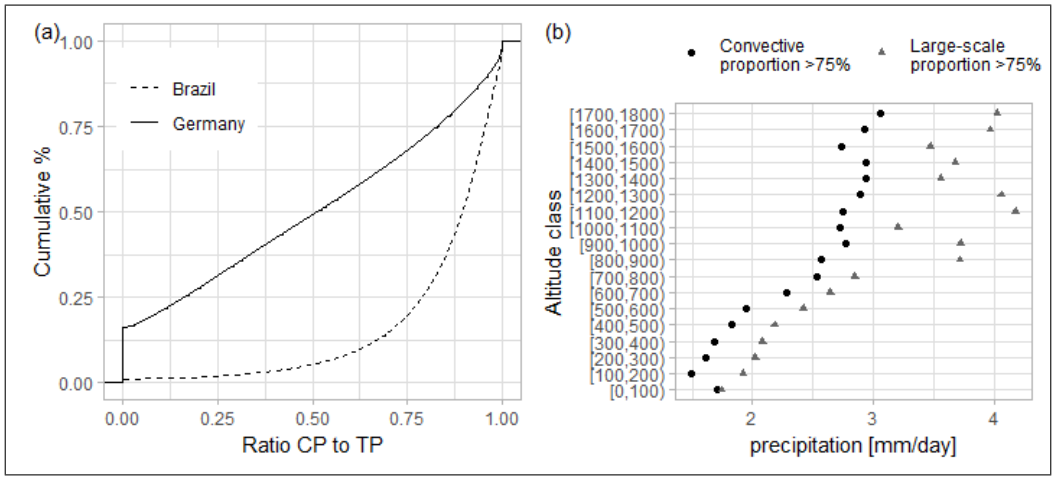

Figure 6: Cumulative distribution of the CP to TP ratio for the 2001-2002 period for both study areas (left) and average daily precipitation for altitude classes for precipitation with a proportion of CP $>$ 75% or CP $<$ 25%, averaged for 1981 to 2010.

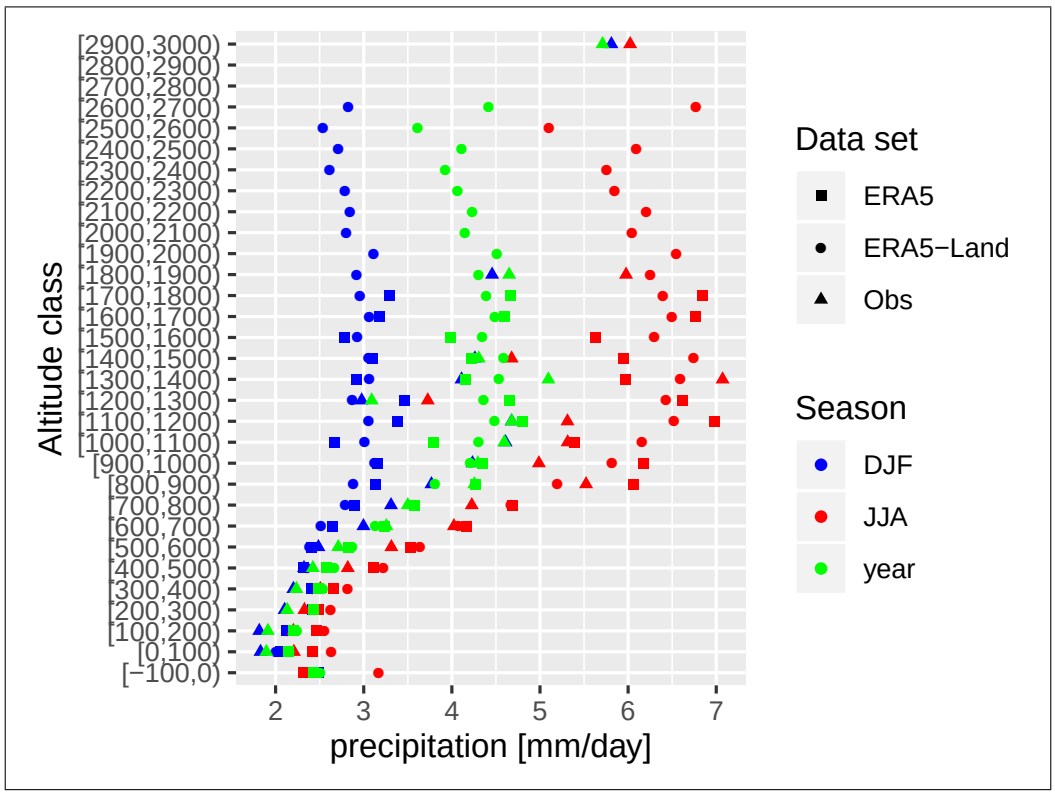

Figure 7: Distribution of altitude and average daily precipitation for met. stations, ERA5, and ERA5-Land in Germany for summer, winter, and the annual average.

Fig. A.2) when plotting the precipitation values against a DEM remapped to the IMERG grid. However, since elevation is not part of the IMERG data and has no influence in the IMERG algorithm, elevation cannot be used for an altitude adjustment. To a certain extent, the correction with GPCC data in the final run product incorporates some elevation information, but the effort to infer corresponding altitude information would be very elaborative and it is unclear if this would be expedient. Therefore, IMERG-F was not considered further for the altitude

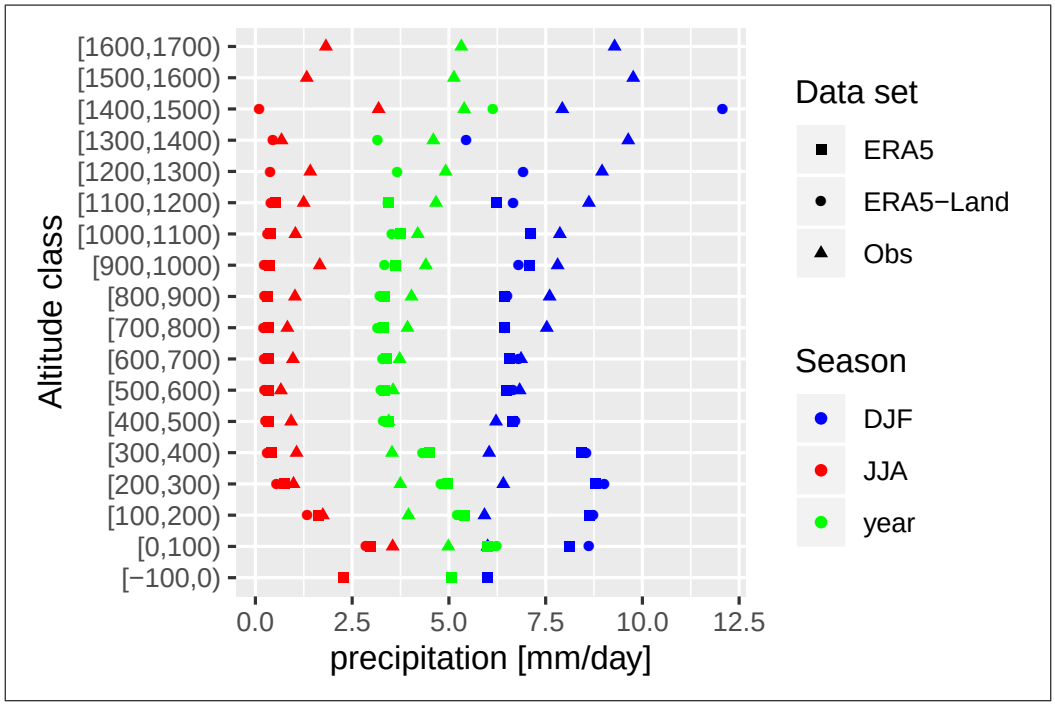

Figure 8: Distribution of altitude and average daily precipitation for met. stations, ERA5, and ERA5-Land in Brazil for summer, winter, and the annual average.

adjustment.

Since it would be impossible to reproduce the evaporation rates calculated in ERA5 because not all of the necessary variables are part of the accessible output, (e.g., precipitation flux at CBH) and furthermore, other values based on several dependencies could not be reproduced, the idea is to account for the effect of altitude statistically. It is argued that due to averaging several grid cells into one altitude class, in any case underestimated topographic effects such as windward and leeward effects are even more balanced. First, a linear regression with the LTM precipitation data of ERA5 (1981 - 2010) and ERA5-Land (2001 - 2010) was used to find the slope of how much precipitation increases per altitude. The results from the linear model are listed in Table 2. ERA5 has altitudes that are approximately between 0 m and 1800 m, in the Germany study area, and the minimum number of cells per altitude class in this range is three whereby for classes up to 1000 m there is no class with less than 10 grid cells per class. The slope of the linear regression reveals a precipitation increase of $0.15\,\mathrm{mm}/100\mathrm{m}$. Using ERA5-Land and likewise limiting the altitude classes to 1800 m shows a similar slope, though the coefficient of determination increases from 0.82 (ERA5) to 0.89 (ERA5-Land), referred to as ERA5-Land (a) in Table 2. The better values probably result from the higher number of grid cells in ERA5-Land, (e.g., the minimum number of grid cells in any altitude class is 23). When altitude classes of ERA5-Land are not limited but the only condition is more than three grid cells per class, the highest class reaches 2500 m, and is referred to as ERA5-Land (b) in Table 2. For this setting, the linear model shows a decreased R^2 of 0.79 and a lower slope of $0.11\,\mathrm{mm}/100\mathrm{m}$. The lower slope and lower R^2 can be explained by the ceasing of the relationship between altitude and precipitation from approximately 1000 m.

Table 2: Results from linear and non-linear models for long-term average precipitation and altitude. *** indicates p-values < 0.001.

	Model	Slope mm/100m	R^2	Sign.
Linear Model	ERA5-Land (a)	0.16	0.89	***
	ERA5-Land (b)	0.11	0.79	***
	ERA5	0.15	0.82	***
Non-Linear Model	ERA5		0.83	***

All of the regression results were statistically significant and the explained variance is quite high. Thus it seems reasonable to use a linear model to account for altitude differences. The model to adjust a given precipitation value, here daily LTM precipitation $\bar{p}_{model}$ at model height z_{model} to a target altitude z_{target} with model slope S_l in mm/m is:

$$\bar{p}_{lm} = \bar{p}_{model} + (z_{target} - z_{model}) * S_l \tag{4}$$

The adjustment is only done for z_{target} values below the ERA5 CBH. If the target altitude is above the condensation level, unadjusted model precipitation was used just to be consistent in the evaluation. An estimation of precipitation variation in and above the condensation level was not possible. To receive daily values, the model could be directly applied to daily values since the slope is estimated on average daily precipitation, but because the model was fitted to LTM data, daily anomalies were used to generated daily altitude-adjusted values for day i, where $\bar{p}$ is the LTM ERA5 or ERA5-Land precipitation and $\bar{p}_{lm}$ is the corresponding altitude-adjusted LTM value.

$$\begin{aligned} p_{lm,i} &= \frac{p_{model,i}}{\bar{p}_{model}} * \bar{p}_{lm} \\ &= \frac{p_{model,i}}{\bar{p}_{model}} * \bar{p}_{model} + (H_{target} - H_{model}) * S_l \end{aligned} \tag{5}$$

In the two research projects mentioned in this thesis a model was developed to distinguish areas below and above the condensation level to take different evaporation rates into account. The original version of the non-linear model (NLM) assumed a slight reduction of precipitation per increasing distance above the CBH, which was assumed to be the condensation level. However, in this study ERA5 data was used to fit the NLM, and no grid cell was above the average CBH, which is why the model was only fitted for the part below CBH. The ratio between model or target elevation and CBH is used to account for elevation differences embedded in a compound factor. First, $\bar{p}_{model}$ is divided by the denominator of Equation 6 to reduce $\bar{p}_{model}$ to the sea level elevation. Second, $\bar{p}_{model}$ at sea level is multiplied by the numerator to adjust

the precipitation to the target elevation.

$$\bar{p}_{nonlm} = \bar{p}_{model} * \frac{1 + F * (z_{target}/\overline{CBH}_{target})^U k}{1 + F * (z_{model}/\overline{CBH}_{model})^U k} \tag{6}$$

When other or better information about the CBH exist, they could be used in the numerator of eq. 6. In this study ERA5 CBH is used in the numerator and the denominator. The factors F and U_k were fitted to the LTM data using non-linear least-square estimates, which resulted in $F = 1.4375$ and $U_k = 0.8619$. For the mean CBH only days with precipitation were considered. The results of the model fit are listed in Table 2, which shows that the relationship between altitude and precipitation in ERA5 is already well explained by a linear model, and the NLM does not add much to the explained variance. In contrast to the linear model, the NLM is applied directly to daily values (Eq. 7), suggesting that daily CBH is highly variable. When on a certain day with precipitation the corresponding CBH was below z_{target} instead of applying the NLM, the original value was used.

$$p_{nonlm,i} = p_{model,i} * \frac{1 + F * (z_{target}/CBH_{target,i})^U k}{1 + F * (z_{model}/CBH_{model,i})^U k} \tag{7}$$

At least from two different points of view, the adjustment is not linear. First, the adjusted precipitation is the product of modelled precipitation and the adjustment factors and therefore depends on the modelled precipitation itself. This means that for otherwise constant values, higher precipitation rates receive a higher adjustment. This contradicts the assumption of less evaporation on high precipitation intensities. Second, the fit to the LTM data results in a higher rate of change near CBH compared to further from the same. Additionally, adjusting towards the CBH results in a larger change than away from CBH because the larger z_{target}/CBH_{target} the stronger the effect of exponent U_k. While precipitation theoretically never disappears or evaporates completely in the NLM because reduction is based on a quotient it does evaporate completely in the linear model (LM). However, in the evaluation all precipitation values below $0.1\,\mathrm{mm/day}$ are set to 0, and using daily anomalies for the daily LM adjustment likewise results in an incomplete evaporation.

Taking another look at the CBH and its effect on precipitation in Germany, one can admit that the average CBH is closer to the model topography in elevated areas, such as over the Eifel, Hessisches Bergland, Erzgebirge, or the Alps. In these areas, which are also regions of increased precipitation (cf. Fig. 4), the average CBH is roughly 500 to 1000 m above the model orography. In the remaining area CBH is notably higher than 1000 m above model orography. A spatial visualisation of the average CBH above model orography can be found in Figure A.3. For three stations in Germany – Hamburg-Fuhlsbüttel in the north Germany plain; Brocken, a centrally located mountain of roughly 1100 m; and the 3000 m summit Zugspitze the highest mountain in Germany located in the south – precipitation is compared to the CBH on wet days

in Figure 9. Only CBH classes with more than 30 events were considered in that figure that likewise shows three Brazilian stations – Altamira, which is near the equator in central Pará; Sinop, approximately 350 m a.s.l, south of Amazonia in central Mato Grosso; and Alto Paraíso de Goiás in northern Goías, which is roughly 1000 m a.s.l. For all stations except Hamburg a strong decrease of precipitation with increasing CBH was found. Additionally, in Hamburg, precipitation decreases with higher CBH on an annual basis. Looking at the Brazilian stations, it is apparent that two southern stations there is only a little rainfall during austral summer with elevated CBH and no clear pattern. Although there is a general pattern without further analysis it is impossible to say how much of this effect should be attributed to evaporation and how much to the fact that a higher CBH is cooler and has a lower ability to hold precipitable water. Additionally, the decrease at all stations is rather steep and would result in a much steeper slope than estimated by the linear regression against the altitude.

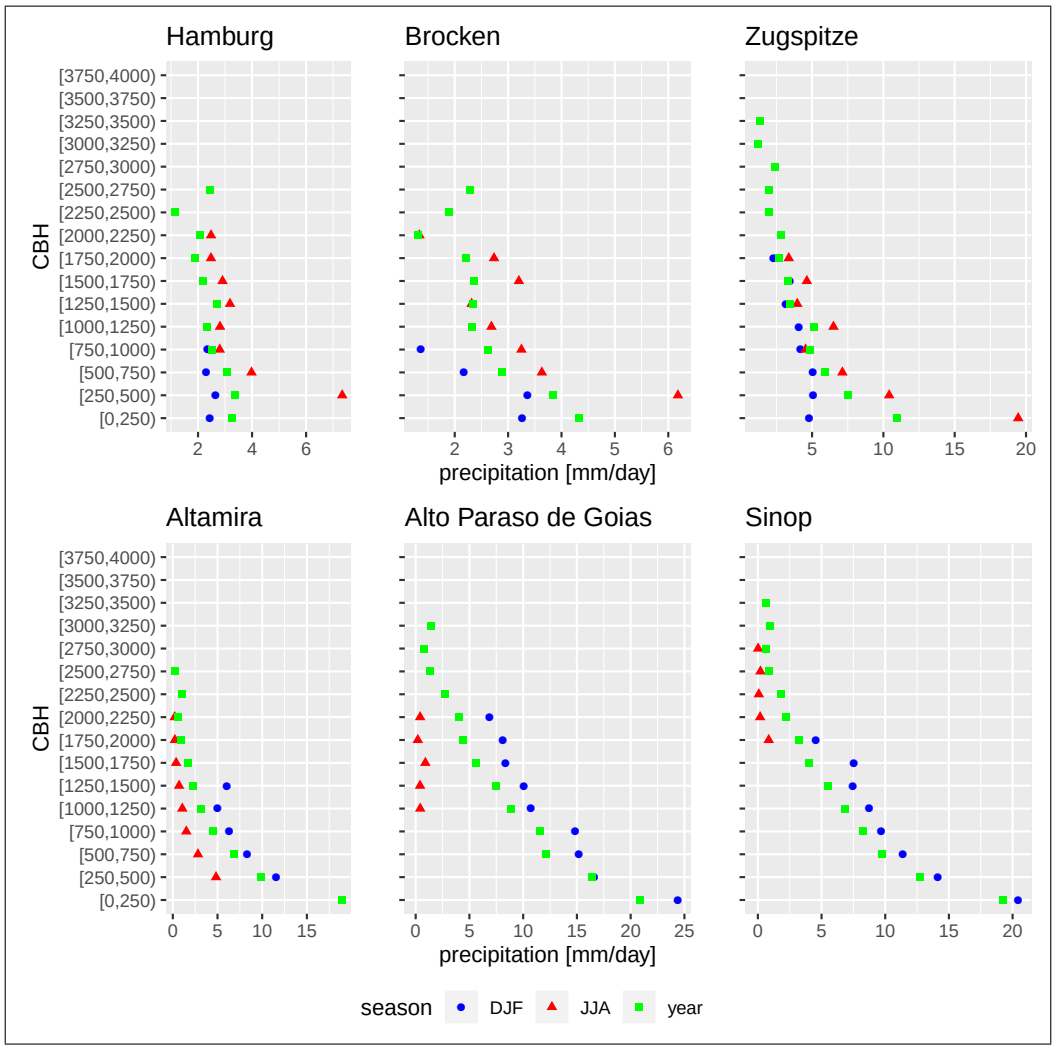

Figure 9: Averaged daily mean precipitation from 2001 to 2010 for classes of cloud base heights. The upper row shows three stations in Germany and the lower row three stations in the Brazilian study area for summer, winter, and the annual averages.

4.3 Verification Measures for Binary Events

In addition to verification measures for continuous variables, some verification measures for binary events were used and are briefly introduced below. Not all of the mentioned indices are used in the evaluation, but they partly depend on each other, so they are mentioned for reasons of completeness. Hits or true positives are tp, false alarms or false positives are fp, misses or false negatives are fn, correct rejections or true negatives are tn, and the sample size is n. The base rate or event probability s produces the unconditional probability of the observed occurrence of the event, here precipitation, and is thus not a measure of the performance.

$$s = \frac{tp + fn}{n} \tag{8}$$

The probability of modelled rainfall is expressed by the probability of occurrence PO:

$$PO = \frac{tp + fp}{n} \tag{9}$$

and can also be written as $PO = (1 - s)F + s * H$ where H is the hit rate, sometimes also called POD, which indicates the proportion of correctly forecasted events (O et al., 2017; Mason, 2003).

$$H = \frac{tp}{tp + fn} \tag{10}$$

The false alarm rate (F), sometimes also called probability of false detection, is the ratio of incorrectly modelled non-occurrence of events.

$$F = \frac{fp}{fp + tn} \tag{11}$$

F is different from the false alarm ratio (FAR), which is the proportion of modelled occurrences that do not correspond with an actual event. Because FAR can vary between 0 and $1 - S$, it does not necessarily carry any information about skill (Mason, 2003).

$$FAR = \frac{fp}{tp + fp} \tag{12}$$

The frequency bias frequency bias (FB) is the ratio between modelled and observed rain-

fall events:

$$FB = \frac{tp + fp}{tp + fn} \tag{13}$$

Alternatively, FB can be expressed as $FB = \frac{1-s}{s}F + H$. A perfect model would have values of $FB = 1$, $H = 1$ and $F = 0$, while a model with no skill to reproduce event occurrences would have $H = F$ with $FB = H/s$ or $FB = F/S$.

A further measure is the proportion correct (PC), which is

$$PC = \frac{tp + tn}{n} \tag{14}$$

or $PC = (1 - s)(1 - F) + s * H$.

Additionally, the Heidke Skill Score (HSS) and CSI are frequently used. Heidke Skill Score takes into account the proportion of forecasts that would have been correct by chance and adjusts the PC by this effect.

$$HSS = \frac{PC - E}{1 - E} \tag{15}$$

with

$$PC = \frac{tp + tn}{n} \tag{16}$$

or $PC = (1 - s)(1 - F) + s * H$ and

$$E = \frac{tp + fn}{n}\frac{tp + fp}{n} + \frac{fp + tn}{n}\frac{fn + tn}{n} \tag{17}$$

The CSI is expressed as follows:

$$CSI = \frac{tp}{tp + fp + fn} \tag{18}$$

(Mason, 2003). It can also be written as $CSI = H/(1 + (F(1 - s))/s$ with a perfect skill of $CSI = 1$, when $H = 1$ and $F = 0$, the minimum value for CSI is 0. However, the zero skill can lie between 0 ($H = F = 0$) and $S(H = F = 1)$ (Mason, 2003). Because the CSI is not affected by tn (the same is true for H, FAR, and POD), it has often been used to measure the performance when event occurrence is low. However, it depends strongly on s. Dealing with rare events can be problematic because the frequency of tn can be much larger than the other three elements when every non occurrence of a rare event is correctly forecasted. On the other hand, when non-occurrence is rare, the model should be rewarded for correct rejections, which is for instance not the case for CSI (Mason, 2003). Ferro and Stephenson (2011) have

criticised the existing verification measures as deficient for rare events and introduced two extremal dependency indices that are independent from the base rate. Nevertheless, HSS (and CSI) are still frequently used also in studies with seasonal precipitation patter and thus the HSS is used in this thesis.

5 Results

Before the results are discussed in detail, it should once again be noted that the altitude adjustment was not applied in the Brazilian study area, and it was not applied to IMERG-F data. Moreover, the comparison of the LTM values is based on the period 2001 to 2010 for IMERG-F and ERA5-Land and 1981 to 2010 for ERA5. For the evaluation of daily values, the period 2001 to 2010 was used for all datasets.

5.1 General Results for Germany

The spatial distribution of annual and seasonal precipitation in Germany is quite similar in ERA5 and ERA5-Land, but some marked differences exist in IMERG-F (c.f. Fig. 4). During winter, IMERG-F is generally wetter, which is likewise visible for the annual values. Moreover, the spatial distribution seems rather smooth and does not have any features, which stands in contrast to the ERA reanalyses. Figure 10 shows a temporal comparison of weekly accumulated precipitation at the Hamburg-Fuhlsbüttel and Brocken stations. During some weeks, larger differences occurred, such as, in Hamburg around Week 33 in 2006, and at Brocken in Week 47 of 2010 at Brocken. Moreover, especially for Hamburg, many weeks in ERA5-Land had lower precipitation values than in IMERG-F, which becomes clearly visible if one looks at the winter weeks 46 through 51. Moreover, there is a small data gap at the beginning of 2002, as already mentioned.

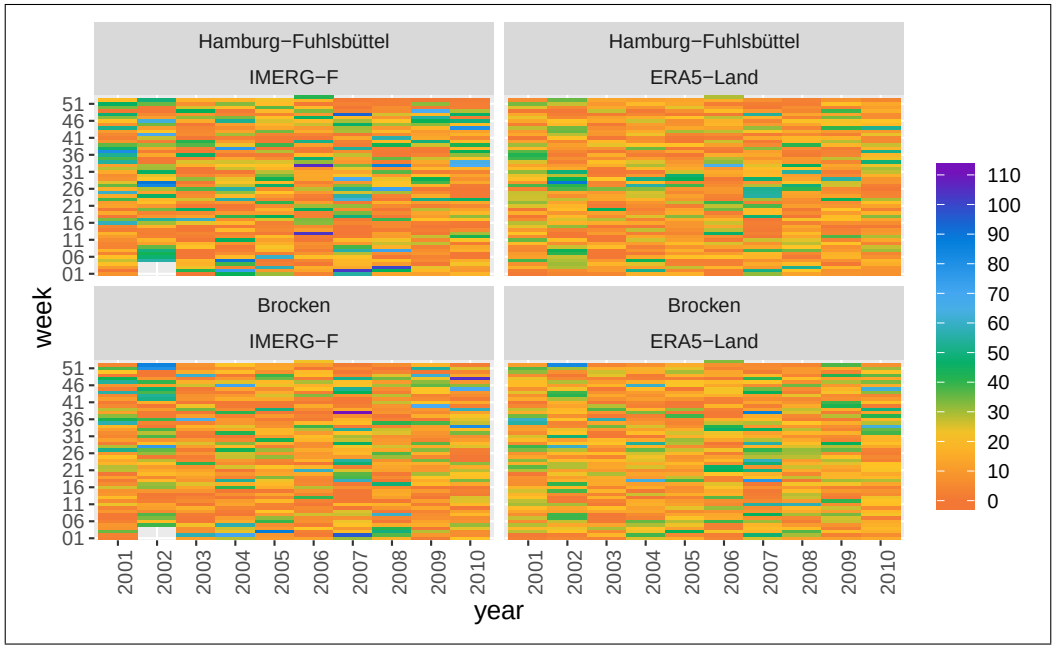

Figure 10: Temporal comparison of weekly aggregated IMERG-F and ERA5-Land precipitation [mm/week] for the Brocken and Hamburg-Fuhlsbüttel stations in Germany.

Focusing on the comparison with LTM met. station data, the three gridded precipitation products had higher daily precipitation values than at the 1504 stations used in Germany. The overestimation was higher in IMERG-F than in the ERA reanalyses. Because the majority of

stations are located in terrain with low elevations virtually no effect of the altitude adjustment is visible from the average values. Looking at the 14 stations above 1000 m, a negative bias occurred for all datasets, with the highest (negative) bias for IMERG-F. The altitude adjustment had a positive effect on both ERA reanalyses and reduced the bias. Because station elevation itself does no say anything about elevation differences between model and station elevation, it is distinguished between stations above the model orography, called Type 1 stations, and stations below model orography, called Type 2 stations. If not otherwise stated model orography refers to the ERA5 elevation and Type 1 (2) stations have a difference of at least 250 m. The rather high difference was chosen to have distinct elevation differences, presuming more differentiated differences between Type 1 and Type 2 stations.

For Type 1 (2) stations the bias was negative (positive) over all datasets other than IMERG-F for Type 2 stations, which were also slightly negative. The adjustment of ERA5 to the elevation of ERA5-Land did not approximate the biases because these were quite similar before, and the altitude adjustment led to a larger spread. Nevertheless, for both station types the adjustment to station elevation reduced the overall bias. Unfortunately, the number of stations with notable elevation differences from model orography is rather low at 26 (22) Type 1 (2) stations. Table 3 lists the described values with the elevation difference. The part in parentheses in Table 3 indicates the LM or the NLM and an adjustment to the ERA5-Land (ERA5-L) or station (Stn.) elevation.

Table 3: Long-term average daily precipitation values and the corresponding biases for different subsets in Germany where elevation differences between model orography and station elevations are at least 250 m. ERA5-L refers to ERA5-Land.

	All Stations	Stations above 1000 m	Type 1 Stations	Type 2 Stations
		average P		
Observation	2.3	4.5	3.8	4.0
		bias		
IMERGF-F	0.4	-1.3	-0.9	-0.5
ERA5-Land	0.2	-0.6	-0.6	0.3
ERA5	0.2	-0.6	-0.6	0.6
ERA5 (LM, ERA5-L)	0.2	-0.3	-0.4	0.4
ERA5 (LM, Stn.)	0.2	0.0	0.0	0.1
ERA5-L (LM, Stn.)	0.2	-0.3	-0.4	0.1
ERA5 (NLM, ERA5-L)	0.2	-0.3	-0.4	0.3
ERA5 (NLM, Stn.)	0.2	0.1	0.1	-0.1

Investigating daily values in Germany generally led to the same pattern and improvement. In Table 4, P.1 (precipitation) and bias.1 denote values that consider the entire times series, including observed dry days where the model produced precipitation, (i.e. false positive events). For bias.2 and MAE.2, only those days are considered where the observed precipitation was above a given threshold while P.2 was calculated separately for the observations and modelled

data when the given threshold was exceeded. If not otherwise stated, bias refers to bias.1. For P ≥ 0.1 mm over all stations the modelled average daily precipitation (P.1) was higher than observed at the stations, and P.2 for both ERA reanalyses was lower than the observed P.2. In both cases, IMERG-F had higher P values than the observation and than the reanalyses. The higher bias.1 compared to bias.2 for P ≥ 0.1 mm, indicates more modelled than observed wet days and a good agreement of observed and modelled P for the days where P actually occurred. For higher rainfall intensities of P ≥ 10.0 mm) (roughly the 90^{th} percentile over all daily observation), which are listed in the upper right side in Table 4, reanalysis precipitation had a strong negative bias. Likewise, the bias of IMERG-F became negative, but not as strongly as that of ERA5 and ERA5-Land. Furthermore, when only days with P ≥ 10.0 mm are considered for each of the respective datasets, IMERG-F still overestimates the average rainfall, and ERA5 and ERA5-Land show lower average precipitation.

Taking into account elevation and elevation differences, (lower row of Tab. 4), there is a general underestimation of modelled precipitation for stations above 1000 m. Likewise, Type 1 stations underestimated average daily P, while Type 2 stations overestimated the same. In all three elevation-considering cases, IMERG-F had the highest bias. The bias of ERA5-Land was slightly below the bias of ERA5, and the altitude adjustment for the ERA reanalyses reduced the bias in all cases. For stations above 1000 m and ERA5 and LM adjustments to station elevation, the bias became 0, which mainly represents the adjustment of the LM. The bias of ERA-Land remained slightly negative after the altitude adjustment, and adjusting ERA5 to the elevation of ERA5-Land did not result in a more comparable bias. Correcting the elevation bias of the gridded precipitation by the LM took the model bias, both bias.1 and bias.2, closer to 0 than the NLM, though at Type 2 stations the effect of the NLM was greater, which turned the positive bias into a slightly negative one. The MAE.2 was not notably affected by the altitude adjustment. The reason for the resemblance between stations above 1000 m and the Type 1 stations partly result from 11 stations that are part of both subsets, but no Type 2 station is located above 1000 m.

Table 4: Mean daily precipitation, biases and MAE, for different subsets in Germany. Mean daily precipitation p.1 and bias.1 include dry days in averaging, p.2 and bias.2 include only those days, when the precipitation threshold is met for observations.

	All Stations, P $\geq$0.1 mm					All Stations, P $\geq$10.0 mm				
	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2
Met. Stn. Observation	2.3		4.6					16.6		
IMERG- F	2.7	0.4	5.1	0.3	4.1			19.1	-2.7	10.0
ERA5-Land	2.5	0.2	3.6	0.0	3.0			14.9	-5.7	7.4
ERA5	2.5	0.2	3.6	0.0	3.0			14.9	-5.7	7.4
ERA5 (LM, ERA5-L)	2.5	0.2	3.6	0.0	3.0			14.9	-5.7	7.4
ERA5 (LM, Stn.)	2.5	0.2	3.6	0.0	3.0			14.9	-5.8	7.4
ERA5-L (LM, Stn.)	2.5	0.2	3.6	0.0	3.0			14.9	-5.8	7.4
ERA5 (NLM, ERA5-L)	2.5	0.2	3.6	0.0	3.0			14.9	-5.8	7.4
ERA5 (NLM, Stn.)	2.5	0.2	3.5	-0.1	3.0			14.8	-6.0	7.5

	Stations above 1000 m P $\geq$0.1 mm (n=14)					Type 1 stations P $\geq$0.1 mm (n=26)					Type 2 stations P $\geq$0.1 mm (n=22)				
	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2
Met. Stn. Observation	4.6		8.2			3.9		6.8			3.9		7.4		
IMERG- F	3.2	-1.4	5.7	-3.0	6.4	2.9	-1.0	5.4	-2.1	5.4	3.5	-0.4	6.0	-1.4	5.8
ERA5-Land	4.0	-0.5	5.5	-1.4	5.1	3.3	-0.6	4.5	-1.4	4.2	4.4	0.5	6.0	0.3	4.7
ERA5	3.9	-0.7	5.4	-1.7	5.1	3.2	-0.7	4.4	-1.6	4.2	4.5	0.6	6.2	0.5	4.9
ERA5 (LM, ERA5-L)	4.2	-0.4	5.7	-1.2	5.1	3.4	-0.5	4.7	-1.1	4.2	4.3	0.4	5.9	0.1	4.7
ERA5 (LM, Stn.)	4.5	0.0	6.2	-0.6	5.2	3.8	-0.1	5.2	-0.5	4.3	4.0	0.1	5.5	-0.3	4.6
ERA5-L (LM, Stn.)	4.3	-0.3	5.9	-1.0	5.1	3.5	-0.3	4.9	-0.9	4.2	4.1	0.3	5.7	-0.1	4.6
ERA5 (NLM, ERA5-L)	4.2	-0.4	5.7	-1.2	5.1	3.4	-0.5	4.7	-1.1	4.2	4.2	0.3	5.7	-0.1	4.6
ERA5 (NLM, Stn.)	4.2	-0.4	5.7	-1.3	5.1	3.5	-0.4	4.8	-1.0	4.3	3.8	-0.1	5.2	-0.8	4.5

Regarding the seasonality in the ERA data, a close to 0 but slightly positive bias.2 was found for the winter season (DJF), and a slightly negative bias.2 was found for summer (JJA). IMERG-F had a strong positive bias in the winter season that was reduced to a slightly positive bias during summer; hence, it had the same seasonal tendency as the reanalyses. Furthermore, the known pattern in the correlation coefficient (r) (Pearson correlation coefficient) was found. Although the variations are small, ERA5 and ERA5-Land had a higher correlation with the observation during winter (r $= 0.7$) than during summer (r $= 0.5$), and it was the other way around for IMERG-F with r $= 0.5$ during winter and (r $= 0.6$) for the summer season.

The previous results already suggest that the number of wet days is overestimated in the gridded precipitation data, and thus it is not surprising that the FB of both reanalyses is notable above 1, which indicates an overestimation of wet days, while the FB of IMERG-F is rather low (c.f. Tab. 5). There is a low seasonality with the average FB being slightly higher during summer ($\approx +0.1$) than during winter (≈ -0.1) for all datasets. The HSS was approximately $+(-)0.1$ during winter (summer), compared to the annual values; for the ERA reanalyses and IMERG-F the signs were reversed but the rate of change was the same with a higher HSS during summer. The annual values of CSI are comparable with those of the winter season, while during summer IMERG-F had a higher CSI of $\approx +0.1$ and the reanalyses had lower values of ≈ -0.1. The altitude adjustment did not cause any changes. For a threshold of P $\geq$10.0 mm the FB for ERA5 and ERA5-Land is close to 0, but it is higher for IMERG-F with a value of 1.7. Moreover, CSI and HSS were notably reduced to approximately 0.3 and 0.5, respectively. The differences between Type 1 and Type 2 stations are small, with a FB for Type 2 stations that is comparable to the annual values over all stations and a lower FB at Type 1 stations (≈ -0.1). Additionally, IMERG-F had a FB of 1.0 (1.1) for Type 1 (2) stations. Noteworthy differences in CSI or HSS do not exist for any of the gridded datasets.

Table 5: Performance measures for binary events for Germany and Brazil, including all used stations, with P $\geq$0.1 mm.

	Germany			Brazil		
	FB	CSI	HSS	FB	CSI	HSS
IMERGF-F	1.05	0.61	0.68	2.53	0.37	0.24
ERA5-Land	1.41	0.66	0.88	3.45	0.37	0.17
ERA5	1.40	0.66	0.50	3.44	0.37	0.17

5.2 Results for Selected Stations in Germany

In addition to the average value over various stations, a specific look at single stations seemed necessary to obtain more insight into the results. Therefore, nine stations were selected; three Type 2 stations, two stations with rather low elevation differences between station altitude and ERA5 orography, and four Type 1 stations. An overview including the locations of these stations can be found in Figure A.1 and Table B.1. All of the following figures show the station

in a sorted manner, from Type 2 to Type 1 stations.

Therefore, looking to the biases of the nine stations in Figure 11, an alternate picture was found. It seems that ERA5 overestimated P at the Type 2 stations, at least as long as the elevation difference was large enough. At the Type 1 stations there might be a tendency towards underestimation in the results from all stations, but Figure 11 also shows that there are stations that are distinct above the ERA5 orography where the bias of the reanalyses is approximately 0. Furthermore, the effect of the altitude adjustment depends on the individual station. At Marktschellenberg (Type 2), for instance, the already existing underestimation of P from ERA5 became more pronounced after the altitude adjustment. Moreover, the finding from the previous section that the effect of the NLM seemed to be greater than that of the LM at Type 2 stations was clearly visible again at the selected stations. This is a results from the effect of the exponent U_k and the fraction z/CBH. However, the issue of higher biases is certainly not limited to Type 2 stations but also occurred at Type 1 stations such as Feldberg/Schwarzwald. At stations with distinct underestimations of P like Kahler Asten or Brocken, bias was reduced after altitude adjustment. At Brocken, two other effects were clearly visible. The difference between the ERA5-Land orography and station altitude is roughly $350\,\mathrm{m}$ smaller than for the ERA5 orography. Additionally, to the smaller difference, there is a lower slope in the LM for ERA5-Land. Because the original bias was quite similar, both factors together resulted in smaller bias for ERA5 after the altitude adjustment. It is interesting to note that the bias at Wendelstein was very small even though the station altitude is roughly $1000\,\mathrm{m}$ above the ERA5 orography. Moreover, while seasonality does not change the sign of the bias for most stations, at least not when the value is not too small, the stations Feldberg/Schwarzwald and Wendelstein in particular demonstrate a distinct seasonality where the sign changes between summer (positive bias) and winter (negative bias).

Regarding r, RMSE, and SD, shown in Figure 12, ERA5-Land are very close together, and the correlation of the two with the observations is higher than for IMERG-F at all selected stations, while the RMSE is lower. The correlation of the nine stations does not vary much and is r avlous of approximately 0.6 in line with the r found for all stations. The correlation between IMERG-F and the observation is somewhat more variable. Furthermore, the SD of IMERG-F is higher than the original SD of the reanalyses at all nine stations. The altitude adjustment reduced the SD at Type 2 stations while increasing the SD at Type 1 stations. In the same way, RMSE is reduced and increased but the RMSE of both reanalyses is somewhat lower than for IMERG-F. However, for Marktschellenberg and for Brocken, the SD is distinctly below the observed values. The Taylor diagrams show also the larger effect of the NLM at Type 2 stations compared to the LM, while it is the other way around at Type 1 stations.

Next to seasonal differences in the bias, disparities with respect to the rain rates exist. The quantile-quantile plots in Figure 13 show that precipitation events with low intensities were overestimated in all datasets. The number of events with higher precipitation values was mostly underestimated in ERA5 and ERA5-Land. The underestimation became more pronounced for higher P intensities. At Garmisch-Partenkirchen (Type 2), for instance, the

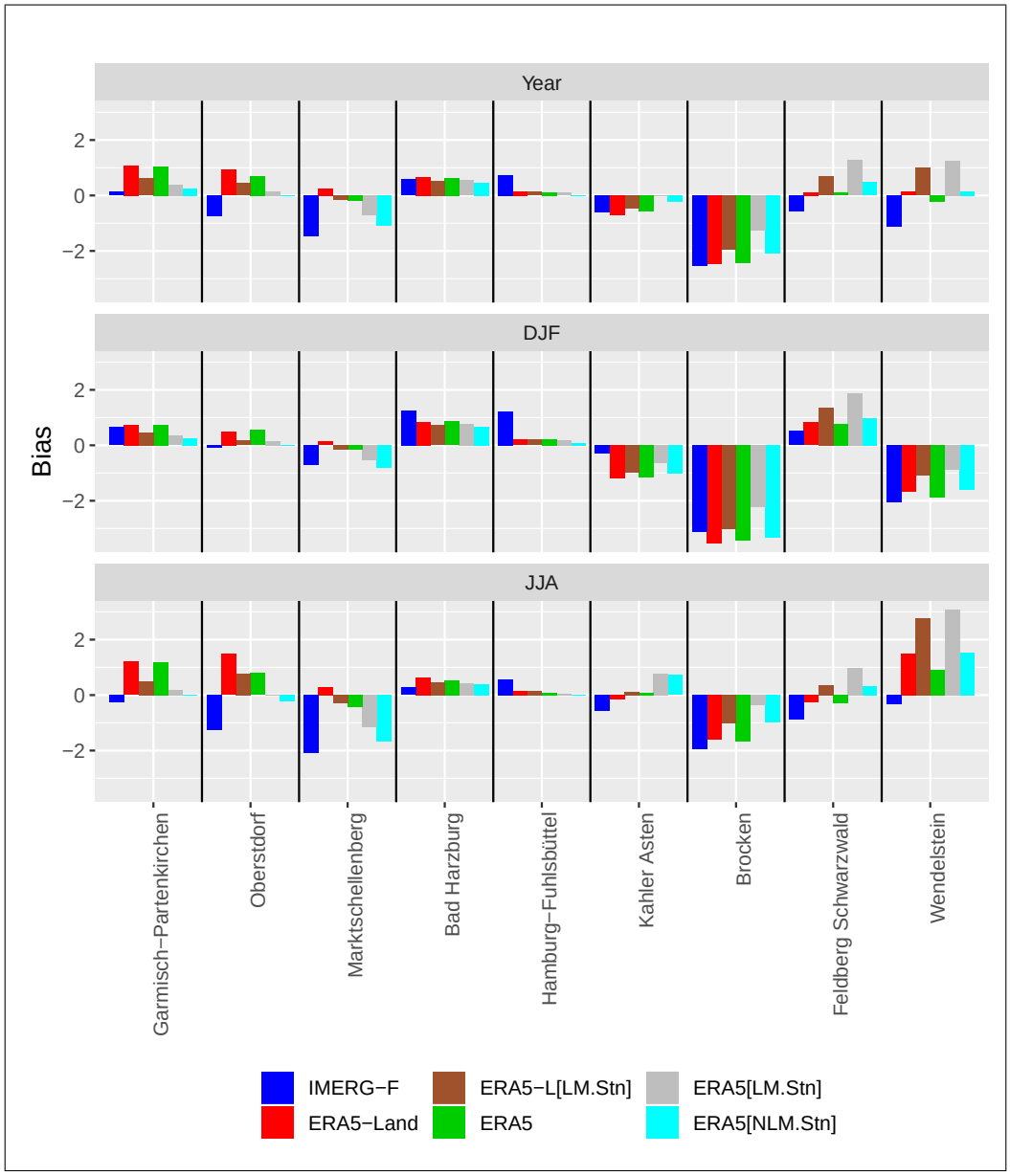

Figure 11: Biases of the three raw precipitation datasets and the altitude adjusted ERA5 and ERA5-Land data for nine example stations in Germany annually summarised as well as for summer and winter season.

bias was positive and altitude adjustment reduced the daily precipitation which led to a smaller bias, but the reduction of already underestimated occurrence (of high precipitation values) led to a even higher spread in the quantile-quantile (Q-Q) plots. For Type 1 station, e.g. Wendelstein, the already overestimated occurrence of low daily precipitation was increased due to the altitude adjustment, while the occurrence of high daily P intensities was increased and approximated the diagonal.

Finally, a brief look at event identification for different P thresholds, P $\geq$0.1 mm (the detection limit in the met. stations and IMERG-F), P $\geq$1.0 mm (approximately the 25% quantile (Q_{25}) over all German stations), P $\geq$ 10 mm (the 75% quantile over all stations) and, P $\geq$30 mm for extreme precipitation events. As expected, for most of the nine stations the FB was above

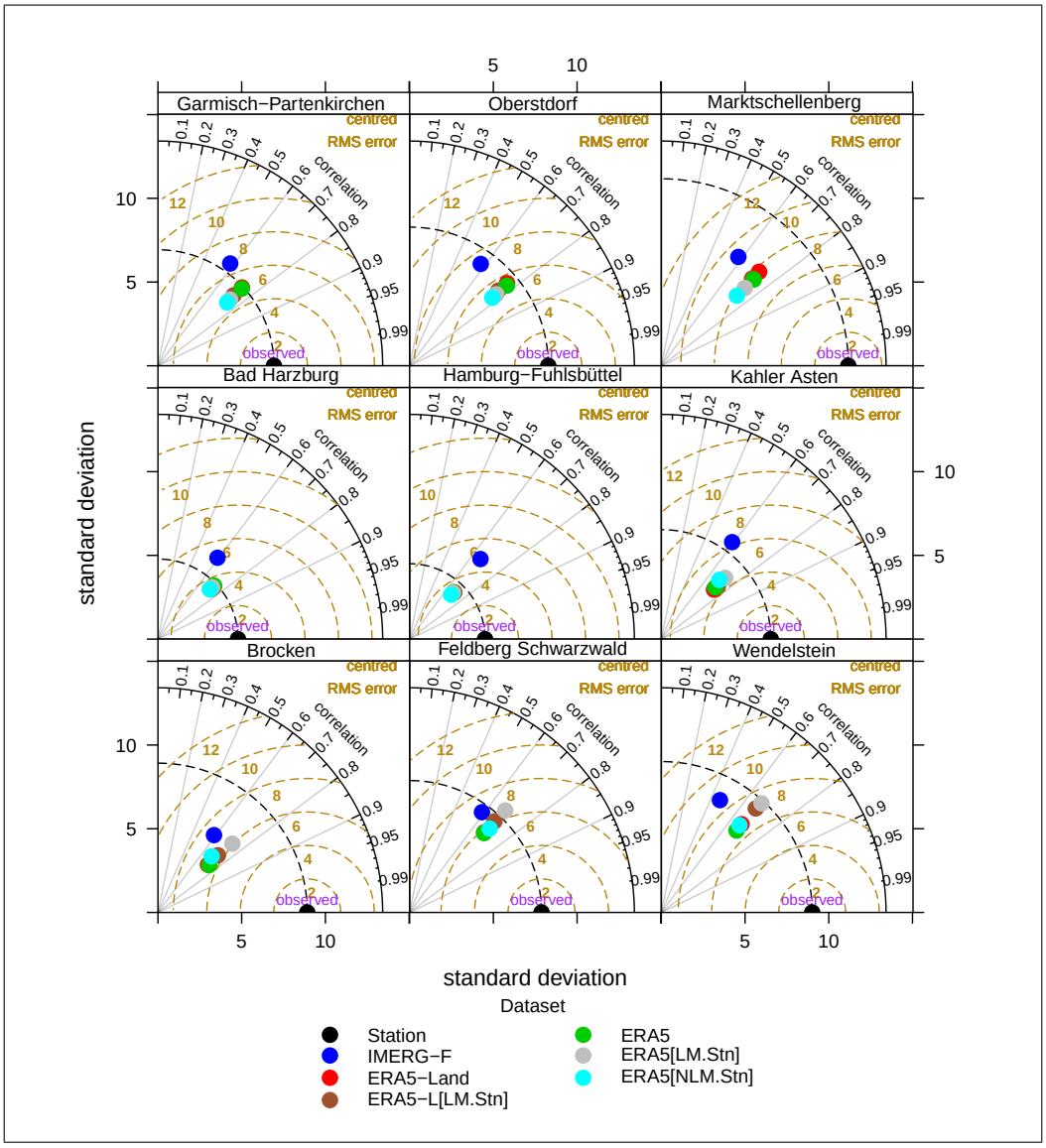

Figure 12: Taylor Diagrams with SD, r, and RMSE for the nine selected stations in Germany.

1 without noteworthy changes between P $\geq$0.1 mm and P $\geq$1.0 mm. At Bad Harzburg and Hamburg-Fuhlsbüttel, the two stations with only small elevation differences, FB was markedly higher than at the other stations. However, whether the altitude adjustment leads to a higher or lower FB depends more on the individual station than on the elevation difference. Looking at H or POD, one sees lower values for IMERG-F than for the reanalyses. Figure 11 shows that the influence of altitude adjustment on H became notable for higher daily precipitation. This effect corresponds to the findings regarding the Q-Q plots. However, H was reduced for Type 2 stations, and increased for Type 1 stations. In the same way, FAR was increased (reduced) for Type 1 (2) stations after the altitude adjustment. The generally decreased H and increased FAR values for higher P thresholds suggest a low ability of the models to detect the actual events. The very low H and high FAR for both reanalyses in Hamburg-Fuhlsbüttel for P $\geq$30 mm should mainly result from only 18 such extreme precipitation events. As a consequence, the HSS for ERA5 and ERA5-Land in Hamburg-Fuhlbüttel is very low considering

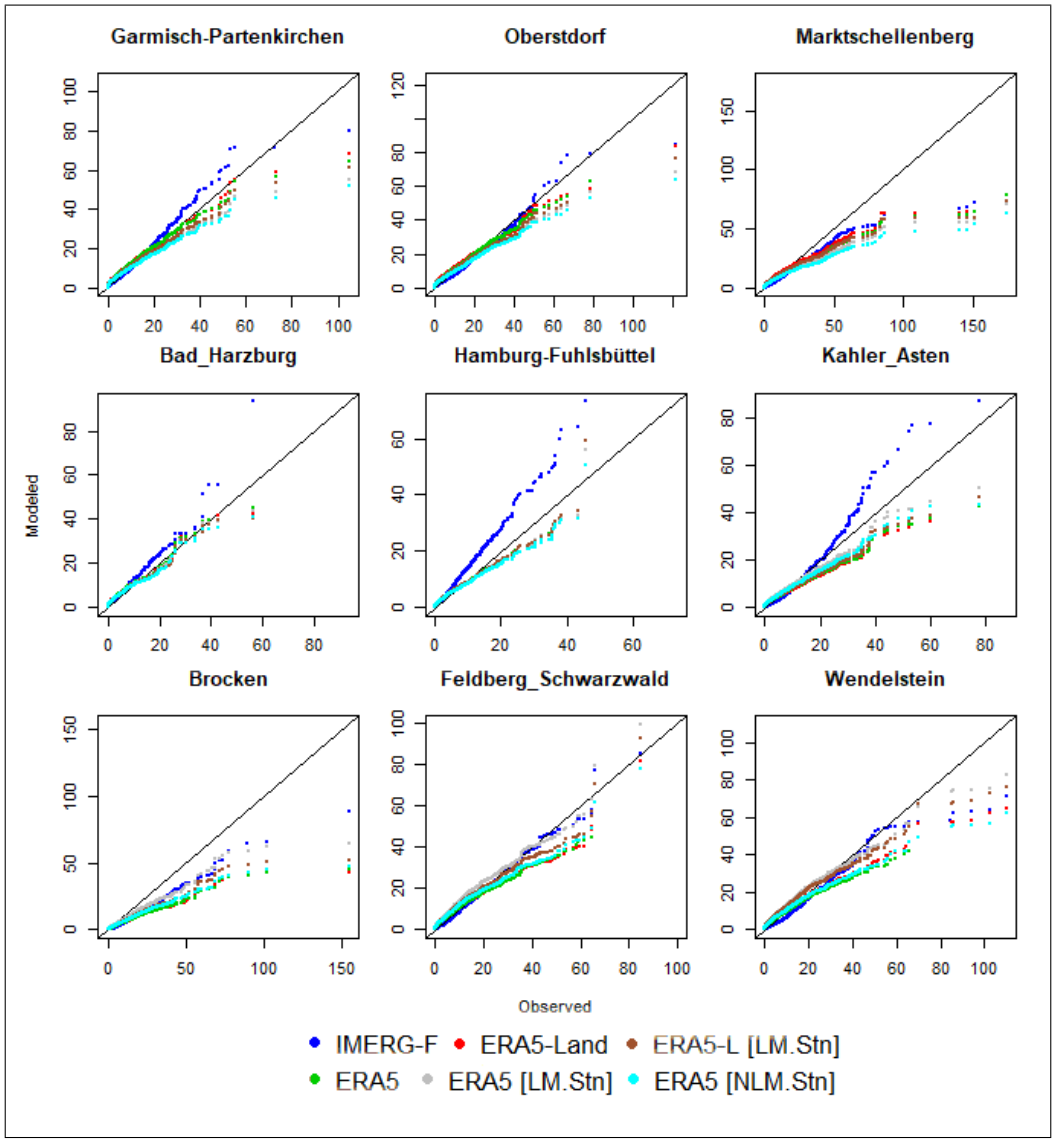

Figure 13: Quantile-quantile plots for the nine German stations.

precipitation events with P $>=$ 30 mm/day. For the lower P intensities, the HSS of the gridded datasets is relatively similar, although IMERG-F has somewhat lower values, which also remained for higher P intensities, but HSS generally decreased with higher P intensities. The effect of altitude adjustment on HSS was rather small, because the increase and decrease of H and FAR due to the adjustment levelled each other out. However, the higher impact of the NLM on Type 2 stations and the higher effect of the LM on Type 1 stations are again visible.

5.3 General Results for Brazil

The spatial patterns of the three gridded precipitation products are in good agreement in the Brazilian study area. One interesting feature of higher rainfall in IMERG-F, which does not appear in the reanalyses can be seen in austral summer at approximately 5°S and 50°W (c.f. Fig. 5). When comparing the gridded data with the 472 stations used in the Brazilian study area, a positive bias (bias.1) prevails for IMERG-F, while bias.1 for ERA5 and ERA5-

Land is slightly negative. Considering only the days where the threshold was met at the met. stations, bias.2 is clearly negative for all three datasets with a more distinct bias for the reanalyses. This pattern of biases remains for P $\geq$30 mm/day, whereby bias.2 is, according to the threshold, relatively large. Considering elevation and elevation differences, as shown in Table 6, a positive bias for stations above 1000 m can be found. Likewise, the distinction between Type 1 and Type 2 stations shows the known pattern from Germany with a negative bias.1 for the reanalyses and a strong positive bias for IMERG-F. In contrast to Germany the bias for the reanalyses remained negative for the Type 1 stations, albeit less distinct. Though close to 0, the bias.1 for IMERG-F remained positive. While P values in Germany were higher at Type 2 stations, in the Brazilian study area P were higher at Type 1 stations. However, bias.2 is negative for all datasets and for the three height-dependent distinctions, but bias.2 is higher for Type 2 than for Type 1 stations. Additionally, in the Brazilian study area, no station above 1000 m is a Type 1 station. Moreover, the number of stations was low, which makes a general statement difficult.

Table 6: Mean daily precipitation, biases and MAE, for different subsets in Germany. Mean daily precipitation p.1 and bias.1 include dry days in averaging, p.2 and bias.2 include only those days, when the precipitation threshold is met for observations.

	All Stations, P $\geq$0.1 mm (n=472)					All Stations, P $\geq$30.0 mm, (n=472)				
	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2
Met. Stn. Observation	4.0		11.7					47.3		
IMERG- F	4.3	0.28	7.1	-3.7	10.0			46.4	-27.0	31.3
ERA5-Land	3.9	-0.09	4.9	-5.3	9.8			43.5	-35.5	36.4
ERA5	3.9	-0.09	4.9	-5.3	9.8			43.6	-35.5	36.4

	Stations above 1000 m (n=15)					Type 1 stations P $\geq$0.1 mm (n=11)					Type 2 stations P $\geq$0.1 mm (n=3)				
	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2	P.1	bias.1	P.2	bias.2	MAE.2
Met. Stn. Observation	3.8		9.7			4.6		10.1			2.8		10.0		
IMERG- F	3.8	0.01	6.8	-3.0	8.2	5.4	0.71	7.8	-2.3	8.9	2.8	0.06	6.3	-3.5	8.4
ERA5-Land	3.9	0.11	5.1	-2.9	8.2	4.4	-0.28	5.3	-3.8	8.5	2.7	-0.1	3.5	-4.5	8.0
ERA5	3.9	0.08	5.1	-3.0	8.2	4.4	-0.28	5.3	-3.8	8.5	2.8	-0.03	3.6	-4.4	8.1

Although it is not representative, at least averaged over the available stations in both study areas, the FB is much higher in the Brazilian study area for both reanalyses datasets and for IMERG-F (c.f. Tab.5). Furthermore, in the Brazilian study area, the FB is approximately $+1.0$ during austral summer compared to the annual value, while it is roughly -1.0 during austral winter. Even though the seasonal change is much greater than in Germany, the pattern is the same. Another coincidence is the lower FB for Type 1 compared to Type 2 stations, though the number of stations is low at 3 and 11, respectively. The performance measures CSI and HSS are generally lower in the Brazilian study area, and seasonal differences are more pronounced with lower values in austral winter. However, one should keep in mind that several uncertainties exist in the Brazilian observations, such as long series of $0\,\mathrm{mm}$ reports.

5.4 Results for Selected Stations in Brazil

In addition to those in Germany, nine stations were selected in the Brazilian study area with the same distinction between Type 1 and Type 2 stations, but the elevation differences are generally not as pronounced as in Germany, and they are likewise less marked at the selected station. An overview of the location as well as some meta data can be found in Figure A.1 and Table B.1.

The station Belterra, has a high positive bias (c.f. Fig. 14), which might have resulted from erroneous measurements. While no clear pattern regarding the bias between Type 1 and Type 2 stations is visible, the biases at the northern stations Itaituba and Belterra are continuously positive. During austral winter, the bias is virtually 0 at all stations except for Belterra, where the high positive bias is at least reduced during austral winter. During austral summer, the biases are correspondingly higher. Moreover, at virtually all stations IMERG-F better agrees with the number of events of observed events than ERA5 or ERA5-Land does.

The diagrams in Figure 15 show that IMERG-F differs explicitly from the reanalyses and has more comparable values with the observations not only regarding the bias but also often in terms of SD and r. While the RMSE is higher than at the German stations for the three datasets, r is clearly lower, with values approximately between 0.3 and 0.6. Furthermore, while at most of the nine stations r is comparable among the datasets, at Itaituba and Belterra IMERG-F has distinctly better values. These are the northern stations previously mentioned in the context of the positive biases. Additionally, at these two stations the RMSE is somewhat higher than at the remaining stations; moreover, both reanalyses effectively have SDs below the observed values and below IMERG-F.

The previously explained tendency that IMERG-F rainfall better represents the rainfall in the Brazilian study is additionally reflected in the Q-Q diagrams shown in Figure 16. The diagrams indicate that the reanalyses underestimated the occurrence of higher daily rainfall intensities at all stations except Belterra, while IMERG-F was in better agreement with the observations. Like in Germany at most os the selected stations in the Brazilian study area IMERG-F had a higher occurrence of high amount rainfall events than ERA5 and ERA5-Land.

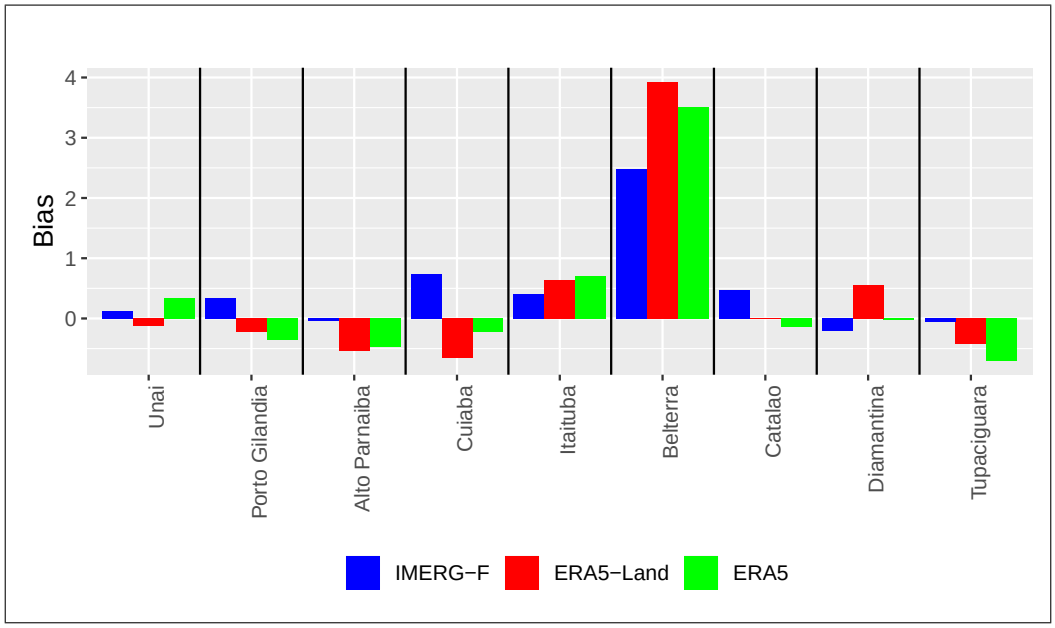

Figure 14: Bias for the selected Stations in the Brazilian study area summarised as annual, summer and winter precipitation.

Both reanalyses datasets produced very similar results, though at some stations such as Porto Gilandia the differences between the two reanalysis products became notable with higher rainfall intensity.

To measure the ability of the gridded datasets to meet the occurrence of rainfall events with different intensities at the selected stations, P thresholds of P $\geq$0.1 mm (the detection limit of IMERG-F and the observations), P $\geq$3.0 mm (approximately Q_{25} of all Brazilian stations used), and P $\geq$30.0 mm (approximately Q_{90}) over all observations are used. Unsurprisingly, the FB got smaller for higher thresholds, and for P $\geq$3.0 mm the FB was approximately 1 apart from the northern stations Itaituba and Belterra where the number of wet days is still notably overestimated by the reanalyses. For P $\geq$30.0 mm the reanalyses include fewer events than observed at the stations (except for Belterra), and the FB of IMERG-F is distinctly smaller. Looking at H and FAR it can be noted that for lower rainfall intensities the reanalyses have higher H as well as higher FAR values. For P $\geq$30.0 mm there was a drastic decrease in H for all three datasets. Nonetheless, H is higher for IMERG-F for the high rainfall intensities. The higher FAR values for the reanalyses remain for the high rain intensity, as shown in Figure A.5. In line with this the HSS is lower for high rain intensities, whereby the HSS is higher for IMERG-F compared to the reanalyses. Interestingly, at Belterra the HSS is quite low for the lower thresholds. In Itaituba the HSS is also very low for the reanalyses but not for IMERG-F. The low performance of ERA5 and ERA5-Land at these two stations results from the correction of tp and tn by chance, which is the term E in the HSS. However, IMERG-F generally produced higher and thus better HSS values over all three precipitation thresholds. In addition, the choice of performance measure has some impact on the evaluation results. For instance, while the HSSs in Itaituba for IMERG-F and ERA5 are 0.49 and 0.03, respectively,

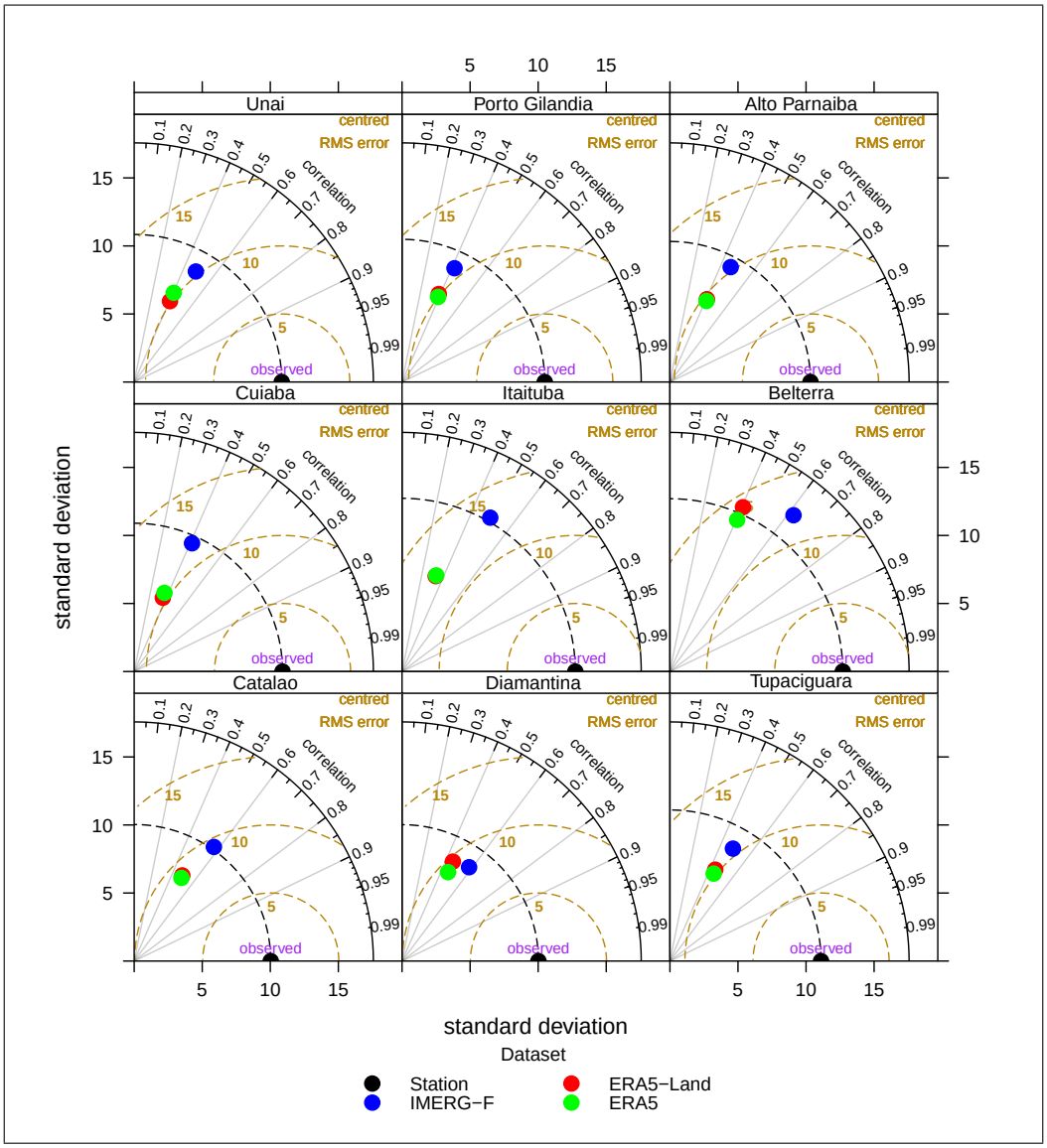

Figure 15: Taylor Diagrams with SD, r, and RMSE for the nine selected stations in Brazilian study area.

the extremal dependency index is 0.22 and 0.57 for $P \geq 0.1\,\text{mm}$ and thus draws another picture of the performance in this special case.

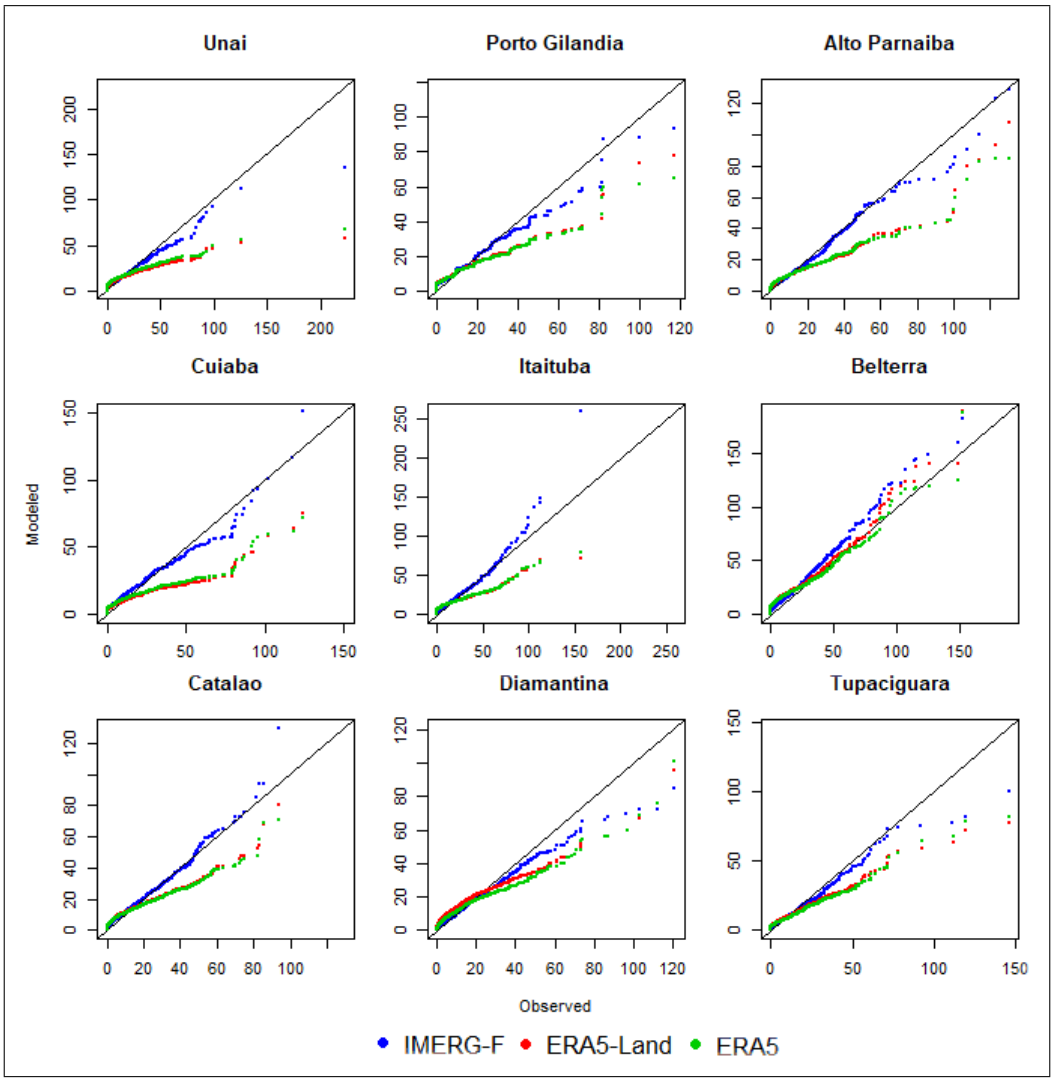

Figure 16: Quantile-quantile plots for the nine selected stations in the Brazilian study area.

6 Discussion and Conclusion

From the results above, one can see that on the one hand the reanalyses provide better results in the German study area in relation to the satellite-based IMERG-F estimates, while on the other hand IMERG-F is better in the tropics. This is true applies for both the continuous measures and the verification measures for binary events. Thus, most of the findings of this thesis are in good agreement with earlier studies that used older reanalyses and satellite precipitation, or IMERG-F data. For example, there is the underestimation of heavy rainfall in Brazil by Sun et al. (2018), which was also found here for most stations, and the overestimation of summer rainfall in Brazil for IMERG-F found by Rozante et al. (2018), which is likewise still prevalent in the actual IMERG-F v06, even though the bias during austral summer is only slightly positive at $0.1\,\mathrm{mm/day}$. The seasonal performance differences for reanalyses and satellite precipitation with worse (better) performance during summer for reanalyses (satellite precipitation) were likewise found for other areas (e.g. Asong et al. 2017).However, as also became evident, results from an evaluation with met. stations strongly depend on the stations themselves, because of the strong spatial and temporal variability of precipitation. Therefore, the comparison of specific results as the underestimation of precipitation by IMERG-F across Northwest Europe found by Sun et al. (2018) can not be affirmed without limitations. While the average bias across all 1504 stations in Germany is slightly positive, at six out of the nine selected stations, the annual bias is negative. Because when averaged over all station the overestimation also appears during summer, the summer overestimation found by Sun et al. was also found here. Although the findings from Zolina et al. (2004) are somewhat older, the higher correlation of reanalyses and observations during winter compared to summer is still prevalent in the ERA5 and ERA5-Land in both study areas. Moreover, the underestimation of high-precipitation events was found at most of the selected stations in both study area. The high r values of roughly 0.9, that have been found by (de Leeuw et al., 2015) for ERA-Interim in England and Wales were not reached in this study. The strong overestimation of winter precipitation in IMERG-F compared to radar-based precipitation estimates in Germany (Ramsauer et al., 2018) is still prevalent in IMERG-F v06, where the average daily P during winter is $2.9\,\mathrm{mm}$ compared to the observed $2.1\,\mathrm{mm}$. The positive (negative) bias during summer (winter) found by Betts et al. (2019) for Canada was also found for Germany when the bias was calculated for observed wet days. Considering all days, the bias was slightly positive in both seasons in Germany.

The non-presence of an altitude-precipitation relationship that was found for satellite-based precipitation estimates in both research areas in this study was also partly found by Arakawa and Kitoh (2011). They found a lower relationship for satellite precipitation merged with observation than for gridded observation-based datasets in low and middle latitudes, though they found a comparable relationship between those to types of datasets in high latitudes. However, the use of IMERG-F and ERA5 still offers some scope for more detailed investigations.

Because this thesis aimed to compare the best available data, the quality index that is part of the IMERG-F was not considered, but it would be interesting to look at differences with respect to the quality index and the incorporated sensor data (e.g. IR and MW). Moreover, because monthly GPCC data is used to correct IMERG-F precipitation, it could be examined if the performance of IMERG-F depends on the quality of the GPCC data. This point is especially strengthened since the GSMaP product from JAXA with daily correction outperformed IMERG-F in northern Brazil (Rozante et al., 2018). For ERA5, it would be interesting to use the ensemble data to see if the error lies within the ensemble spread, for example.

The results of this study indicate that an elevation-dependent bias or performance difference exists, as has often been stated (Kim et al., 2017; Prakash et al., 2018; Chen et al., 2016; Asong et al., 2017; Hirpa et al., 2010; Guo et al., 2017), but more precisely, bias and performance rather depend on the elevation difference between the station and the model. In both study areas the average bias was negative (positive) for Type 1 (2) stations, whereby IMERG-F shows the same direction of change between Type 1 and 2 stations, but does not necessarily have the same sign as ERA5 and ERA-Land. However, the findings from Section 4.2 and the station type-dependent bias were reasons to try to estimate the model bias corrected for the elevation difference. Overall, the effect of the altitude adjustment was positive, although notably different within each station subset. Though the averaged bias after altitude adjustment was smaller, it does not necessarily have to be; for instance, at stations with a high bias but a low elevation difference, the bias was higher after the adjustment. The altitude adjustment only marginally affected the measure for binary events and approximately levelled itself out, because increased daily precipitation resulted in a higher number of both hits and false alarms.

Considering the method of altitude adjustment and Brunsdon et al. (2001), who found a steep increase of precipitation with altitude in Great Britain and who furthermore argued that this was related to a lower number of stations and the distributions of the same, as well as Daly et al. (1994), who have suggested estimating topographic effects from a rather coarser resolution, the attempt to use model output to assess the effects of elevation seems justified. Nevertheless, limiting the LM to certain altitudes or limited regions altered the slope by a considerable amount. This suggests the need for a method that considers the local conditions. This could also help to apply the altitude adjustment in the tropics. Regarding the Brazilian study area the convective proportion of TP was continuously high and no relation could be found between altitude and rainfall, there was no substantial difference in the precipitation-elevation relationship for high or low convective proportions in Germany. Thus, the differentiation of types and proportions based on ERA5 might be insufficient. Additionally, other studies have confirmed the limited relationship between CP and elevation. For instance, Sokol and Bližňák (2009) have differentiated between convective and non-convective rainfall on the rainfall duration during summer in the Czech Republic and have not found a relationship between convective rainfall and elevation. In keeping with this, Ragette and Wotawa (1998) have ascribed low evaporation in the tropics to the high humidity. However, de Leeuw

et al. (2015) differentiated ERA-Interim CP and large-scale precipitation in Great Britain and found no significant differences between modelled and observed precipitation based on the precipitation type.

Because the approach of estimating evaporation from model output was deficient in the tropics and because its impact in Germany varied spatially, meaning from station to station, it seems advisable to bring the method to a more physically based approach. This could be done by incorporating temperature and humidity profiles, for instance, since in ERA5 or the IFS the evaporation depends on the saturation deficit. Another benefit of incorporating other variables would be that one could attempt to estimate the evaporation for single events instead of using long-term average values. This is in keeping with Shige and Kummerow (2016) and Shige et al. (2013), who have tried to distinguish wind-induced rainfall and used different rainfall profiles. Damseaux et al. (2019) have tried to tackle the problem of poor topographically representation in regional climate models by different DEM generalisation methods, with only slight success for rain shadow effect in Patagonia.

The step from trying to achieve better evaporation estimates to considering other topographic effect is not large and has been made by Funk et al. (2003), for instance. Although, other studies (e.g. Gerlitz et al. 2015 and Karger et al. 2017) used topographic indices, such as the Wind Exposition Index (Böhner and Antonić, 2009), its impact on precipitation at the (coarse) resolution of ERA5 and ERA5-Land is hard to estimate; indeed, using linear regression a negative relationship was found for the German study area in this study. Thus, to determine the effect of such topographic indices, other datasets and other methods than the ones used here would be necessary.

If more topographic effects are considered, the improvement of the horizontal resolution of modelled precipitation or the improved representation of modelled precipitation at the station location is often aimed for. The incorporation of topographic effects in such a statistical downscaling remains difficult, because the estimation of the effect remains difficult since high-density networks would be necessary. On the other hand, such effects may explain some of the variations within the subsets of Type 1 or Type 2 stations. Nonetheless, the problem of too many wet days is not solved by solely incorporating more topographic effects. Instead the deviating distributions would need to be corrected like in Volosciuk et al. (2017). However, only a certain fraction of the precipitation variability can be explained by the topography and in areas without complex topography the opportunities to improve gridded precipitation data in a statistical manner are probably very limited. Nonetheless, since there has already been an attempt to use machine learning to resolve moist convection Gentine et al. (2018), there are new possibilities to improve modelled rainfall.

Based on this thesis, it can be concluded that the performance of reanalyses and satellite-based precipitation estimates prevalent in earlier datasets still can be found in the most recent datasets, namely IMERG-F v06, ERA5, and ERA5-Land. Moreover, the assumption of an elevation-dependent bias due to elevation differences can be verified using met. stations in

Germany. In general, the bias was reduced when elevation differences and a bias existed at the same time. It would be good to shift the altitude adjustment to a more physically based approach for two reasons. First, event-based and local differences could be considered, and second, the method could also be applied in the tropics where the evaporation of precipitation is rather low. However, because the altitude adjustment does not aim to correct for other topographical effects than elevation, it is not useful to reduce the bias and the frequency bias of gridded precipitation at certain station locations.

Bibliography

Abel, S. J. and Boutle, I. A. (2012). An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Quarterly Journal of the Royal Meteorological Society*, 138(669):2151–2162.

Adler, R., Huffman, G., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present). *Journal of Hydrometeorology*, 4(6):1147–1167.

Agência Nacional de Águas (2019). HIDROWEB - Sistema de Informações Hidrológicas.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., de Rosnay, P., and Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better? *Hydrology and Earth System Sciences*, 22(6):3515–3532.

Alvares, C. A., Stape, J. L., Sentelhas, P. C., de Moraes Gonçalves, J. L., and Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22(6):711–728.

Alvarez, M. S., Vera, C. S., Kiladis, G. N., and Liebmann, B. (2016). Influence of the Madden Julian Oscillation on precipitation and surface air temperature in South America. *Climate Dynamics*, 46(1-2):245–262.

Arakawa, O. and Kitoh, A. (2011). Intercomparison of the relationship between precipitation and elevation among gridded precipitation datasets over the Asian summer monsoon region. *Global Environ. Res*, 15(2):109–118.

Arvor, D., Dubreuil, V., Ronchail, J., Simões, M., and Funatsu, B. M. (2014). Spatial patterns of rainfall regimes related to levels of double cropping agriculture systems in Mato Grosso (Brazil). *International Journal of Climatology*, 34(8):2622–2633.

Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., Prat, O. P., Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R., and Prat, O. P. (2015). PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, 96(1):69–83.

Asong, Z. E., Razavi, S., Wheater, H. S., and Wong, J. S. (2017). Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against Ground Precipitation Observations: A Preliminary Assessment. *Journal of Hydrometeorology*, 18(4):1033–1050.

Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55.

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R., and Wood, E. F. (2018). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences Discussions*, 23(1):207–224.

Beck, H. E., Van Dijk, A. I., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and De Roo, A. (2017). MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, 21(1):589–615.

Bengtsson, L. and Shukla, J. (1988). Integration of space and in situ observations to study global climate change. *Bulletin of the American Meteorological Society*, 69(10):1130–1143.

Beria, H., Nanda, T., Singh Bisht, D., and Chatterjee, C. (2017). Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale. *Hydrology and Earth System Sciences*, 21(12):6117–6134.

Berndt, C. and Haberlandt, U. (2018). Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. *Journal of Hydrology: Regional Studies*, 15(October 2017):184–202.

Betts, A. K., Chan, D. Z., and Desjardins, R. L. (2019). Near-Surface Biases in ERA5 Over the Canadian Prairies. *Frontiers in Environmental Science*, 7:129.

BfN (n.d.). Land use in Germany in 2014. `https://www.bfn.de/en/service/facts-and-figures/the-utilisation-of-nature/land-use-overview/land-use-in-germany.html`. Last checked on Oct 20, 2019.

Böhner, J. and Antonić, O. (2009). Chapter 8 Land-Surface Parameters Specific to Topo-Climatology. In Hengl, T. and Reuter, H., editors, *Developments in Soil Science*, volume 33, chapter 8, pages 195–226. Elsevier.

Brunsdon, C., Mcclatchey, J., and Unwin, D. J. (2001). Spatial variations in the average rainfall–altitude relationship in Great Britain: an approach using geographically weighted regression. *International Journal of Climatology*, 466:455–466.

Ceccherini, G., Ameztoy, I., Hernández, C. P. R., and Moreno, C. C. (2015). High-resolution precipitation datasets in South America and West Africa based on satellite-derived rainfall, enhanced vegetation index and digital elevation model. *Remote Sensing*, 7(5):6454–6488.

Chen, C., Chen, Q., Duan, Z., Zhang, J., Mo, K., Li, Z., and Tang, G. (2018). Multiscale comparative evaluation of the GPM IMERG v5 and TRMM 3B42 v7 precipitation products from 2015 to 2017 over a climate transition area of China. *Remote Sensing*, 10(6):1–18.

Chen, F., Li, X., Chen, F., and Li, X. (2016). Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China. *Remote Sensing*, 8(6):472.

Coelho, C. A., Cardoso, D. H., and Firpo, M. A. (2016a). Precipitation diagnostics of an exceptionally dry event in São Paulo, Brazil. *Theoretical and Applied Climatology*, 125(3-4):769–784.

Coelho, C. A., de Oliveira, C. P., Ambrizzi, T., Reboita, M. S., Carpenedo, C. B., Campos, J. L. P. S., Tomaziello, A. C. N., Pampuch, L. A., Custódio, M. d. S., Dutra, L. M. M., Da Rocha, R. P., and Rehbein, A. (2016b). The 2014 southeast Brazil austral summer drought: regional scale mechanisms and teleconnections. *Climate Dynamics*, 46(11-12):3737–3752.

Copernicus CCS (2017). ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate.

Copernicus Climate Change Service (C3S) (2019). C3S ERA5-Land reanalysis . Copernicus Climate Change Service. `https://cds.climate.copernicus.eu/cdsapp#!/home`. Last checked on Oct 20, 2019.

Daly, C., Neilson, R. P., and Phillips, D. L. (1994). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, 33:140–158.

Damseaux, A., Fettweis, X., Lambert, M., and Cornet, Y. (2019). Representation of the rain shadow effect in Patagonia using an orographic-derived regional climate model. *International Journal of Climatology*, pages 1–15.

de Leeuw, J., Methven, J., and Blackburn, M. (2015). Evaluation of ERA-Interim reanalysis precipitation products using England and Wales observations. *Quarterly Journal of the Royal Meteorological Society*, 141(688):798–806.

Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., and Ropelewski, C. F. (2007). Validation of satellite rainfall products over East Africa's complex topography. *International Journal of Remote Sensing*, 28(7):1503–1526.

DWD (2018). Climate Data Center - Historical daily precipitation observations for Germany, version v006, 2018.

DWD (n.d.). Vieljährige Mittelwerte. `https://www.dwd.de/DE/leistungen/klimadatendeutschland/vielj_mittelwerte.html`. Last checked on Oct 20, 2019.

Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., Harada, Y., Yasui, S., Miyaoka, K., Takahashi, K., Kamahori, H., Kobayashi, C., Endo, H., Soma, M., Oikawa, Y.,

and Ishimizu, T. (2011). The Japanese 55-year Reanalysis “JRA-55”: An Interim Report. *Sola*, 7:149–152.

ECMWF (2016a). Part I: Observation. Technical report, ECMWF, Shinfield Park, Reading.

ECMWF (2016b). Part IV : Physical Processes. Technical report, ECMWF, Shinfield Park, Reading.

ECMWF (2018). ERA5 continuity 2009/2010. `https://confluence.ecmwf.int/pages/viewpage.action?pageId=100045763`. Webpage last modified on Jun 21, 2018. Last checked on Oct 20, 2019.

ECMWF (2019a). ECMWF - ERA5-Land data documentation. `https://confluence.ecmwf.int/display/CKB/ERA5-Land+data+documentation`. Webpage last modified on Oct 07, 2019. Last checked on Oct 20, 2019.

ECMWF (2019b). What are the changes from ERA-Interim to ERA5? `https://confluence.ecmwf.int//pages/viewpage.action?pageId=74764925`. Webpage last modified on May 02, 2019. Last checked on Oct 20, 2019.

ECMWF (n.d.a). ERA5 data documentation. `https://confluence.ecmwf.int/display/CKB/ERA5+data+documentation`. Webpage last modified on Nov 05, 2019. Last checked on Nov 06, 2019.

ECMWF (n.d.b). Forecast User Guide - Large Scale Precipitation. `https://confluence.ecmwf.int/display/FUG/9.5+Large+Scale+Precipitation`. Last checked on Oct 20, 2019.

Ferro, C. A. T. and Stephenson, D. B. (2011). Extremal dependence indices: Improved Verification measures for deterministic forecasts of rare binary events. *Weather and Forecasting*, 26(5):699–713.

Fick, S. E. and Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12):4302–4315.

Funk, C., Michaelsen, J., Verdin, J., Artan, G., Husak, G., Senay, G., Gadain, H., and Magadazire, T. (2003). The collaborative historical African rainfall model: Description and evaluation. *International Journal of Climatology*, 23(1):47–66.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2:150066.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper,

C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., Zhao, B., Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., and Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock? *Geophysical Research Letters*, pages 5742–5751.

Gerlitz, L., Conrad, O., and Böhner, J. (2015). Large-scale atmospheric forcing and topographic modification of precipitation rates over High Asia a neural-network-based approach. *Earth System Dynamics*, 6(1):61–81.

Gerold, G., Couto, E. G., Madari, B. E., Jungkunst, H. F., Amorim, R. S. S., Hohnwald, S., Klingler, M., de Almeida Machado, P. L. O., Schönenberg, R., and Nendel, C. (2018). Carbon-optimised land management strategies for southern Amazonia. *Regional Environmental Change*, 18(1):1–9.

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228:113–129.

Guo, H., Bao, A., Ndayisaba, F., Liu, T., Kurban, A., and De Maeyer, P. (2017). Systematical Evaluation of Satellite Precipitation Estimates Over Central Asia Using an Improved Error-Component Procedure. *Journal of Geophysical Research: Atmospheres*, 122(20):10,906–10,927.

Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, 34(3):623–642.

Hersbach, H. and Dee, D. (2016). ERA5 reanalysis is in production | ECMWF.

Hirpa, F. A., Gebremichael, M., Hopson, T., Hirpa, F. A., Gebremichael, M., and Hopson, T. (2010). Evaluation of High-Resolution Satellite Precipitation Products over Very Complex Terrain in Ethiopia. *Journal of Applied Meteorology and Climatology*, 49(5):1044–1051.

Hong, Y., Hsu, K.-L., Sorooshian, S., Gao, X., Hong, Y., Hsu, K.-L., Sorooshian, S., and Gao, X. (2004). Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *Journal of Applied Meteorology*, 43(12):1834–1853.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5):701–722.

Huffman, G., Adler, R., Bolvin, D., and Nelkin, E. (2010). The TRMM Multi-Satellite Precipitation Analysis (TMPA). In Gebremichael, M. and Hossain, F., editors, *Satellite Rainfall Applications for Surface Hydrology*, page 327. Springer.

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., Sorooshian, S., Tan, T., and Xie, P. (2018a). Algorithm Theoretical Basis Document (ATBD) Version 5.2 NASA - NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) - Algorithm Theoretical Basis Document (ATBD) Version 5.2. *National Aeronautics and Space Administration (NASA)*, IMERG Algo(February):1–31.

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

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K.-L., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J., and Xie, P. (2019a). Algorithm Theoretical Basis Document (ATBD) Version 06 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG). *National Aeronautics and Space Administration (NASA)*, pages 1–34.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Stocker, E. F., and Tan, J. (2018b). V05 IMERG Final Run Release Notes. Technical report, NASA.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Stocker, E. F., and Tan, J. (2019b). IMERG V06 Release Notes. Technical Report June, NASA.

Huntingford, C., Zelazowski, P., Galbraith, D., Mercado, L. M., Sitch, S., Fisher, R., Lomas, M., Walker, A. P., Jones, C. D., Booth, B. B. B., Malhi, Y., Hemming, D., Kay, G., Good, P., Lewis, S. L., Phillips, O. L., Atkin, O. K., Lloyd, J., Gloor, E., Zaragoza-Castells, J., Meir, P., Betts, R., Harris, P. P., Nobre, C., Marengo, J., and Cox, P. M. (2013). Simulated resilience of tropical rainforests to CO2-induced climate change. *Nature Geoscience*, 6(4):268–273.

Jarvis, A., Reuter, H., Nelson, A., and Guevara, E. (2008). Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database.

JAXA (2016). GPM Data Utilization Handbook second Edition. Technical report, JAXA.

Joseph, R., Smith, T. M., Sapiano, M. R. P., Ferraro, R. R., Joseph, R., Smith, T. M., Sapiano, M. R. P., and Ferraro, R. R. (2009). A New High-Resolution Satellite-Derived Precipitation Dataset for Climate Studies. *Journal of Hydrometeorology*, 10(4):935–952.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*, 5(3):487–503.

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., Potter, G. L., Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L. (2002). NCEP–DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, 83(11):1631–1644.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific Data*, 4:170122.

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., and Deutschländer, T. (2013). Monitoring of climate change in Germany – data, products and services of Germany's National Climate Data Centre. *Advances in Science and Research*, 10:99–106.

Kessler, E. (1969). On The Distribution And Continuity Of Water Substance In Atmospheric Circulation. *Meteorological Monogrpahs*, 10(32).

Khanna, J., Medvigy, D., Fueglistaler, S., and Walko, R. (2017). Regional dry-season climate changes due to three decades of Amazonian deforestation. *Nature Climate Change*, 7(3):200–204.

Khodadoust Siuki, S., Saghafian, B., and Moazami, S. (2017). Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products. *International Journal of Remote Sensing*, 38(2):558–571.

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

Kilian, M. (2017). *Climate variability and potential future climate change in southern Amazonia: sensitivity of the hydrological cycle to land use changes*. Dissertation, Universität Hamburg.

Kim, K., Park, J., Baik, J., and Choi, M. (2017). Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia. *Atmospheric Research*, 187:95–105.

Kleist, D. T., Parrish, D. F., Derber, J. C., Treadon, R., Wu, W.-S., Lord, S., Kleist, D. T., Parrish, D. F., Derber, J. C., Treadon, R., Wu, W.-S., and Lord, S. (2009). Introduction of the GSI into the NCEP Global Data Assimilation System. *Weather and Forecasting*, 24(6):1691–1705.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan. Ser. II*, 93(1):5–48.

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

Liebmann, B. and Allured, D. (2005). Daily Precipitation Grids for South America. *Bulletin of the American Meteorological Society*, 86(11):1567–1570.

Liebmann, B. and Mechoso, C. R. (2011). 9 . THE SOUTH AMERICAN MONSOON SYSTEM. In Chang, C.-P., Ding, Y., Lau, N.-C., Johnson, R., Wang, B., and Yasunari, T., editors, *The Global Monsoon System: Research and Forecast*, chapter 9, pages 137–157. World Scientific Series on Asia-Pacific Weather and Climate, 2 edition.

Marengo, J. A., Nobre, C. A., Tomasella, J., Oyama, M. D., Sampaio de Oliveira, G., de Oliveira, R., Camargo, H., Alves, L. M., Brown, I. F., Marengo, J. A., Nobre, C. A., Tomasella, J., Oyama, M. D., de Oliveira, G. S., de Oliveira, R., Camargo, H., Alves, L. M., and Brown, I. F. (2008). The Drought of Amazonia in 2005. *Journal of Climate*, 21(3):495–516.

Marengo, J. A., Torres, R. R., and Alves, L. M. (2016). Drought in Northeast Brazil - past, present, and future. *Theoretical and Applied Climatology*, pages 1–12.

Mason, I. B. (2003). Binary Events. In Jolliffe, I. T. and Stephenson, D. B., editors, *Forecast Verification - A Practioners Guide in Atmospheric Science*, chapter Binary Eve, pages 37–76. Wiley, Chichester, 1 edition.

McGregor, G. R. and Nieuwolt, S. (1998). *Tropical Climatology*. John Wiley & Sons, Chichester, second edi edition.

Minder, J. R., Mote, P. W., and Lundquist, J. D. (2010). Surface temperature lapse rates over complex terrain: Lessons from the Cascade Mountains. *Journal of Geophysical Research Atmospheres*, 115(14):1–13.

Muñoz Sabater, J. (2019). First ERA5-Land dataset to be released this spring. Technical report, ECMWF, Reading.

Müller-Westermeier, G. (2006). *Wetter und Klima in Deutschland*. Hirzel Verlag, Stuttgart, 4 edition.

National Centers for Environmental Information (NCEI) (n.d.). NCEP/NCAR Reanalysis-1/ Reanalysis-2. `https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/reanalysis-1-reanalysis-2`. Last checked on Oct 20, 2019.

Nepstad, D., McGrath, D., Stickler, C., Alencar, A., Azevedo, A., Swette, B., Bezerra, T., DiGiano, M., Shimada, J., Seroa da Motta, R., Armijo, E., Castello, L., Brando, P., Hansen, M. C., McGrath-Horn, M., Carvalho, O., and Hess, L. (2014). Slowing Amazon deforestation through public policy and interventions in beef and soy supply chains. *Science*, 344(6188):1118–23.

Nobre, P., Malagutti, M., Urbano, D. F., de Almeida, R. A. F., and Giarolla, E. (2009). Amazon Deforestation and Climate Change in a Coupled Model Simulation. *Journal of Climate*, 22(21):5686–5697.

O, S., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J., and Petersen, W. A. (2017). Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrology and Earth System Sciences*, 21(12):6559–6572.

Olauson, J. (2018). ERA5: The new champion of wind power modelling? *Renewable Energy*, 126:322–331.

Oliveira, R., Maggioni, V., Vila, D., Morales, C., Oliveira, R., Maggioni, V., Vila, D., and Morales, C. (2016). Characteristics and Diurnal Cycle of GPM Rainfall Estimates over the Central Amazon Region. *Remote Sensing*, 8(7):544.

Prakash, S., Mitra, A., AghaKouchak, A., Liu, Z., Norouzi, H., and Pai, D. (2018). A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region. *Journal of Hydrology*, 556:865–876.

Ragette, G. and Wotawa, G. (1998). The evaporation of precipitation and its geographical distribution. *Physics and Chemistry of the Earth*, 23(4):393–397.

Ramsauer, T., Weiß, T., and Marzahn, P. (2018). Comparison of the GPM IMERG Final Precipitation Product to RADOLAN Weather Radar Data over the Topographically and Climatically Diverse Germany. *Remote Sensing*, 10(12):2029.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., and Gratzki, A. (2013). A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorologische Zeitschrift*, 22(3):235–256.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A.,

da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., Woollen, J., Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Julio Bacmeister, Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, 24(14):3624–3648.

Roy, S. B. and Avissar, R. (2002). Impact of land use/land cover change on regional hydrometeorology in Amazonia. *Journal of Geophysical Research*, 107(D20):8037.

Rozante, J., Vila, D., Barboza Chiquetto, J., Fernandes, A., and Souza Alvim, D. (2018). Evaluation of TRMM/GPM Blended Daily Products over Brazil. *Remote Sensing*, 10(6):882.

Sawyer, J. S. (1956). THE PHYSICAL AND DYNAMICAL PROBLEMS OF OROGRAPHIC RAIN. *Weather*, 11(12):375–381.

Schamm, K., Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Schneider, U., Schröder, M., and Stender, P. (2014). Global gridded precipitation over land: A description of the new GPCC First Guess Daily product. *Earth System Science Data*, 6(1):49–60.

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

Seth, A., Rojas, M., Liebmann, B., and Qian, J. H. (2004). Daily rainfall analysis for South America from a regional climate model and station observations. *Geophysical Research Letters*, 31(7):3–6.

Sharifi, E., Steinacker, R., and Saghafian, B. (2016). Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sensing*, 8(2).

Shige, S., Kida, S., Ashiwake, H., Kubota, T., and Aonashi, K. (2013). Improvement of TMI Rain Retrievals in Mountainous Areas. *Journal of Applied Meteorology and Climatology*, 52(1):242–254.

Shige, S. and Kummerow, C. D. (2016). Precipitation-Top Heights of Heavy Orographic Rainfall in the Asian Monsoon Region. *Journal of the Atmospheric Sciences*, 73(8):3009–3024.

Shimizu, M. H., Ambrizzi, T., and Liebmann, B. (2017). Extreme precipitation events and their relationship with ENSO and MJO phases over northern South America. *International Journal of Climatology*, 37(6):2977–2989.

Silva, V. B. S., Kousky, V. E., Shi, W., and Higgins, R. W. (2007). An Improved Gridded Historical Daily Precipitation Analysis for Brazil. *Journal of Hydrometeorology*, 8(4):847–861.

Sokol, Z. and Bližňák, V. (2009). Areal distribution and precipitation – altitude relationship of heavy short-term precipitation in the Czech Republic in the warm part of the year. *Atmospheric Research*, 94:652–662.

Spengler, R. (2002). The new Quality Control- and Monitoring System of the Deutscher Wetterdienst. In *Proceedings of the WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation*, Bratislava. Deutscher Wetter Dienst.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K. L. (2018). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Reviews of Geophysics*, 56(1):79–107.

Tang, G., Behrangi, A., Long, D., Li, C., and Hong, Y. (2018). Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products. *Journal of Hydrology*, 559:294–306.

Tang, G., Ma, Y., Long, D., Zhong, L., and Hong, Y. (2016). Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. *Journal of Hydrology*, 533:152–167.

Tatsch, J. (2019). inmetr: Historical Data from Brazilian Meteorological Stations in R.

Thorne, P. and Vose, R. (2010). Reanalyses suitable for characterizing long-term trends. *Bulletin of the American Meteorological Society*, 91(3):353–362.

Trenberth, K. E., Koike, T., and Onogi, K. (2008). Progress and Prospects for Reanalysis for Weather and Climate. *Eos, Transactions American Geophysical Union*, 89(26):234.

Trenberth, K. E. and Olson, J. G. (1988). An Evaluation and Intercomparison of Global Analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts. *Bulletin of the American Meteorological Society*, 69(9):1047–1057.

Urraca, R., Huld, T., Gracia-Amillo, A., Martinez-de Pison, F. J., Kaspar, F., and Sanz-Garcia, A. (2018). Evaluation of global horizontal irradiance estimates from ERA5 and COSMO-REA6 reanalyses using ground and satellite-based data. *Solar Energy*, 164:339–354.

Volosciuk, C., Maraun, D., Vrac, M., and Widmann, M. (2017). A combined statistical bias correction and stochastic downscaling method for precipitation. *Hydrology and Earth System Sciences*, 21(3):1693–1719.

Weischet, W. and Endlicher, W. (2000). *Regionale Klimatologie Teil 2 - Die Alte Welt*. Teubner Verlag, Stuttgart, Leipzig.

Werth, D. and Avissar, R. (2002). The local and global effects of Amazon deforestation. *Journal of Geophysical Research*, 107(D20).

Whiteman, C. D. (2000). *Mountain meteorology : fundamentals and applications*. Oxford University Press, Oxford, New York, 1 edition.

WMO (2019). WMO-GCOS-ECV.

Xavier, A. C., King, C. W., and Scanlon, B. R. (2015). Daily gridded meteorological variables in Brazil (1980-2013). *International Journal of Climatology*, 36(6):2644–2659.

Xie, P. and Arkin, P. A. (1997). Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. *Bulletin of the American Meteorological Society*, 78(11):2539–2558.

Yamamoto, M. K. and Shige, S. (2015). Implementation of an orographic/nonorographic rainfall classification scheme in the GSMaP algorithm for microwave radiometers. *Atmospheric Research*, 163:36–47.

Yamamoto, M. K., Shige, S., Yu, C.-K., Cheng, L.-W., Yamamoto, M. K., Shige, S., Yu, C.-K., and Cheng, L.-W. (2017). Further Improvement of the Heavy Orographic Rainfall Retrievals in the GSMaP Algorithm for Microwave Radiometers. *Journal of Applied Meteorology and Climatology*, 56(9):2607–2619.

Yoon, J.-H. and Zeng, N. (2010). An Atlantic influence on Amazon rainfall. *Climate Dynamics*, 34(2-3):249–264.

Yuan, F., Zhang, L., Win, K., Ren, L., Zhao, C., Zhu, Y., Jiang, S., Liu, Y., Yuan, F., Zhang, L., Win, K. W. W., Ren, L., Zhao, C., Zhu, Y., Jiang, S., and Liu, Y. (2017). Assessment of GPM and TRMM Multi-Satellite Precipitation Products in Streamflow Simulations in a Data-Sparse Mountainous Watershed in Myanmar. *Remote Sensing*, 9(3):302.

Zolina, O., Kapala, A., Simmer, C., and Gulev, S. K. (2004). Analysis of extreme precipitation over Europe from different reanalyses: A comparative assessment. *Global and Planetary Change*, 44(1-4):129–161.

Appendices

A Figures

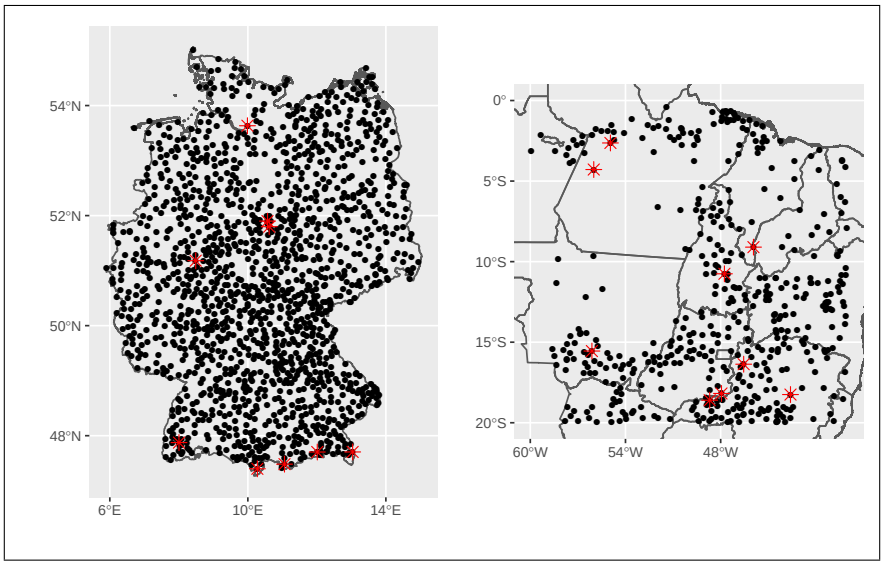

Figure A.1: Stations considered in this work (black dots) and 9 selected stations for Germany (left) and the Brazilian study area (right), respectively.

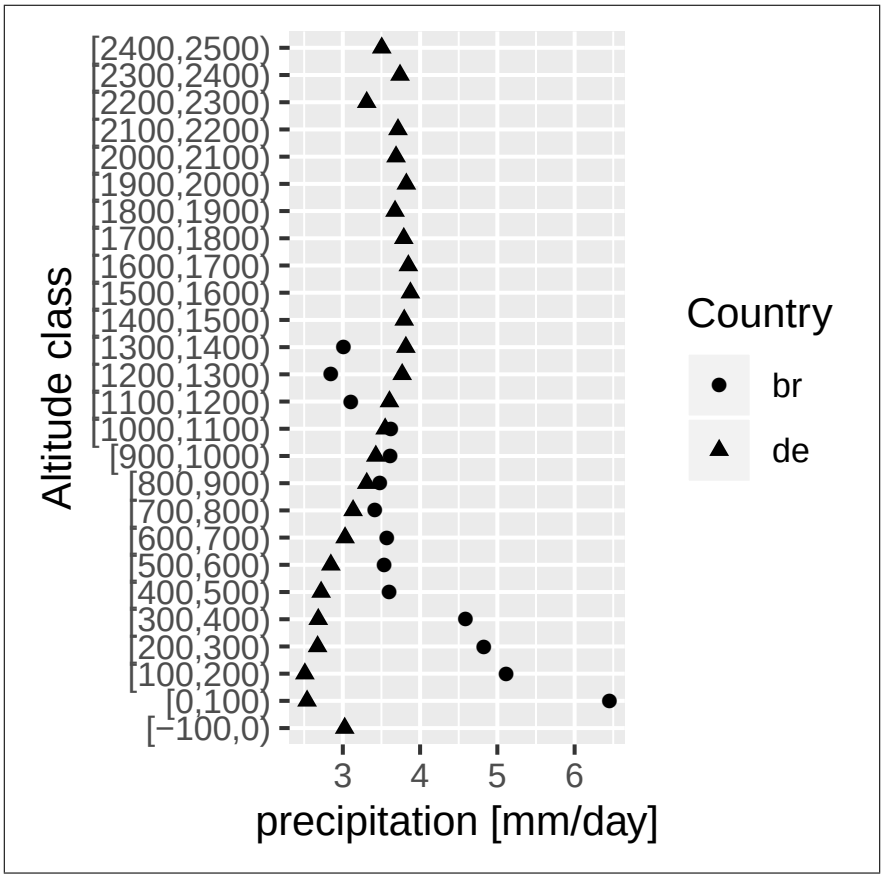

Figure A.2: Distribution of altitude and average daily precipitation from IMERG-F in Germany and Brazil.

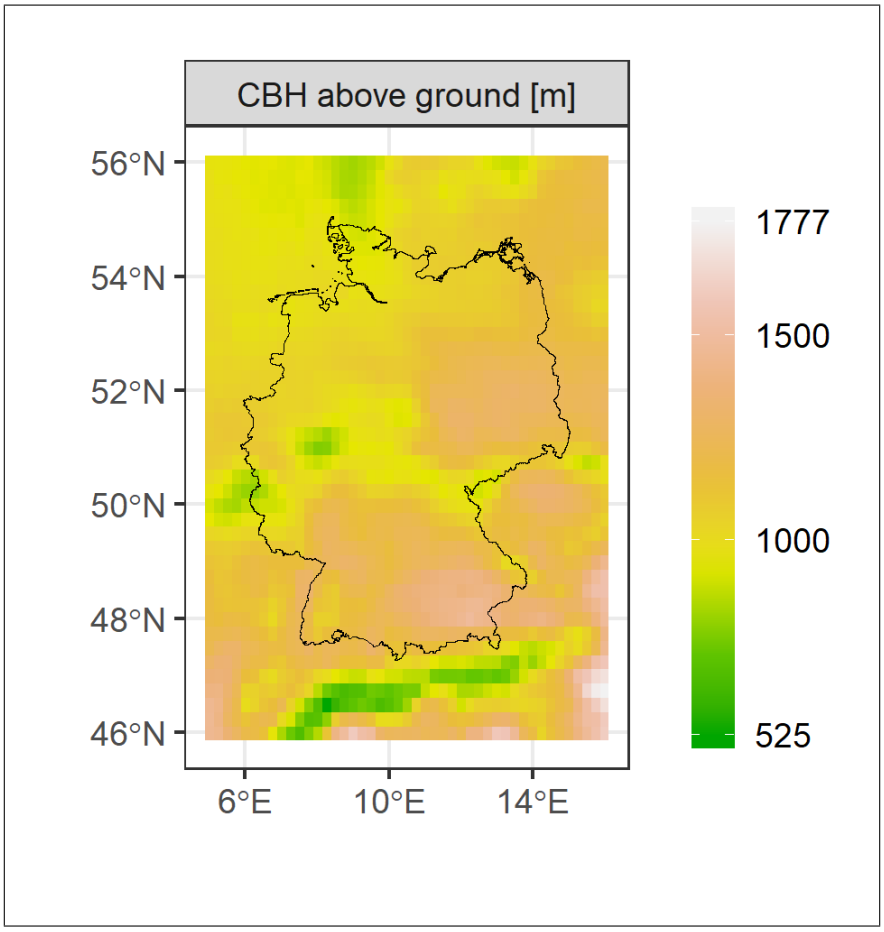

Figure A.3: Average ERA5 CBH (1981-2010) above model topography in Germany.

68

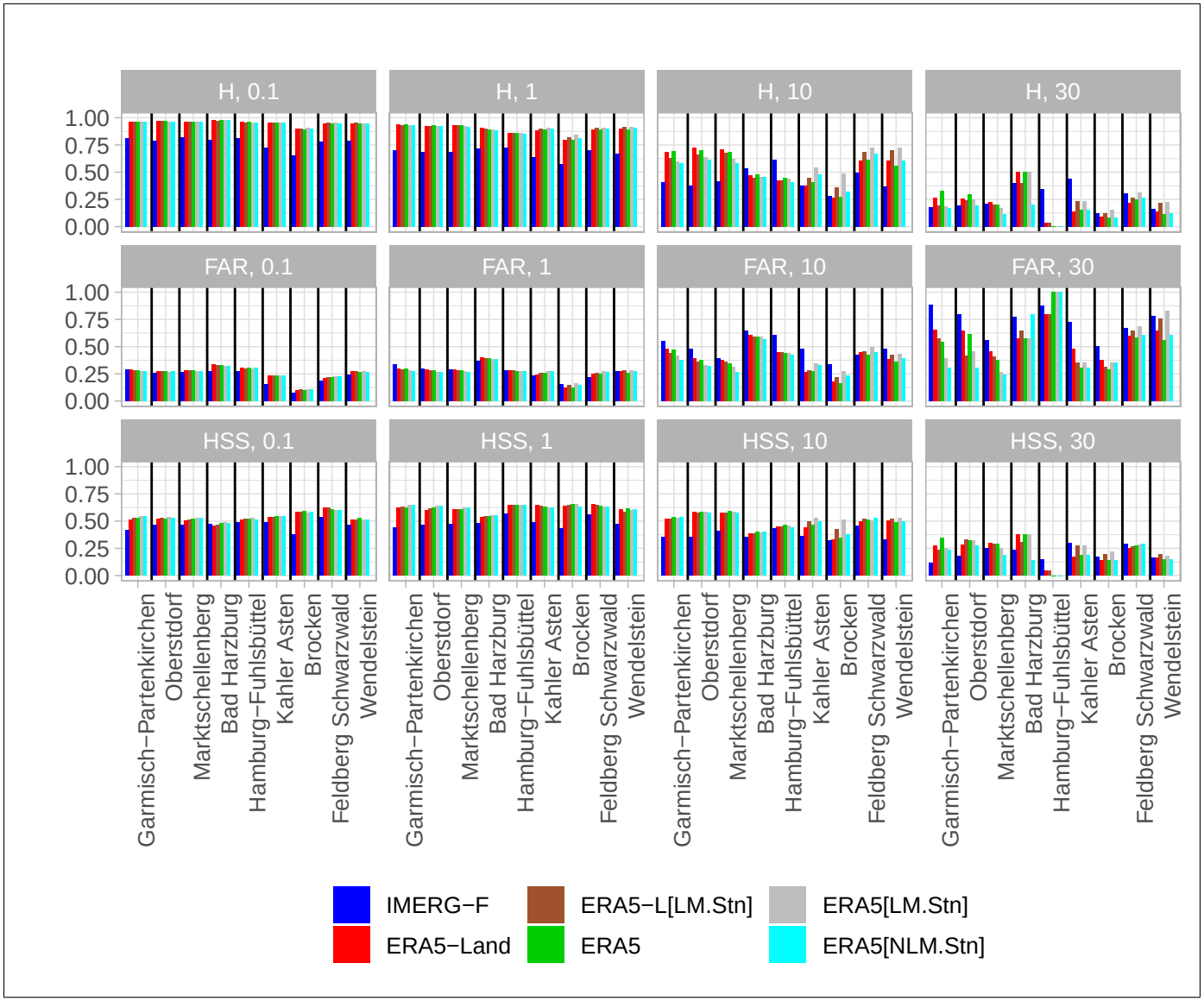

Figure A.4: Hit Rate or POD, FAR and HSS for different precipitation thresholds (0.1, 1.0, 10.0 and 30.0 mm/day) for the the nine selected German stations.

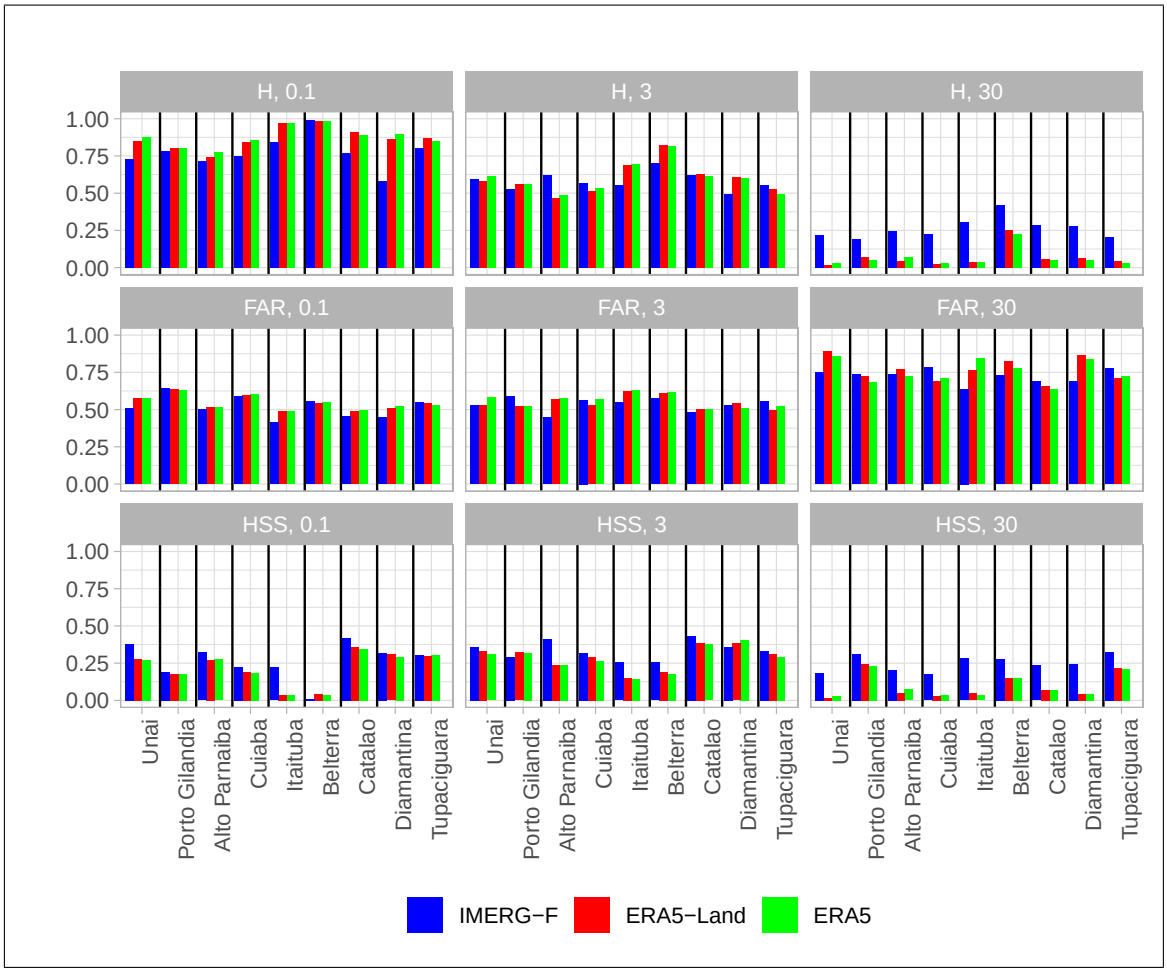

Figure A.5: Hit Rate or POD, FAR and HSS for different precipitation thresholds (0.1, 3.0, and 30.0 mm/day) for the the nine selected in the Brazilian study are.

B Tables

Table B.1: Metadata for the selected stations in Germany and Brazil. The bracketed elevation value is taken from 250 m SRTM data.

Name	Altitude	Lat.	Lon.	Altitude ERA5 Topo.	Altitude ERA5-L Topo.	Difference Altitude-ERA5 Topo.	Missing [%]
Garmisch-Partenkirchen	719	47.48	11.06	1157	1130	-438	0.2
Oberstdorf	806	47.40	10.28	1182	1235	-376	0
Marktschellenberg	501	47.71	13.04	835	865	-334	2.5
Harzburg,Bad	201	51.90	10.57	255	307	-54	58.3
Hamburg-Fuhlsbüttel	14	53.63	9.99	26	15	-12	0
KahlerAsten	839	51.18	8.49	462	644	377	0
Brocken	1134	51.80	10.62	354	684	780	0
Feldberg/Schwarzwald	149	47.88	8.00	681	974	809	0
Wendelstein	1832	47.70	12.01	814	1061	1018	0
Unaí	460	-16.37	-46.55	834	814	-374	0.1
Porto Gilandia	220	-10.76	-47.77	410	374	-190	40.2
AltoParnaíba	285	-9.10	-45.93	418	349	-132.95	0
Cuiabá	151	-15.55	-56.12	246	184	-94.66	12.3
Itaituba	45	-4.28	-56.00	66	32	-21	0
Belterra	176	-2.63	-54.95	65	114	110.74	0
Catalão	840	-18.18	-47.95	724	805	116.47	0.1
Diamantina	1296	-18.25	-43.60	1065	1171	231.12	1
Tupaciguara	(905)	-18.60	-48.69	665	862	240	37.5

Versicherung an Eides statt

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel: "Evaluation of ERA5, ERA5-Land, and IMERG-F precipitation with a particular focus on elevation-dependent variations – A comparative analysis using observations from Germany and Brazil" selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, 12. November 2019

.....

Tobias Kawohl