Universität
Augsburg
University

Spatiotemporal Variability of Precipitation

Measurements - Simulations - Limitations

Dissertation
zur Erlangung des Doktorgrades an der
Fakultät für Angewandte Informatik
der Universität Augsburg

vorgelegt von

Andreas Wagner

2017

Erstgutachter: Univ.-Prof. Dr. Harald Kunstmann
Zweitgutachter: PD Dr. Christoph Beck
Tag der mündlichen Prüfung: 27.11.2017

Abstract

State-of-the-art distributed hydrological models require high quality and high resolution precipitation input in space and time. This applies especially to urban areas with sewage systems, a high degree of soil sealing and limited retention areas. Weather radar systems largely fulfill these requirements and are chosen in this work as a very promising device for precipitation measurement. If datasets of precipitation measurements show a poor quality, their time series are too short or they are not even available, then the simulation of precipitation by state-of-the-art Regional Climate Models (RCM) may close this gap. Nowadays, high resolution simulations are able to provide similar resolutions in space and time as radar data do.

The present work investigates the performance and the limitations of radar measurement and high resolution RCM simulations regarding the spatiotemporal variability of precipitation. The specific objectives are the investigation of systematic effects in radar data and the development of statistical correction algorithms to improve precipitation estimates from radar data. The main innovation regarding the investigation of radar measurements is the long term observation of systematic effects in radar products of the German Met. Service (DWD) based on accumulated radar images. The net effect of these systematic deficiencies are analyzed and used to develop new statistical post-correction schemes which are also applicable to composite radar products. In addition, the added value of high resolution RCM modeling with WRF (Weather Research and Forecasting Model) including the influence of key land surface properties on simulation results of precipitation is investigated. As new features, a new WRF setup for high resolution simulations in Germany is identified. Based on different simulations, the influence of topography, convection parametrization and the application of an Urban Canopy Model (UCM) with respect to the added value of high resolution simulations is evaluated.

In order to identify systematic effects in single radar data as well as additional effects in the German radar composite, a multi-annual analysis of radar data of DWD from 2000 to 2009 is performed. These effects are described and analyzed in this work regarding their dependences on time and reflectivity level. The most striking feature is the dependence of radar measurements on the altitude. On an annual basis, a robust linear relationship can be observed for all radar systems, whereas for shorter temporal scales a higher variability become apparent. Besides effects such as overshooting, partial beam filling and growing range-bin sizes, the dependence of the reflectivity on the state of aggregation of water drops and thus the dependence on temperature mainly cause this feature. Consequently, the winter half year is more affected, not only by a reduction of the reflectivity signal but also by an absence of precipitation echoes. Further systematic effects are caused by clutter remnants, shading behind obstacles and the way of compositing single radar data.

Subsequently, the identified effects and their dependences are then implemented into statistical corrections to eliminate these effects. Four modules are developed to capture these effects namely the altitude correction, the correction of spokes due to shading effects, the adjustment to rain gauge data and clutter correction. The correction schemes are robust, produces reliable precipitation patterns and show significant improvements compared to uncorrected radar data for the annual time scale. With shorter time steps the spatiotemporal variability of precipitation increases significantly and thus the correction algorithm becomes more error-prone. However, the RMSE values for hourly and 5-minute rain amounts compared to available rain gauge data are improved by the applied corrections, both in space and time.

To tackle the objectives of the high resolution RCM simulations, several simulations in two investigation areas in northern Germany (22000 km²) and southern Germany (15000 km²) with grid sizes of 15 km, 5 km and 1 km are performed from June 2005 to May 2006. The simulation results are compared to the new, above mentioned corrected radar dataset and to interpolated rain gauge data. According to the obtained results, a 15 km grid size is hardly able to represent precipitation patterns adequately, in particular in complex terrain, whereas significant improvements of simulation results are observed for grid sizes of 5 km and 1 km. Smaller grid sizes than 5 km are only obviously superior for higher rain intensities. The influence of convection parametrization based on the 5 km grid size becomes obvious for simulated spatial precipitation patterns and the dry hour probability. Nevertheless, a similar quality can be achieved with and without convection parametrization. The impact of the topography becomes apparent in complex terrain, where the localization of precipitation maxima are better realized with smaller grid sizes, but the maxima are overestimated, simultaneously. An improvement of simulated spatial precipitation patterns can be achieved by the additional application of an UCM, although its influence is minor compared to model resolution or convection parametrization.

In general, both, weather radar and high resolution RCM simulations have a high potential to capture the spatiotemporal variability of precipitation. For the weather radar, individual radar image corrections as well as post-corrections have to be applied to minimize systematic effects. High resolution simulations are able to represent the diurnal cycle of precipitation as well as the structure and the intensity distribution of precipitation well. The simulation of the exact locations of precipitation maxima and the temporal accordance with validation datasets is moderate and must be improved in future studies. In this respect, simulations cannot compete with radar data.

Zusammenfassung

Aktuelle diskretisierte hydrologische Modelle erfordern einen Niederschlagsinput von hoher Qualität und hoher räumlicher und zeitlicher Auflösung. Dies gilt insbesondere für Stadtgebiete mit Abwassersystemen, einem hohen Versiegelungsgrad und begrenzten natürlichen Rückhalteflächen. Wetterradarsysteme werden den genannten Anforderungen gerecht und daher in dieser Studie als vielversprechendes Niederschlagsmesssystem verwendet. Sollten die Niederschlagsmessungen jedoch eine unzureichende Qualität aufweisen, ihre Zeitreihen zu kurz sein oder Messungen komplett fehlen, können Regionale Klimamodelle (RCM) diese Lücke schließen. Heutzutage können hochaufgelöste Simulationen ähnliche Auflösungen in Raum und Zeit wie Radardaten liefern.

Die vorliegende Arbeit untersucht Leistungsfähigkeit und Grenzen von Radarmessung und hochaufgelösten RCM-Simulationen hinsichtlich der räumlich-zeitlichen Variabilität des Niederschlags. Konkret werden systematische Fehler und Effekte in Radardaten betrachtet und statistische Korrekturalgorithmen entwickelt, um die Niederschlagsbestimmung aus Radardaten zu verbessern. Die wichtigste Innovation hinsichtlich der Untersuchung von Radarmessungen ist die langfristige Beobachtung systematischer Effekte bei Radarprodukten des Deutschen Wetterdienstes (DWD) auf Basis akkumulierter Radarbilder. Der Nettoeffekt dieser systematischen Mängel dient der Entwicklung neuer statistischer Post-Korrektursysteme, die auch für Kompositradarprodukte gelten. Darüberhinaus wird der Mehrwert der hochaufgelösten RCM-Modellierung mit WRF (Weather Research and Forecasting Model) einschließlich des Einflusses der wichtigsten Landoberflächeneigenschaften auf die Niederschlags-Simulationsergebnisse untersucht. Als Neuerungen werden zum einen ein neues WRF-Setup für hochaufgelöste Simulationen in Deutschland identifiziert. Zum anderen, wird basierend auf verschiedenen Simulationen, der Einfluss der Topographie, der Konvektionsparametrisierung und der Anwendung eines Stadtmodells in Bezug auf den Mehrwert hochauflösender Simulationen ausgewertet.

Die Analyse von Radardaten basiert auf Daten des DWD von 2000 bis 2009, um systematische Effekte in Einzelradardaten sowie zusätzliche Effekte im deutschen Radarkomposit zu identifizieren. Diese Effekte werden in dieser Arbeit hinsichtlich ihrer Abhängigkeit von Zeit und Reflektivitätsniveau beschrieben und ausgewertet. Auffällig ist die Abhängigkeit von Radarmessungen von der Messhöhe. Auf jährlicher Basis zeigt sich für alle Radarstandorte eine robuste lineare Beziehung, während bei kürzeren zeitlichen Skalen eine höhere Variabilität sichtbar wird. Dies beruht auf der Abhängigkeit der Reflektivität vom Aggregatzustand des Wassers und damit der Abhängigkeit von der Temperatur. Daneben tragen die Messung in Höhen, in denen die Radarvoxel nur noch teilweise oder gar nicht mehr mit Niederschlag gefüllt sind, und die generelle Vergrößerung der Radarvoxel mit der Entfernung vom Radar zu dieser Variabilität bei. Infolgedessen ist das Winterhalbjahr sowohl durch verringerte Reflektivitätssignale als auch durch ausbleibende Niederschlagsechos stärker betroffen. Weitere systematische Effekte werden durch Clutterreste, Abschattungseffekte hinter Hindernissen sowie die Art und Weise der Kompositierung von Einzelradardaten verursacht.

Um diese Effekte zu eliminieren, werden sie und ihre Abhängigkeiten in statistische Korrekturen umgesetzt. Die vier neu entwickelten Module sind: die Höhenkorrektur, die Korrektur von Speichen durch Abschattungseffekte, die Aneichung an Stationsdaten und die Clutterkorrektur. Die Korrekturschemata sind robust, produzieren zuverlässige Niederschlagsmuster und zeigen deutliche Verbesserungen gegenüber unkorrigierten Radardaten auf jährlicher Zeitskala. Bei kürzeren Zeitschritten nimmt die räumlich-zeitliche Variabilität des Niederschlags deutlich zu, womit der Korrekturalgorithmus fehleranfälliger wird. Trotzdem werden die RMSE-Werte für stündliche und fünf-minütige Regenmengen im Vergleich zu den verfügbaren Regenmessdaten durch die Korrekturen sowohl räumlich als auch zeitlich verbessert.

Zum Erreichen der Ziele der hochauflösenden RCM-Simulationen werden von Juni 2005 bis Mai 2006 verschiedene Simulationen in zwei Untersuchungsgebieten in Norddeutschland (22000 km²) und Süddeutschland (15000 km²) mit Rastergrößen von 15 km, 5 km und 1 km durchgeführt. Die Simulationsergebnisse werden mit dem neuen korrigierten Radardatensatz und mit interpolierten Stationsdaten verglichen. Demnach ist eine Rastergröße von 15 km kaum in der Lage, Niederschlagsmuster adäquat, insbesondere im komplexen Gelände, darzustellen, während für Rastergrößen von 5 km und 1 km deutlich bessere Simulationsergebnisse beobachtet werden. Kleinere Rastergrößen als 5 km sind nur bei höheren Regenintensitäten offensichtlich überlegen. Der Einfluss der Konvektionsparametrisierung basierend auf der 5 km Gittergröße wird für simulierte räumliche Niederschlagsmuster und die Trockenzeitwahrscheinlichkeit auf stündlicher Basis deutlich. Trotzdem ist die erzielte Qualität mit und ohne Konvektionsparametrisierung ähnlich. Der Einfluss der Topographie wird im komplexen Gelände sichtbar: hier werden die Niederschlagsmaxima mit kleineren Rastergrößen besser lokalisiert, auch wenn die Maxima gleichzeitig überschätzt werden. Eine Verbesserung der simulierten räumlichen Niederschlagsmuster kann durch die zusätzliche Anwendung eines Stadtmodells erreicht werden, obwohl dessen Einfluss im Vergleich zur Modellauflösung oder Konvektionsparametrisierung gering ist.

Im Allgemeinen haben sowohl das Wetterradar als auch die hochauflösenden RCM-Simulationen ein hohes Potential, die räumlich-zeitliche Variabilität des Niederschlags zu erfassen. Für das Wetterradar müssen individuelle Radarbildkorrekturen sowie Post-Korrekturen angewendet werden, um systematische Effekte zu minimieren. Hochauflösende Simulationen sind in der Lage, den täglichen Zyklus des Niederschlags sowie die Struktur und die Intensitätsverteilung des Niederschlags gut darzustellen. Die Simulation der genauen Lokalisierungen der Niederschlagsmaxima und die zeitliche Übereinstimmung mit Validierungsdatensätzen ist mäßig und muss in zukünftigen Studien verbessert werden. Diesbezüglich können Simulationen nicht mit Radardaten konkurrieren.

Contents

List of Figures

List of Tables

List of Symbols

η	Radar reflectivity of the target in m^{-1}
λ	Wavelength in m
ϕ	Antenna beamwidths in the vertical plane in radians
σ	Radar cross section in m^{2}
θ	Antenna beamwidths in the horizontal plane in radians
a	Z/R constant
alt	Altitude in km
ARW	Advanced Research WRF
b	Z/R constant
BEP	Building Environment Parametrization
$BouLac$	Bougeault–Lacarrère
C	Specific radar constant in W m^{-1}
$CCOR$	Clutter Correction parameter
CDO	Climate Data Operators
$CONRAD$	CONvection in RADar products
$CORINE$	Land use dataset
CPS	Convection Permitting Scale
D	Drop diameter in mm
$D-North$	Investigation area in Northern Germany
$D-South$	Investigation area in Southern Germany
$D01$	WRF domain with 15 km grid size
$D02$	WRF domain with 5 km grid size
$D03N$	WRF domain with 1 km grid size in Northern Germany
$D03S$	WRF domain with 1 km grid size in Southern Germany
dBZ	Radar reflectivity factor in dB
DWD	German Meteorological Service
DX	Single radar product in RVP6-units
$E-OBS$	Gridded dataset based on station data
f_{cor}	Correction factor in km^{-1}
FSS	Fractions Skill Score
G	Antenna gain (ratio)

GCM	General Circulation Model
$GDALWARP$	Image reprojection and warping utility
GF	Grell-Freitas
$Global-WRF$	WRF for modeling on a global scale
GPM	Global Precipitation Measurement
$GSLIB$	Geostatistical Software Library
h	Pulse length of the signal in m
$HWRF$	WRF for simulations of hurricanes
I_o	Observed binary field
I_x	Simulated binary field
IIR	Infinite Impulse Response
K_{DP}	Specific differential phase in deg km^{-1}
KF	Kain-Fritsch
LES	Large-eddy-simulation
LOG	LOG Channel
LSM	Land Surface Models
m_r	Mean slope depending on range in km^{-1}
m_{alt}	Mean slope depending on height in km^{-1}
MET	Model Evaluation Tools
$MM5$	Mesoscale Meterology Model 5; predecessor of WRF
$MSE_{(n)ref}$	Largest obtainable MSE with the given observed and simulated fractions
$MSE_{(n)}$	Mean Squared Error
MYJ	Mellor–Yamada–Janjić
n	Horizontal scale in km
$N(D)$	Amount of drops
$NCUP$	WRF simulations without CUmulus Parametrization at 5 km grid size
NMM	Nonhydrostatic Mesoscale Model
$NOAH-LSM$	NOAH Land Surface Model
$NURB$	WRF simulations without URBan Canopy Model
$O_{(n)}$	Observed fractions
Or	Observed precipitation field
P_r	Received power in W
P_t	Transmitted power in W
PBL	Planetary Boundary Layer
PDF	Probability Density Function
PRF	Pulse-Repetition Frequency
PX	Classified single radar product in RVP6-units
q	Precipitation threshold
QQ	Quantile-Quantile
R	Rain rate in mm
r	Distance between antenna and target in m
RAD	Corrected RADar dataset

$radar$	RAdio Detecting And Ranging
$RADOLAN$	Radar Online Adjustment
$RADU$	Uncorrected RADar dataset
RCM	Regional Climate Model
$REGNIE$	Regionalised precipitation heights
$RMSE$	Root-Mean-Square-Error in mm
$RRTM$	Rapid Radiative Transfer Model
RU	Radar composite product in mm, which is adjusted and merged with rain gauges
RW	Radar composite product in mm, which is adjusted to rain gauges
RX	Composite radar product in RVP6-units
$SLUCM$	Single-Layer Urban Canopy Model
sp	Percentile (spatially)
SQI	Signal Quality Index
STA	Interpolated rain gauge dataset
$SYNOPSE$	Synthetic precipitation time series for the optimal planning and operation of urban drainage systems
t	Percentile (temporally)
TKE	Turbulent Kinetic Energy
UCM	Urban Canopy Model
$UKVmodel$	Met. Office Unified Model
URB	WRF simulations with URBan Canopy Model
$USGS$	United States Geological Survey
$v(D)$	Fall velocity in m s^{-1}
VIL	Vertically integrated liquid water content in kg m^{-2}
VPR	Vertical Profile of Reflectivity
WPS	WRF Preprocessing System
WRF	Weather Research and Forecasting Model
$WRF-CHEM$	WRF for simulations of air chemistry
$WRF-DA$	WRF Data Assimilation
$WRF-HYDRO$	WRF for hydrological modeling
$WSM6$	WRF Single-Moment 6-class scheme
$X_{(n)}$	Simulated fractions
Xr	Simulated precipitation field
Y	Frequency of occurrence of radar reflectivity
Y_0	Frequency of occurrence of radar reflectivity at ground level
YSU	Yonsei University scheme
Z	Radar reflectivity factor in mm^6 m^{-3}
Z_{DR}	Differential reflectivity in dB

1 Introduction

> *... The general problem of representativeness is particularly acute in the measurement of precipitation. Precipitation measurements are particularly sensitive to exposure, wind and topography, and metadata describing the circumstances of the measurements are particularly important for users of the data.*
>
> — WMO [2014]

1.1 Motivation

Precipitation is one of the main components of the water cycle. The knowledge about the spatiotemporal distribution of precipitation is of utmost interest for hydrology, as it is the main driving variable for many processes in hydrology. The precipitation amount, its annual cycle, intensity and duration as well as the type of rain directly influences processes such as runoff generation [Dyck and Peschke, 1995]. State-of-the-art distributed hydrological models are able to operate on daily or hourly resolution and grid sizes of 1 km, but are still limited by the quality of precipitation input [Faures et al., 1995, Fu et al., 2011]. This becomes even more acute e.g. in urban areas with sewage systems, a high degree of soil sealing and limited retention areas. Due to the very short response time of runoff, hydrological models in urban areas have to operate on a time step of about 5 minutes [Berne et al., 2004]. Moreover, precipitation shows a very high temporal and spatial variability during severe thunderstorms that can produce flash floods in urban areas [Zoccatelli et al., 2010]. Since a high damage and hazard potential is associated with these events [Einfalt et al., 2009], warning and nowcasting systems exist, which have to operate on a very high temporal and spatial resolution and require high resolution precipitation input (e.g. CONRAD[1] [Lang, 2001]) .

Therefore, there is an increasing need for high quality and high resolution precipitation data to capture those characteristics, both, in space and time.

In general, there are two fundamentally different approaches to generate precipitation patterns in a high spatial and temporal resolution: purely driven based on measured data and generated by models.

---

[1] CONRAD (CONvection in RADar products) is based on radar data with 5 min time step and 1 km grid size.

Regarding measured data, different techniques exist that provide either precipitation data as single point measurements, path averaged rain rates or pattern estimations. The reference for precipitation measurements usually is the rain gauge. It provides reliable point measurements but reveal a limited spatial representativeness. Statistical interpolation techniques such as kriging [Goovaerts, 1998, 2000, Krige, 1951] or copula-based schemes [Bárdossy and Pegram, 2013] are used to generate areal precipitation patterns from this point measurement. However, the quality of interpolated precipitation fields strongly depends on the network density of rain gauges. Nevertheless, these techniques are hardly capable to reproduce precipitation patterns correctly, especially if single rain cells are not even detected by rain gauges [Mishra, 2013]. A more recent measurement technique is based on microwave data providing path averaged rates of precipitation [Chwala et al., 2012]. It shows a good coverage in inhabited areas. But again interpolation techniques are needed to generate precipitation fields. Observation techniques, that directly measure precipitation patterns, are remote sensing techniques such as satellite or weather radar measurements. According to Yoon et al. [2012] satellite based precipitation fields are often superior to interpolated point measurements even though the detected precipitation echoes have to be transferred to rain rate. The most innovative satellite system namely the Global Precipitation Measurement (GPM) provides a resolution of 5 km and a temporal resolution of up to 3 hours [GPM, 2013, Hou et al., 2014]. Especially the temporal resolution limits its usage in common hydrological models. By contrast, weather radar systems show a spatial resolution of usually 1 ° and about 1 km in range direction and a temporal resolution between 3 and 5 minutes [Seltmann, 1997]. Additionally, the coverage of radar systems in industrial countries is excellent and also in many developing countries radar systems exist.

If precipitation datasets show a poor quality, their time series are too short or they are not even available, then the simulation of precipitation by state-of-the-art regional climate models (RCM) may close this gap. Examples, where long time series are needed, which measurements usually cannot provide, are the design and projection of water management works and systems, the design of sewage systems in cities and the simulation of floods with a low recurrence interval (e.g. [Berne et al., 2004]). Furthermore, RCMs are able to additionally calculate certain scenarios or to estimate future development - both, forecasts and future projections. On the one hand, these possibilities, which a model offers, are very attractive, but on the other hand, the temporal and spatial variability of precipitation places high demands on models' performance.

In general, the main challenge of generating a reliable precipitation dataset is to derive small-scale precipitation patterns in a high quality [Beven, 2011]. The high temporal and spatial resolution of radar measurements make them appear as the most promising measuring device for spatial precipitation patterns and to represent the spatiotemporal variability of precipitation. Radar data is regarded to provide reasonable precipitation patterns on a short temporal scale (< 1 h), but not much is known about its performance on a longer temporal scale (months or years). To generate reliable precipitation fields based on radar measurements, a number of corrections have to be applied to detect and eliminate erroneous measurements from a radar image. In contrast, high resolution RCMs are also able to provide precipitation fields with similar resolutions as weather radar data do. Due to the further applications of simulated precipitation fields, models can be regarded as a complement to measurements and their performance at these small scales have to be evaluated.

1.2 State-of-the-art

1.2.1 Quality issues of single and composite radar data

Weather radar systems provide measurements of precipitation echoes till a distance of 50 to 150 km from the radar site with a high temporal and spatial resolution. The radar reflectivity Z, that is derived from the precipitation echoes has to be converted to rain rate R to achieve areal precipitation fields. This is usually performed by applying an empirical Z/R relationship [Battan, 1973]. More information about the basics of weather radar measurements can be found in Chapter 2. The way of measuring, i.e. scanning the circumambient air masses for precipitation from one specific point, can lead to some unwanted effects in radar data which will be presented in the following:

There is a huge variety of sources of errors in data of single radar systems that have to be considered. The uncertainty of deriving rain amounts from radar measurements using a Z/R relationship is well known [Steiner et al., 2004, Uijlenhoet et al., 2003]. Errors of comparable magnitudes may be induced by the influence of the melting layer (Bright Band) affecting huge parts of the radar image [Fabry and Zawadzki, 1995, Franco et al., 2006, Krajewski et al., 2010, Vignal et al., 1999], especially in the temperate zone. Further possible influencing factors for weather radar measurement are shown in Fig. 1.1. Attenuation behind strong convective cells, shading effects of buildings or mountains and artificial effects such as interfering transmitters, the sun and dual PRF (pulse-repetition frequency) second-trip echoes lead to spokes in radar images [Meischner, 2004]. Non-meteorological echoes caused by birds, insects, airplanes,

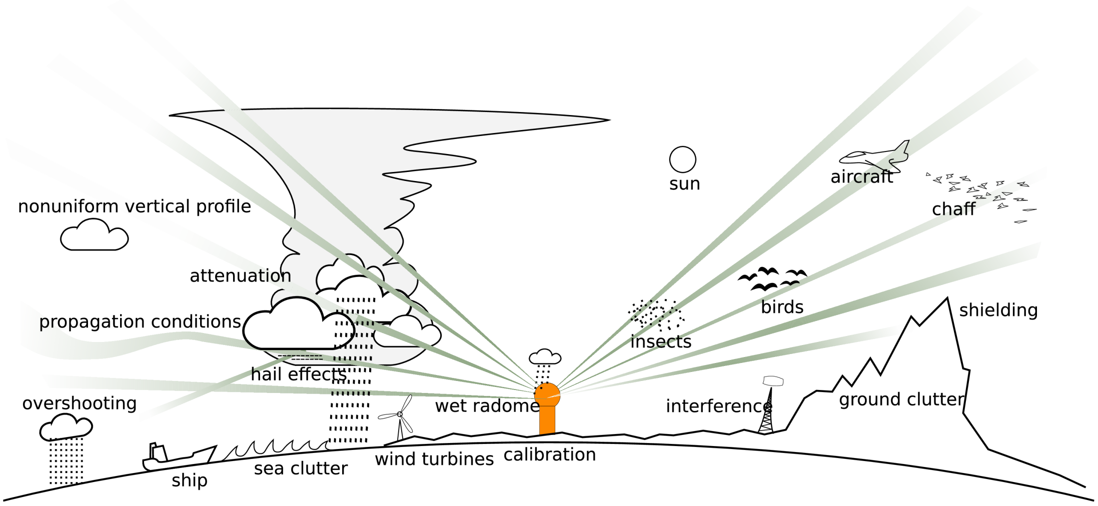

Figure 1.1: Overview of errors which can influence quantitative precipitation estimation using weather radar. Adapted from [Overeem et al., 2009], his Fig. 1.3.

ships or windmills usually affect single pixel or a couple of pixels. Additionally, the geometry of weather radar measurements has a significant influence on the quality and the comparability of radar range-bins within the image. On the one hand, variations that are induced by natural differences of reflectivity and rain amounts at different altitudes, partial beam-filling, overshooting and transition from snow into rain, are intensified by varying beam-width and increasing altitudes of range-bins with distance from the radar site. On the other hand, the probability that a radar pixel is influenced by shading or attenuation increases with distance from the radar site. These main problems and limitations of radar measurements apply to weather radar systems in all countries.

As a consequence, corrections have to be developed and applied before radar data is further processed [Holleman, 2007]. There is a huge variety of correction algorithms for individual radar images. For example, Gabella and Notarpietro [2002] presented a clutter correction scheme, Krämer and Verworn [2008] developed a correction for attenuation and Kitchen et al. [1994]and Koistinen [1991] developed algorithms to correct the Vertical Profile of Reflectivity (VPR). The wradlib libraries [Heistermann et al., 2013] provide a compilation of correction algorithms in Python.

The demand for high quality radar data increased in the recent years due to new application fields. In particular, large area analyses and the assimilation of weather radar data into numerical weather prediction models have increased the importance of composite radar products. To provide e.g. a nationwide precipitation product, the radar images of single radar systems have to be merged. This new product is called radar composite. The generation of composites covering large geographical areas such as several northern European countries in NORDRAD and BALTRAD [BALTRAD, 2014, Collier, 1992] or entire Europe [OPERA, 2017] creates the need for consistent data quality and comparable input data. The COST program [COST, 2017] e.g. helps to meet these standards. Meanwhile, weather services operate nationwide networks (e.g. Finland, Spain, France, United Kingdom, Germany, etc.) with their own philosophy, scan strategy or correction schemes. In most countries the composite is created based on volume data (several sweeps with different beam elevations) where the range of each radar system is covered by more than one elevation. Thus, it is possible to bypass obstacles. In Germany, the composite for quantitative means is generated based on the terrain-following precipitation scan to provide compositing products as frequently as possible (see Chapter 2).

However, generating a composite image is not straightforward, since many additional effects produced by the compositing algorithm have to be taken into account [Seo et al., 2013]. The main source of error emerges from the compositing algorithm itself. Range-bins of the single radar measurements in polar coordinates from different radar sites have to be merged and projected to the Cartesian system of the composite. Another possible source of error in composite data is non-identical radar hardware. But even if the same radar systems are used, the radar availability and calibration can differ slightly. This may lead to spatial differences in the composite depending on the radar site that provides the measurement.

A further problem of composite images are erroneous measurements coming from a single radar which are directly transferred to the composite. That is the reason why many statistical corrections are usually performed on single radar data. Differences between two single radar areas[2] are only corrected by the adjustment to rain gauges when the reflectivities are transferred to the rain rate. But calibration techniques are meanwhile often used that are independent from the radar site and should provide almost identical calibrations. Various calibration techniques exist in this respect, e.g. using the sun [Holleman et al., 2010] or a setup of distrometers combined with a vertical weather radar beam [Frech, 2013]. A correction algorithm that was developed particularly for radar composites is presented by Koistinen et al. [2003]. They developed a VPR correction scheme for the Finnish weather radar network.

For single time steps or short periods (e.g. days), a thorough usage of correction algorithms on single radar images usually provides a good data basis even for areal precipitation fields of radar composites. Weak spatial deviations or small systematic differences within the radar image are usually negligible on these scales. But for a long-term use of radar data, systematic deviations can accumulate and may therefore cause major errors [Fairman et al., 2015, Wagner et al., 2012]. According to Fairman et al. [2015] less is known about the performance of radar precipitation composite as input to water balance models over multi-annual scales, but an influence of these effects are probable. Regarding the length of the time series, radar data cannot compete with a common precipitation climatology, which is usually based on a multi-decadal time series of rain gauges (e.g. Beck et al. [2005]). Nevertheless, those long-term radar time series are attractive for a long-term statistical analysis of precipitation patterns to prepare e.g. a radar based precipitation 'climatology' (radar climatology) [Carbone and Tuttle, 2008, Fabry et al., 2013, Fairman et al., 2015, Overeem et al., 2009, Tabary et al., 2012]. Many authors make an effort to correct one specific physical effect in single radar images and rely only on these, even for radar climatology (e.g. Overeem et al. [2009]). All of these long-term radar data investigations aim at analyzing precipitation patterns. These analyses are only additionally used in terms of creating static clutter maps [Meischner, 2004], but not for the identification of any other systematic deviations. In general, systematic deviations on a longer temporal scale are not considered for correction so far.

1.2.2 High resolution modeling of precipitation

The simulation of precipitation can be achieved by stochastic models, dynamic models as well as hybrids of both [Keefer, 2003]. Many stochastic approaches usually simulate only a few variables at one or a few locations whereas dynamical models can provide a variety of meteorological variables on a regular grid in a consistent way. Regarding the interest on the spatial distribution of precipitation amounts in this work, dynamical models are one of the most promising tools. Computation devices nowadays provide the opportunity to perform high resolution simulations using dynamical models similar to weather radar systems.

Dynamical models are designed to emulate physical processes and describe them in a mathematical way. Current General Circulation Models (GCMs) used in global climate modeling, often use grid-spacing larger than 100 km [Taylor et al., 2012] or even just below 100 km. Nested Regional Climate Models (RCMs) are able to reduce this resolution to the lower kilometer range or even below that. The finer the model grid, the more bio-geophysical variabilities such as land use and topography can be considered within model simulations

---

[2] A single radar area is an area where only one radar site provides measurements.

[Prein et al., 2015]. However, it requires shorter calculation time steps to achieve stable simulations and therefore longer computation times and higher storage capacities. Furthermore, the total size of the investigation area has to be reduced compared to coarser grid sizes to achieve reasonable computation times. The quality of simulation results depends on how well processes are described in the model at the specific scale. Finer spatial resolutions lead to faster-growing errors and these small scales become unpredictable more quickly in dynamical models [Lorenz, 1969, Mittermaier and Roberts, 2010, Zhang et al., 2003]. Many sub-grid scale processes such as the representation of deep convection are usually parametrized. The gap between explicitly resolved flows and implicitly parametrized flows – called the ’gray zone’ – has to be taken into account. With finer resolutions these gray zones are often reached. Undesired side-effects may occur that deteriorate simulation results [Zhou et al., 2014]. As a consequence, it is much more challenging to receive stable and reliable simulation results at finer grid resolutions. On the contrary, parametrization of critical sub-grid processes, which is regarded as a major source of uncertainties in climate modeling, may become redundant [Knight et al., 2007, Prein et al., 2013a, Sanderson et al., 2008].

One of the first high resolution simulations in the Convection Permitting Scale (CPS; 4 km) was done in 2000 by Grell et al. [2000]. They used a dynamical downscaling approach with a minimum grid size of 1 km in southern Bavaria. They showed significant differences regarding precipitation patterns and maxima between high resolution simulations and coarser grid sizes in complex terrain. But they used a very small domain, where a self-development of precipitation is hardly possible. Further early simulations in the CPS, also in southern Bavaria, were performed by Kunstmann et al. [2004] and Kunstmann and Stadler [2005]. A coupling with hydrological models was realized to benefit from this high resolution meteorological output. The following years, only a few studies were conducted for high resolution simulations, most of them in complex terrain to benefit from the higher resolution of topography at small grid sizes regarding the representation of spatial patterns and temporal performance. Examples are presented in [Baldauf et al., 2011, Hohenegger et al., 2008, Weusthoff et al., 2010]. With the progress in computation capacities in recent years, a number of high resolution simulations in the CPS were published and presented in Prein et al. [2015]. For instance, Kendon et al. [2012] used one configuration of the Met. Office Unified Model (UKV model) with grid sizes of 12 km and 1.5 km to evaluate differences in simulation results regarding the duration and spatial extent of hourly rainfall in Great Britain over 20 years. Prein et al. [2013a,b] used the WRF model (Weather Research and Forecasting Model) in the Alps and in Colorado to analyze the added value of higher resolutions (3 km / 4 km vs. 10 km / 12 km) regarding heavy rainfall. Ban et al. [2014] utilized the COSMO-CLM model with 12 km and 2.2 km grid sizes in the Alps for 10 years also with the intention to achieve an added value due to smaller grid sizes.

Fosser et al. [2015] performed simulations with the COSMO-CLM model in Baden-Württemberg with 7 km and 2.8 km resolution over 30 years. To evaluate the simulation results of 7 km grid size in the gray zone of convection, they used another 10 year simulation with 7 km grid size and without parametrization of deep convection. They presented significant improvements for the simulation without deep convection parametrization at the 7 km grid size.

The largest added value in these investigations was observed during convective events and for sub-daily periods. In winter months the added value of high resolution simulations was rather restricted to the better orographic resolution [Prein et al., 2013b]. For daily time-scales the results were varying. Ban et al. [2014] showed improvements due to smaller grid sizes only for the highest precipitation quantiles in mountainous regions whereas Brisson et al. [2015] presented a general added value for higher resolutions.

Another aspect that becomes more important with higher resolutions is the representation of the land use besides the topography. Tölle et al. [2014] showed the impact of land use by a simple soil vegetation-atmosphere model. Urban Canopy Models (UCM) can additionally be used to simulate urban specific effects such as the heat island effect, e.g. Wouters et al. [2013] compared this effect for temperature for different grid sizes in Paris.

1.3 Objective and Research Questions

The overall objective of this thesis is the elaboration of opportunities and limitations of the generation of precipitation fields regarding the spatiotemporal variability of precipitation. Therefore, weather radar data and high resolution dynamic modeling are evaluated.

The first part of the thesis addresses the quality of single radar and composite radar products provided by the German Met. Service (DWD) on a long temporal scale. Based on this analysis, the following specific research questions are explored in detail.

(i) **Which types of systematic effects in radar data do exist and what is their influence on data quality?**

Non-meteorological spatial patterns and systematic spatial variations within radar data shall be identified on a large temporal scale. The dependences of these effects on reflectivity level and time will be evaluated and interpreted.

(ii) **How can systematic effects in radar data be corrected and which quality can be achieved in this way?**

After identifying systematic effects in radar data, a statistical post-correction scheme will be developed to correct those effects. Then, the improvements due to these corrections have to be analyzed and the influence of the temporal resolution on the performance of such statistical correction algorithms regarding the quality of radar derived precipitation data will be evaluated.

The second part examines the ability of RCMs to generate high resolution precipitation fields in space and time.

(iii) **What is the added value of high resolution modeling?**

First a suitable model configuration for a high resolution nested RCM simulation in Northern and Southern Germany will be identified. The obtained simulation results with different grid sizes will be analyzed regarding the representation of temporal and spatial characteristics of precipitation. One particular aim is to elaborate whether to omit convection parametrization in the gray zone of convection[3] or to apply new parametrization schemes.

(iv) **What is the impact of key land surface properties such as topography or a more complex representation of urban areas on the simulation results?**

Besides physical constraints of simulations, the influence of the topography on the results of high resolution simulations of precipitation will be investigated based on two different investigation areas. Furthermore, the influence of additional Urban Canopy Models (UCM) on the simulation results of precipitation will be analyzed.

(v) **Are weather radar measurements and high resolution Regional Climate Models (RCMs) able to represent the spatiotemporal variability of precipitation in a suitable way?**

The assets and drawbacks of weather radar measurements and high resolution RCMs for small and large temporal and spatial scales will be evaluated. Additionally, the ability of both datasets to represent the distribution of rain intensities on the highest available temporal and spatial resolution will be compared.

1.4 Innovations

Weather radar

- This comprehensive analysis of systematic limitations in weather radar images of DWD is a novelty. The impact of the dependence of radar measurements on altitude and the huge influence of the compositing algorithm (see Chapter 3) on a long temporal scale have been unknown, at least quantitatively, so far.
- Based on the aforementioned outcomes, a novel statistical post-correction scheme has been designed and developed. The idea to use the net effect of a summation of limitations has been developed within this framework as well as the selection of important effects and their transfer into several correction modules.

The above mentioned innovations, explicitly for the single radar of Munich, are published in the following peer-reviewed paper:

Wagner, A., Seltmann, J., and Kunstmann, H. (2012). Joint statistical correction of clutters, spokes and beam height for a radar derived precipitation climatology in southern Germany. Hydrol. Earth Syst. Sci., 16: 4101–4117.[4]

Dynamic modeling with WRF

- The setup of the WRF model configuration in a high resolution in Northern and Southern Germany is new.
- In this work, a comprehensive combination of high-resolution simulations using WRF

---

[3] The gray zone of convection represents grid sizes between approx. 4 to 10 km, where common convection parametrization schemes are not designed for but which are too coarse for an explicit calculation of convection.

[4] The post-correction scheme is solely developed, implemented and validated. The main results and figures are adopted from this paper and can be found in Chapter 3 and in Chapter 4.1. Only the validation is modified to meet more strict statistical requirements such as split-sampling.

with different grid sizes, modules, parametrizations and topographies is applied. In contrast to all mentioned investigations in Chapter 1.2, which only focus on one specific investigation area, the influence of topographies on the performance of high resolution simulations can be evaluated by using two different investigation areas with varying topographies. In this way the differences between the results of Ban et al. [2014] and Brisson et al. [2015] regarding daily precipitation values in different areas can be analyzed.

- The gray zone of convection is circumvented in most investigations. Only Fosser et al. [2015] used such a grid size, but they only analyzed the influence of convection parametrization on the diurnal cycle of precipitation. Additionally, they applied a convection parametrization scheme that is not designed for such small grid sizes. Here, the advantages and disadvantages of state-of-the-art convection parametrization schemes are investigated due to the interest of these grid sizes for long-term climate simulations.
- An urban model is applied to better represent cities. The overall effect for the entire domain and not only for urban pixels is aimed at. The urban models are applied even for coarser grid resolutions to evaluate their influence on simulation results. This is a novelty for longer CPS simulations.

1.5 Outline

The outline is as follows:

Chapter 1 includes the motivation for the present work and the formulation of research questions and innovations.

Chapter 2 presents the basic information about weather radar measurements. It also gives an overview about radar products and observational datasets used for adjustment and validation of radar measurements.

Chapter 3 comprises the statistical analysis of single radar data and composite radar data. The individual systematic error sources are identified and analyzed. Research question (i) is evaluated in this Chapter in detail.

Chapter 4 includes the development of correction schemes for the specific systematic effects that are identified in Chapter 3. First, the schemes for single radar data are developed and their quality is analyzed. Afterwards, these schemes are extended so that they can also be applied for composite radar data. Finally, the schemes are modified to be applicable also for higher temporal resolutions. A validation with rain gauge data is performed for each of the three versions of the correction scheme. Research question (ii) is answered in this Chapter.

Chapter 5 reveals theoretical aspects of dynamical modeling with WRF. First, an overview of the WRF model and the study area as well as the applied setup is given. Then urban modeling within WRF is explained and the practical implementation is described. In addition, preliminary results of the testing process to derive a satisfactorily WRF setup for high resolution simulations in Northern and Southern Germany are additionally briefly summarized.

Chapter 6 is the main chapter of the analysis of high resolution WRF simulations with regard to the representation of the spatiotemporal variability of precipitation. These evaluations have been performed within the SYNOPSE project[5]. First, the model simulations and the observational datasets that are used for comparison are presented. Then, the evaluation methods to determine the model performance in space and time are described in detail. Finally, the outcomes of the WRF high resolution simulations are presented and discussed in comparison to several high resolution observation datasets. Research questions (iii) and (iv) are answered in this Chapter.

Chapter 7 opposes and discusses the results of the evaluations of weather radar data and simulation results with respect to the spatiotemporal variability of precipitation. Research question (v) is answered in this Chapter.

Chapter 8 presents the conclusion and the outlook of this work regarding the research questions. Furthermore, suggestions for improvements and possible prospective investigations are presented.

---

[5] The high resolution WRF simulations have been performed within the framework of the SYNOPSE project (Synthetic precipitation time series for the optimal planning and operation of urban drainage systems). This project is funded by BMBF and aims at generating high resolution, reliable and long-lasting synthetic precipitation time series for the three German cities Braunschweig, Freiburg and Hamburg for the planning and optimization of urban sewage systems [Krämer et al., 2016, NAWAM-INIS, 2014]. Long-term reanalysis and climate simulations with WRF over 30 years on a 5 km grid are disaggregated and downscaled with the help from high resolution WRF simulations with 1 km grid size and stochastic methods. Therefore, the WRF setup for both simulations are similar. Further information about the performance of the long-term WRF simulations can be found at Wagner and Kunstmann [2016].

2 Radar basics and data for the evaluation of weather radar measurements

2.1 Radar basics

2.1.1 Principles of radar measurements

Weather radar systems are radar units (RAdio Detecting And Ranging) which detect weather data. The term weather radar or radar is used in the following for the most common type of weather radar namely the radar used to locate precipitation.

The basic principle of operation of weather radar systems is to emit signals and to detect the reflected echo. Therefore, the transmitter of weather radar systems send electromagnetic energy pulses using a cavity (magnetron or klystron) tube connected by a waveguide to a parabolic antenna. The electromagnetic wave is bundled by the antenna reflector to propagate into the desired direction and detect particles in the atmosphere. All targets produce a diffuse reflection. A small portion of the energy is reflected (“scattered”) and returns to the radar set. This returned energy is called the radar echo. The highly sensitive receiver has to amplify this echo to enable the signal processor to digitize the signal and process it further. A duplexer switches between transmitting and receiving so that only one antenna has to be used. Usually, a radome covers these devices as protective measures.

The exact location of the detected particle is calculated by pointing direction and propagation time of the electromagnetic wave. The signal strength is a measure of rain intensity if hydrometeors are detected.

Most weather radar systems use the Doppler effect to achieve velocity data about targets at a certain distance. The motion of the target in radial direction modifies the frequency of the returned signal. This frequency shift is an accurate measurement of the radial component of the target’s velocity. In this way, wind products based on weather radar data can be deduced. The analysis of the Doppler spectrum can furthermore be used to separate moving echoes from non-moving echoes such as ground clutter [Seltmann, 1997].

The new generation of radar systems is polarized. They simultaneously transmit and receive horizontal and vertical polarized electromagnetic waves. In this way, the calculation of rain rate from reflectivity measurements and the hydrometeor identification can be improved [Bringi and Chandrasekar, 2001, Zrnić and Ryzhkov, 1999].

More detailed information about radar principles can be found in the relevant radar textbooks (e.g. Atlas [1990], Battan [1973], Rinehart [2004], Sauvageot [1992]).

Radar equation

A detailed derivation of the radar equation can be found at Collier [1989]. Here, only a brief description is presented in the following: The averaged received power from a point target at the center of a radar beam as a function of the range and radar cross section of a target can mathematically be described in Eq. 2.1 ([Rinehart, 2004]):

$$P_r = \frac{P_t \cdot G^2 \cdot \lambda^2 \cdot \sigma}{64 \cdot \pi^3 \cdot r^4} \tag{2.1}$$

where P_r in W is the returned and P_t in W the transmitted power, G the antenna gain (ratio), λ in m is the wavelength of the radar, r in m the distance between antenna and target and σ in m^2 the radar cross section. The antenna gain describes the relation between the radiation intensity achieved by the bundling of the antenna and by an isotropic propagation. For a distributed target such as precipitation, that fills the radar beam, Eq. (2.1) can be written as follows:

$$P_r = \frac{P_t \cdot G^2 \cdot \lambda^2 \cdot \theta \cdot \phi \cdot h}{1024 \cdot \pi^2 \cdot \ln 2} \cdot \frac{\eta}{r^2} \tag{2.2}$$

where θ in radians and ϕ in radians are the antenna beamwidths in the horizontal and vertical planes, h in m is the pulse length of the signal, and η in m^{-1} is the radar reflectivity of the target. It is assumed that the scattering volume is uniformly filled and that the antenna pattern has a Gaussian shape. The radar equation provides the radar reflectivity η as the measurable property of precipitation. It depends on the sizes and concentration of the hydrometeors and their state of aggregation.

Three different wavelengths of weather radar systems are commonly used: X-band (0.024 - 0.0375 m), C-band (0.0375 - 0.075 m) and S-band (0.075 - 0.15 m), which mainly differ in their attenuation characteristics.The Rayleigh scattering approximation can be applied because each single drop is significantly smaller than the wavelength of the radar system:

$$\eta = \frac{|K|^2 \cdot \pi^5}{\lambda^4} \cdot Z \tag{2.3}$$

where $|K|^2$ is a dielectric factor, equal to 0.93 for water and 0.2 for ice and Z in $\mathrm{mm}^6\ \mathrm{m}^{-3}$ is the radar reflectivity factor of precipitation.

The radar specific variables are summarized in one specific radar constant C ($\mathrm{W\ m}^{-1}$) which is defined as:

$$C = \frac{P_t \cdot G^2 \cdot \theta \cdot \phi \cdot h \cdot \pi^3}{1024 \cdot \ln 2 \cdot \lambda^2} \tag{2.4}$$

Combining Eq. 2.2, Eq. 2.3 and Eq. 2.4 results in the simplified radar specific Eq. 2.5, where the relation between the returned power P_r and the radar reflectivity factor Z is shown:

$$P_r = C \cdot Z \cdot \frac{|K|^2}{r^2} \tag{2.5}$$

Derivation of precipitation parameters

Both the radar reflectivity factor Z and the rain rate R in mm depend on the amount of rain drops $N(D)$ and the drop diameter D in mm. They can be expressed as different moments of the drop size distribution (DSD) [Rinehart, 2004]. The radar reflectivity factor Z is the 6th moment of the DSD:

$$Z = \int_0^\infty N(D) \cdot D^6 dD \tag{2.6}$$

The rain intensity R is approximately the 3rd moment of the DSD:

$$R = \frac{\pi}{6} \int_0^\infty D^3 \cdot v(D) \cdot N(D) dD \tag{2.7}$$

with $v(D)$ as the fall velocity in m s^{-1} depending on the drop diameter D.

The radar reflectivity factor Z is usually expressed in dB due to its wide range of values.

$$dBZ = 10 \cdot log_{10}(Z/Z_0) \tag{2.8}$$

with $Z_0 = 1$ mm^6 m^{-3}.

The drop diameter as well as the amount of drops must be known to solve these equations. The reliable determining of both variables in real rain is neither mathematically nor by measurements achievable. As a consequence, an empirical approach is usually applied that creates the relationship between radar reflectivity and rain intensity, the so called Z/R relationship.

$$Z = a \cdot R^b \tag{2.9}$$

Both constants a and b also depend on drop size and drop diameter and are therefore highly variable in space and time. A wide range of Z/R relationships exist and the constants for snow, light rain and heavy rain differ significantly. But even for a specific type of rain or rain intensity the eligible constants vary and affect the calculation of rain intensities. As a consequence, radar data is often adjusted to rain gauges [Collier, 1989].

In the following, the term reflectivity is used for the radar reflectivity factor expressed in dB.

2.1.2 Weather radar measurements at the DWD

History

75 years ago, in 1941, the first weather radar system was built in the USA. But it took nearly 40 years before the era of nationwide weather radar networks started in the 1980s [Atlas, 1990,

Whiton et al., 1998]. In Germany, plans for a nationwide network started to be realized as a result of the dramatic hailstorm in Munich in 1984. The whole area of re-united Germany has gradually been covered with finally 16 weather radar systems until the year 2000, operated by DWD. At Hohenpeissenberg, south of Munich, the DWD runs an additionally research radar system to test certain configurations, scans and new hardware and software. All of these radar systems are monostatic C-band weather radar systems which provide spatially distributed reflectivity measurements of precipitation patterns. Further explanations about historical facts and the technical functionality can be found in Seltmann [1997]. The following years, all sites have been refitted with Doppler measurements. This offers the opportunity to eliminate clutter in reflectivity measurements very efficiently and additionally enables the measurement of radial wind speed. Between 2005 and 2010 there were only minor technical changes of the radar network in Germany. An overview of modifications of weather radar measurements or change of products within the RADOLAN framework (Radar Online Adjustment) can be found in DWD [2017] and DWD [2016b]. From 2011 on the entire German radar network has been upgraded and the coverage of Germany by radar systems has been optimized. Therefore, four radar sites have been shifted to another location and a 17th radar site has been installed. Figure 2.1 gives an overview of DWD's radar network with the maximum used ranges (a) before 2010, (b) between 2010 and 2011 and (c) the current network. However, the greatest improvement of the new generation of radar systems is the use of a dual-polarization technique.

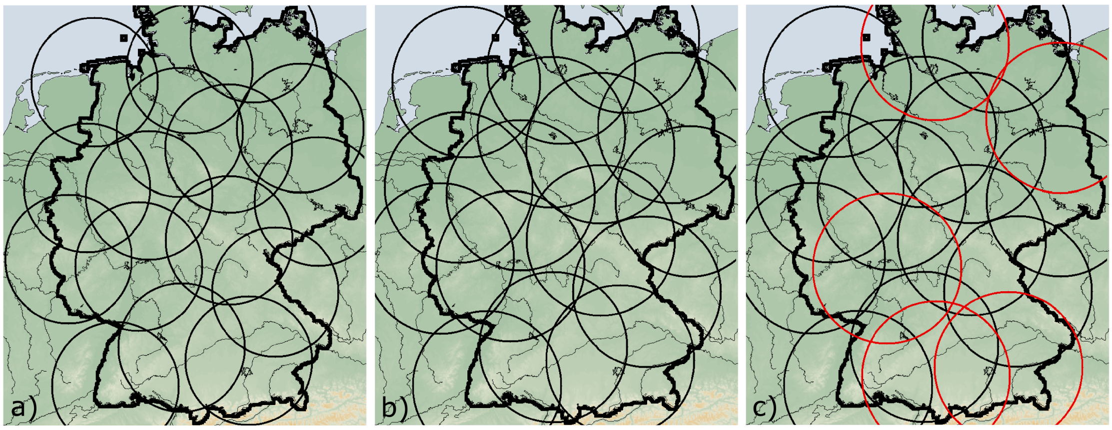

Figure 2.1: German radar network (DWD) a) until 2010 with a maximum used range of 128 km and 16 radar sites, b) same as a) but with a maximum used range of 150 km from 2010 on and c) current network with a maximum used range of 150 km and 17 radar sites. The new and shifted radar sites are marked in red color.

Scan strategy

The DWD runs two different scan algorithms to produce a variety of radar products called the volume scan and the precipitation scan. Until 2012, the volume scan consisted of 23 elevations between 0.5° and 37° to get three-dimensional information of precipitation echoes. 18 scans with a maximum range of 128 km are Doppler scans with Dual-PRF (pulse-repitition-frequency) to increase the unambiguous velocity. The other five intensity scans with low elevations provided only qualitative measurements up to a maximum range of 256 km. The entire volume scan took about 15 minutes to complete all circulations. With the new radar network, the scan strategy for the volume scan has been modified to receive more frequent volume information: every 5 minutes the entire volume is sampled based on 10 elevations ranging between 0.5° and 25° with a maximum range of 120 km or 180 km (depending on the elevation angle). The precipitation scan is a particular feature for Germany. It is a terrain-following Doppler scan every 5 minutes used for quantitative radar products. Until 2010 its maximum evaluated range was 128 km, then it has been extended to 150 km range. The radar aperture angle is 1° with a range resolution of 1 km. Consequently, each radar sweep provides 46080 measurements based on a maximum range of 128 km (360 x 128) with gradually increasing range-bin sizes (voxels) with distance from the radar site [DWD, 2017, Seltmann, 1997].

Radar products

There is a huge variety of radar products in DWD [DWD, 2016a]. In general, radar products based on the volume scan are largely intended for applications where the knowledge of the vertical extension or structure of precipitation is of high importance, such as vertically integrated liquid water content VIL, severe convection or hail detections. Radar products based on the precipitation scan are mainly used for quantitative products of precipitation, as the name suggests. A milestone for quantitative radar products in Germany has been the initiation of the project RADOLAN [Bartels et al., 2004] conducted between 1997 and 2004. Diverse methods of radar adjustment have been developed and applied which has led to a number of quantitative products with a high emphasis attended to the quality of radar data. Additionally, the compositing of radar data from different radar systems to produce nationwide products as well as the continuous storage of radar data have been implemented and realized within this framework in 2005. For the German radar composites, data from the precipitation scan of individual radar sites is used. The maximum criterion is applied to merge those measurements at different altitudes.

We divide radar products into primary and secondary products. The primary radar products are base or quality controlled radar data based on the pre-processed data provided by the signal processor. Secondary radar products are based on primary radar products and then further processed on the basis of complex algorithms or of additional data such as rain gauge data to provide additional and improved information. The nowcasting product CONRAD (CONvection in RADar products) [Lang, 2001] is a typical example of the latter product: convective cells including hazards such as intense precipitation, wind shear and the probability of hail are identified, tracked and interpolated into the immediate future (1 or 2 hours) only on the basis of radar data. It exploits the advantages of radar data such as the detection of small and heterogeneous precipitation patterns and wind and prepares the results for the user in a convenient way. The last couple of years, a completely new range of products have been developed enabled by the new dual-polarization technique that uses the differences between the horizontal and vertical direction of the backscattered electromagnetic waves.

New parameters such as the specific differential phase K_{DP}[1] (deg km^{-1}) and the differential reflectivity Z_{DR}[2] (dB) are calculated and provide information on liquid water content and drop sizes. The combination of existing and these new parameters offer a variety of possibilities to improve radar data and the calculation of rain rate [Bringi and Chandrasekar, 2001]. For DWD, a new post-processing quality control workflow [Werner, 2014] and a new hydrometeor classification [Steinert, 2014] have been developed, for instance.

2.2 Radar products and validation data

The presented study is performed for single radar data by the example of the Munich weather radar as well as for the German radar composite. The total evaluation period is 2000 to 2009. This period is chosen because the requirements are on the one hand a long period of time and on the other hand a homogeneous data basis as much as possible unchanged by modifications such as scan strategy or radar site re-locations.

Different primary radar products that are based on the precipitation scan are analyzed. All of them undergo the usual corrections within the signal processor, e.g. Doppler filtering, clutter correction, speckle remover, and thresholding for noise (LOG) and signal quality (SQI) [Seltmann, 1997]. No further corrections that might modify the spatial precipitation patterns such as adjustment are applied to these reflectivity products. Additionally, point measurements of rain gauges and interpolated rain gauge data are used for the purpose of evaluation and adjustment.

2.2.1 Single radar

In a first step, the Munich weather radar and the pertinent single radar products PX and DX are examined. The Munich radar is situated 15 km to the north of the city of Munich. The investigation area is a circle of 100 km around the site (see Fig. 2.2). Annual rain amounts range from 700 mm in the northern part of the radar site to 1500 mm in the alpine upland. Even higher rain amounts are measured in the Alps. With respect to data quality, the Munich radar is located at a challenging site: a number of obstacles including the distant Alps, the close city of Munich and a nearby hill to the north-east produce shadowing effects. Therefore, the beam elevation angle of the terrain-following scan varies between 0.8° and 2.1°, which results in higher altitudes a.s.l. of the radar beam for the high beam elevation angles. The altitudes for all radar pixels are shown in Fig. 2.2 to indicate the consequences of different beam elevation angles per azimuth. In the south-eastern part the altitude only increases to a maximum of 2.5 km at a distance of 100 km from the radar site, whereas in the north-eastern part the radar beam reaches maximum altitudes of 4.2 km. This variation will become apparent in the statistical analysis and has to be taken into account for the statistical corrections. In 2006, an optimization of this variation was implemented, which also has to be paid attention to. The Munich weather radar was dopplerized in 2004. Since then a much better clutter suppression (see Fig. 4.6a) is in place because of Doppler-filters. Additionally, the range of the Alps leads to clutter and shading effects in the southern part of the radar image.

---

[1] K_{DP} is the difference in phase shift between horizontal and vertical polarized returned energy.

[2] Z_{DR} is the ratio of reflected horizontal and vertical power returns.

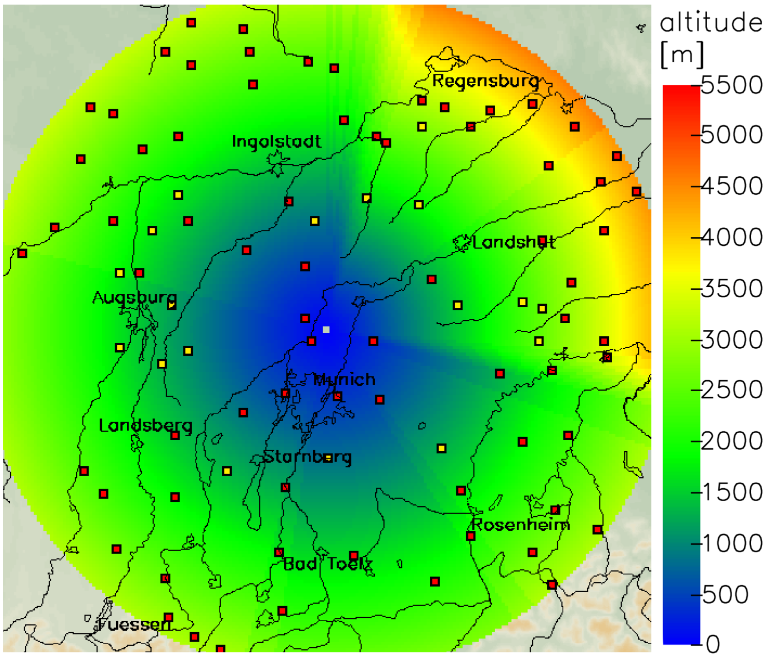

Figure 2.2: Mean altitudes of the near-surface precipitation scan of the Munich weather radar (2000–2006), overplotted by locations of the rain gauges used for comparing rain amounts. The yellow color indicates rain gauges used for adjustment and the red color shows stations used for validation.

Single radar product PX

The so-called PX-product with six reflectivity levels (see Tab. 2.1) and a spatial resolution of 1×1 km^2 shows the longest time series starting in 2000. It includes a threshold for noise. The lower limit of reflectivity level 1 is 7 dBZ in the summer and 1 dBZ in winter. This radar product for the period 2000 to 2006 is the basis for the analysis of disturbances within the radar image as well as for the development of the correction algorithms. The availability of this product for the Munich weather radar is between 75 % and 95 % per year. The evaluation range has been set to 100 km because until 2005 the maximum range of the PX-product has only been 100 km, which results in a total of 2.4×10^{10} pixel-measurements as a profound basis for a statistical analysis.

Reflectivity [dBZ]	< 19	19 - 27.9	28 - 36.9	37 - 45.9	46 - 54.9	> 55
Rain Rate [mm h^{-1}]	< 0.7	0.7 - 3.2	3.2 - 7.5	7.5 - 26.8	26.8 - 80	> 80
Level	1	2	3	4	5	6

Table 2.1: Reflectivity levels of the PX-product. The corresponding rain rates have been calculated by the three-part Z/R relationship in Tab. 3.1.

Single radar product DX

The second radar product used is the DX-product with 256 reflectivity levels and a resolution of 0.5 dB ranging from -31.5 dBZ to 95.5 dBZ. For this analysis the reflectivity range has been limited to the range from 1 dBZ (noise) to 60 dBZ (hail). Its original resolution is 1° in azimuth-direction and 1 km in range. This product is also only based on the near-surface precipitation scan every 5 min, similar to the PX-product. It is ideal for quantitative purposes and is therefore used for adjustment of radar data to rain gauges and for means of validation. It has continuously been stored since 2004 for the Munich weather radar with an availability of about 85 % until 2006 and between 95 % and 98 % afterwards.

Rain gauge data

90 rain gauges in the vicinity of the Munich weather radar with reliable time series of daily precipitation measurements are available for the comparison with the radar DX-product.

Tipping-buckets as well as rain collectors operated by DWD are used. Both are quality controlled according to official quality assurance standards. The temporal resolution is daily. The original database included 140 rain gauges. Only 90 rain gauges cover the whole period from 2004 to 2009, showing only a few missing values and fulfilled the following additional quality controls: The rain gauges are manually quality controlled for any bias including limit exceedance of daily, monthly and annual mean values as well as compared to time series of adjacent rain gauges.

Split sampling of stations for adjustment and for validation is used to guarantee independence. The rain gauges are separated into three groups: one group for the adjustment of radar data and two groups for validation purposes. 36 gauges exist within a distance of 30 to 70 km from the radar site representing the area of the most reliable radar measurements. 18 of the ones evenly distributed are used for the adjustment of radar data and represent the first group. The second group consists of 57 of the remaining 72 rain gauges within the whole radar coverage excluding those rain gauges which are situated in regions of radar-pixel interpolation such as parts of the alpine region or the city of Munich. The third group includes all 72 rain gauges used for validation. In Fig. 2.2 the locations of the rain gauges used for adjustment are indicated by yellow dots and those for validation by red dots.

2.2.2 Radar composite

The investigation area of the radar composite is the whole of Germany with the Alps in the very southern part, the Central German Uplands in the southern and mid parts and the Northern Lowland with the coasts of the Baltic Sea and the North Sea. The German radar composite is a mosaic of 16 weather radar systems of the German Meteorological Service.

Single radar product PX

Regarding the investigation of disturbances in the radar composite, the PX-product is used to analyze the influence of beam-broadening and increasing altitude of radar range-bins with distance from the radar site and permanent clutter effects for each of the remaining 15 weather radar systems. By analogy with the investigation of the Munich weather radar, the evaluation period is 2000 to 2009 and the maximum range is set to 100 km. The availability of the PX-product was about 80-85 % before 2006 and over 90 % afterwards for all radar sites. The advantage of the additional use of single radar images is that the complete radar image of one radar site can be investigated so that even gradual deviations may become apparent. Furthermore, it may be clarified if artificial patterns in composites can be explained by clutter effects in single radar images.

Radar composite RX

The RX-product with 256 classes (-31.5 dBZ to 95.5 dBZ) has a resolution of 0.5 dB. It is based on the terrain-following precipitation scan of up to 16 radar systems and displays a high temporal and spatial resolution (5 minutes; 1x1 km^2) projected on a 900x900 Cartesian grid. This product has been continuously available since 2004/2005. The evaluation period for this product lasts from 2005 to 2009.

In Fig. 2.3(a) the whole coverage of the RX radar composite is illustrated. The maximum ranges of each of the 16 contributing radar systems are marked by black circles, the location of the sites is represented by its abbreviation in white letters. The colors indicate the mean altitude of each composite pixel above the radar site. For the comparison of two radar systems the altitude of each radar site has to be taken into account, but it is neglected in this figure.

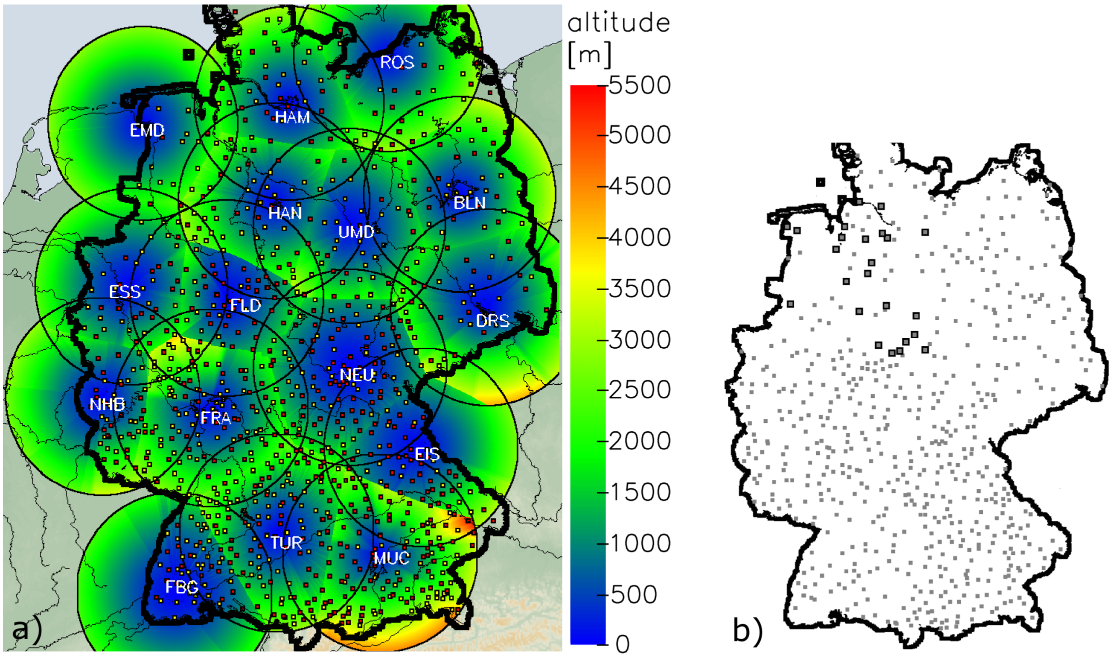

Figure 2.3: (a) Mean altitudes of the near-surface precipitation scan of the 16 contributing radar sites for the German radar composite, overplotted by locations of the rain gauges used for comparing rain amounts for the evaluation (red) and for the adjustment (yellow) on an annual basis. The location of the radar sites are represented by its abbreviation in white letters. They are explained in Tab. 2.2; (b) locations of the rain gauges used for the evaluation of hourly radar data (gray dots) and for 5-minute radar data (larger black dots).

The red points mark the precipitation gauges that are used for the evaluation. The yellow points represent additional rain gauge locations used for the adjustment step of the correction algorithm.

Table 2.2 gives an overview of important local specifics of the 16 contributing German radar sites to the German radar composite RX. The abbreviation and name of the radar site is used to identify the radar site. The altitude of the site and the minimum and maximum radar elevation at a distance of 128 km from the radar site allows the interpretation of differences within the composite. A huge difference between the minimum and maximum radar elevation is an indication of limited comparability within the radar image. Whereas a high absolute value of the radar elevation probably leads to partial beam filling and overshooting.

abbreviation	name	altitude [m]	MIN [km] (128 km)	MAX [km] (128 km)
HAM	Hamburg	46	2.6	2.8
ROS	Rostock	36	2.1	2.9
EMD	Emden	58	2.7	2.7
HAN	Hannover	81	2.1	3.2
UMD	Ummendorf	185	2.0	3.5
BLN	Berlin	80	2.8	2.8
ESS	Essen	180	2.6	2.8
FLD	Flechtdorf	623	2.4	2.9
DRS	Dresden	262	1.9	3.7
NEU	Neuhaus	873	1.7	2.3
NHB	Neuheilenbach	585	2.0	3.0
FRA	Frankfurt	146	2.5	4.7
EIS	Eisberg	799	1.8	2.4
TUR	Türkheim	765	2.1	3.1
MUC	München	511	3.0	5.6
FBG	Feldberg	1517	1.2	1.9

Table 2.2: Site-specific characteristics of the 16 contributing radar sites of the German radar composite. The abbreviation, the name, the altitude of the site and of the minimum and maximum radar elevation at a distance of 128 km from the radar site are shown.

Rain gauge data

For the validation and adjustment of long-term radar data throughout Germany, monthly data of 1260 rain gauges (tipping-buckets and rain collectors) from DWD have been available. These data are quality controlled according to DWD quality standards. Not all of them cover the whole period from 2005 to 2009 or accomplish additional quality controls (see Chapter 2.2.1). A total between 771 and 1182 rain gauges meets the above criteria, differing from year to year. Those 516 rain gauges that cover the whole investigation period are used for validation. The residual rain gauges are used for the adjustment of radar data (see Fig. 2.3a).

The validation of short-term radar data correction is mainly performed on hourly and 5-minute values. Therefore, 926 rain gauges in Germany with an hourly resolution are available for the time span 2005 to 2007, but only 642 of them are selected due to quality reasons or availability. Rain gauges with a 5-minute resolution have been available only in Lower Saxony for the same time span. 33 of the 68 rain gauges cover the whole period and meet the additional quality standards. In Fig. 2.3(b), the locations of the chosen rain gauges with an hourly resolution (gray dots) and those with a 5-minute resolution (larger black dots) are shown.

REGNIE

REGNIE is a dataset of daily interpolated rain gauge values on an approx. 1 km^2 grid for the area of Germany. The underlying principle is to interpolate anomalies on a long-term average. Therefore, the long-term monthly background fields are produced based on rain gauge data by multiple linear regression taking into account the geographical longitude and latitude as well as direction and amount of exposition. The ratios of daily values and monthly values (anomalies) at the location of the rain gauges are then interpolated distance weighted on the whole grid of Germany [REGNIE, 2017]. It is available for the years 1931 to the present day. About 2000 quality controlled rain gauges contribute to the REGNIE dataset since 1995. In this evaluation, REGNIE data is used for the daily adjustment of radar composite RX data in Chapter 4.

3 Statistical analysis of disturbances in weather radar measurements

3.1 Methods for the identification and analysis of disturbances in radar data

3.1.1 Basic concept

The presented approach is a pattern analysis of weather radar data on a long temporal scale. Radar images, both rain amounts and frequencies of occurrence for each reflectivity level of the PX-product, are accumulated for several years. A three-part Z/R relationship (see Tab. 3.1) is used to calculate rain amounts from radar reflectivities [Bartels et al., 2004]. For light and moderate rain the database is large enough to rely on these results, whereas for heavy precipitation the database is too scarce to confide in only these results. Therefore, the findings for light and moderate rain are used to support the results of heavy precipitation's analysis. Especially for extreme rainfall, radar data have a high potential to accurately represent the high heterogeneity of precipitation patterns [Overeem et al., 2010, Pedersen et al., 2008, Rudolph et al., 2011, Wagner et al., 2006].

Reflectivity [dBZ]	< 36.5	36.5 - 44	> 44
a	125	200	77
b	1.4	1.6	1.9

Table 3.1: Three-part Z/R relationship used to calculate rain rate from DX radar products.

Disturbances and other systematic effects in these accumulated images are identified, grouped and analyzed with respect to different scenarios such as reflectivity level and time. Thereafter, the possible causes for these effects are elaborated and explained to comprehend these patterns and to deduce systematic dependences that can be used for a statistical correction algorithm.

The statistical correction aims at improving the quality on average. The advantage of such a statistical correction is that no single disturbance variables, but only their overall impact has to be taken into account. The disadvantage is closely related: assumptions are necessary, whether a certain pattern is a measuring or processing bias or naturally induced as its source is often not distinctive or a mixture of different issues. The main assumption of this approach is that for light and moderate rain the average frequency of occurrence of the associated radar reflectivities with the same distance from the radar site should be almost equal at all distances. Consequently, for a certain reflectivity level, the median of the frequencies of occurrence of one ring of range-bins with the same height should be almost equal to the median for all other heights. Systematic variations such as a mean decrease of frequencies with height are regarded as bias. Variations within such a ring are regarded as naturally induced. Certain geographical characteristics resulting in different meteorological situations with different rain amounts are not specifically taken into account but remain in the database. Conversely, abrupt variations of frequencies of occurrence in space indicate errors in data basis.

3.1.2 Step-by-step approach

Patterns and disturbances in radar data are investigated in a four-step statistical analysis. The first three steps mainly refer to the evaluation of single radar data. The forth step describes the extension of the analysis to composite radar data.

Step 1: Separation of corrupted and uncorrupted pixels

The first step of the statistical analysis is to separate uncorrupted from obviously corrupted pixels in accumulated single radar images (see Fig. 3.1). Two groups of corrupted pixels become apparent: The first one comprises clutter pixels mainly caused by non-meteorological echoes while the second one is made up of spokes due to beam blockage by obstacles near the radar site. Basically, each type of clutter or disturbance can be identified as long as its appearance is conspicuous on a longer temporal scale. Figure 3.1(a) gives an overview of corrupted and uncorrupted radar pixels for reflectivity level 1 of the Munich weather radar. The red color represents pixels which are affected by clutter or clutter correction and the yellow color indicates spokes. The 'city clutter' within a distance of 40 km to the south of the radar site and the 'mountain clutter' in the south at distances of over 70 km from the radar site can easily be separated because of the underlying topography. For these two clutter types, a certain area including corrupted pixels of the same source is visually defined. It is an area where uncorrupted pixels form the majority of pixels. The blue color in Fig. 3.1(a) indicates the area of mountain clutter. Pixels in this area have comparable beam heights and distances from the radar site. Corrupted and uncorrupted pixels within these areas have to be separated. This separation has to be performed only once and includes manual work: For each area of correction, thresholds of frequencies of occurrence are used to separate those pixels which are obviously corrupted from the rest. Additionally, a safety margin of 2 km is established around the corrupted pixels to mark those pixels which are likely to be influenced by clutter. For the residual pixels a histogram of frequencies of occurrence is calculated. The uncorrupted pixels (comparison group) show comparable frequencies of occurrence and, therefore, form a distinctive peak in the histogram. Pixels which differ from this distribution can be separated manually if the pre-selection of uncorrupted pixels is difficult (see Fig. 3.1b). As a last step, the final separation is realized by the analysis of an empirical distribution of frequencies of occurrence, where its 95 % percentile marks the range of uncorrupted pixels (see Fig. 3.1c).

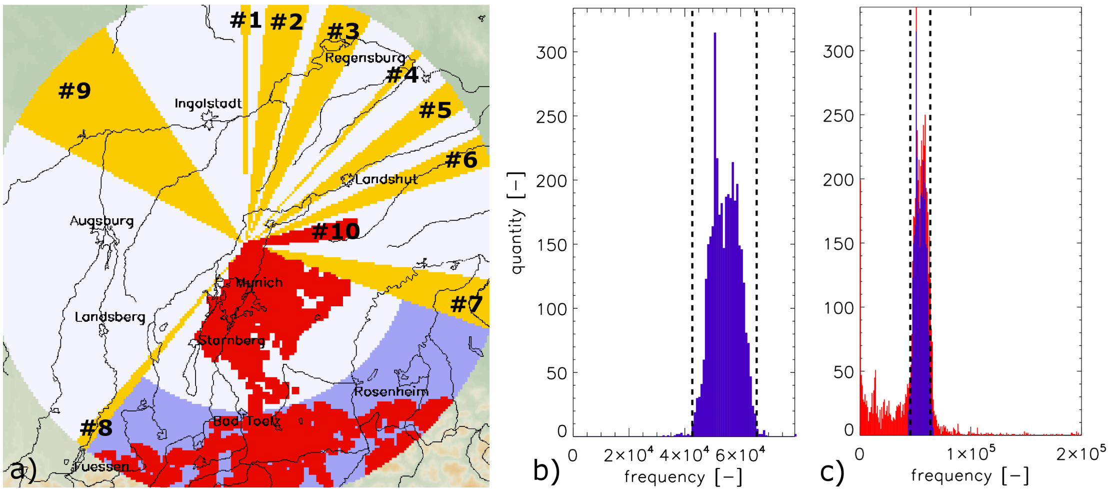

Figure 3.1: (a) Overview of clutter and disturbances for reflectivity level 1 of the Munich weather radar including clutter pixels (red), spokes (yellow) and the comparison area of uncorrupted radar pixels for the mountain clutter (blue); (b) histogram of frequencies of occurrence of radar reflectivities of probably uncorrupted pixels with a manual separation (dashed lines) for mountain clutter (blue area in a); (c) histogram of frequencies of occurrence of radar reflectivities of all pixels in the blue area in a (red); the dashed lines mark the final separation of corrupted and uncorrupted pixels based on the 95 % interval of uncorrupted pixels in (b) (blue).

As reflectivity level 1 shows the highest amount of corrupted pixels, the classification for all levels is based on level 1. This can be checked visually.

For the separation of spokes from uncorrupted pixels all pixels of one azimuth are treated jointly as their source of error is the same. If the median of the frequency of occurrence of radar reflectivities of adjacent azimuth angles differs by more than 10 % this usually suggests the presence of a spoke, but can also easily be checked visually. The sources of clutter may differ between radar sites. Wind mills or shipping routes are further sources of error in radar data. The way of separating corrupted and uncorrupted pixels remains the same in principle.

Step 2: Investigation of uncorrupted pixels (variations with height)

All uncorrupted radar pixels are inspected whether an influence of the increasing beam height and -width with distance from the radar site exists (variation with altitude). Measurements at low ranges from the radar site should be comparable to those at far ranges on average. Any dependences can be linearly addressed and are regarded as bias. Therefore, the median of frequencies of occurrence of radar reflectivities in each altitude class (one class per 100 m) is calculated and plotted versus height. Plotting all frequencies of occurrence of radar reflectivities against height would lead to a dense point cloud as naturally induced precipitation patterns usually show variations, whereas the median provide the average behavior with height. Many patterns and disturbances in radar images depend on the reflectivity level. Therefore, three reflectivity levels are chosen, which represent the variety of precipitation intensities:

- level 1 for light rain (and snow)
- level 3 for moderate rain
- level 5 for heavy rain

Step 3: Investigation of corrupted pixels (clutter and spokes)

For each group of corrupted pixels ('city clutter', 'mountain clutter' and 'spokes') the median of the frequency of occurrence for each radar reflectivity level is compared to the median of the corresponding frequency of occurrence of adjacent undisturbed radar pixels (comparison group). In this way, the patterns of identified corrupted pixels are differentiated with regard to their situation, the circumstances they depend on and the reason for their occurrence.

Step 4: Investigation of compositing effects

The last step is the analysis of composite data. The corrupted and uncorrupted areas are adopted from single radar analysis. Where appropriate, the only measurements used are the ones where all single radars contribute to the composite to remove effects of radar availability. The investigation of effects of different coordinate conversion methods as well as the comparison of rain amounts between adjacent radar sites and between single radar areas and overlapping areas is mainly qualitatively performed. To derive the preferred allocation of pixels to one radar system within overlapping areas, the dependence of frequencies of occurrence of radar reflectivities or rain amount as a function of height is used, based on measurements. Similar to single radar measurements, the median of frequencies of occurrence of radar reflectivities of uncorrupted pixels in each distance class (one class per 5 km) within one overlapping area is calculated and plotted versus height for each contributing radar system.

The results of the statistical analysis and, first and foremost, the deduced systematic behavior (e.g. variation with altitude) are the basis for the correction scheme presented in Chapter 4.

3.2 Results of the analysis of disturbances in radar data

3.2.1 Single radar

The results for the Munich weather radar are presented due to the fact that this radar site is very challenging and illustrates many systematic deficiencies of radar data very clearly. Additionally, the high level of data availability and the long time series are further aspects for choosing this radar site.

Figure 3.2 gives an overview of the frequency of occurrence of radar reflectivities for level 1 (light rain), level 3 (moderate rain) and level 5 (heavy rain). The three main types of clutter and disturbances (city clutter, mountain clutter and spokes) become obvious within the vicinity of the Munich weather radar. A comparison of these three radar images reveals conspicuous differences of the values of the clutter disturbances. The uncorrupted radar pixels show a significant decrease of the frequency of radar reflectivities close and far from the radar site for all radar reflectivities (variations with height).

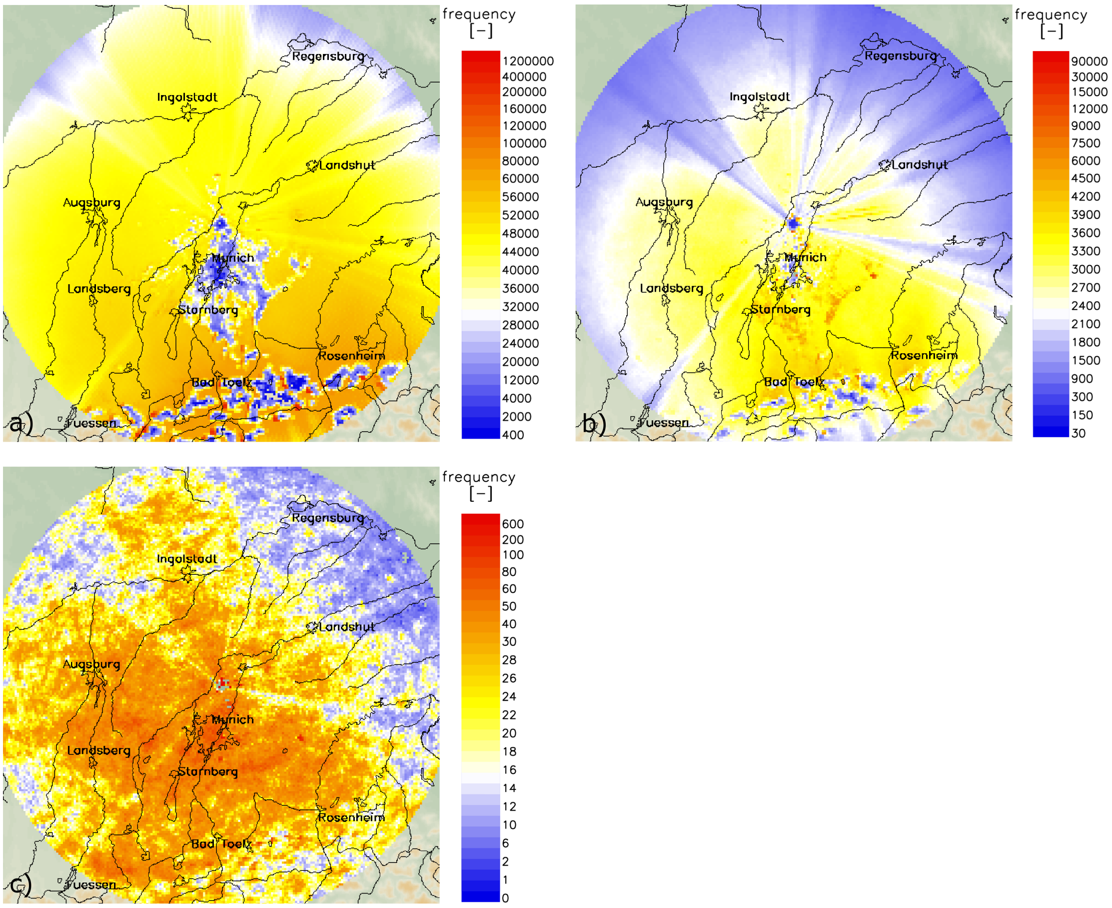

Figure 3.2: Uncorrected frequencies of occurrence of radar reflectivity level 1 (a), level 3 (b) and level 5 (c) of the Munich weather radar from 2000–2006 (PX data). The scale is 200×200 km^2.

Variations with height

Theory

The classified uncorrupted radar pixels show significant differences of frequencies of occurrence of radar reflectivities close to and far from the radar site. These discrepancies are attributable to the interaction of measuring effects and natural variations of reflectivity with height. The main measuring effects are the increasing range-bin height and size with altitude above the radar site. Increasing range-bin sizes lead to intensified averaging of small-scale rain structures for example from convective cells. The measurements at different altitudes of the Vertical Profile of Reflectivity (VPR) can result in even higher variations, depending on whether the radar detects liquid water, snow or partly melted ice within the bright band. Additionally, rain with low vertical extent may lead to only partial beam filling or to overshooting at greater distances from the radar. Temperature and precipitation type are the influencing factors that dominate the vertical structure of rain. In summer, convective rainfall occurs more frequently, which has a larger vertical extent and strong reflectivity cores aloft causing positive

VPR gradients. For the United Kingdom, Hand [1996] showed idealized vertical reflectivity profiles for the cell stages of cumulonimbus clouds. For most stages reflectivities between the cloud base and the midcloud level are considerably larger than those below the cloud base [Overeem et al., 2009]. So there is a variety of sources why variations at different ranges from the radar can occur in measurement. Simultaneously, these VPRs are highly variable even within one radar image. However, in this approach the net effect of these different profiles per year is analyzed to de-bias the statistics on average. Accordingly, a climatological mean vertical profile for each reflectivity level is deduced in Fig. 3.3 from all available radar measurements. These profiles can be regarded as the summation of the large amount of highly variable VPRs per year within the coverage of the Munich weather radar.

Dependence on reflectivity level

The following figures show the behavior of the median of the frequency of occurrence of radar reflectivities with altitude, separated into classes of height (100 m). The lowest and the highest altitudes should be neglected for the interpretation of the mean behavior as they are only supported by a very small amount of pixels. Figure 3.3 gives an overview of the behavior of the frequency of occurrence of radar reflectivity levels 1, 3 and 5 with height. For level 1 to 3 an increase of the frequency of occurrence of pixels with height becomes obvious below 1 km height; the increase seems to be very variable. Above 1 km height, a steady decrease of the frequency of occurrence of pixels for light and moderate rain can be observed. For higher reflectivities a decrease for all heights is shown. For reflectivity level 1 the frequency of occurrence of pixels decreases by 12.9 % per 1 km difference in altitude according to Fig. 3.3(a) above 1 km height. For higher rain intensities the decrease is even larger, but seems to be constant at 20.3% per 1 km difference in altitude (Fig. 3.3b, c). Even though the beam elevation angle over azimuth varies to a great extent, the fluctuations of the decrease with height are small and therefore negligible.

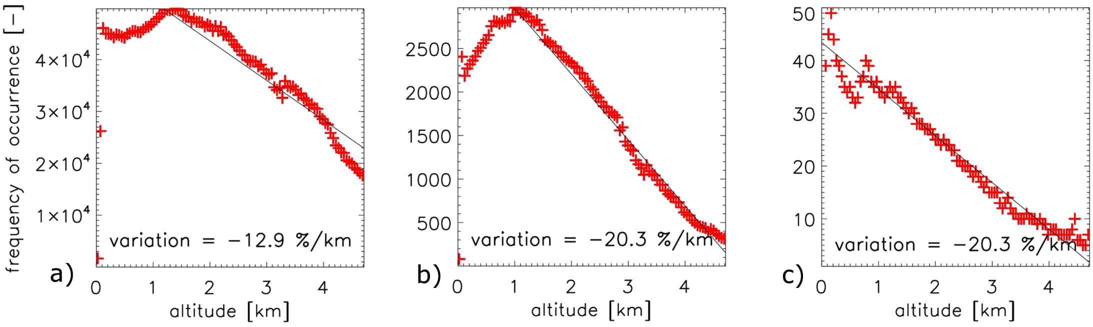

Figure 3.3: Characteristics of the median of the frequency of occurrence of uncorrupted pixels with height for equidistant classes of altitude for the reflectivity levels 1 (a), 3 (b) and 5 (c) of the Munich weather radar from 2000–2006 (PX data).

The decrease of reflectivity frequencies with reflectivity level varies for different radar sites. For example, the decrease with height for the Hamburg weather radar is also linear but different for each level, ranging from 10 % km^{-1} height for level 1 to 16 % km^{-1} height for level 3 and 25 % km^{-1} height for level 5 (not shown here). Several influencing factors may be responsible for this varying behavior: High reflectivities can be associated with small scale convective rain types showing some high vertical homogeneity. But often this rain type cannot be adequately resolved and leads to higher decreases at far ranges from the radar site than at low ranges. So the higher reflectivity values depend strongly on the range-bin size. Snow and light rain only show low vertical extensions of rain echoes but usually a high horizontal homogeneity. So the dependence of the reflectivity value on height is important. Partial beam filling and overshooting results for higher altitudes. Consequently, high beam elevations (e.g. at the Munich weather radar) may affect the decrease of small reflectivities with a higher impact than the decrease of higher reflectivities. Local differences in meteorological or radar site specific conditions such as the altitude of the radar site compared to the environment may play a certain role as well. But the analysis of the data derived from the 14 residual radar systems does not allow an unambiguous interpretation.

The presented results are based on a mixture of different types of rain, ranging from strong convective cells to snow. So a temporal separation into months may display different results.

Dependence on time

Figure 3.4 shows the same analysis of level 3, but separately for the months January, April, July and October. The extreme decrease of the frequencies of occurrence in January can be explained by a large proportion of low reflectivity levels in snow and a tendency of typically low vertical extensions of rain.

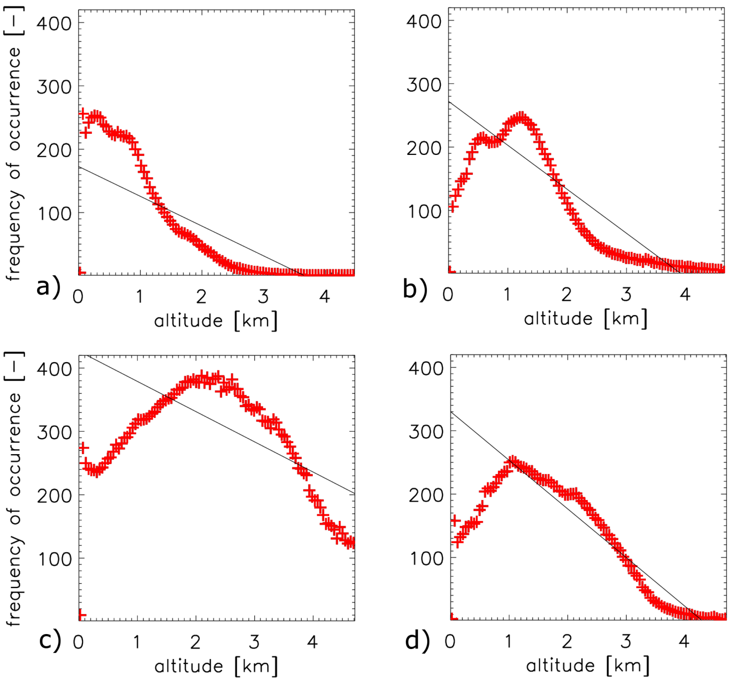

Figure 3.4: As Fig. 3.3, but only for reflectivity level 3 separated into months showing January (a), April (b), July (c) and October (d).

Figure 3.4(b) reveals a remarkable characteristic in April. The typical decrease of frequencies of occurrence starts after a short increase at 1.5 km altitude. In July the 'peak' is shifted to higher altitudes of about two or three kilometers. The figure for October is comparable with April but with a more constant decrease with height.

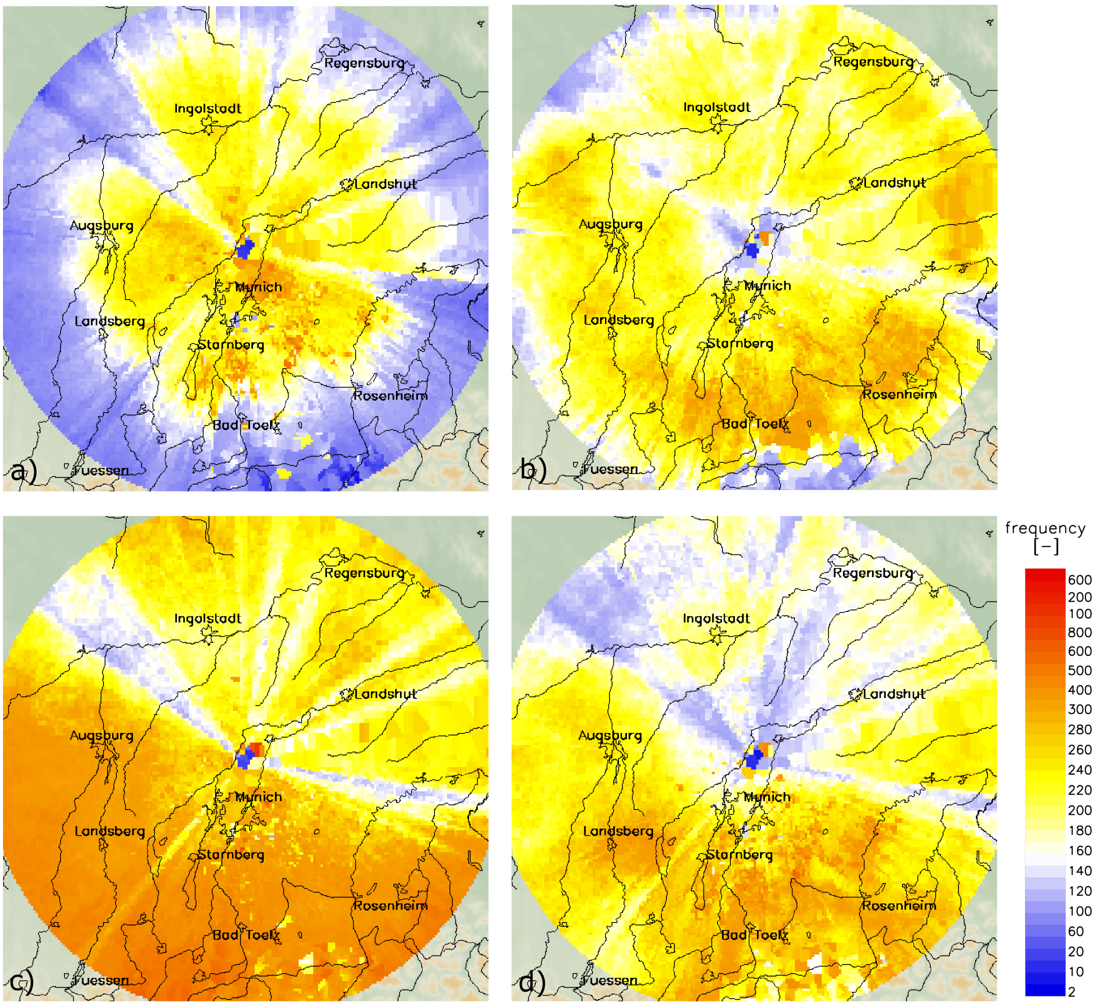

Figure 3.5: Spatial distribution of the frequency of occurrence of radar pixels of level 3 of the Munich weather radar with height from 2000-2006 for January (a), April (b), July (c) and October (d) where all rays are reduced to the same maximum height.

To evaluate whether this feature is the result of a meteorological situation or a measuring effect, Fig. 3.5 shows the distribution of radar reflectivities with height above the radar site instead of range from the radar site as usual. Due to the terrain-following elevation angle as a function of azimuth, range-bins with the same altitude are measured at different distances from the radar. To refer all measurements to the same height above radar, range-bins with altitudes higher than 2.5 km (lowest maximum height for all rays) are neglected. For all months in Fig. 3.5 the respective main decrease with altitude is almost equal for each azimuth. Additionally, a very uniform ring of higher frequencies becomes obvious especially in April even though the corresponding pixels are derived at distances ranging from 30 to 50 km from the radar site. This ring is shifted to lower altitudes in winter and higher altitudes in summer (see also Fig. 3.4). Both facts indicate that temperature is the key indicator. One reason for this maximum of frequencies at a certain height is a naturally induced monthly variable height of maximum precipitation. The vertical extension of clouds is very low in winter which results in frequent overshooting at higher altitudes, while in summer the maximum rain occurs at much higher altitudes reflecting the higher vertical extension of clouds in the warmer season. Thus, besides the measuring effects, the transition from snow to rain seems to play an important role. Furthermore, a considerable maximum in the VPR is usually an indication for the bright band that is typical for stratiform rain events. The rain clouds must have a certain vertical extension with corresponding radar reflectivities of level 2 or 3 to develop a measureable bright band reaching reflectivities which may be more than 10 dB higher than reflectivities of pure rain. This explains why the bright band effect becomes obvious most clearly at reflectivity level 2 and especially level 3. In any case, a certain amount of measurements is shifted to the next reflectivity level. The frequency of occurrence in a higher class is much lower than in a lower class. So if only a small percentage of reflectivities is increased, an effect in the frequency of the higher class will result. Thus, it becomes possible to observe a bright band effect even with only six reflectivity classes. There are further indications why the bright band might be at least partly responsible for the peak of frequencies of occurrence at a certain altitude, like the coincidence of the mean zero degree level per month with the altitude of the maximum frequency of occurrence.

Clutter

Theory

The influence of clutter in the accumulated radar images can be seen by a sharp gradient between single pixels or pixel groups and the adjacent radar pixels. Two counteracting processes influence the accumulated rain amounts or frequencies of occurrence of radar reflectivities, both attributed to clutter. The first one is insufficiently corrected clutter resulting in higher reflectivity levels and counts. The second one is thresholded data resulting in lower reflectivity levels and counts. Since 2004 Doppler IIR-filtering (Infinite Impulse Response) is used for clutter correction in all radar systems at the DWD. This correction provides good results for obstacles eliminating signals with a velocity of about 0 m/s in the Doppler spectrum. The relative power difference between the filtered and the unfiltered time series is calculated and then subtracted from the output reflectivity. If this clutter correction parameter (CCOR) is large, the retrieval of rain echoes fails and the measurement is discarded. As a consequence, the influence of clutter decreases with higher reflectivities [Seltmann, 1997].

City clutter

First, the city clutter effects are analyzed and displayed in Fig. 3.6. The figure of level 1 (Fig. 3.6a) for city clutter shows a significant influence resulting in smaller frequencies of occurrence. This is probably due to clutter correction in the signal processor where too much of the signal is discarded. For radar reflectivities of higher levels (3 and 5) in Fig. 3.6(b, c) no further interference of clutter or clutter correction can be observed. So, the city clutter is dominant only for light rain, but may still be present at higher reflectivities.

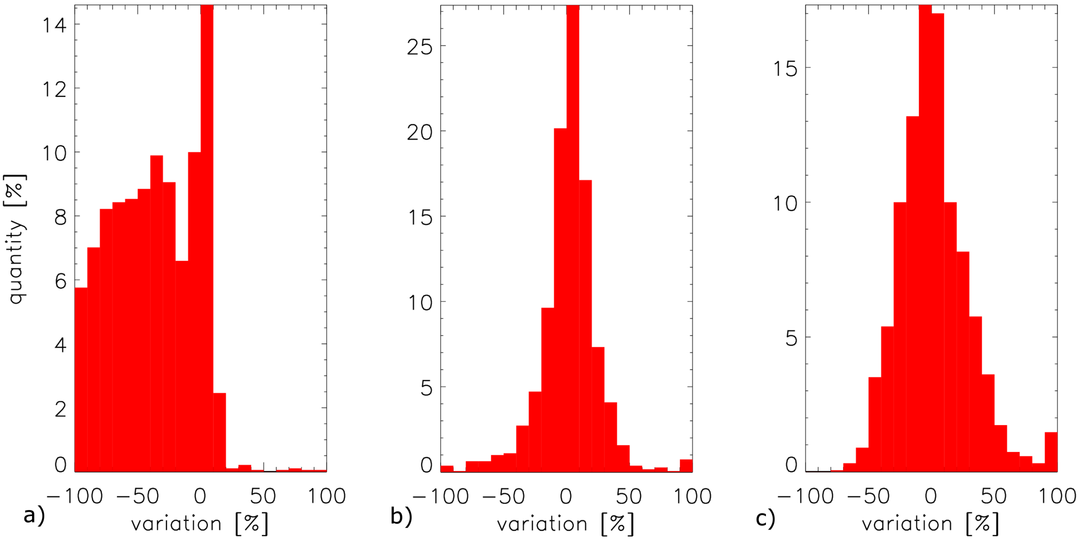

Figure 3.6: Difference between the frequency of occurrence of city clutter pixels and the median of uncorrupted pixels nearby for the reflectivity levels 1 (a), 3 (b) and 5 (c) of the Munich weather radar from 2000–2006 (PX data) in percent.

Mountain clutter

According to Fig. 3.7(a), the mountain clutter for light rain displays the same characteristics as the city clutter. For low rain intensities clutter is corrected, which results in lower frequencies of occurrence. The obstacles of mountain clutter lead to higher reflectivities than the city clutter. So for higher rain intensities the retrieval of precipitation is better than for low rain intensities but the influence of mountain clutter remains and may lead to higher reflectivities due to remaining clutter. It seems that for level 3 (Fig. 3.7b) the influence of clutter is minimized, but it is likely that the remaining clutter and the corrected clutter balance each other. For level 5 (Fig. 3.7c) the portion of remaining clutter dominates. But there are still obstacles such as the ’Zugspitze’ mountain with reflectivities of 60–70 dBZ. A retrieval of rain amount is not possible there.

The Doppler filters are quite effective, but still both insufficiently corrected clutter[1] and thresholded data affect the frequencies of occurrence. Clutter that is produced by moving echoes such as insects, birds, wind turbines or ships cannot be corrected efficiently by the Doppler correction method. Additionally, ships have different cross sections, so the produced clutter is variable. Compared to buildings, clutter from ships is not a permanent phenomenon. So the frequencies of occurrence of radar reflectivities of each level play an important role. For high frequencies the ship-clutter should be negligible whereas for low frequencies it can become a dominant pattern.

---

[1] Due to imperfect filtering.

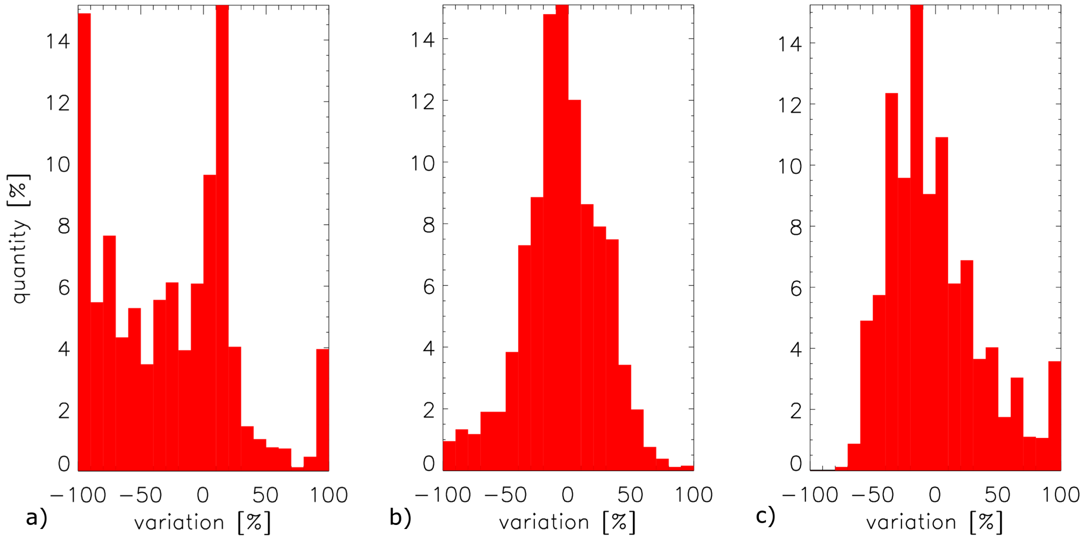

Figure 3.7: As Fig. (3.6), but for mountain clutter for level 1 (a), 3 (b) and 5 (c).

Spokes

There are different sources and types of spokes in radar images. Positive spokes usually originate from other emitters like the sun or from second-trip-echoes under dual PRF [Meischner, 2004]. But the frequency of these spokes is usually too small to become dominant on a longer temporal scale. The spokes in Fig. 3.1 and Fig. 3.2 are negative spokes caused by shading effects behind buildings (beam blockage) near the radar site. Only part of the power is available for measurements behind these obstacles resulting in lower apparent reflectivities. Over range, this affects all pixels of one azimuth in a similar way. So for detection, pixels of one azimuth should be regarded jointly. An overview of the ten apparent spokes is given in Fig. 3.1(a), numbered consecutively. In Fig. 3.8, these ten main spokes are analyzed (clockwise; starting at the 12 o'clock position).

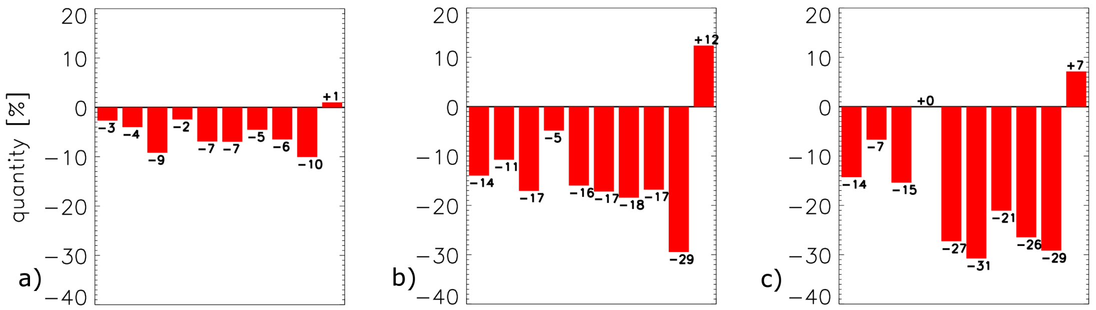

Figure 3.8: Difference between the median of the frequency of occurrence of pixels within ten obvious spokes and the median of uncorrupted pixels nearby for the reflectivity levels 1 (a), 3 (b) and 5 (c) of the Munich weather radar from 2000–2006 (PX data).

So each column of Fig. 3.8 shows the difference in percent between the median of the frequency of occurrence of radar reflectivities of all pixels of one spoke and the median of the corresponding frequencies of unaffected pixels around (approx. 10 km). For level 1 a slight underestimation under the influence of the spokes can be observed. The class width of level 1 is larger than that of the other levels and only part of the radar beam is blocked. Therefore, most measurements in the spoke will still fall in class 1. For higher rain amounts the underestimations are larger indeed, but they stay more or less stable for reflectivity levels 2 to 4 (not all shown here). Only the red spoke #10 differs completely from the others in displaying increased frequencies of occurrence of radar reflectivities instead of decreased frequencies. This is not astonishing, as its source is not shading but clutter from airplanes approaching Munich airport.

3.2.2 Radar composite

Measurements from different radar sites with varying beam elevations and different radar site altitudes have to be merged for compositing. Original single radar range-bins in polar coordinates have to be transferred to Cartesian or geodetic coordinates. In areas where single radar beams from different radar sites overlap, a strategy must be found how the different measurements are combined. Consequently, some additional systematic variations and effects in composite data are to be expected which are analyzed in the following.

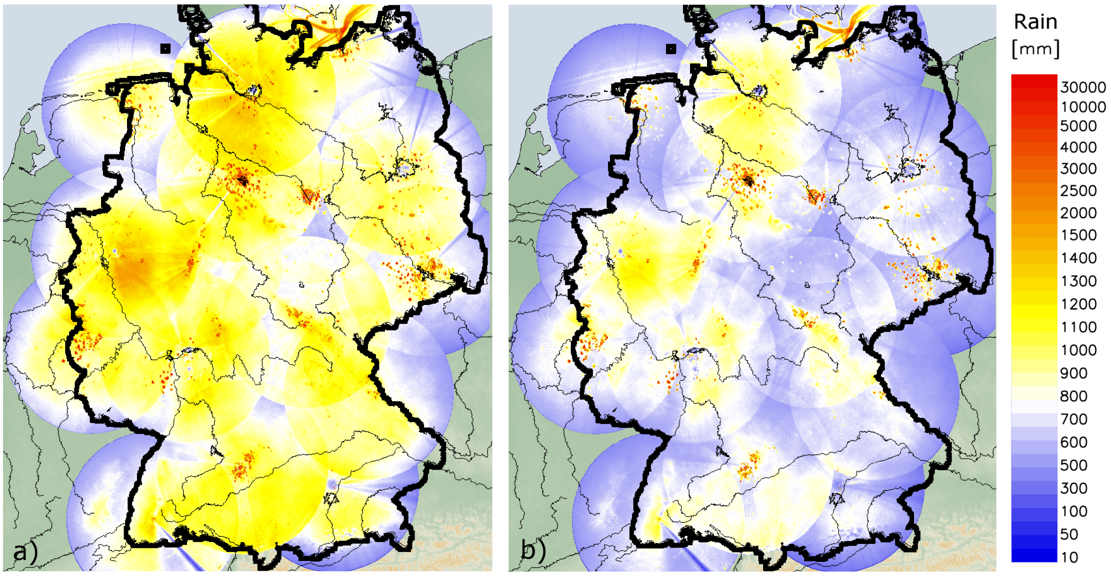

Figure 3.9: Uncorrected annual rain amounts for Germany based on radar composite data RX from 2005-2009 for all measurements (a) and for only those measurements where all 16 radar systems contribute (b). The rain amounts are calculated by the three-part Z/R relationship in Tab. 3.1. The scale is 775 x 900 km^2.

There are two main reasons why the RX-product is chosen to be analyzed as an example for German radar composite data. The first one is that this primary radar product is one of the basic radar products of DWD where only the corrections within the signal processor are performed (see Chapter 2). So this product includes unmodified systematic measuring effects and shortcomings. Secondly, some secondary radar products such as the cell-tracking product CONRAD [Lang, 2001] are based on this product. If the results of these secondary products are going to be corrected, the effects and disturbances within the RX data should be known or even a correction scheme may be transferred from the primary radar product to the secondary one.

Figure 3.9(a) shows the mean annual rain amount for Germany based on the RX composite data from 2005 to 2009. Some particular features and anomalies become apparent that do not originate from precipitation. High 'rain amounts' due to the influence of clutter close to the radar sites and negative spokes around radar sites are visible. Near the coasts of the North and Baltic Sea, lines of clutter are produced by ships. All of the radar systems reveal a greater or lesser decrease of rain amounts with distance from the radar site. In addition, there are some compositing shortcomings. Adjacent radar systems e.g. the Emden radar in the very north-west of Germany and the Hamburg radar to the east of the Emden radar, show significant differences in rain amounts. Furthermore, the boundaries of the overlapping areas of several radar systems are visible and there is the tendency of higher rain amounts in all overlapping areas compared to single radar areas.

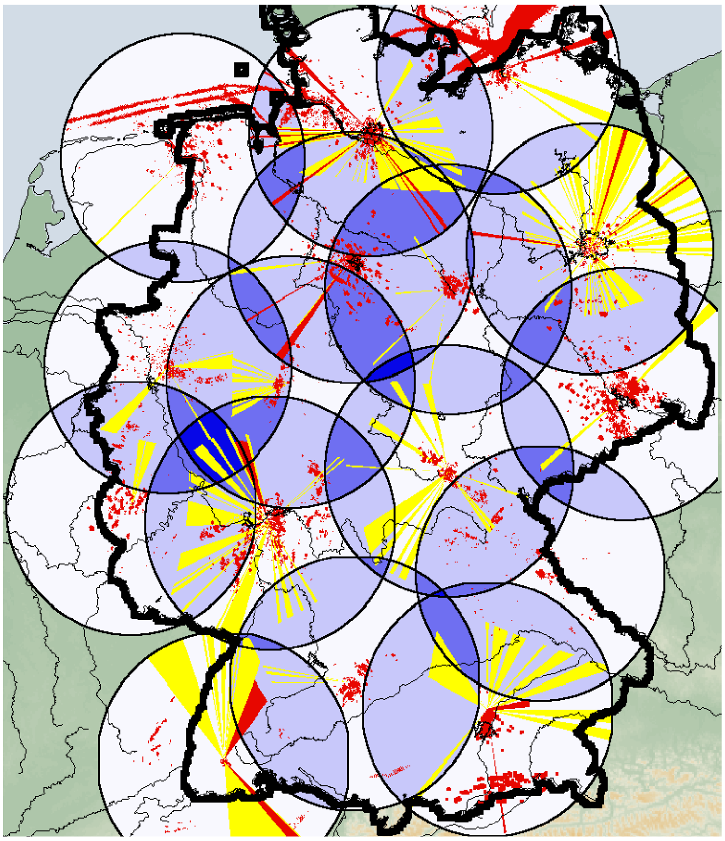

Figure 3.10: Overview of disturbances within the radar composite product RX including clutter pixels (red), spokes (yellow) and the overlapping areas of several radar systems in different blue colors revealing the varying number of contributing radar systems.

Figure 3.10 gives an overview of the main types of disturbances in the RX composite data. Single radar areas are marked white whereas overlapping-areas are marked in blue colors. The yellow color indicates the spokes which still include precipitation patterns. The red color marks clutter. The identification of disturbances within the radar composite is based on the evaluation of disturbances within the PX-product from the 16 single weather radars.

The effects that are already apparent in single radar images are discussed exemplarily in this chapter based on the analysis of the Munich radar data. In the following, the patterns and disturbances are analyzed and evaluated that are caused by the compositing algorithm, based on the RX composite data from 2005 to 2009.

Effects of the coordinate conversion method

If several range-bins of one or more radar systems match a composite pixel, the maximum criterion is applied in DWD to decide which range-bin is used. Until 2006 the so-called *push*-procedure[2] leads together with the maximum criterion to the extension of high range-bin values to up to 9 Cartesian pixels [Weigl and Winterrath, 2010]. Figure 3.11(a) shows the annual rain amounts for the single radar area of the Munich weather radar in 2006 based on the RX composite. In Fig. 3.11(b) the annual rain amounts for 2006 based on the DX single radar product is opposed.

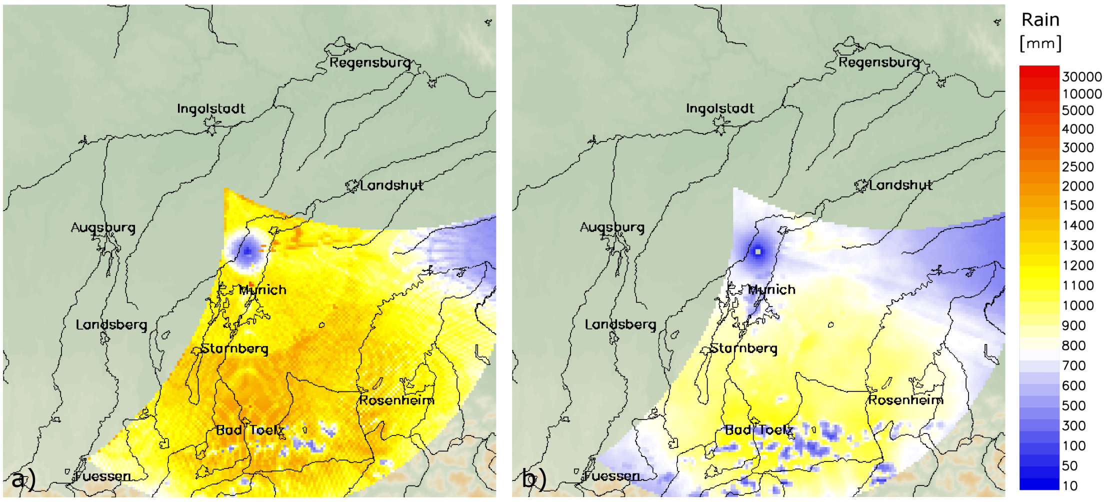

Figure 3.11: Uncorrected annual rain amounts for 2006 for the single radar area of the Munich weather radar based on the RX-product (a) and on the single radar DX-product (b).

The section of the RX-product reveals a significant higher magnitude of rain amounts than the DX section. The main patterns are identical but ring structures at distances of approximately 9 km, 25.5 km and 81 km become apparent in the RX section. This is the same for other radar sites. Additionally, small-scale variations of rain amounts subject to the position of the radar beam compared to the Cartesian grid are likely; they are responsible for the grainy or pixelized patterns in the left image. In 2005, the same effects can be found (not

---

[2] For the *push*-procedure all range-bin values in the overlapping 1 x 1 km^2 area are considered and the maximum value is chosen.

shown here). Since 2007 the *pull*-procedure ensures better results as only one fixed range-bin is used for each composite pixel. Since 2007 there is hardly any difference between single radar areas within the RX composite and the corresponding DX-sections (not shown here).

Higher rain amounts in overlapping areas

The varying availability of radar data from different radar sites is obviously one reason for higher rain amounts in overlapping areas. The availability for each radar system is usually over 90 %. In order to determine whether further influencing factors exist, only those composite data are accumulated, where all radar systems contribute. The total of 518690 measurements decreases by 29 % to 369320 measurements in this way, even when most of the time only one radar system is missing. Figure 3.9(b) shows the mean annual rain amount for these measurements. No significant differences of the precipitation patterns become obvious compared to Fig. 3.9(a). Again, the overlapping areas show higher rain rates than expected from the adjacent single radar areas and the sharp boundaries are visible. The maximum criterion seems to be a plausible explanation for these patterns. This criterion might reduce a probable underestimation in these outer areas of single radar measurements, but hides the natural decline of rain amounts (or of frequencies of occurrence of radar reflectivities) with distance from the radar site, simultaneously.

Variation with altitude in overlapping areas

The altitude dependence of composite pixels in overlapping areas has to be evaluated separately because pixels with different dependences on altitude (different radar sites) are merged. For each measurement and each pixel within the overlap, the maximum criterion decides which radar system provides the range-bin value. On average, a preferred allocation of a radar system for each pixel results. Again the decrease of the frequencies of occurrence of radar reflectivities with altitude is used to realize this allocation which has already been explained in Chapter 3.1.2. The intersection of two radar systems usually leads to a characteristic behavior of the frequencies of occurrence of radar reflectivities with distance from the radar site due to the interaction of single radar calibration, range-bin height and dependence of reflectivity on altitude.

An example of an arbitrary overlapping area shall illustrate this behavior: Figure 3.12 shows the overlapping area of two contributing radar systems for level 1, level 3 and level 5. The black crosses mark the frequency of occurrence of these three reflectivity levels with distance from each radar site. The red crosses indicate their median for equidistant classes (5 km). Pixels within the range of negative gradients imply an allocation to the considered radar system whereas positive gradients mean the opposite. The turning-point marks the transition, where it is statistically uncertain which radar contributes most frequently to the composite pixel under consideration. For overlapping areas with three or four contributing weather radar systems the interpretation of the diagrams is similar. Fig. 3.13 reveals an example for an overlapping area of three radar systems for level 3.

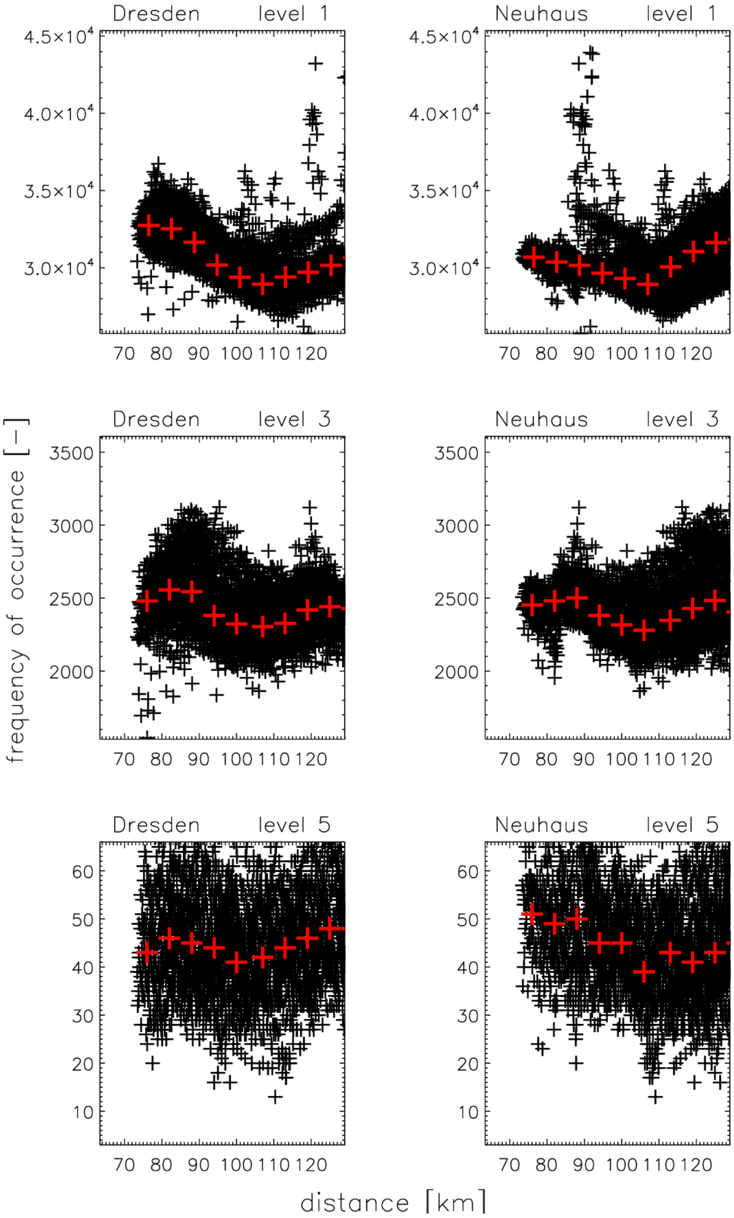

Figure 3.12: Characteristics of the frequency of occurrence of uncorrupted pixels as a function of distance from the radar site for the reflectivity levels 1 (top), 3 (middle) and 5 (bottom) of the overlapping area of the radars Dresden (left) and Neuhaus (right) for the years 2005, 2006 and 2009, overplotted by the corresponding median of equidistant classes of distance (red).

These gradients vary depending on the availability of the contributing radar systems and many overlapping areas are too small to derive robust values for these gradients. Fortunately, the turning points (area of transition) in Fig. 3.12 and Fig. 3.13 represent stable patterns for all levels. Knowing the areas of transition for all of the 76 overlapping areas, a map of ’core-competence’ for each composite pixel is derived. Figure 3.14 shows this map. Pixels with the same color show a statistical derived allocation to the same radar site of more than 50 %.

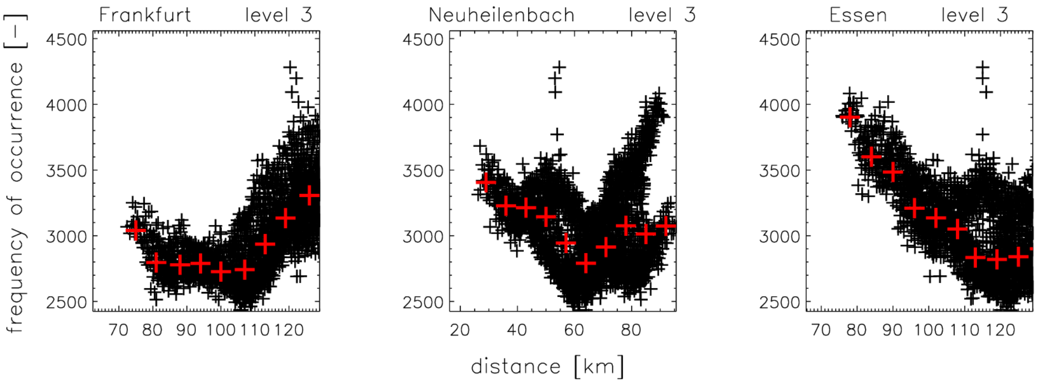

Figure 3.13: Characteristics of the frequency of occurrence of uncorrupted pixels as a function of distance from the radar site for the reflectivity level 3 of the overlapping area of the radars Frankfurt (left), Neuheilenbach (middle) and Essen (right) for the years 2005, 2006 and 2009, overplotted by the corresponding median of equidistant classes of distance (red).

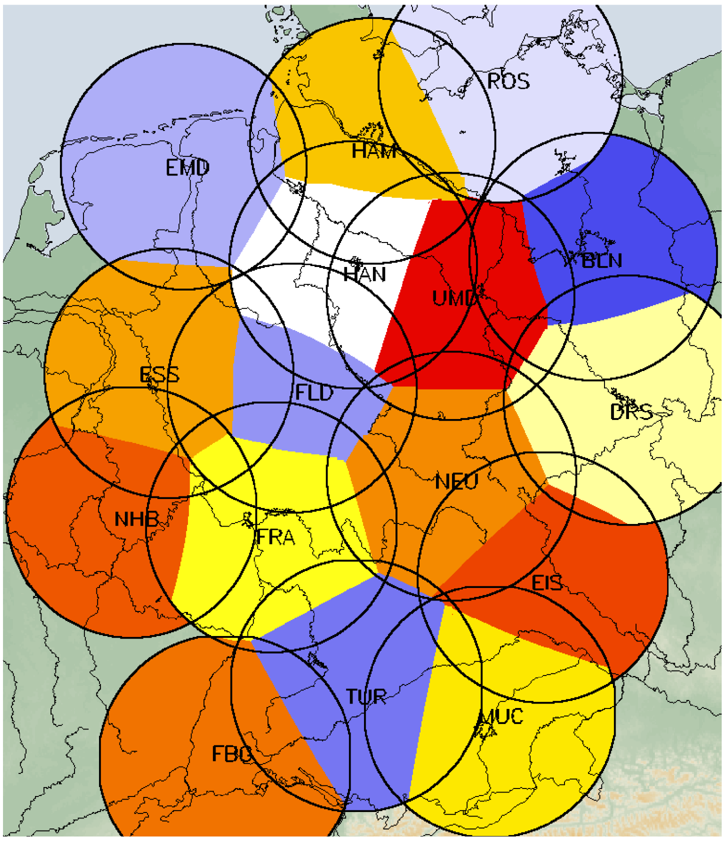

Figure 3.14: Mean allocation of radar pixels to radar sites within the radar composite RX. The abbreviations are explained in Tab. 2.2.

3.3 Discussion

Individual single radar images[3] usually reveal plausible precipitation patterns for the whole radar coverage. On a long temporal scale such as months or years some significant systematic effects in radar data such as clutter influence, shading effects and the decrease of rain amounts with distance from the radar site, become apparent. In this investigation only the net effect of the identified limitations is analyzed. Clutter effects can lead to both underestimations due to discarded values and overestimations due to clutter remnants. The observed relationship of clutter influence on reflectivity level reflects the sensitivity of the Doppler correction scheme in the signal processor. This is true for all clutter effects. Some pixels are permanently affected by clutter, e.g. because of obstacles. Moving non-meteorological echoes cannot be removed by Doppler filters. Even sophisticated pattern identification schemes (e.g. [Gabella and Notarpietro, 2002]) may fail to remove them entirely as differences between clutter pattern and precipitation pattern are often weak. On a longer temporal scale these effects usually are visible if it is a permanent or at least a longer lasting phenomenon. Static clutter maps are therefore often used and may be accompanied by clutter identification schemes based on longer temporal scales.

Obstacles near radar sites can cause negative spokes. If such spokes still include precipitation patterns the reduction of the signal due to the shading effect is weak and usually cannot be observed in individual single radar images. Influencing factors such as refraction due to certain air stratifications (anomalous propagation [Meischner, 2004]) can influence the magnitude of values in the spoke but are rare phenomena.

The main measuring effect concerning all pixels, is the altitude dependence of frequencies of occurrence of radar reflectivities. This effect arises from the natural variability of precipitation with height on one hand side and the impact of beam-widening and increasing beam-height above ground with increasing range as well as the dependence of reflectivity on the state of aggregation on the other hand side. A stable linear behavior of this effect can be deduced on an annual basis whereas higher variations are observed on a monthly time scale.

The analysis of accumulated composite data reveals the same limitations as single radar data plus additional compositing effects. The *push*-procedure that was used until 2006 clearly reveals the impact such an algorithm can show on precipitation patterns. The significant overestimation of rain amounts can be corrected by adjustment to rain gauges. Conversely, the interference of the algorithm in spatial precipitation patterns is to value worse because the main added value of weather radar data compared to other measurements is the spatial pattern. The maximum criterion that is used for compositing in overlapping areas also shows side effects. Overlapping areas are usually at far ranges from the radar sites where a tendency of underestimation of rain amounts exists. The maximum criterion increases rain amounts and can lead to a better accordance with rain gauge data. At the same time, the scheme produces sharp gradients at the boundaries of single radar areas and overlapping areas and disturbs real rain patterns in this way on a longer temporal scale. Additionally, it is not possible to trace back the origin of a pixel value in overlapping areas which massively impede any

---

[3] Single radar images of single time steps.

cause-based correction there. The above mentioned side effects are specific for the German radar composite and the underlying algorithm. As mentioned in Chapter 1, most countries create the composite based on a volume scan. In this way pixels at comparable altitudes can be merged, spokes can often be avoided at the expense of higher range-bin heights and lower range-bin altitudes for certain areas can usually be used. But the usage of different beam heights may lead to further inconsistencies in precipitation patterns and the temporal availability of a volume scan is lower than for a single scan. Despite those differences the main measuring effects evaluated in this chapter count for all weather radar measurements.

Consequences for a correction scheme

The analysis of disturbances in single radar images is based on a classified product with six reflectivity classes. Higher resolution data would certainly be desirable for detailed investigation, but consequently the data base would be scarce if more than six classes are analyzed. Fewer classes were fit based on safe statistics. With this data base, it is possible to estimate the average bias due to beam elevation angles, clutter, and beam shading.

As the drop size and its state of aggregation highly vary with time and depend on air mass and temperature, two ways to realize an altitude correction seem reasonable. The first one is to correct each single radar image by the vertical profile of reflectivity, which is very sensitive, difficult and not available in retrospect available for radar products of DWD. The other one is to perform a mean correction, which is presented here. The linearity of the decrease with height is reasonable for investigations on an annual basis. For shorter periods it still seems to be a plausible approximation as long as the melting layer does not influence the range of radar range-bin heights. If so, the bright band effect cannot adequately be corrected with a linear approach. The bright band should be negligible at least for heavy precipitation in convective cells as convective cells usually have a large vertical homogeneity of rain. There are some indications for that effect within this analysis, but the frequency of occurrence of such high reflectivity levels is too low to be statistically robust.

The described clutter is a permanent or a sporadically occurring phenomenon. Therefore, the analysis of clutter with time is neglected, but a dependence on reflectivity level become apparent. A correction of the affected pixels only seems to be reasonable by the use of interpolation techniques.

For the correction of spokes, a differentiation between spokes with and without rain patterns is useful. This is done visually. If reliable rain patterns are still apparent, they can be used for correction, otherwise interpolation techniques have to be used again. The enhancement of those areas is easy to realize. It is important on a longer temporal scale to preserve the continuity of rain patterns and it seems plausible to regard it as a static feature.

The current compositing algorithm of DWD reveals significant deficits with regard to producing homogeneous rain patterns on a long temporal scale. Corrections are necessary to get rid of systematic variations within the composite. Any limitation in single radar data, such as disturbances of clutter and spokes and the systematic variation with altitude that remains in the dataset even when corrections are applied, are transferred to composite data. Additional compositing effects, especially due to the maximum criteria lead to further systematic differences within rain patterns.

The grainy structure with ring effects due to the *push*-procedure until the end of 2006 cannot be translated back to original data, so these effects remain in the data. But the related higher rain amounts can be corrected and may lead to different adjustment factors than for the rain amounts from 2007 on. The obvious differences of rain amounts among different single radar areas also suggest different adjustment factors for the respective radar sites. According to the above results in this chapter, the overlapping areas have to be handled separately from single radar areas. This is true for adjustment and for altitude corrections. Furthermore, it is not possible to derive reliable dependences on altitude in overlapping areas, only the turning points are stable features. This impedes a post-correction for altitude of these areas. The development and the validation of the correction schemes based on these findings are presented in the following Chapter 4.

4 Development and validation of a post-correction algorithm for radar data

4.1 Single radar data on an annual basis

4.1.1 Modules of a correction algorithm

Based on this knowledge of disturbance processes a modular concept to correct these deviations has been designed. The correction algorithm consists of four modules: 1) the altitude correction, 2) a correction of spokes which still include precipitation patterns, 3) the correction of clutter affected pixels and 4) the adjustment to rain gauge data. The most effective order of the modules to achieve best results is used: The altitude correction (module 1) eliminates the effects of different elevation angles of the terrain-following scan so that the adjustment of spokes (module 2) is based on a more homogeneous data basis. For the following modules, the number of uncorrupted and corrected pixels influences the quality of the correction as those pixels are used for clutter correction (module 3) and for the adjustment to rain gauges (module 4). To realize a reflectivity dependent clutter correction, this module has to be performed before the adjustment. Module 4 is not based on frequencies of occurrence of radar reflectivities from the PX-product but on the comparison of rain amounts of rain gauges and the radar DX-product. So, modules 1, 2 and 3 are based on frequencies of occurrence of radar reflectivities and module 4 on rain amounts.

The separation of corrupted and uncorrupted radar pixels described in Chapter 3.1.2 is similar to the one described here, using an empirical distribution technique for each reflectivity class. As the frequencies of occurrence of radar reflectivity levels 5 and 6 are too small to be statistically analyzed, the selection of corrupted pixels of level 4 is adopted for these heavy rain levels.

Figure 4.1 gives a visual impression of the impact and the gained quality that is achieved by the following correction scheme. Based on reflectivity level 3[1], the result after each step of correction (module) is presented.

---

[1] Also see Fig. 3.2(b).

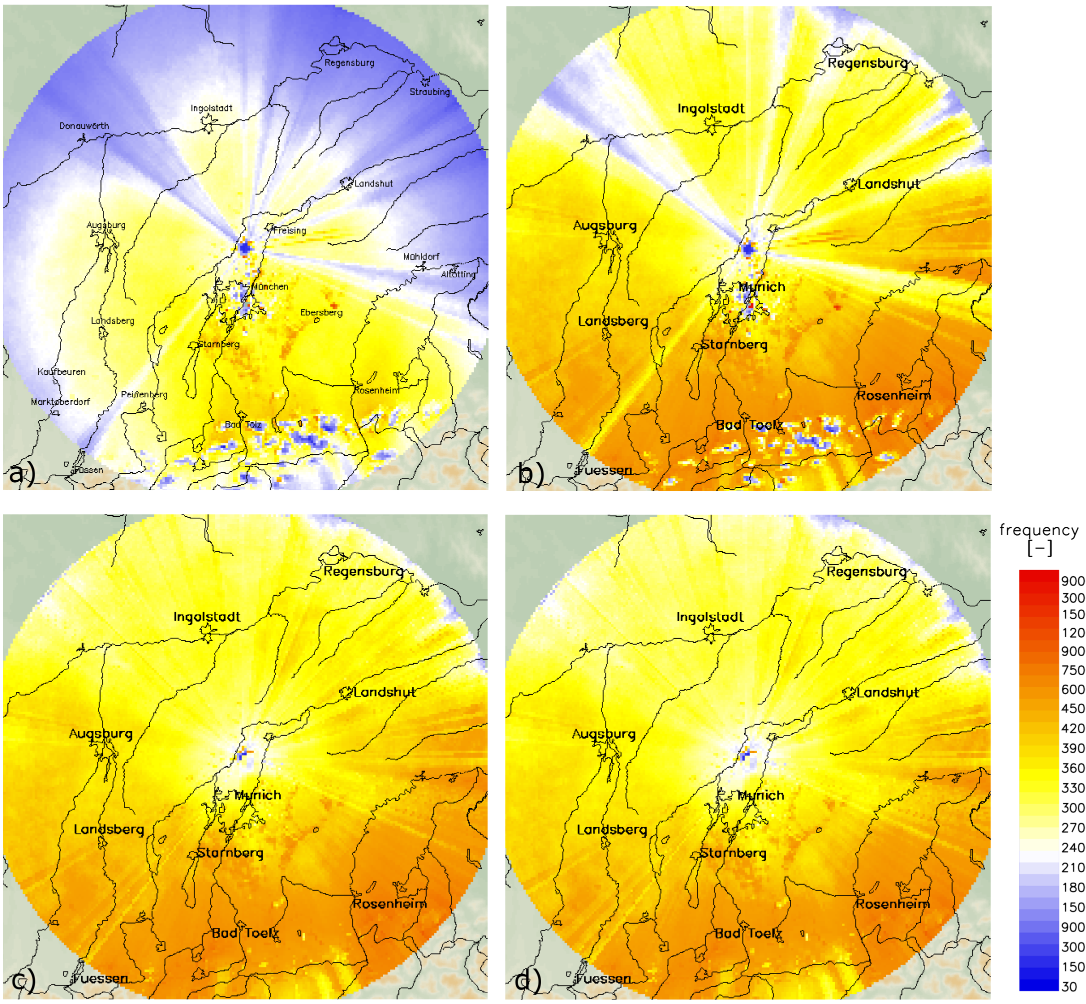

Figure 4.1: Result of the statistical correction of the frequencies of occurrence of reflectivities (level 3) of the Munich weather radar from 2000-2006: (a) uncorrected, (b) altitude correction, (c) altitude correction and correction of spokes and clutter, (d) overall correction.

Module 1: Altitude correction

The altitude correction is applied to all pixels, even the uncorrupted pixels. In Chapter 3 a dependence of the frequency of occurrence of radar reflectivities on beam elevation angles and on distance from the radar has been found. Based on these statistical results, a mean correction factor is derived to provide comparable results for all beam elevation angles. Therefore, as the beam elevation angle varies between 0.8° and 2.1° over azimuth, six mean classes of comparable beam elevation angles are built. Each reflectivity level is analyzed separately. Above 1 km height, Fig. 3.3 shows a linear decrease of frequencies of occurrence with height suggesting a linear regression model. Thus, the frequencies of occurrence of radar reflectivities at each height will be adjusted to the ground value Y_0 (see Eq. 4.3).

First, the mean slope m_r in km^{-1} (depending on range r in km) of a linear regression for each reflectivity level and each elevation angle class ϵ is calculated with range r as predictor.

Then the relation between this mean slope and the elevation angle for each level is analyzed. The multiplication of this mean slope with the elevation angle shows approximately equal results even for reflectivity levels 2 to 5. So the regression equation for the frequencies of occurrence of radar reflectivities Y that applies to all pixels can be formulated as follows, where Y_0 is the calculated frequency of occurrence of radar reflectivities at ground level:

$$Y = m_r \cdot r \cdot \epsilon + Y_0 \tag{4.1}$$

The altitude alt in km is approximately also a function of range r and elevation angle ϵ:

$$alt = tan(\epsilon) \cdot r \tag{4.2}$$

So the regression equation with the mean slope m_{alt} in km^{-1} (depending on height alt) can be formulated as follows:

$$Y = m_{alt} \cdot alt + Y_0 \tag{4.3}$$

The correction factor f_{cor} in km^{-1} is the quotient of m_{alt} and Y_0 and describes the mean variation of the frequency of occurrence of radar reflectivities with height.

$$f_{cor} = \frac{m_{alt}}{Y_0} \tag{4.4}$$

The equation for correcting single pixels can be described by using the altitude alt of each radar pixel:

$$Y_0 = \frac{Y}{alt \cdot f_{cor} + 1} \tag{4.5}$$

For the Munich weather radar the correction factor is calculated to be f_{cor} = -0.203 km^{-1} for moderate and heavy precipitation.

The decrease of the frequency of occurrence of radar reflectivities (level 3) with height of all radar pixels in Fig. 4.2(a) is compared to only the median of each altitude class (Fig. 4.2b)[2] for the same data basis. The variation is similar to the correction factor f_{cor} which is calculated from the regression equation (see Eq. 4.4). Figure 4.2(a) reveals a large dispersion including natural variations of radar reflectivities, but still the decrease of reflectivity detections with height becomes obvious in both figures. The different behavior of the frequency of occurrence of radar reflectivities at a height below 1 km altitude is neglected because of two reasons: First, the effect of the altitude correction near the radar site is small. Second, the full correction algorithm is generated especially for convective rain events and heavy rain, where this effect does not occur. Using the altitude correction for the values in Fig. 4.2(b) means to adjust all frequencies of occurrence of radar reflectivity of level 3 to approx. 3800 (see regression line) by using the correction factor f_{cor} = -0.203 km^{-1} and Eq. 4.5.

The plain correction of altitude increases the frequency of occurrence of radar reflectivities.

---

[2] Same as Fig. 3.3(b).

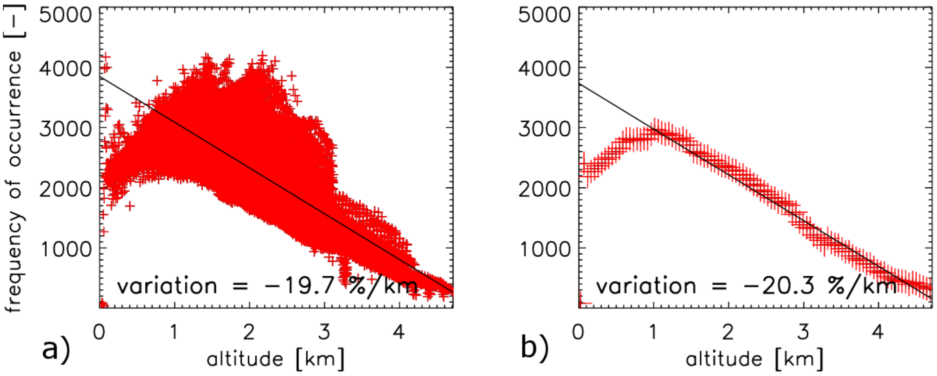

Figure 4.2: Characteristics of the frequency of occurrence of uncorrupted pixels with height for the reflectivity level 3 of the Munich weather radar from 2000-2006 (PX data) of all radar pixels (a) and of the median for equidistant classes of altitude (b).

The maximum of the frequencies of occurrence of radar reflectivities is not at ground level, but the altitude correction is based on this linear regression line and accordingly increases the frequencies to ground level. The reasons for this maximum have already been discussed in Chapter 3. Whether this approach leads to an overrating of rain amounts also depends on the Z/R relationship, but is in any case corrected by the adjustment (module 3).

The part of the altitude correction has to be extended for light rain to apply the correction scheme to total rain amounts. Especially snow in the winter months with smaller reflectivities than rain causes the decrease of the frequency of occurrence of light rain pixels with altitude. Additionally, especially in winter the class width of level 1 is larger than that of levels 2 to 5, thus many measurements even at higher altitudes will still fall in class 1. Instead of the correction factor of f_{cor} = -0.203 km^{-1} for moderate and heavy rain a factor of f_{cor} = -0.129 km^{-1} for light rain is derived. So the observed decrease of the frequency of occurrence of radar pixels with height is smaller for light rain.

Figure 4.1(b) shows the result for level 3, which obviously reveals significant improvement. The result of the altitude correction is a homogeneous distribution of frequencies of occurrence including rain patterns with areas of higher and lower frequencies of occurrence. The mean frequencies of occurrence are not dependent on the position of the radar any more as range dependence is not considered natural but taken as bias. But still clutter influence remains; the second and third correction steps should solve this problem.

Module 2: Spokes

One main aim of the correction scheme is to preserve the reliable natural precipitation patterns. Within negative spokes a part of the transmitting power of the radar-beam is shaded, which leads to an underestimation of reflectivity. Compared to single corrupted pixels, the patterns within spokes are still reliable, in principle. Only if beam blockage is very large the spoke pixels are classified as clutter. This determination is visually done. The remaining pixels jointly undergo a mean adjustment, as each pixel of one spoke suffers the same blockage caused by the same obstacle. As a consequence the patterns are preserved. For each 1° azimuth angle of one spoke the median of the frequency of occurrence of one reflectivity class is calculated and compared to the median of the frequencies of all uncorrupted pixels of the adjacent 20 azimuth angles. The ratio of the two medians is calculated, which then serves as a correction factor. These factors vary between 1.1 and 1.5. The two spokes (spoke #10 in Fig. 3.1a) east of the radar are caused by departing or landing planes and must therefore be completely interpolated.

Module 3: Clutter

According to the results of the statistical analysis in Chapter 3, clutter effects are dependent on reflectivity. The lower the reflectivity of rain, the more the influence of clutter dominates. The number of pixels that are corrupted by clutter decreases significantly when the reflectivity increases. Therefore the correction will be parametrized by reflectivity.

If single pixels are affected by clutter the interpolation technique is used, as these pixel values are not reliable and an adjustment of single pixels may lead to a high variability in space for other time spans than the calibration period. The frequency of occurrence for these corrupted pixels is calculated by interpolation by the closer environment (10 to 20 km) of the frequency of occurrence of uncorrupted pixels.

The improvements according to the correction modules 2 and 3 are shown in Fig. 4.1(c).

Module 4: Adjustment

The adjustment is necessary to correct uncertainties related to the calculation of rain amounts from radar measurements. Therefore, a stable adjustment to the mean rain amounts of rain gauges is implemented.

The adjustment is based on the DX-product and represents the fourth module of the statistical correction scheme. Here, rain amounts are compared instead of frequencies of occurrence. So the radar data is converted into rain rate by the three-part Z/R relationship in Tab. 3.1. Then mean annual rain amounts from radar data are compared to rain amounts from rain gauges for the time span 2004-2006.

For the adjustment only the 18 rain gauges within a distance of 30 to 70 km from the radar (group 1) are used (see Chapter 2). Radar pixels in this area are highly reliable as these pixels are neither influenced by city nor mountain clutter, but are still close to the radar site. The rest of the rain gauges (group 2 and group 3) are used in this Chapter for evaluation. So the rain amounts of the 18 rain gauges are compared to the corresponding rain amounts derived from radar reflectivities (9-pixel-value) resulting in a common mean adjustment factor. A locally varying factor would tend to distort measured precipitation patterns radar climatology usually aims at.

This factor has been determined for the corrected radar data as well as for the uncorrected radar data. For the uncorrected data a factor of 1.37 has been determined. This underestimation can be explained by the significant decrease of rain amounts with distance from the radar site due to large beam elevations. The factor of 0.95 for corrected data shows that even before adjustment the accordance between rain gauge data and corrected radar data on an annually basis is good.

The adjustment factor is calculated based on an annual rain amount. To apply the adjustment factor also to frequencies of occurrence of radar reflectivities an assumption has to be made: An integral rain amount (e.g. annual rain amount) is considered to be the product of the rain amount calculated from radar reflectivity times the corresponding frequency. For example, if the adjustment factor is 0.5, the rain amounts calculated from radar reflectivities by the three-part Z/R relationship must be halved. The same quantitative result can also be achieved by halving the absolute frequencies of occurrence for each radar reflectivity class instead. This does not change the relative frequency distribution of radar reflectivity classes, which have so far been normalized to the arbitrary value of Y_0. Consequently, the adjustment factor can be applied to frequencies in order to integrate the rain amount of the respective class, as long as the calculation of rain rate is performed by a static Z/R relationship.

The frequency of occurrence of radar reflectivities is then simply multiplied by this factor in Fig. 4.1(d). After performing all correction steps for the Munich weather radar, the visual inspection reveals now a relatively homogeneous distribution of frequencies of occurrence of radar reflectivities over the whole image but is still conserving meteorologically or geographically induced minima and maxima of rain. The comparable frequencies of occurrence of radar reflectivities within the whole radar image are indications for an effective and successful correction. In order to prove the quality of the correction algorithm a comparison with rain amounts from rain gauges has been performed and results are summarized in the following section.

4.1.2 Evaluation of corrected single radar data

Application of the method

Despite the fact that all modules are stand-alone corrections, it is adviseable to apply the altitude correction only in combination with the adjustment due to its impact on all pixels.

For the following diagrams, the order of correction steps is changed to demonstrate the improvement with each step of correction: gauge adjustment, altitude correction and correction of spokes and clutter affected pixels. The correct magnitude of rain amounts has a significant impact on statistical values such as the Root-Mean-Square-Error (RMSE) so the gauge adjustment is shown first. The presented order in this chapter has no side effects on the results itself.

The annual rain amounts derived from radar measurements (9-pixel-value) are opposed to the annual rain amounts of rain gauges (group 2) of the period 2004 to 2006. The 57 measurements are further subdivided into five groups according to their distance from the radar (every 20 km) representing the likely modification of radar rain amounts with distance from the radar site.

Figure 4.3 and Fig. 4.4 demonstrate each step of correction of radar data in comparison to rain gauges. Figure 4.3 shows scatterplots of mean annual rain amounts of radar and rain gauge measurements. The subdivision into 20-km-classes (range class) is illustrated in columns. Each row represents one step of the correction, starting with the uncorrected data (Fig. 4.3a), adjusted radar data (Fig. 4.3b), additional altitude correction (Fig. 4.3c) and full correction

(Fig. 4.3d). The first column shows the mean annual rain amounts in class 1 (inner 20 km). While two measurements are a small basis for a comparison, the following reasons explain, why they should be analyzed separately. First, the measurements within a few kilometers around the radar site are not very reliable. In addition, some parts of class 1 represent the city of Munich with possible clutter effects. Second, regarding Fig. 3.3, the measurements between 20 and 40 km from the radar site are at the beginning of the descending branch of the frequencies of occurrence. Measurements below 20 km are in the ascending branch (at least for level 1 to level 4), where the regression line of the altitude correction does not represent this behavior. Hence, the results from class 1 have to be handled with care.

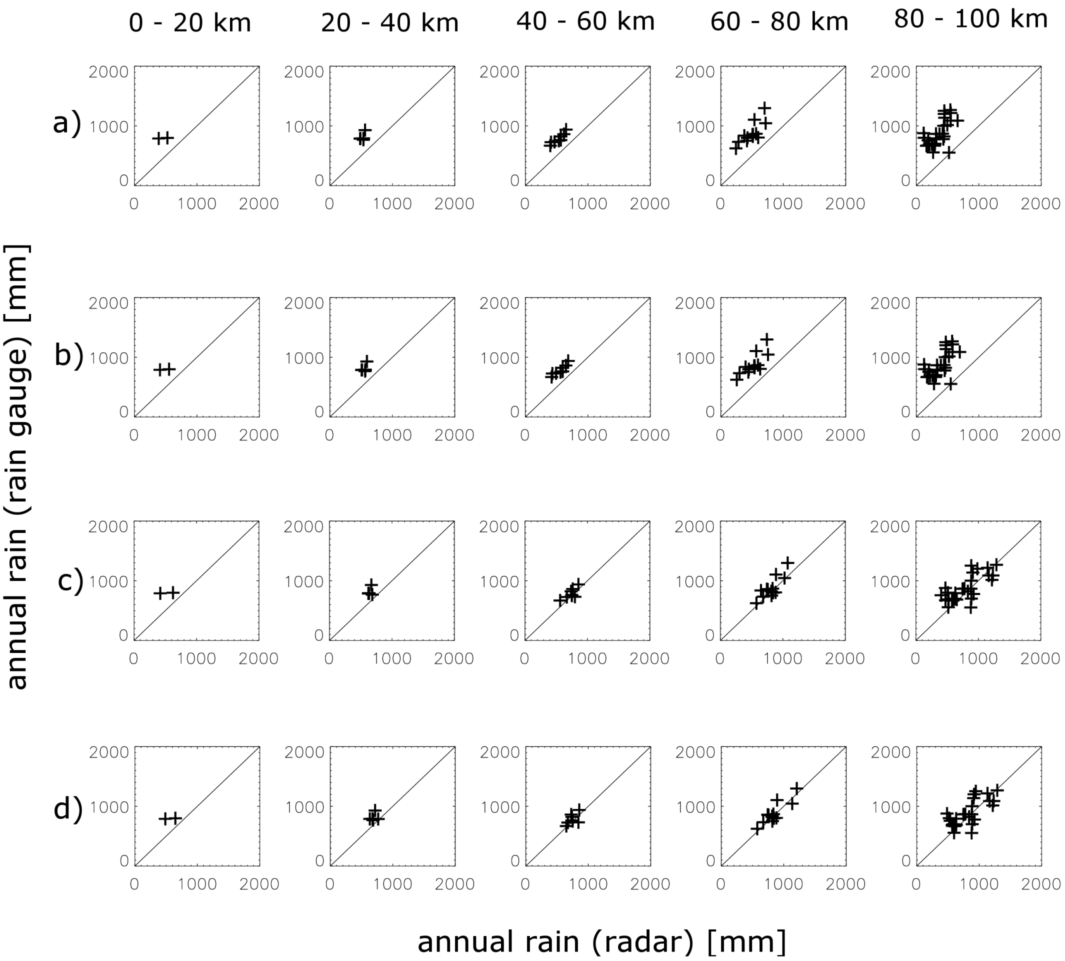

Figure 4.3: Scatterplot of radar and rain gauge pairs of values (group 2) for the statistical correction of reflectivity level 3 of the Munich weather radar from 2004–2006 (DX data). Each row shows one step of the correction algorithm: (a) uncorrected, (b) adjusted, (c) adjusted and altitude corrected, (d) adjusted, altitude correction and correction of clutter and disturbances. The radar and rain gauge pairs of values are subdivided into five classes according to their distance from the radar site. Each column represents one range class: 0–20 km, 20–40 km, 40–60 km, 60–80 km and 80–100 km.

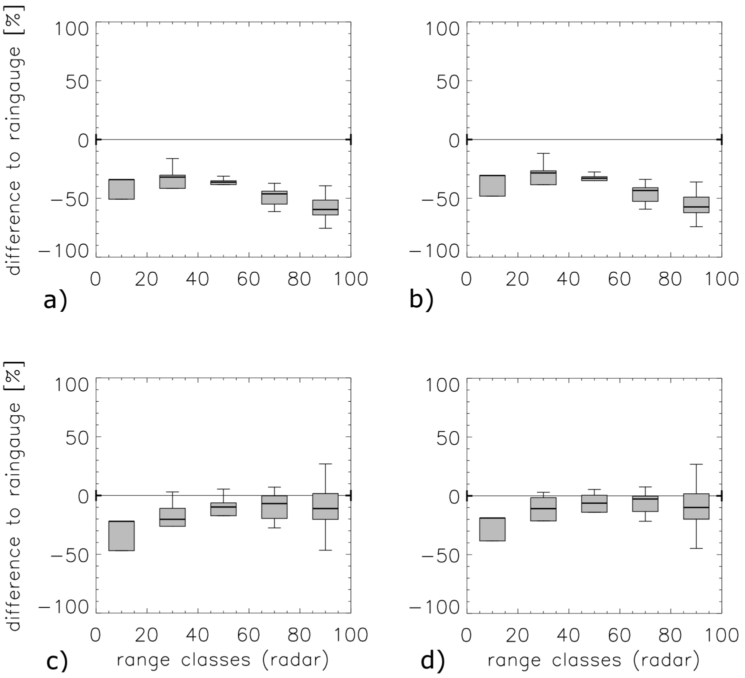

Figure 4.4: Same data as for Fig. 4.3 but box-and-whisker-diagram. The difference in percent between radar and rain gauge data for each range class is shown. The thick bar indicates the median of each range class. The boxes show the deviation of 50 % of all radar and rain gauge pairs of values of one class. The whiskers mark 1.5 times the corresponding interquartile range or, if not reached, the maximum deviation.

Figure 4.4 shows the same comparison in a box-and-whisker-diagram with five boxes starting with the 0-20 km-class. The boxes and whiskers mark the difference in percent of radar measurements in comparison with the corresponding rain gauge measurements for each range class.

For all range classes of the uncorrected data in Fig. 4.3(a) and Fig. 4.4(a) an underestimation of annual rain amounts based on DX radar measurements becomes obvious for all pairs of values. Besides this general underestimation a mean gradual decline can also be observed, starting by approx. 30 % in class 2 and increasing continuously to 60 % in class 5 (up to a maximum of 76 %). It can be attested that the decrease of rain amounts with distance from the radar on average is a measuring effect of the radar measurements, whereas the rain amounts based on rain gauges are independent from this distance. The three-part Z/R relationship used to calculate rain amounts from reflectivity measurements is based on long-term measurements. The best consistency of measurements from radar and from gauges can be expected for long-term-measurements, even though systematically induced variations may add up. It is well known from radar hydrology, that for shorter time spans much higher mean differences between radar measurements and gauge measurements occur.

With the statistically derived mean factor adjustment, the radar measurements are shifted to a little higher rain amount level in row b) of Fig. 4.3. For all classes an underestimation of

	uncor (*)	cor-b (**)	cor-ba (***)	cor-bac (****)
RMSE [mm]	448.0	401.8	153.3	140.0

* uncorrected radar data
** corrected radar data (bias / adjusted)
*** corrected radar data (bias / adjusted, altitude)
**** corrected radar data (bias / adjusted, altitude, clutter)

Table 4.1: RMSE for the comparison of radar and rain gauge data in Fig. 4.3 and Fig. 4.4.

rain amounts derived by radar measurements remains. The RMSE decreases from 448.0 mm to 401.8 mm (see Tab. 4.1). Figure 4.4(b) reveals mean underestimations between 25 % and 55 % by radar data. Without the subsequent correction modules this correction only leads to a minor improvement of data quality. Using the adjustment factor of 1.37 for uncorrected data a significant improvement of the RMSE to 269.5 mm becomes obvious (not shown here).

After the altitude correction a significant improvement of the consistency of measurements from radar and rain gauges becomes apparent in Fig. 4.3(c). Especially the higher range classes, where the altitude correction shows the greatest impact on the rain amounts, are improved, both visually and based on statistical values. The RMSE changes from 401.8 mm (adjusted data) to 153.3 mm. The box-and-whisker-diagram supports these results. The deviations of the median of radar rain amounts from the median of gauge rain amounts at each range class decrease from a span of -60 % to -30 % (adjusted data) to a span of -21 % to -8 % (corrected data).

Row (d) of Fig. 4.3 shows the results after the full correction including the correction of spokes and cluttered pixels. The RMSE (140.0 mm) indicates the further improvement as the diagram itself visually does. The maximum deviations of the median of radar rain amounts from the median of rain amounts from rain gauges of each range class decrease to a span of -18 % (range class 1) to -1 %. A comparison of the diagrams in Fig. 4.3(d) with the diagrams in Fig. 4.3(a) shows a significant improvement for all range classes. The greater the distance from the radar site, the higher are the deviations of rain amounts between measurements from radar and rain gauges. But even at a distance of 80 to 100 km (class 5) the maximum range of deviation of mean rain amounts is -44 % to +28 % for all 57 pairs of values. This appears to be a reasonable result, taking into account that measurements from rain gauges are affected by measuring problems with wind and snow or representativeness errors such as different sampling volumes and measurement heights. They may therefore differ from radar measurements. In addition, the geography of the Munich radar coverage is very difficult concerning measurements of rain amounts.

Validation of the method

For the validation of the presented correction algorithm the same comparisons of mean annual rain amounts from radar measurements and rain gauges are used, but for the time span 2007 to 2009 and for all available 72 pairs of values used for validation (group 3). This also includes pairs of values in areas where the rain amounts do not seem reliable or where radar measurements are obviously disturbed by clutter effects and have therefore been interpolated by measurements in the closer vicinity. The very southern part within the range of the Alps usually shows higher rain amounts especially at higher altitudes. The interpolation does not take these effects into account, so a massive underestimation of rain amounts by radar measurements in comparison to rain gauge data in this area is likely. A second problematic area is the inner part of the radar coverage. Major parts of range class 1 cover the city of Munich with clutter problems. The results are presented in Fig. 4.5 and Fig. 4.6.

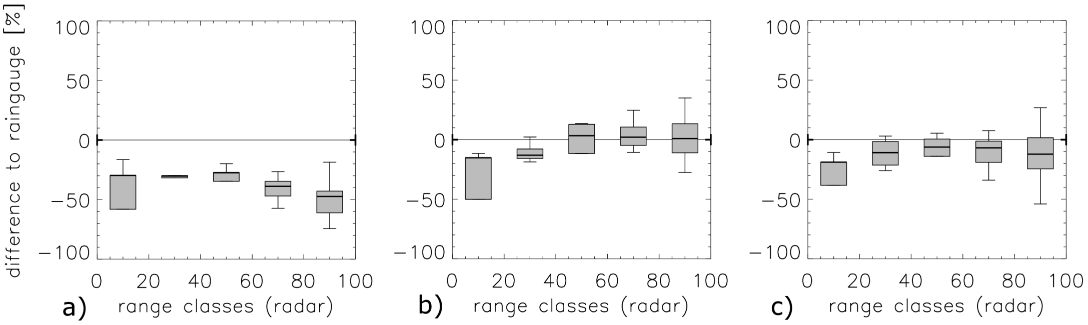

Figure 4.5: Box-and-whisker-diagram showing the difference in percent of all radar and rain gauge pairs of values (group 3) for the statistical correction of reflectivity level 3 of the Munich weather radar – (a) uncorrected (2007–2009, validation period), (b) full correction (2007–2009, validation period), (c) full correction (2004–2006, calibration period) (DX data).

The three diagrams of Fig. 4.5 are comparable to those in Fig. 4.4. Comparisons of rain amounts from uncorrected radar data of the years 2007 to 2009 in Fig. 4.5(a) are opposed to corrected radar data (2007-2009) in Fig. 4.5(b). Figure 4.5(c) shows the identical comparison as in Fig. 4.5(b) but for the years 2004 to 2006.

Figure 4.5(a) reveals a high variance, which is not very astonishing as this comparison also includes questionable pairs of values. The variances in range class 1 reflect the disturbances mainly caused by the inner city of Munich. For the residual classes the modification of rain amounts with distance from the radar becomes clearly apparent. Again a clear underestimation of rain amounts without corrections is shown.

Figure 4.5(b) shows the result of the full correction algorithm. A significant improvement becomes obvious, both visually and based on the RMSE, which decreases from 533.4 mm to 181.3 mm (see Tab. 4.2). As expected, some pairs of values show only poor consistency. Most of them are located in the alpine area. The interpolated radar data is not able to reproduce the high rain amounts measured by rain gauges. The boxes and whiskers of Fig. 4.5(b) show comparable ranges to Fig. 4.4(d). The maximum range of deviation of mean rain amounts is -15 % (range class 1) to +4 % for all 72 pairs of values.

In Fig. 4.5(c) the same 72 pairs of values are compared for the calibration period to get evidence which part of the impairment of the results is caused by the additional pairs of values and which part is caused by the new time span. Figure 4.5(c) shows a higher impairment than Fig. 4.5(b) does. The RMSE of the validation period (RMSE = 181.3 mm) is smaller than the RMSE of the calibration period (RMSE = 258.9 mm). So the impairment of the RMSE of 72

	uncor (*) (2007-2009)	cor-bac (****) (2007-2009)	cor-bac (****) (2004-2006)
RMSE [mm]	533.4	181.3	258.9

* uncorrected radar data
**** corrected radar data (bias / adjusted, altitude, clutter)

Table 4.2: RMSE of mean annual rain amounts of radar and rain gauge data in Fig. 4.5.

pairs of values of the validation period compared to the RMSE of the 57 pairs of values of the calibration period (see Tab. 4.1) is mainly induced by the additional pairs of values.

The correction of radar data in the validation period shows improvements comparable to the calibration period. In conclusion, the validation verifies the possibility of adaption of the correction algorithm for other time spans. For this validation time span the results are even better than those of the calibration time span. This is mainly due to an advanced scan strategy with different beam elevation angles per azimuth in 2006. The installation of a new signal-processor leading to a much better suppression of clutter shows its effects already at the beginning of the calibration period in February 2004.

Figure 4.6 serves as a final visual validation of the results of the full correction algorithm. It shows the mean annual rain amounts derived from DX radar products before (Fig. 4.6a) and after the full statistical corrections (Fig. 4.6b). The correction is based on frequencies of occurrence of DX data and calculated to rain amounts afterwards. The image shows a map of mean annual rain amounts which is very similar to maps of annual rain amounts based on point measurements (not shown here). The highest rain amounts are measured in the Alps and on the fringe of the Alps decreasing to the north. The radar image still reveals some remnants caused by clutter (e.g. city of Munich) or spokes. They can be regarded as weaknesses of the presented correction schemes or effects that are not handled by the schemes. The aim of the correction is to preserve natural patterns as far as possible. Pixels that are minimally affected by shading (spokes) and clutter remain uncorrected in the data base and can lead to minor variations in space. Additionally, the rain patterns of adjusted spokes can slightly differ from the surrounded pixels, when the retrieval of measurements is not entirely possible especially for light rain and low vertical rain echoes. The most conspicuous pattern is the underestimation of rain amounts near the radar site. This area is probably distorted by radar side lobes which point to the ground, produce ground clutter and can affect the measurement of the radar main lobe. Doppler clutter correction removes these effects, but discards measurements at the same time which results in an underestimation of rain amounts in this area. This effect is not explicitly corrected by the correction schemes and seems to be more dominant in the DX-product (polar coordinates) than in the PX-product (Cartesian grid). This can probably be attributed to the fact, that these discarded values are neglected for the transfer of range-bin values in polar coordinates to the Cartesian grid of the PX-product.

Summarized, the correction schemes significantly improve data quality, although there are still effects in radar data which are not handled by the schemes or which are minimized but not entirely removed.

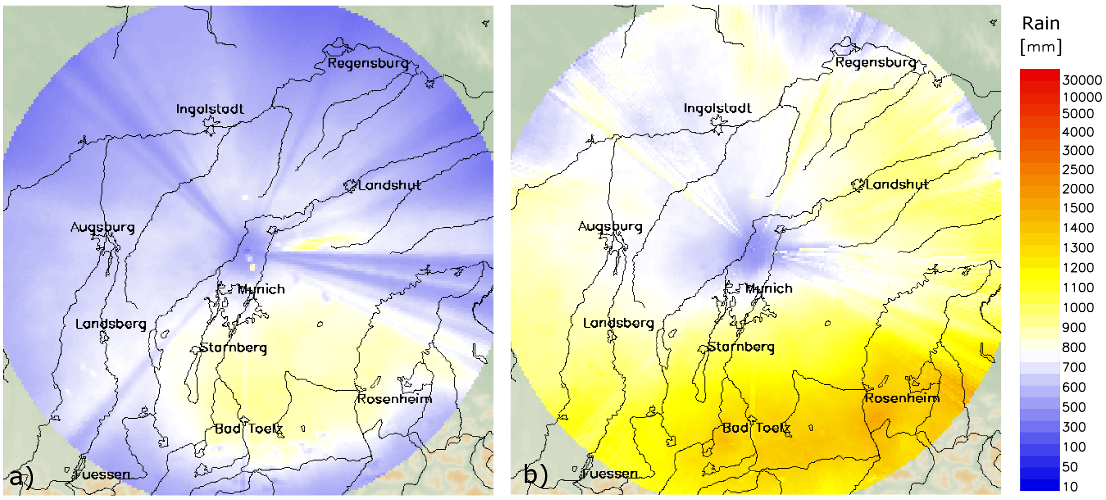

Figure 4.6: Mean annual rain amounts derived from the DX radar product of the Munich weather radar 2007–2009 – (a) uncorrected, (b) corrected.

4.2 Composite radar data on an annual basis

4.2.1 Extension of modules of the correction algorithm

The four correction modules presented in Chapter 4.1 remain the same in principle, but due to the additional compositing effects, some modifications are necessary. The separation of corrupted and uncorrupted pixels is adopted from the statistical analysis in Chapter 3. The altitude correction (module 1) is performed first, followed by the correction of spokes which still include precipitation patterns (module 2) and the adjustment to rain gauge data (module 3). Subsequently, pixels with clutter (module 4) are corrected as the last step of the correction scheme. As these pixels are not used for adjustment in module 3, the best results are achieved by correcting of the surrounding pixels as far as possible. This improved data basis for module 4 outperforms the disadvantage that module 4 is applied to rain amounts which prevent the usage of a different selection of clutter pixels per reflectivity level. Figure 4.7 shows the flowchart of the whole correction schemes for composite data.

As a first step, the analyses of clutter pixels, spokes and dependences on altitude are performed on the PX-product for each of the 16 single radar sites of DWD for the years 2000 to 2006 (insofar as available). Then, the results of the single radar analysis are used as a first rough selection of corrupted pixels. Based on accumulated images of the RX composite, this coarse selection is checked and verified. Therefore, identical tools may be applied to identify corrupted pixels in single and composite radar data (see Chapter 3). The aim is to correct as many pixels as necessary and as little pixels as possible. The verification of clutter pixels, massively corrupted spokes and spokes that still include precipitation patterns but with higher degrees of shading effects, is easy and could even be performed visually. Only spokes with a minimal deviation from the surrounded pixels are more difficult to identify, but they have a minor impact on data quality anyway.

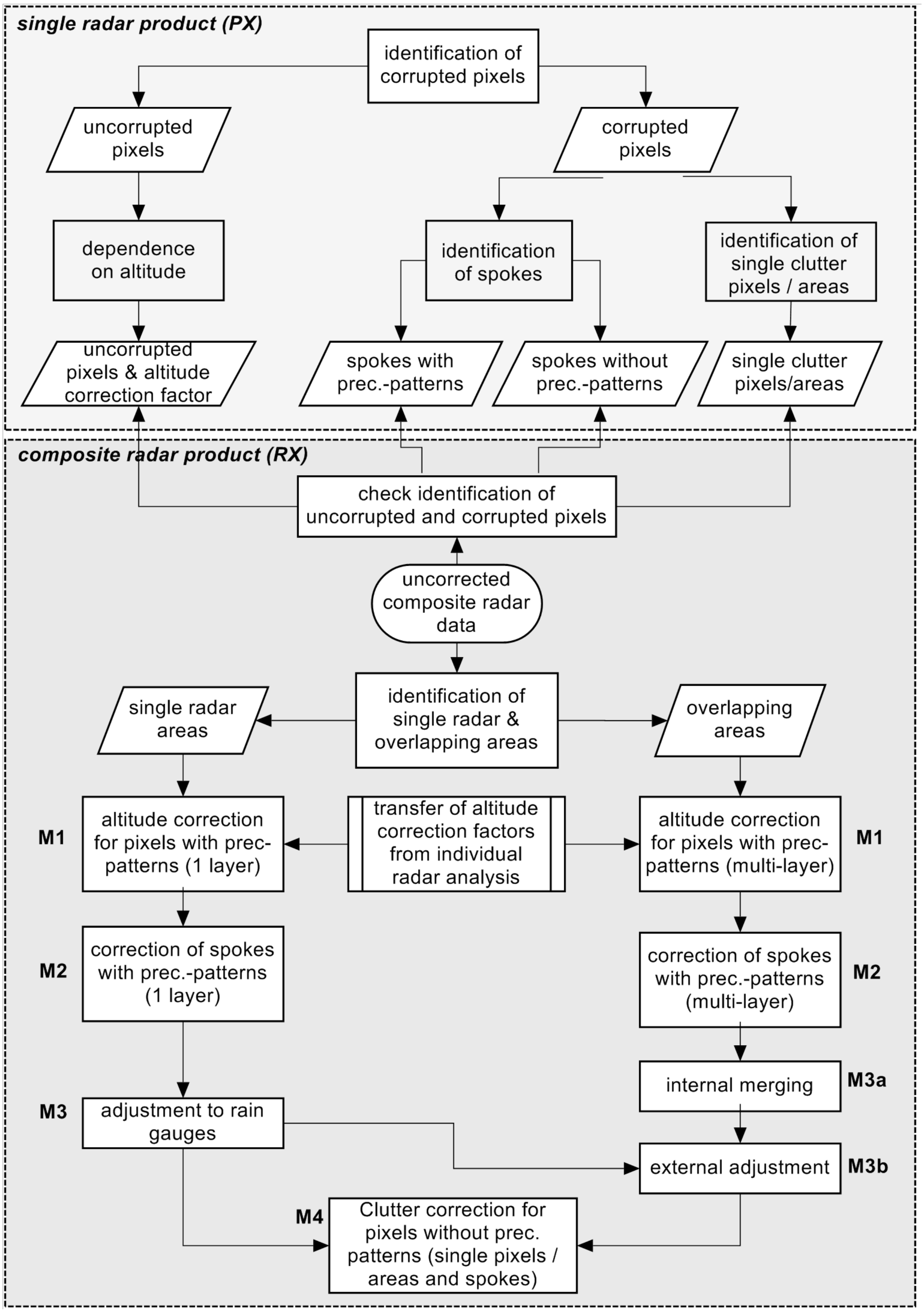

Figure 4.7: Flowchart of the complete post-correction schemes. The modules are labelled M1 .. M4.

Concerning the further course of the correction algorithm, single radar areas and overlapping areas are handled separately in modules 1 to 3, according to the results in Chapter 3. Due to the maximum criterion it is uncertain how measurements of the respective radar sites contribute to composite pixels in overlapping areas. Therefore, a multi-layer approach for overlapping areas is applied, where all of these pixels are corrected independently for each contributing radar site. In module 3 (adjustment) these multi-layers are merged. The correction of clutter (module 4) is performed on a joint data basis as the last step of the correction scheme.

Module 1: Altitude correction

The altitude correction presented in Chapter 4.1 is also used here but with a variable correction factor. For most of the 16 radar systems, the frequency of occurrence of radar reflectivities is not only dependent on altitude, but also on reflectivity level. For the correction, a gradual approach is followed: First, the linear relationship between the frequency of occurrence of radar reflectivity and altitude leads to a correction factor. This correction factor is calculated for all of the six reflectivity levels of the PX-product and for all of the 16 radar sites. Again, the variation of the correction factor with reflectivity in logarithmic scale (dBZ) can be described by a linear relationship for reflectivity level 2 and above. Reflectivity level 1 behaves slightly different from the other levels and requires its own correction factor. This may partially be attributed to the higher influence of snow.

Finally the equations are transferred to the radar composite. If the elevation angles have changed since 2006, or in between 2005 and 2009, the altitude of each pixel needs to be updated, whereas the correction factor remains unmodified.

The correction for each pixel is again based on Eq. 4.5 in Chapter 4.1. For pixels in the single radar area, the correction is similar to the altitude correction of single radar data in Chapter 4.1. For the pixels within overlapping areas the altitude correction has to be performed for each of the contributing radar sites. If e.g. radar beams of three radar sites contribute to an overlapping area, the altitude correction has to be applied three times to each of the corresponding pixels using the respective altitudes and correction factors. Three different values result for each pixel, there.

Module 2: Spokes including rain patterns

The spoke correction proceeds azimuthally by calculating the median of the rain amount or of the frequencies of occurrence of radar reflectivities for each azimuth. The median of the affected azimuth is compared to the adjacent 10 uncorrupted spokes, which results in a correction factor. This correction factor is multiplied by the value of each pixel of the same azimuth.

The correction is performed separately for single radar areas and overlapping areas. For single radar areas the correction is similar to the spoke correction of single radar data in Chapter 4.1. For the overlapping areas the multi-layer approach is applied in turn. Each spoke can be allocated to one radar site. So for each contributing radar site the corresponding spokes are corrected in the above described way. In overlapping areas it is highly likely, that within shaded areas the measurements of other overlapping radar sites have been used. This is applicable to spokes with and without patterns. Even though the transition area within a spoke can vary accordingly to some extent, this is neglected for the correction. This uncertainty has to be accepted as these areas are too small to be analyzed in a reasonable way. The results after using this correction algorithm support this approximation and are presented in this Chapter.

Module 3: Adjustment to rain gauge data

The adjustment to rain gauge data is effected in the third module. Monthly rain gauge data for the time span 2005 to 2009 have been quality checked, aggregated to mean annual values and finally interpolated to the same grid as the composite radar data (see A in the appendix). This new dataset is developed only to guarantee independence by using certain rain gauges solely for adjustment or for validation (split-sampling). Otherwise REGNIE data could also have been used.

The adjustment of radar data for those pixels in single radar areas with a distance between 15 km and 80 km from the radar site, where both data bases provide measurements and where radar data is not affected by clutter, is performed on an annual basis. The area for adjustment is expanded compared to the adjustment of single radar data (30 km to 70 km) to achieve reliable correction factors even for radar sites in the composite with only very small single radar areas. According to the results of the statistical evaluation in Chapter 3, each single radar area has to receive its own adjustment factor. For each single radar area the median of all associated radar pixels and the median for all corresponding interpolated rain gauge pixels are calculated, compared and finally led to an individual correction factor. In this way, radar systems that are calibrated differently are adjusted to the same independent data basis.

The adjustment in overlapping areas is considerably more extensive. The multi-layer data has to be merged to one final dataset (*internal merging*) before the adjustment. Within the adjustment step, the magnitude of rain amounts has to be corrected and simultaneously the obvious boundaries between single radar areas and overlapping areas should be minimized (see Fig. 3.9).

The *internal merging* aims at intersecting the multilayer results in the overlapping areas avoiding gradients. For each pixel in the overlapping area, a percentage allocation to each contributing radar system is realized, based on the map in Fig. 3.14. As a first approximation, the transition area between two radar systems is set to a percentage of 50 % for each radar system, whereas the pixels with the minimum distance to either radar site are set to a percentage of 100 %. The percentages of the residual pixels of the radar systems are derived by distance-weighted interpolation. These proportions are regarded to be static as long as the scan strategies do not change. Individual corrections of two or more radar systems are homogenized by using these proportions.

In case of a percentage of 100 % affiliation to one radar site for the boundary of an overlapping area, this boundary should not be visible in Fig. 3.9(b). So the *internal merging* is obviously not able to correct these discrepancies at the boundaries. The basic idea for the adjustment of overlapping areas is to compare the boundary areas of corrected and uncorrected areas in order to realize the adjustment and minimize the discrepancies at the boundaries (*external adjustment*). Therefore, a step-by-step adjustment to already adjusted areas is performed. First, the overlapping areas of two radar systems are considered. The boundaries of overlapping areas and adjacent, already adjusted single radar areas are compared. Hence, the medians of 2 to 3 rows of pixels at the boundaries are calculated and compared. In this way, a correction factor is derived for each of these boundaries. Finally, these correction factors are calculated for the whole overlapping area using the static proportions derived previously for the *internal merging*. For the adjustment of three overlapping radar systems the single radar areas and the recently adjusted overlapping areas of two radar systems are the basis of comparison to derive the correction factors, etc. So, the correction factor includes the common adjustment to rain gauges, the varying percentage of contributing radar systems at the boundaries of the overlapping areas and probably additional compositing effects due to the maximum criterion. The correction factor directly depends on the data that is to be corrected and therefore it is not static.

Module 4: Clutter

The last module of the correction scheme is the distance-weighted interpolation of clutter affected pixels by surrounded pixels. It is performed on the entire dataset. It comprises all pixels that are massively corrupted, which means that the actual rain patterns are not visible or are not reliable anymore. Regarding radar climatology, a temporal interpolation is hardly possible, so only a spatial interpolation is implemented considering all uncorrupted pixels within an area of 10 to 20 km around the clutter affected pixels.

4.2.2 Evaluation of corrected composite radar data

Temporally and spatially split sampling is used for the evaluation of the correction algorithm. The time series of composite radar data from 2005 to 2009 is splitted into two parts for the evaluation of the correction algorithm (temporal split sampling). The application period covers the years 2005, 2006 and 2009. The validation period is based on the data from 2007 and 2008. The aim of this temporal separation is to receive two comparable time periods regarding age. In this way it shall be ensured that similar results for the validation period are achieved not only because of the newer data with perhaps higher data quality. There is no obvious temporal development in the complete time series except the ring patterns caused by the *push*-technique, which only become obvious in the years 2005 and 2006. Data from 516 rain gauges, which are continuously available, quality proofed and not used for adjustment, are the basis of comparison (spatial split sampling). On an annual basis, the rain gauge data are opposed to the rain amount derived from the spatially corresponding 9-pixel-value of the radar measurements. The consistency between rain gauge data and radar data is shown in the scatterplot in Fig. 4.8 and the Root-Mean-Square-Error (RMSE) between those two datasets is calculated. Additionally, the spatial distribution of annual rain amounts derived from uncorrected and corrected radar data is opposed to interpolated rain gauge data (basis for adjustment).

Application period (2005, 2006 and 2009)

The data of the application period is used to check the correction factors of the altitude correction and the identification of corrupted pixels for the radar composite as well as to calculate the radar site specific adjustment factors for this period. According to Fig. 4.8(a), a clear relation between radar data and rain gauge data is reasonable, but the dispersion is very high for the application period. A significant overestimation of annual rain amounts based on the radar data compared to rain gauge data becomes apparent. An RMSE of 759 mm per year is calculated. Applying all correction modules except the adjustment leads to an RMSE of 459 mm (not shown). The already high rain amounts derived from radar data are further increased by the altitude correction to minimize the systematic differences between low and far ranges from the radar site. So, the consistencies between some pairs of values are initially further deteriorated. Applying a radar site specific adjustment factor improves radar data quality significantly (see Fig. 4.8b) according to an RMSE of 161 mm and high consistencies of pairs of values in the scatterplot.

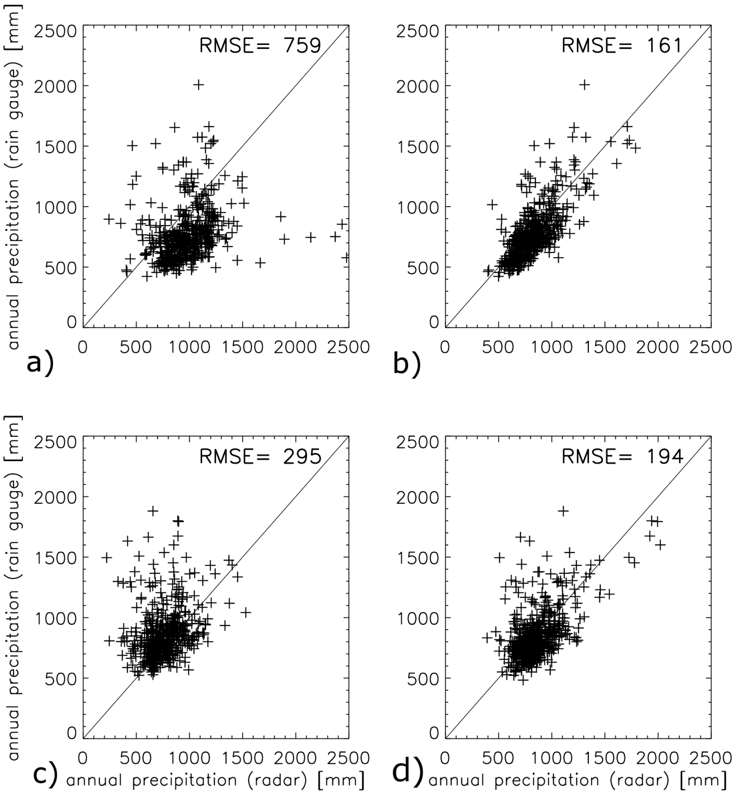

Figure 4.8: Scatterplots of radar and rain gauge pairs of values for the annual rain amounts before (a) and after the correction (b) of the RX radar product for the years 2005, 2006 and 2009 and the respective plots for the years 2007 and 2008 (c, d). For the years 2007 and 2008 no adjustment is performed.

Validation period (2007 and 2008)

For the validation period there is no obvious overestimation of annual rain amounts from radar data. Also the dispersion in the scatterplot (Fig. 4.8c) is much lower, resulting in an RMSE of only 295 mm per year for uncorrected radar data. The main reason for these differences compared to the application period may originate from the *push*-technique which was used for compositing the single radar images until 2006. This technique induces spatial inconsistencies and overestimations (see Chapter 3) and may lead to further deviations of radar data and rain gauge data. As a consequence, adjustment factors derived from the application period cannot be applied to data from the validation period. Additionally, the varying availability of radar site measurements also argues against a temporal transfer of adjustment factors for composite data in general. Due to the already high consistency of radar data and rain gauge data and due to the fact that the very strict validation criteria to use temporal and spatial split sampling shall be maintained, an adjustment factor will not be used for the validation period. As a consequence, the validation is only performed for correction modules 1, 2 and 4. The result after applying all correction modules except the adjustment leads to an RMSE of 194 mm, which is already comparable to the RMSE of the full correction scheme for the application period (161 mm). According to Fig. 3.9 some radar sites are conspicuous due to relatively high (radar Hamburg) or low rain amounts (radar Feldberg) compared to radar data of adjacent radar sites. As a consequence, a further improvement would be achieved when calculating adjustment factors.

According to these results, the correction algorithm is also applicable to other time periods, except the adjustment factors which should be calculated for the respective data basis. The spatial distributions of rain amounts in Fig. 4.9(c, d) support this finding. The uncorrected radar data (Fig. 4.9a, b) shows a variety of clutter effects, measuring effects and differences between radar sites. After the correction, a homogeneous distribution of rain amounts in Fig. 4.9(c, d) results, eliminating disturbances efficiently. Minor disturbances may remain but the high consistency of rain amounts derived from radar data and from rain gauge data in Fig. 4.9(e, f) becomes obvious. Certain differences between those two datasets are probable as the quality of rain gauge data and the quality of its interpolation results are also limited.

Subsequently, a more in-depth analysis of corrected and uncorrected radar data and the corresponding interpolated rain gauge data for the whole period between 2005 and 2009 has been evaluated. For the correction of radar data, the complete correction scheme including radar site specific adjustment factors for the time span 2005 to 2009 has been applied. Box-and-whisker-diagrams are used to show the relative difference in percent of the annual rain amount derived from radar data and interpolated rain gauge data for all spatially corresponding pixels. In addition, the datasets are sub-divided spatially according to what radar site the data originated from and whether a pixel lies within an area of overlapping radar systems or within a single radar area. The allocation of pixels in overlapping areas to a certain radar site is realized according to Fig. 3.14. In that way, 16 plots, each containing 4 box-and-whisker-diagrams, are created. On the left hand side of each plot in Fig. 4.10 the pixels within single radar areas of uncorrected and corrected pairs-of-values are shown and on the right hand side the uncorrected and corrected pairs-of-values of pixels within areas of overlapping radar

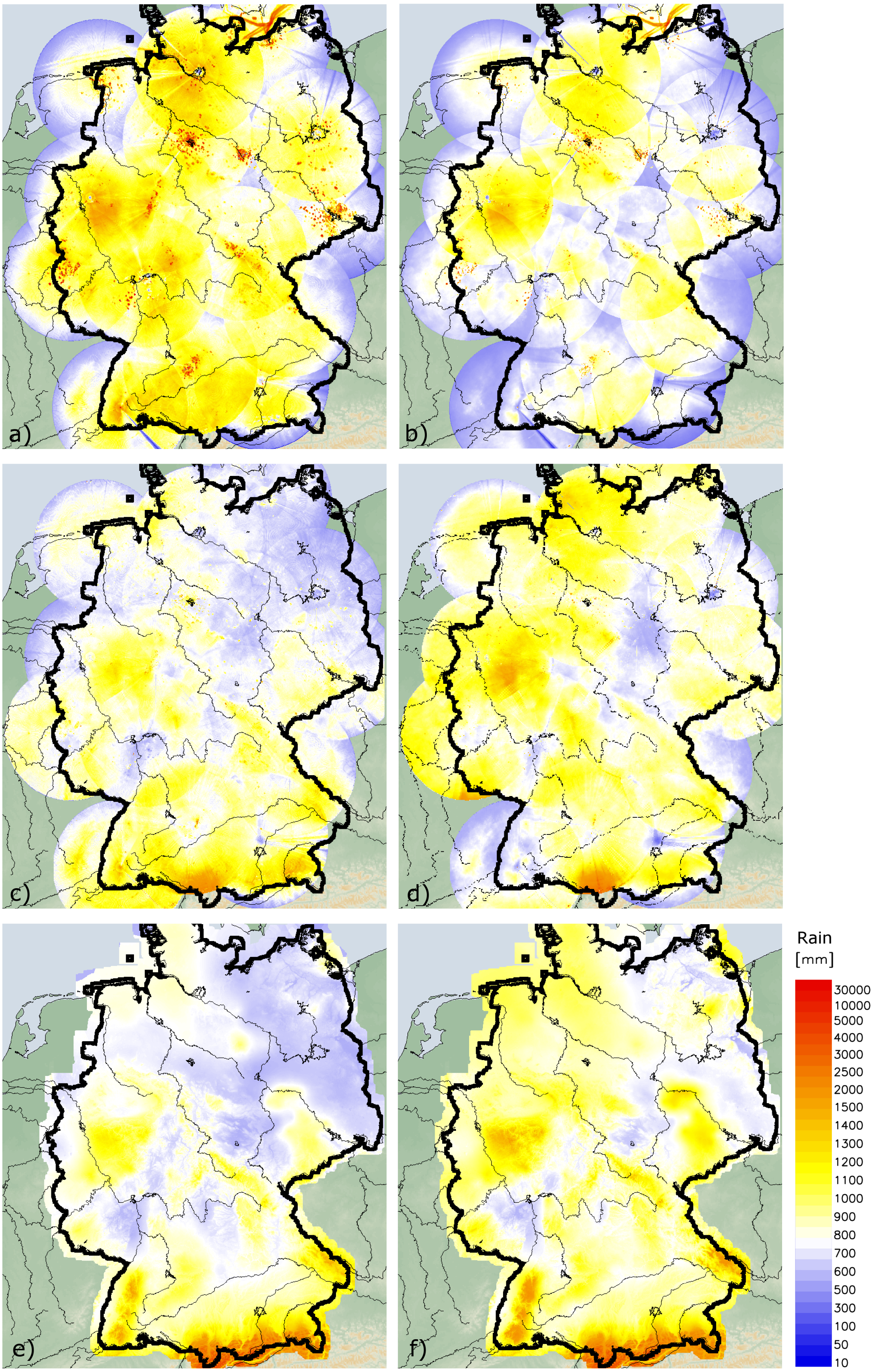

Figure 4.9: Annual rain amounts for Germany based on radar composite data RX for uncorrected radar data (a, b), corrected radar data (c, d) and based on gauge data (e, f) for the years 2005, 2006 and 2009 (left) and for the years 2007 and 2008 (right). For the years 2007 and 2008 no adjustment is performed.

systems can be seen. The median is presented as a thick black line within a gray box. The gray box indicates the central 50 % of pairs-of-values.

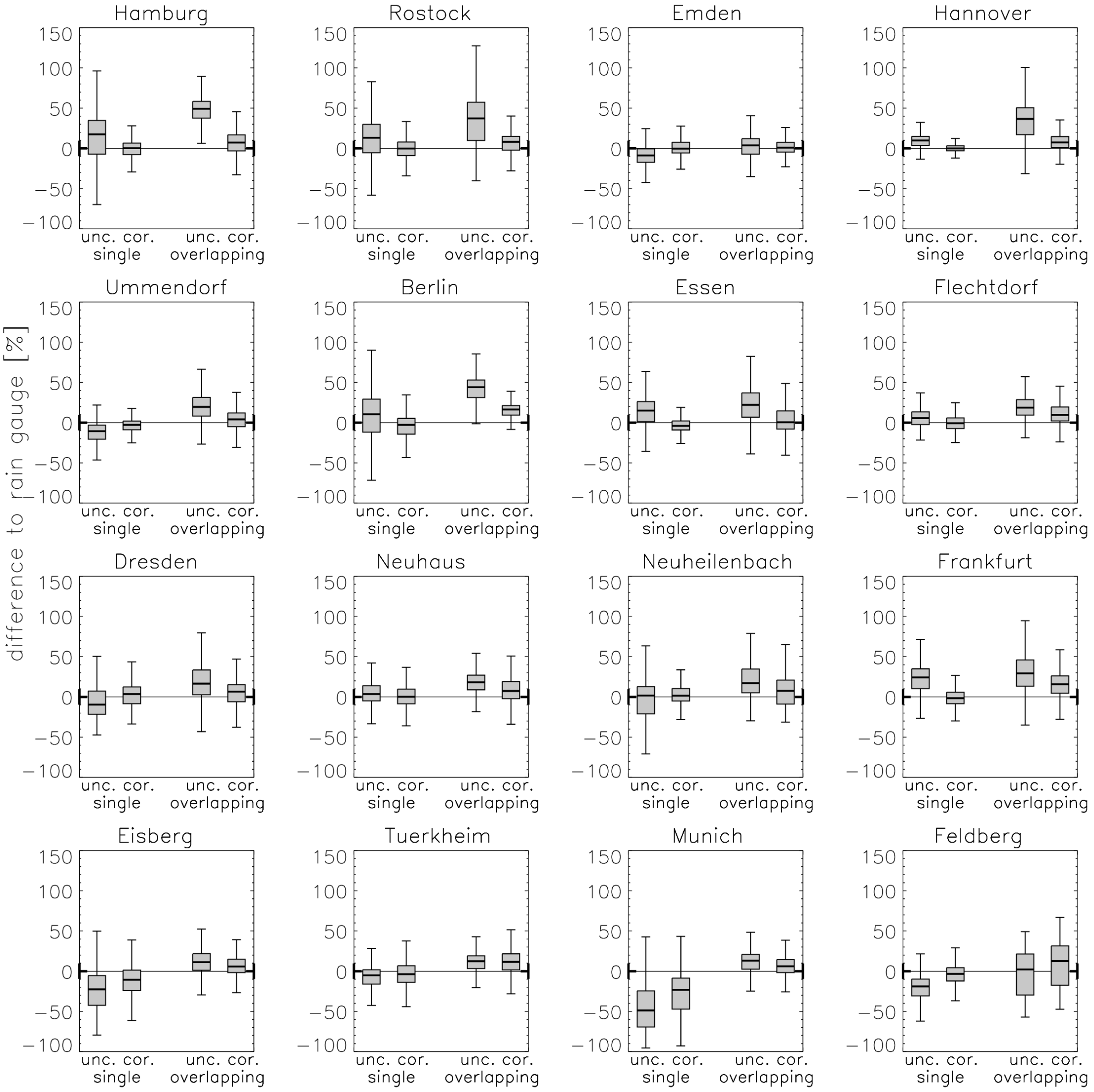

Figure 4.10: Box-and-whisker-diagrams of all corresponding pixels of Fig. 4.9 separated for each radar site for the time span 2005-2009. The difference between radar data and rain gauge in percent is shown. The thick bar indicates the median of all differences. The boxes show the deviation of 50 % of radar and rain gauge pairs of values. The whiskers mark 1.5 times the corresponding interquartile range or, if not reached, the maximum deviation. The first two boxes of each diagram represent uncorrected and corrected pixels of the single radar of each radar site and the last two boxes are uncorrected and corrected pixels of corresponding overlapping areas.

Both overestimations and underestimations of annual rain amounts based on the uncorrected radar data values in single radar areas become apparent. The main reason for these varying results seems to be the ability of radar systems to detect precipitation at low altitudes for the whole radar coverage. The radar systems in the southern part of Germany have to cope with complex terrain or mountains. So either the radar location is at a higher altitude (e.g. the Feldberg weather radar) or the beam elevations have to be enlarged (e.g. the Munich weather radar). In both cases an underestimation of actual rain amounts is probable. Radar systems in the northern part of Germany are usually located at low altitudes and can use low beam elevations and consequently show slight overestimations of rain amounts. In urban areas such as Hamburg, Essen, Frankfurt and Berlin the same tendency can be observed but the reasons for this behavior are rather challenging radar calibrations or clutter influences. But clutter effects alone cannot explain the differences between adjacent radar sites. Almost the complete central 50 % of radar pixels (gray box) within the single radar area of the Emden radar site show lower discrepancies to rain gauge data than the central 50 % of radar pixels of the radar site of Hamburg. The plots clearly highlight these differences among radar sites. But it has to be taken into account that the size of the single radar area for each radar site in the radar composite varies.

Regarding the pixels in overlapping areas, an overestimation of rain amounts by the radar data is shown, although the frequencies of occurrence of radar reflectivities and hence the rain amounts decrease with range-bin height, especially within the outer parts of single radar images. As explained earlier, the maximum criterion used for compositing single radar images enhanced by the higher data availability seems to be responsible for this effect.

The median of annual rain amounts of pixels within single radar areas of corrected radar data should only show minor deviations from the corresponding rain amounts of rain gauge data because of the adjustment step. Differences can be assigned to interpolated radar pixels due to clutter effects, which are neglected for the correction step of adjustment. Only the Eisberg and the Munich weather radar still show underestimations in complex terrain such as the Bavarian Forest, the Bohemian Forest and the mountain range of the Alps.

Pixels within overlapping areas are only adjusted at the boundaries to already adjusted areas. Nevertheless, the median of annual rain amounts of radar data and of rain gauge data shows a high consistency there. The Berlin and the Frankfurt radar sites mark the maximum deviations between both datasets (+ 17 %). For the Berlin weather radar the density of rain gauges and the number of pixels within the respective overlapping areas are low. The Frankfurt weather radar is in general a very challenging site partly covering several low mountain ranges. Furthermore, differences of rain patterns between interpolated rain gauge data and measured radar data are probable.

4.3 Composite radar data for smaller time steps

4.3.1 Modification of the correction algorithm

In Chapter 3 statistical analyses of accumulated radar data of single radar products and composite radar products are performed on a long temporal scale of several years. Influencing factors and temporal or spatial dependences of radar measurements are evaluated. So far, the correction algorithm is applied in the case that for each correction module almost stable conditions are achieved. So, the question is how to modify the correction scheme to be applicable to a shorter time step without neglecting any systematic effects established in these chapters. In the following a brief description of a modification for each correction module will be given.

Module 1: Altitude correction

The variation of the altitude dependence is the most challenging part. On average, a mean annual cycle for the altitude dependences for each radar site can be derived. But due to the strong dependence on temperature this annual cycle can only be an approximation to reality. A second, more flexible algorithm has to be used to consider more real-time conditions.

The modified altitude correction itself is similar to module 1 in Chapter 4.2 but with a daily varying correction factor. Therefore, monthly altitude dependences are derived from PX data for each radar site and subsequently transformed to a three month moving average to minimize random fluctuations. In this way, the variations of the annual cycle are derived based on the average behavior of reflectivity level 2 and level 3. The dependence on reflectivity level is adopted from the annual correction in Chapter 4.2. For reflectivity level 1 a separate annual cycle is calculated. As a last step the correction factors are linearly interpolated between consecutive months to achieve daily correction factors. As an example, the mean monthly varying altitude correction factors of reflectivity level 3 for the Hamburg weather radar (a) and the Munich weather radar (b) are presented in Fig. 4.11. The altitude correction factors are almost zero in summer, whereas in winter they reach unrealistic high values for certain altitudes. According to Fig. 3.4 in Chapter 3 the altitude dependence for the Munich weather radar in January is only valid for altitudes lower than 2.5 km. As a consequence an empirical threshold of 8-times the original value is determined for the whole composite to prevent overcorrection. The modified altitude correction and the associated assumptions are a rough approximation and it is advised not to apply it without the 'fine-tuning' presented in module 3 (adjustment).

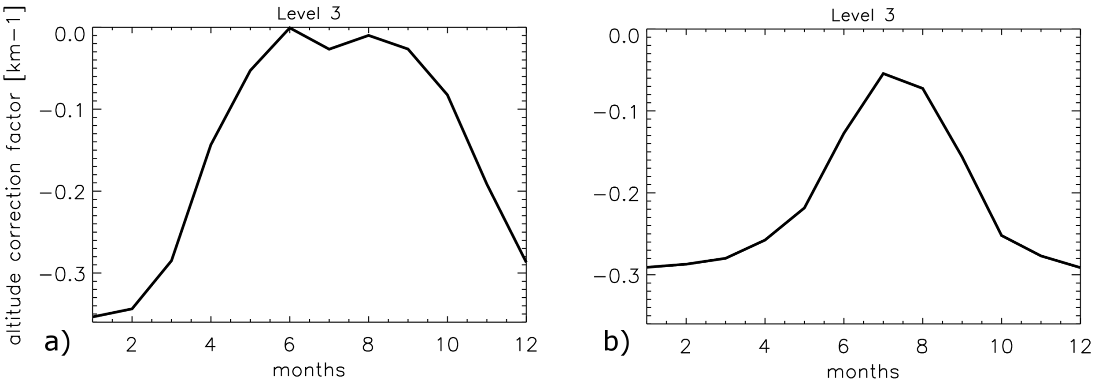

Figure 4.11: Monthly varying altitude correction factor of reflectivity level 3 of 2000-2006 for the weather radar Hamburg (a) and Munich (b). A three month moving average is used to achieve stable values. The values in between are linearly interpolated.

Module 2: Spokes including rain patterns

The correction of spokes including rain patterns is adopted unmodified from the composite correction on an annual basis. Annual correction factors for each spoke are calculated and then applied to each individual composite image. The assumption is that the cause for a spoke is a shading effect of obstacles and therefore a permanent phenomenon.

Module 3: Adjustment to rain gauge data

The adjustment step provides individual correction factors for each radar pixel on a daily basis. REGNIE data is used as a comparator. First, the following steps are performed on a monthly basis to provide backup correction factors. Then, these steps are performed on daily data, but only where the median of the respective data base is at least 0.5 mm. Otherwise, the correction factors from the monthly data are applied. This should prevent unrealistic correction factors due to the lack of reliable rain measurements. The focus is laid on the robustness of the correction scheme, rather accepting underestimations of correction factors than overcorrections.

In principle, the basic structure of adjustment to REGNIE data is similar to the correction module presented in Chapter 4.2. Certain areas are directly adjusted to interpolated rain gauge data based on areal means (median) and adjacent areas are adapted using a comparison of boundaries pixels' mean. The main impact of the altitude correction is at far ranges where in most cases the overlapping areas can be found. As mentioned above, additional feedback mechanisms have to be included to weaken inappropriate altitude correction factors. This is hardly possible when single radar areas are adjusted and overlapping areas are adapted. So the other way round is used here. First, the 28 overlapping areas of two radar sites are adjusted to REGNIE data resulting in one correction factor for each overlapping area. This correction factor is then transferred to single radar areas by comparing the medians of 2 to 3 rows of pixels at the boundaries of these adjacent areas. The resulting individual correction factors of up to six boundaries of one single radar area are then distance weighted interpolated to the entire single radar area.

A second feedback mechanism is used to prevent overcorrection in the single radar areas. Due to the fact that the availability of radar measurements in overlapping areas is equal or higher than in single radar areas, the correction factor transferred from overlapping areas to single radar areas must be smaller or equal to the one calculated separately for single radar areas by adjustment to REGNIE data. If this is not correct, the distance weighted correction factors mentioned above are reduced according to the magnitude of the new correction factors of each single radar area. The adjustment of single radar areas is completed then.

As a last step, the resulting correction factors at the boundaries between single radar areas and overlapping areas of two radar sites are transferred back to overlapping areas and are finally distance weighted applied to radar data.

For the correction of overlapping areas of three or four radar systems, the adjustment at the boundaries is used, which has already been presented in Chapter 4.2.

Module 4: Clutter

The identification of clutter pixels is modified comparably to the annual correction. The previous selection of clutter pixels marks those pixels, where on an annual basis significant differences to adjacent radar pixels become apparent. Some of these pixels are not permanently affected by clutter. The ones that are corrupted only for short periods are not selected. The identification of clutter pixels should also be more flexible with a higher temporal resolution of radar data to correct. The algorithm of Gabella and Notarpietro [2002] for individual radar images is applied. It is a two-part identification algorithm based on echo continuity and minimum echo area to separate meteorological (rain) and non-meteorological echoes such as ground clutter. Its source code can be found in the wradlib-libraries [Heistermann et al., 2013]. Additionally, some clutter effects can hardly be identified in single radar images, therefore thresholds for accumulated images on a daily (300 mm) and monthly scale (1000 mm) are used for the identification of corrupted pixels.

As a final step, the identified clutter pixels are interpolated by inverse distance weighting based on the eight nearest uncorrupted pixels for each time step.

4.3.2 Performance of corrected composite radar data

The modified correction algorithm is still based on static features. The altitude correction is calculated on a monthly basis and then interpolated to daily values. All in all, these are poor prerequisites for the efficient correction of 5-minute values. But nevertheless, there are several reasons, why to apply it: This correction algorithm will not compete with any other algorithm that is deliberately made to correct single radar images. It should be regarded as an additional possibility to improve radar data quality. An alternative correction algorithm that corrects the systematic compositing effects of the German radar composite is not known. One aim of the temporal disaggregation of the correction modules is to show the limits of such statistical correction schemes. But at the same time it may also highlight the shortcomings of radar data on a smaller temporal scale.

Daily REGNIE data which is used for the adjustment of radar data is based on quality proofed rain gauge information of DWD. Any rain gauge data of DWD is probably not independent from this gridded dataset. This fact prevents a proper validation of the correction scheme with rain gauge data. The main focus of the following investigation is on the performance of hourly and 5-minute corrected radar composite data. Corrected and uncorrected radar data with a temporal resolution of 5 minutes for the time span June 2005 to May 2007 is opposed to rain gauge measurements. The whole time span is additionally separated into annual periods or seasons depending on the evaluation. Precipitation data from 642 rain gauges located all over Germany are used for the hourly comparison. 5-minute rain gauge data has only been available from 33 rain gauges in Lower Saxony. The RMSE is chosen to evaluate the consistency of radar measurements and data from rain gauges. The RMSE can be considered temporally and spatially. 'Temporally' means the calculation of the RMSE as a mean value over all rain gauges per each time step. Certain rain events can be highlighted and the seasonal development of the RMSE can be seen in this way. The spatial consideration results in a mean RMSE for a certain period of time for each location of associated rain gauges. The identification of areas and locations of high and low quality of radar data in relation to rain gauges is thereby enabled. Single time steps or rain events with huge differences in rain rate between rain gauge data and radar data can influence the RMSE significantly. The same applies to pairs of values with high discrepancies such as in shaded areas. Therefore, the percentiles 100 % (all values), 98 % and 90 % are calculated for the temporal and spatial consideration to assess the quality of the RMSE values and also to receive a mean consistency of radar data and rain gauge data neglecting outliers.

Monthly rain amounts

Figure 4.12 shows the monthly rain amounts for January 2006 (left) and July 2005 (right) for uncorrected (a, b) and corrected radar data (c, d) and REGNIE data (e, f). The radar images are accumulated images of 5-minute measurements. For uncorrected data only the three-part Z/R relationship is used whereas for the corrected data the whole correction scheme is performed on individual radar composites.

The correction of radar data in winter is challenging due to shallow vertical extension of precipitation echoes, snow measurements that actually require certain Z/S relationships[3] to calculate reliable rain amounts and influencing bright band effects. The uncorrected radar data of January consequently reveals enormous differences in rain rate with altitude. Corrected radar data is able to reduce these effects significantly, but boundaries between single radar areas and overlapping areas might emerge more obviously. These effects may occur and accumulate, if precipitation fields are very heterogeneous and weak or the altitude correction factors are not able to reduce the altitude effect adequately. The image of corrected radar data still shows shortcomings, but the concern that the huge altitude correction factors may impair original radar data is not confirmed.

In July, the influence of the altitude dependence shrinks to a minimum but the magnitude of rain amounts is overrated significantly due to the way of compositing (see Chapter 3). Corrected radar data in summer reduce systematic measuring effects efficiently and show a high similarity with REGNIE data. The rain patterns based on radar data might even be superior to the dotty patterns of REGNIE data, but this is not evaluated here. In general, an increase of the quality of corrected radar data from winter months to summer months can be expected. In the transition periods, bright band effects in single images that cannot adequately be represented by a linear correction function can influence data quality. For a more detailed analysis shorter time steps of radar data measurements have to be investigated.

Hourly rain amounts

Table 4.3 gives an overview of temporally and spatially averaged RMSE values for the whole investigation period separated into seasons and certain percentiles. Such as for the spatial percentile $sp98$ the best and the worst 1 % of stations regarding the RMSE are neglected. $t98$ is similar regarding the temporal consideration and $sp-t98$ represents the 98 % percentile of the RMSE both temporally and spatially considered. The upper table presents the values for uncorrected radar data, the table in the middle those for corrected radar data and the values

---

[3] Reflectivity snow relationships [Atlas, 1990].

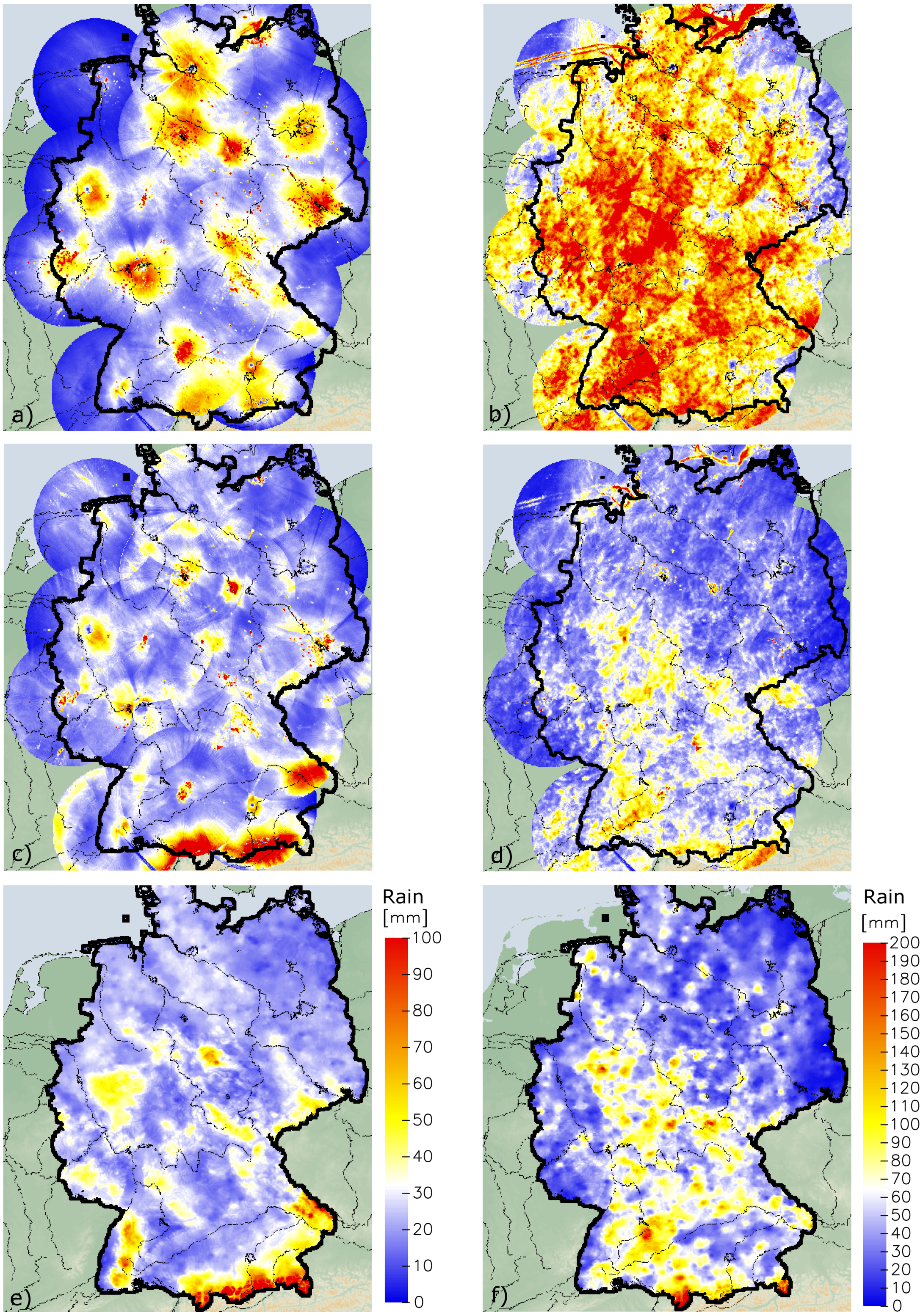

Figure 4.12: Monthly rain amounts for January 2006 (left) and July 2005 (right) for uncorrected (a, b) and corrected (c, d) radar composite data and REGNIE data (e, f). The entire correction scheme is performed on 5-minute values.

		summer		autumn		winter		spring	
		05	06	05	06	05/06	06/07	06	07
sp-t100	ruc	0.72	1.08	0.67	0.55	0.98	0.38	0.96	0.40
sp98	ruc	0.62	0.86	0.50	0.38	0.34	0.34	0.56	0.38
sp90	ruc	0.59	0.82	0.40	0.36	0.26	0.32	0.49	0.37
t98	ruc	0.46	0.69	0.43	0.40	**0.79**	0.26	**0.82**	0.20
sp-t98	ruc	0.31	0.39	0.20	0.19	0.23	0.21	0.34	0.19

		summer		autumn		winter		spring	
		05	06	05	06	05/06	06/07	06	07
sp-t100	rc	0.50	0.57	0.35	0.33	0.37	0.35	**7.48**	0.35
sp98	rc	0.49	0.56	0.32	0.32	0.23	0.34	0.48	0.35
sp90	rc	0.47	0.55	0.30	0.30	0.21	0.32	0.34	0.34
t98	rc	0.19	0.19	0.13	0.14	0.15	0.20	0.22	0.16
sp-t98	rc	0.19	0.18	0.12	0.13	0.14	0.19	0.20	0.16

		summer		autumn		winter		spring	
		05	06	05	06	05/06	06/07	06	07
sp-t100	diff	0.31	0.47	0.49	0.40	0.62	0.10	-6.78	0.12
sp98	diff	0.21	0.35	0.36	0.17	0.32	0.00	0.13	0.10
sp90	diff	0.20	0.34	0.25	0.18	0.18	0.00	0.31	0.10
t98	diff	0.58	0.73	0.70	0.66	0.81	0.25	0.74	0.19
sp-t98	**diff**	**0.39**	**0.53**	**0.38**	**0.31**	**0.39**	**0.10**	**0.41**	**0.16**

Table 4.3: RMSE for the comparison of radar and rain gauge data on an hourly basis for the eight seasons between June 2005 and May 2007. A separation is realized for different percentiles of space (sp) and time (t) and uncorrected radar data (upper table), corrected radar data (middle table) and their deviation in percent (lower table). $sp-t98$ means that the best and worst 1 percent of stations and time steps are neglected. Described values are printed in bold type.

in the lower table are the relative difference of RMSE values of upper and middle table.

The values in the columns imply that the influence of spatial percentiles is lower than those of the temporal ones. But it has to be taken into account, that the rain probability is approximately 10 % for hourly values. Neglecting 1 % of the higher RMSE values results in a reduction of about 10 % of time steps with rain amounts above 0 mm. The divergent behavior for uncorrected radar data in winter 2005/2006 and spring 2006 suggests that a systematic variation between uncorrected radar data and rain gauge data exists that affects certain pairs of values but it can be minimized by the presented correction scheme. The RMSE values change with percentiles, but the general behavior of the RMSE based on uncorrected and corrected radar data remains unchanged. Only for spring 2006 the huge RMSE of 7.48 for $sp100$ is significantly reduced for the other percentiles. The reason for this high RMSE value is a series of massively corrupted measurements of the Berlin radar on the 2nd of April 2006. The correction scheme is not able to eliminate this corrupted area but rather impairs this effect by transferring it to the adjacent overlapping area. For the residual RMSE values an improvement becomes obvious in all seasons. The highest RMSE values occur in summer even for corrected radar data. Although, radar measurement in the summer season is most reliable due to the low influence of the altitude dependence, the high rain amounts and high spatial heterogeneity seem to be superior and cause these higher RMSE values. Likewise, the RMSE values in spring 2006 are almost similar due to high rain amounts. Autumn and winter months with lower precipitation amounts show lower RMSE values. The $sp-t98$-values in the lower table show improvements between 31 % and 53 % until autumn 2006 for the presented correction scheme. In winter 2006/2007 and spring 2007 the improvements decrease below 20 %. The main reason for this behavior is the improved data quality of uncorrected radar measurements due to the new compositing algorithm.

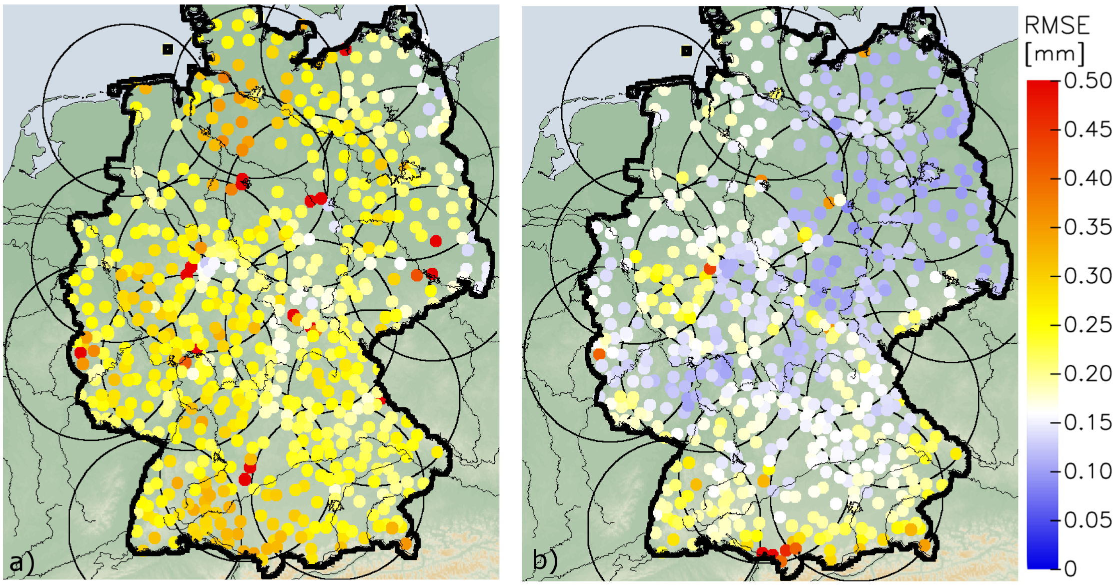

Figure 4.13: Mean spatially distributed RMSE values of hourly rain amounts of each rain gauge and corresponding uncorrected (a) and corrected radar measurement (b) from June 2005 to May 2007. The 98 % percentile of all time steps is used here.

A more in depth evaluation of spatially distributed RMSE values is given in Fig. 4.13 for the whole evaluation period June 2005 to May 2007. The mean RMSE of each rain gauge and corresponding uncorrected (a) and corrected radar measurement (b) for the 98 percentile is presented to identify spatial differences and anomalies. For the uncorrected radar measurements RMSE values between 0.2 and 0.4 are most common but no clear regional differences become apparent. Only those pairs of values near the Hamburg radar in the northern part of Germany and the Türkheim radar in the southern part show slightly higher RMSE values. Particularly noticeable are the high RMSE values of about 0.5 and more for pairs of values situated next to radar sites. Radar measurements are not very reliable there due to side lobes or clutter effects. The altitude dependence of radar measurements cannot be identified within these spatially distributed RMSE values. Even for a single season such as the winter months this dependence is not clearly reflected (not shown here). The reason is that the outer parts of most radar systems cover neighboring countries where rain gauge data are not available or they are overlapped by other radar systems. Additionally, regional differences in rain amounts also influence the spatial distribution of RMSE values. For the corrected radar measurements the RMSE values for almost all pairs of values are significantly reduced to values between 0.1 and 0.25. Those pairs of values near radar sites are also improved. Only at the northern fringe of the Alps the RMSE values remain unchanged or can even show slight deteriorations such as east of the Lake Constance. The main patterns of RMSE values for corrected radar data represent the mean annual distribution of rain amounts with high values near the Alps and near the low mountain ranges and low values in the eastern part of Germany.

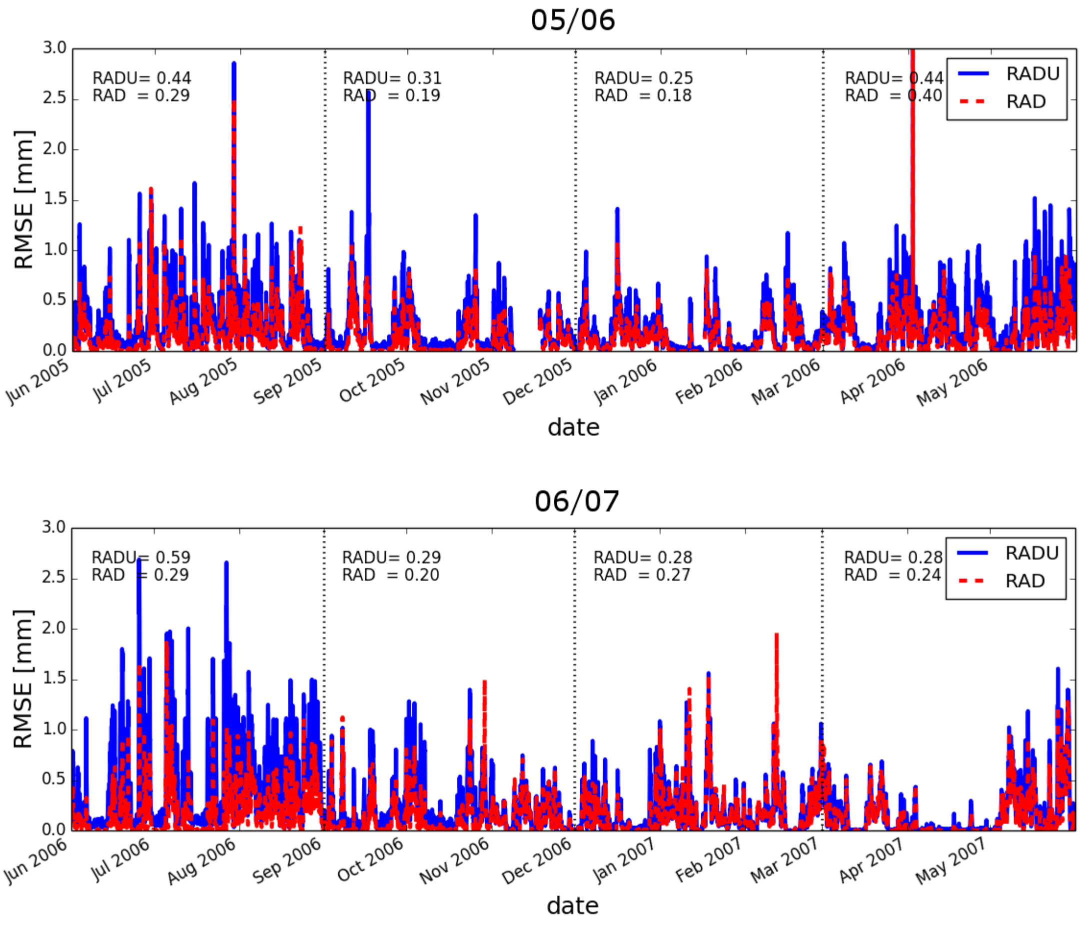

Figure 4.14: Temporal evolution of the mean RMSE values of hourly rain amounts for the comparison of all available pairs of rain gauge data and corresponding uncorrected radar data (blue) as well as corrected radar data (red) from June 2005 to May 2007. The 98 % percentile of stations is used here. The mean RMSE value for each season is shown for uncorrected (RADU) and corrected radar data (RAD).

Besides the spatially distributed RMSE values, the temporal distribution of the mean RMSE value for entire Germany is presented in Fig. 4.14. Therefore, the 98 % percentile of the RMSE values are chosen to reduce the influence of pairs of values such as those next to the radar sites with high RMSE values. The RMSE values of uncorrected radar data is plotted in blue color, overplotted by RMSE values of corrected radar data in red color for the year June 2005 to May 2006 (top) and June 2006 to May 2007 (bottom). In general, the mean RMSE values in different seasons support the findings in Tab. 4.3 with more and higher RMSE peaks in summer and spring than in autumn and winter. For the first year (Fig. 4.14top) an improvement of the RMSE values for nearly all time steps become apparent, both for high and low RMSE values. Even for spring 2006, only the very high RMSE values on the 2nd of April 2006 indicate a deterioration when applying the correction scheme. For the first half-year of the second year (Fig. 4.14bottom) a similar behavior of RMSE values becomes obvious. After that, the improvements by the correction schemes are getting smaller due to the modification of the compositing algorithm. In individual cases, especially in autumn and in winter, short term deteriorations become apparent. Nevertheless, an improvement of RMSE values on an hourly time scale results when using the correction schemes even after 2006. The differences of the RMSE values of uncorrected and corrected radar data are even higher when using all stations and time steps (100 % percentile; not shown here).

Figure 4.15 shows Quantile-Quantile-plots (QQ-plots) for each season to identify systematic differences in the distribution of rain amounts. The best and the worst 1 % of pairs of values are removed (98 % percentile). The blue color marks the uncorrected radar measurements compared to rain gauge data and the red color marks the corrected ones. Here, the relative differences are important and not the absolute values, so the scales vary for all plots. In summer (see Fig. 4.15a, e), rain structure for higher rain rates is well represented for uncorrected radar data, but for intensities below 10 mm/h a systematic overestimation becomes apparent. For corrected radar data, the accordance for rain intensities below 5 mm/h is very good, whereas for higher intensities an underestimation results. In order to understand these differences, the correction schemes have to be regarded: The altitude correction and the bias adjustment show the main impact and affect all pixels. For the summer months, the altitude correction is negligible and the rain amounts are overestimated by the uncorrected radar dataset (see Fig. 4.12b). Consequently, the bias adjustment has to reduce these values and is therefore responsible for the main differences. As the rain distribution for high rain rates is well represented for uncorrected radar data and the bias correction is performed with a correction factor where all rain intensity values are equally scaled, an underestimation for high rain rates results for corrected radar data (see Fig. 4.15a, e). The overestimation of total rain amounts in summer by uncorrected radar data is mainly caused by an overestimation of smaller rain intensities and a lower dry hour probability (0.84 (0.83) vs. 0.91 (0.92)). For the autumn months (see Fig. 4.15b, f), a similar tendency can be observed but less distinct. For uncorrected radar data a tendency of overestimation for low rain rates is still shown. In the winter months (see Fig. 4.15c, g), an overestimation for corrected radar data and higher rain amounts results, whereas an overestimation for uncorrected radar data is observed for the first year and an underestimation for the second year for these rain rates. Most rainfall is

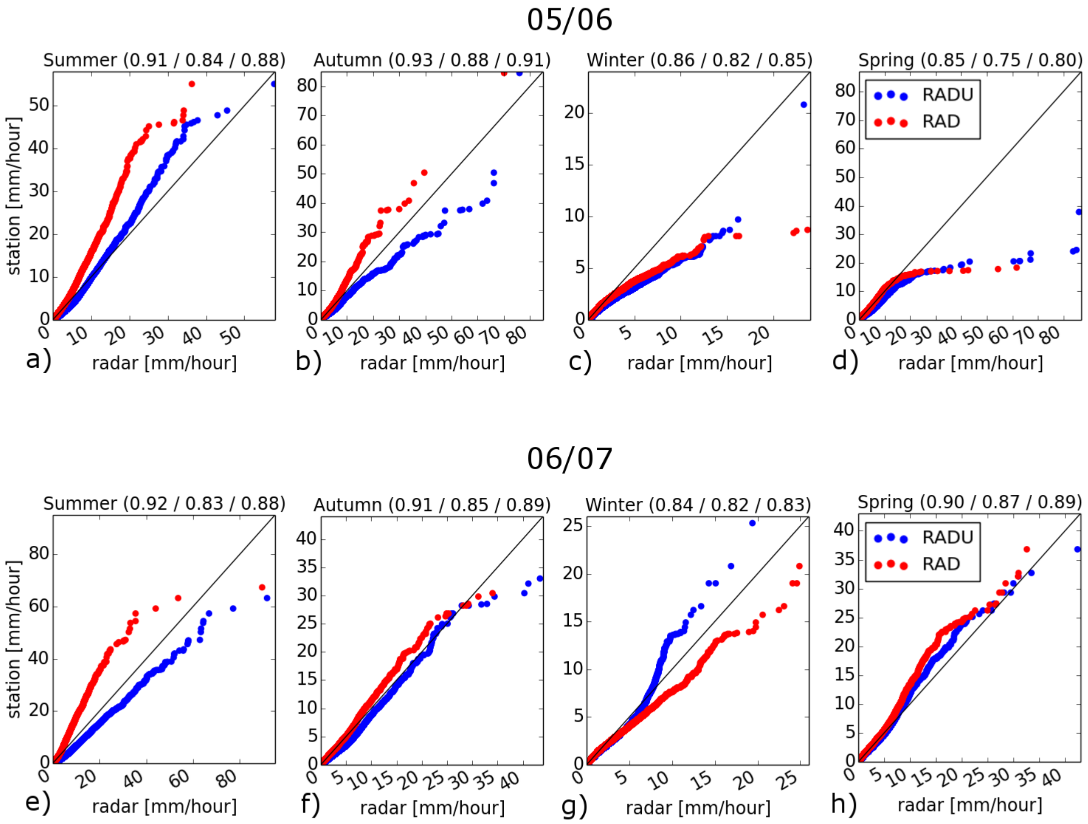

Figure 4.15: Quantile-Quantile-plots of hourly rain amounts for the comparison of rain gauge data and uncorrected radar data (blue) as well as corrected radar data (red) for each season between June 2005 and May 2007. The dry hour probabilities are added in the title of each plot in brackets (rain gauges / uncorrected radar data / corrected radar data). The 98 % percentile of stations is used.

produced by rain intensities below 3 or 5 mm/h, where both uncorrected and corrected data reveals a good accordance with rain gauge rain structure. In spring (see Fig. 4.15d, h), the overestimation of small rain intensities by uncorrected radar data in 2006 nearly vanishes for 2007 due to the new compositing algorithm. Corrected radar data provides measurements with a very good accordance of rain structure compared to rain gauge data for rain intensities below 10 mm/h data for both years. In summary, the old compositing algorithm (*push*-procedure) leads to an overestimation of radar measurements especially for lower rain intensities whereas for higher intensities the accordance is reasonable. As a consequence, the uniform adjustment to REGNIE data within the correction scheme may lead to an underestimation of high rain amounts especially in the summer months. Most rainfall occurs there and the major portion is produced by small and medium rain intensities such as below 5 or 10 mm/h. In this range, the accordance of the distribution of rain amounts of corrected radar data and rain gauge data is very good for all seasons.

5-Minute rain amounts

The evaluation of 5-minute data should give a brief overview of data quality and general suitability of the correction scheme. The temporal distribution of rain amounts based on radar data is mainly prescribed by uncorrected radar data values. Thus, no major deviations from the analysis of hourly data can be expected. The temporal distribution of the mean RMSE of 33 pairs of values of rain gauges and corresponding radar pixels including the mean seasonal RMSE in Fig. 4.16 gives an overview of the performance of data quality based on uncorrected and corrected radar data in Lower Saxony. The rain structure for 5-minute measurements is analyzed based on QQ-Plots in Fig. 4.17.

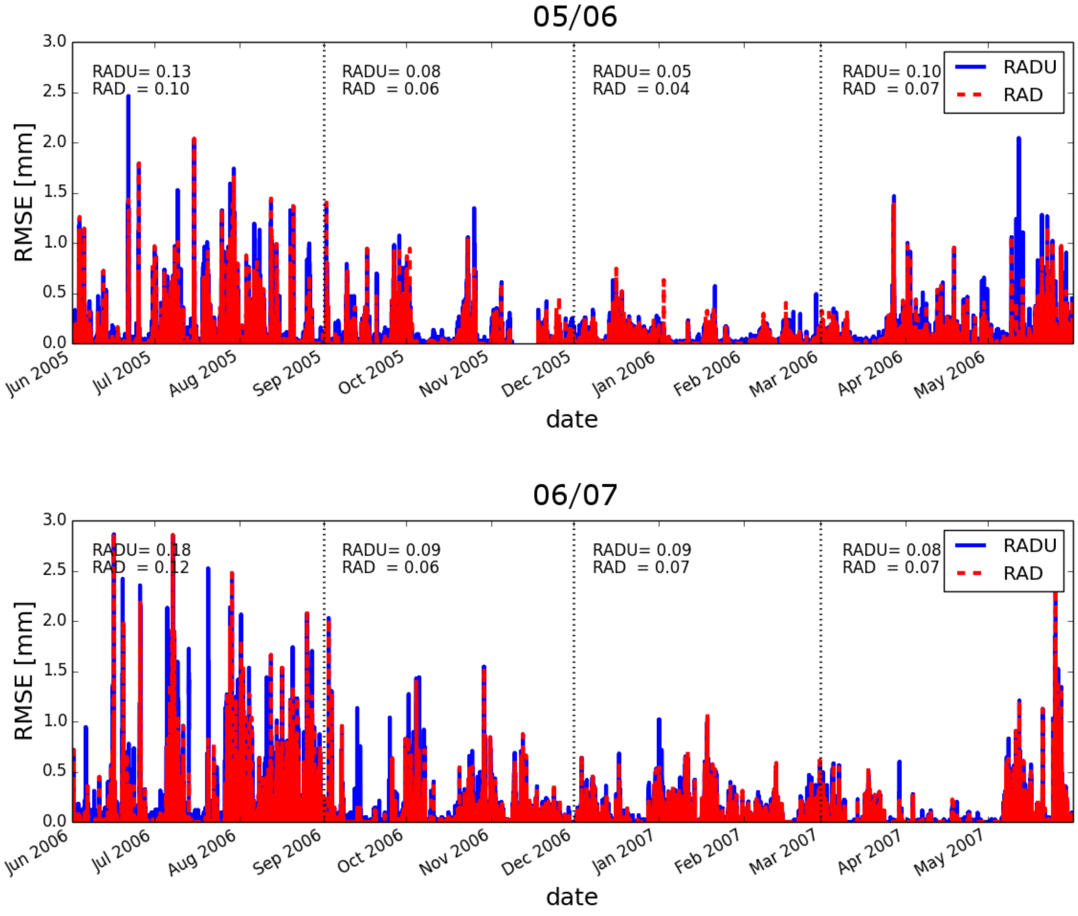

Figure 4.16: Temporal evolution of the mean RMSE values of 5 min rain amounts for the comparison of all available pairs of rain gauge data in Lower Saxony and corresponding uncorrected radar data (blue) as well as corrected radar data (red) between June 2005 and May 2007. The 98 % percentile of stations is used here. The mean RMSE value for each season is shown for uncorrected (RADU) and corrected radar data (RAD).

Figure 4.16 is analogous to Fig. 4.14 (hourly rain amounts) showing the RMSE values of uncorrected radar data in blue color, overplotted by RMSE values of corrected radar data for the 98 % percentile. The results are very similar to those of hourly data for the whole of

Germany. The highest RMSE values are calculated for the summer months, whereas for the rest of the year only minor seasonal differences become apparent. The seasonal RMSE for uncorrected radar data from autumn to spring is about 0.08. For summer, the RMSE is about twice as high. The correction scheme reduces the seasonal RMSE between 20 % and 30 % until autumn 2006. Afterwards, the RMSE is reduced between 15 % and 17 % due to the new composting algorithm. For all seasons and years, the RMSE values are reduced by applying the correction scheme. Deteriorations due to the correction scheme are only present in individual cases despite the massive altitude correction factors applied in the winter months.

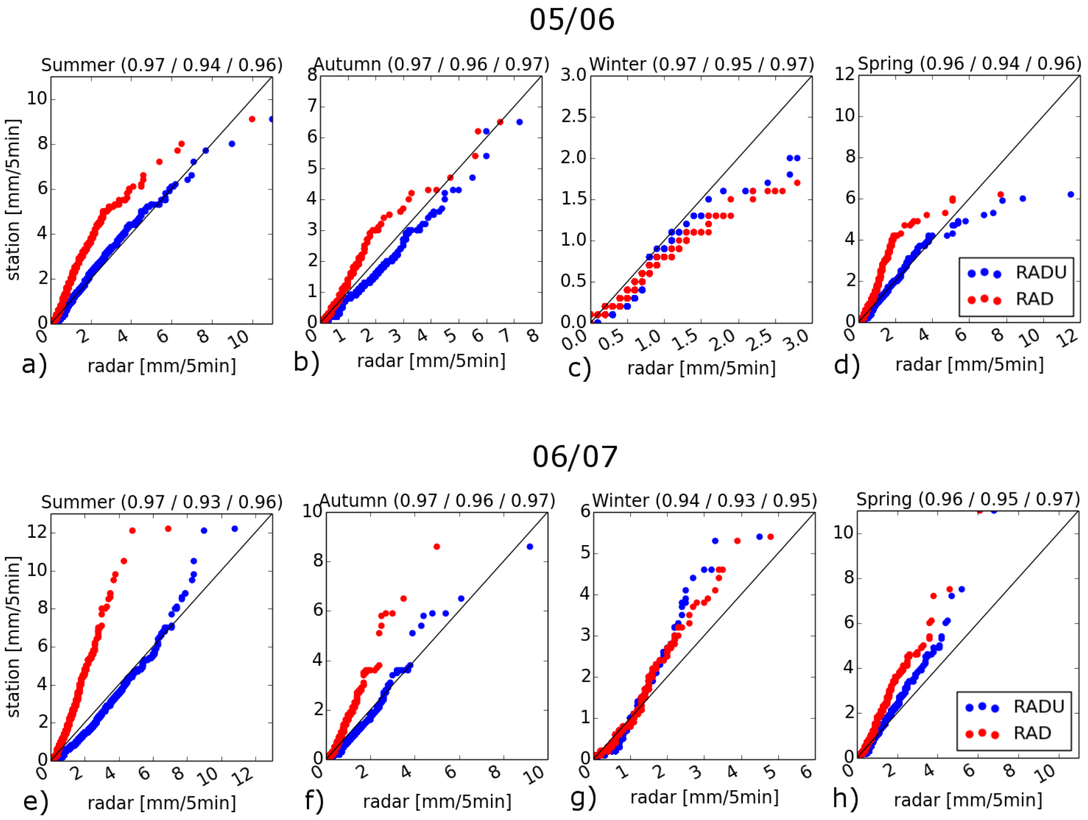

Figure 4.17: Quantile-Quantile-plots of 5-minute rain amounts for the comparison of rain gauge data and uncorrected radar data (blue) as well as corrected radar data (red) for each season between June 2005 and May 2007 in Lower Saxony. The dry 5-minute probabilities are added in the title of each plot in brackets (rain gauges / uncorrected radar data / corrected radar data). The 98 % percentile of stations is used.

For the preparation of the QQ-Plot the radar data values are converted to the same resolution of 0.1 mm as the rain gauges provide. The eight images revealing the different seasons of two years are similar to Fig. 4.15 but show 5-minute data of the 33 stations in Lower Saxony. The distribution of rain amounts of uncorrected radar values and rain gauges above 1 mm match very well. Slightly lower maximum values are observed for the period of the new compositing algorithm starting end of 2006. The values below 1 mm however show the same behavior as the hourly values for low rain intensities do and overrate the measurements of rain gauges. As a consequence, the overestimation of total rain amounts based on uncorrected radar data results from this huge amount of low intensity measurements. The strength of corrected radar data lies especially in this range of low intensity values with very good accordance compared to rain gauge data. Only for the winter months (see Fig 4.15c, g) an overestimation of rain gauge precipitation amounts for these low rain intensities becomes apparent. But higher rain intensities are rather underestimated. The mean rain probability for 5-minute rain gauge data and corrected radar data is about 3 %, whereas the probability of uncorrected radar data is significantly higher. But this is at least partly attributable to the rain gauges' resolution of 0.1 mm.

4.4 Discussion

The presented simple but robust post-correction scheme for single radar data improves data quality significantly. This successful transfer of analysis results to the correction scheme shows that the identified effects and their dependences on signal power and time are interpreted correctly. It is shown that systematic effects affect data quality essentially.

The extension of the correction scheme to composite radar data including the transfer of analyses' results of composite data into the correction scheme is successfully implemented on an annual scale. The main limitations are satisfactorily being corrected, except the impact of the *push*-procedure. So the data quality is improved.

At this stage, the altitude correction may not be applied at shorter temporal resolutions than one year. The parts of the correction algorithm concerning clutter and spokes and the adjustment are in principle also suitable for smaller time scales or even for single radar image corrections. Although only a mean correction is realized where the current meteorological conditions are not considered, the extent of the statistical correction depends on the previous correction procedures. The better the single radar images are corrected in advance, the fewer statistical corrections are necessary afterwards. But relying only on corrections of individual radar images usually results in recognizable errors in radar climatology.

However, the applicability of an annually based correction scheme is obviously limited. The correction modules are then modified to be applicable to shorter time steps to demonstrate the limits of such a statistical correction and to point out the significance of such systematic effects and also its variations. The modified version applied to 5-minute radar data and has to capture a much higher variability and heterogeneity of precipitation patterns. This applies particularly to the altitude correction, which primarily impacts radar patterns. The altitude correction on a daily basis is fairly robust for radar images which are filled with uniform rain amounts. Very heterogeneous rain patterns and only partial coverage with rain echoes is challenging. The adjustment of single areas and overlapping areas at their boundaries is error-prone with very heterogeneous distributions of precipitation echoes. Additionally, the correction factors are calculated on a daily scale. Any intra-daily variations are not captured. These are examples where the correction scheme may approach its limits and overcorrections may follow. But the correction algorithm offers mechanisms to reduce those effects such as the adjustment to daily rainfall data. Therefore, these effects are rarely observed.

The downscaling of statistical correction factors is possible as long as the natural variability can also be reproduced. A variation of the altitude correction factors based on the current data is desirable but tests with heterogeneous precipitation patterns lead to highly variable correction factors in space. This is why currently the altitude correction factors are not modified. Only the impact of inappropriate correction factors is changed by reducing the adjustment factors. A pre-condition for reliable results based on this correction scheme is an almost complete coverage of the regarded areas with rain patterns with a magnitude of 5 mm or above at best. The accumulation time to achieve these conditions may vary significantly. The use of a linear altitude dependence even at short temporal scales may be subordinated compared to bright band effects that cannot be corrected by this approach.

Another aspect that influences the quality of simulation results is the degree of limitations of uncorrected radar data. The problems of overshooting and partial beam filling especially in the colder months are already mentioned. This leads to the question whether the maximum correction of eight times the original calculated precipitation value produces reasonable corrections. If range-bins at lower altitudes detect rain most of the winter time and higher altitudes detect snow, this factor seems reasonable. Also the results of the 5-minutes analysis in Lower Saxony[4] are reasonable. But for higher altitudes such as at far ranges of the Munich weather radar it is unlikely that each rain event is detected. So, the rain amounts of the residual events have to be increased to equalize this effect. An indication for such an effect is assumed in radar data in winter due to the overestimations of corrected hourly radar data for higher rain amounts (see 4.15). These overestimations seem to be necessary to achieve the daily values of REGNIE. If it is not possible or not reasonable to achieve these daily values an underestimation results which is observed according to Fig. 4.10. As a consequence, radar data in winter can significantly be corrupted so that a correction is not able to generate reliable precipitation patterns for certain areas.

As the algorithm is partly based on static correction coefficients or static spatial maps, the question about the impact of changing measuring settings may arise. The correction coefficients of the altitude correction are insensitive to a modification of the scan strategy as long as the beam elevations remain in almost the same range. Such a modification of the scan strategy can influence the static map of the allocation of pixels to a single radar system. But a readjustment can easily be done. In 2010, the maximum evaluated range of single radar measurements was extended from 128 km to 150 km. This modification would require elaborated adjustments: The size of the overlapping areas as well as the total amount of overlapping areas change and therefore have to be re-analyzed.

Despite the presented limitations in radar data, the advantage of measurements compared to simulations regarding precipitation patterns, is the detection of precipitation at several places or even spatially distributed. Simulations of Regional Climate Models (RCMs) have to produce these patterns as a result of humidity excess and only use boundary conditions (e.g. of reanalysis data) for the investigation area (domain). Therefore, the generation of precipitation

---

[4] Rain gauges with a 5-minute resolution have only been available in Lower Saxony.

patterns by RCMs seems to be more challenging than based on measurements, not least because of the close interaction between precipitation, land use and also other meteorological parameters such as radiation. State-of-the-art RCMs are able to provide similar resolutions in space and time as radar data do. In the next chapter, the added value of such high resolution simulations in Germany will be evaluated.

5 Adaptation of the WRF model

5.1 Model description

The Weather Research and Forecasting Model (WRF) is a state-of-the-art mesoscale numerical weather prediction system that is applicable to atmospheric research as well as to operational forecasting. It comprises two dynamical solvers called NMM (Nonhydrostatic Mesoscale Model) and ARW (Advanced Research WRF). WRF-NMM focuses on numerical weather prediction. The WRF-ARW involves a variety of possible applications such as for idealized atmospheric conditions, regional modeling with real data (observations, analysis), air chemistry (WRF-CHEM), hurricanes (HWRF), hydrological modeling (WRF-HYDRO) and modeling on a global scale (Global-WRF). The selection of possible meteorological applications spans a range from meters to thousands of kilometers. The complete WRF model system additionally provides the WRF Preprocessing System (WPS), the WRF Data Assimilation (WRF-DA) system for 3-D and 4-D variables, and further post-processing and visualization tools. An overview is given in the ARWUsersGuide [Wang et al., 2015].

The WRF model is fully compressible, nonhydrostatic and mass conserving with a terrain-following hydrostatic pressure coordinate in vertical direction. 2nd to 6th order advection schemes are used horizontally and vertically within the model. The Runge-Kutta 2nd and 3rd is applied for the time integration schemes. A time-splitting technique separates the fast processes related to acoustic and gravity waves from slower motions in the prognostic equation. WRF allows nesting of model domains with sub-kilometer grids with the help from downscaling techniques to pass information from processes of larger scales to finer ones (one-way-nesting) and back (two-way-nesting). For more details see Skamarock et al. [2008].

In this work, WRF-ARW is applied as a regional model (RCM) in real mode to downscale global atmospheric reanalysis data. The model requires two- and three-dimensional fields that describe the initial atmospheric state and land surface characteristics. Lateral and lower boundary values that vary with time are adopted from the global fields. Static values such as land use or soil properties are stored as fields and tables and have to be pre-processed in advance. Certain physical driver modules are used to account for physical processes that modify the current atmospheric state. In WRF, various schemes exist that are clustered in the following compartments: shortwave and longwave radiation physics, microphysics, the representation of the Planetary Boundary Layer (PBL) and the cumulus parametrization. The interaction between different land use types and the lower atmosphere is realized within Land Surface

Models (LSM). For instance the NOAH-LSM, calculates relevant physical processes for each grid cell and provides surface sensible and latent heat fluxes and surface skin temperatures representing lower boundary conditions [Chen and Dudhia, 2001, Ek et al., 2003]. It accounts for soil temperature and moisture, soil drainage, snow coverage and runoff, but only vertical fluxes are assumed.

5.2 WRF setup and study area

The setup retrieval for WRF simulations including domain size, nesting strategies, parametrizations and physical driver modules is time-consuming, but also the best possible guarantee for reliable simulation results. The first part of the WRF setup is the definition of the domains for which simulations are performed as well as the nesting strategy. Both aspects significantly influence simulation results along with the location of the domain. The second part is the selection of modules and the setting of parameters that control the simulations. Usually both processes are iterative.

Domain size and nesting strategy

The domain size widely determines computation time even though parallelization of calculation steps within WRF is most efficient for huge grid sizes. The aim is to limit grid sizes, especially for small areas of interest. According to de Ela et al. [2002] and Vannitsem and Chomé [2005] small grid sizes only degrade large-scale features and their variability. At best, they downscale coarse-grid results to finer resolutions. But self-development of e.g. convection processes can hardly be expected, there. Brisson et al. [2015] emphasize the significance of the domain size to allow the spatial spin-up of processes such as convection over the graupel-phase. However, large domains can cause undesired effects at the boundaries due to differences of RCM simulations and lateral boundary conditions from the input data or the coarse grid [Leduc and Laprise, 2009, von Storch, 2005]. These differences may arise due to the fact that different physics in RCM and driving data such as cloud water content or Reynolds number of flow exist [Prein et al., 2015].

Prein et al. [2013a] show that there are only minor differences between one-way-nesting and two-way-nesting. Similar results from simulations based on the same setup with and without finer resolutions can only be achieved, choosing one-way nesting[1]. Prein et al. [2015] gives an overview of utilized nesting steps to downscale from General Circulation Model (GCM) or reanalysis data to high resolution simulations at km-scales. Small steps (1:3 or 1:5) may guarantee that the underlying physics remain almost the same. Larger steps save computation time and 'error-prone' grid resolutions can be avoided. Brisson et al. [2015] show improvements of simulations at 3 km grid resolution avoiding an intermediate 7 km grid resolution which is in the 'gray zone' of convection parametrization (see 'selection of modules and parametrization').

---

[1] There is no feedback from fine resolution to the coarser one.

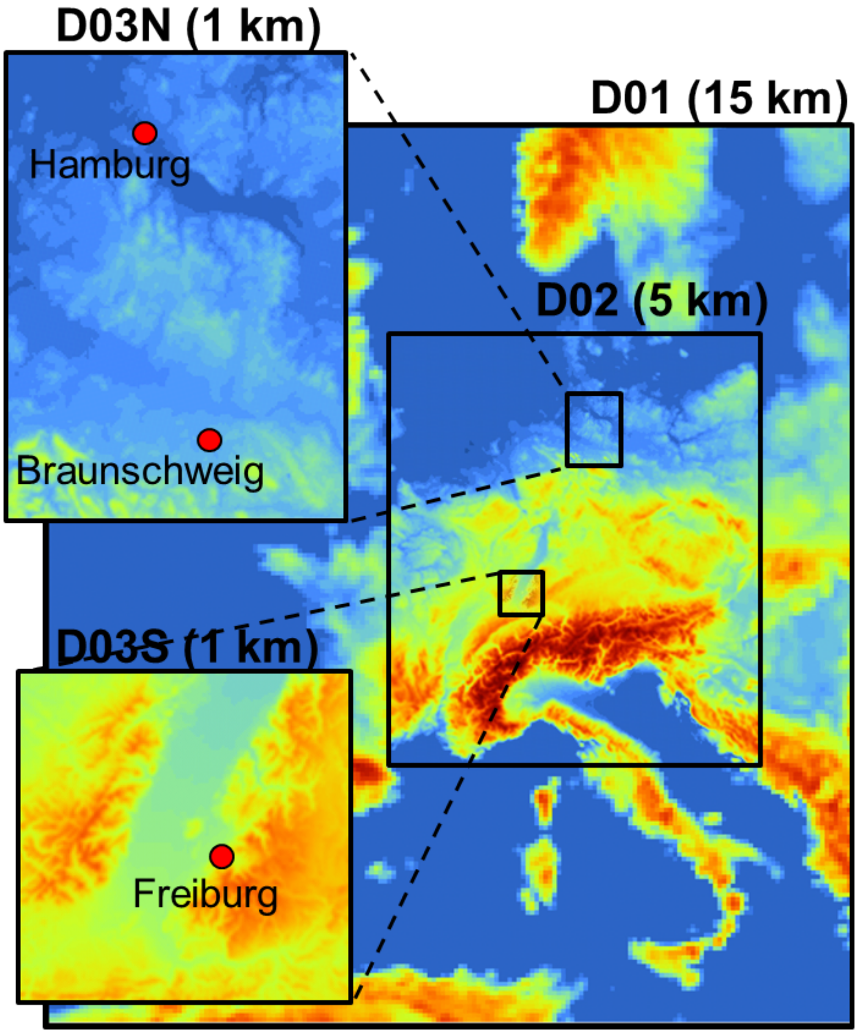

Figure 5.1: Overview of the 4 WRF domains to downscale ERA-Interim data. *D*01 is based on a grid size of 15 km, *D*02 on a grid size of 5 km and *D*03*N* and *D*03*S* show a spatial resolution of 1 km.

Domain setup for the study area

As already mentioned in Chapter 1, the main focus of the analysis within the SYNOPSE project are the cities of Hamburg and Braunschweig in the northern part of Germany and Freiburg in the southern part with a desired grid resolution of 1 km. Therefore, one domain is defined in the northern part of Germany including Hamburg and Braunschweig with 160 x 220 pixels, called *D*03*N* in the following (see Fig. 5.1). The area is situated in the North German Plain, bordered by the low-mountain-range Harz in the south. The location of this domain has been determined with respect to limit boundary effects for both cities and to provide an area where spatial spin-up is enabled. Additionally, borders do not cut through any mountain ranges or areas that are known source regions of important phenomena for simulation results [Rummukainen, 2010, Warner et al., 1997]. The same aspects are considered for the southern domain with Freiburg in the center, called $D03S$. The domain has 150 x 150 pixels and includes the Vosges in the western part, the upper Rhine valley in the middle and the Black Forest in the eastern part. The upper Rhine valley is one of the warmest regions in Germany with a high tendency of convection during summer. In general, the precipitation amount in this southern domain is higher and more dominated by orographic effects and heat convection than the northern domain that shows a more maritime regime.

ERA-Interim reanalysis data [Dee et al., 2011] serves as input for the presented RCM simulations every 6 hours with a resolution of about 80 km. For downscaling 80 km drive data to a fine grid of 1 km resolution, two intermediate steps are used. The 15 km domain (180 x 201 pixels) is used to downscale the large scale patterns of the driving model to the WRF RCM (called $D01$). The 5 km grid size domain is called D02 (240 x 279 pixels). It already provides high resolution simulation results for much larger areas than the two 1 km domains. These small nesting step ratios (5.3-3-5) are used to minimize possible variations at the boundaries between different domains due to scaling effects. Additionally, the same distribution of 42 eta levels[2] are applied for each domain. An overview of the four domains is given in Fig. 5.1. The altitude for each pixel based on the respective grid size is presented.

Selection of modules and parametrizations

The possible variations of settings within WRF are vast and a comprehensive testing is difficult. A very thorough engagement with the principle options, the meaning and especially the limitations of different modules within WRF as well as the study of experiences of other groups with WRF in similar regions or with similar issues help to reduce eligible setup options. Nevertheless, intensive testing of different alternatives is indispensable. The testing criteria to achieve an appropriate WRF setup have mainly been the good performance of the precipitation output and additionally the temperature at 2 m. Mean annual or seasonal deviations of WRF results of the 15 km or 5 km domain compared to REGNIE data [REGNIE, 2017] (precipitation) and E-OBS data [Haylock et al., 2008] (temperature) lead to an initial selection of WRF scheme combinations. Higher resolution comparisons (1 km) also with station data regarding the temporal performance, diurnal cycle or spatial patterns for shorter periods such as days or weeks help to fine-tune the selection. Undesired feedback often occurs when changing a scheme. Testing results are not presented within this framework. Some interesting findings of the testing procedure are briefly described at the end of this Chapter. In the following, the most important settings and also their limitations are discussed.

To further reduce any differences among large scales, spectral nudging [von Storch et al., 2000] is applied for the coarse 15 km domain. This downscaling technique constrains the large atmospheric conditions within a RCM to the global model; as a consequence the propagation of large scale atmospheric patterns into the RCM is better realized. Temperature, horizontal wind, humidity and geopotential height can be 'nudged'. Feser [2006] and Meinke et al. [2006] show improvements for near-surface temperature and precipitation in large scale models[3] on account of this technique. Regarding its application on finer grids, not much is known so

---

[2] Vertical levels.

[3] Models with grid sizes larger than 10 km.

far [Prein et al., 2015]. The goal of this investigation is to adopt the large scale patterns from ERA-Interim, while the fine scale patterns shall be a own development of WRF.

Concerning the radiation modules, six longwave schemes and eight shortwave schemes are available. The RRTM scheme (longwave radiation) and the Dudhia scheme (shortwave radiation) are selected after testing a variety of options.

The microphysics scheme has to be suitable for high-resolution simulations. The WRF Single-Moment 6-class scheme (WSM6) is chosen after testing. It includes ice, snow and graupel processes that are important for reliable simulation of convection [Brisson et al., 2015].

The Planetary Boundary Layer (PBL) scheme is responsible for the parametrization of turbulent fluxes (eddies) which are regarded as one of the key challenges for high resolution simulations [Prein et al., 2015]. The largest eddies can reach the kilometer-scale (1 to 2 km). Parametrizations are designed for grid sizes that are significantly larger (at least 4 or 5 km) than the largest eddies. An explicit large-eddy-simulation (LES) is usually below 100 m. The gap between these scales is the gray zone of the PBL, also called the 'terra incognita' [Wyngaard, 2004]. The majority of turbulent energy remains unresolved at these scales, but the partly resolving of large eddies may lead to undesired effects. Langhans et al. [2012] show rain amounts that increase with grid scale between 0.5 km and 4.4 km. Zhou et al. [2014] call such an effect 'grid dependent convection' and explain it with a higher heat flux in the gray zone than in LES. Ching et al. [2014] analyze the principle suitabilities of PBL schemes for this gray zone. Nonlocal closure schemes such as the Yonsei University scheme (YSU) show weaker influences of large eddies, as the potential temperature profile is kept more neutral, than the Turbulent Kinetic Energy (TKE) schemes such as Mellor–Yamada–Janjić (MYJ) [Janjic, 1994] and Bougeault–Lacarrère (BouLac) [Bougeault and Lacarrère, 1989]. But it is more desirable to apply parametrizations that consider the effects of those large eddies rather than to suppress them.

A very error-prone subgrid process is the parametrization of deep convection [Hohenegger et al., 2008]. Assumptions that are necessary to consider the interacting processes occurring between micro scale and synoptic scales are rather crude [de Rooy et al., 2013]. Additionally, convection parametrization schemes interplay with microphysics, PBL and radiation schemes and may cause nonlinearities [Prein et al., 2015]. As a result, the convection parametrization may lead to shortcomings. The convection parametrizations are designed for grid cells larger than about 10 km, whereas an explicit resolution of convective processes is feasible below 4 km. The gray zone in between is often circumvented [Prein et al., 2015]. The deterioration of 3 km simulation results by an intermediate 7 km grid [Brisson et al., 2015] is probably caused by undesired effects in the gray zone of convection parametrization. To overcome this problem, new schemes such as the Grell-Freitas (GF) scheme are developed and implemented in WRF. This scheme tries to smooth the transition to cloud-resolving scales [Grell and Freitas, 2014] and it is used for the simulations of the 5 km and 15 km domain. For simulations with 1 km grid size cumulus parametrization is turned off.

Table 5.1 summarizes the applied schemes of the most important compartments of the final WRF setup.

Compartment	Scheme
Shortwave radiation	Dudhia
Longwave radiation	RRTM
Planetary Boundary Layer	YSU
Microphysics	WSM6
Surface layer	MM5 similarity
Land surface	NOAH-LSM
Cumulus Parametrization	GF

Table 5.1: Chosen WRF-ARW setup for the regional simulations.

Temporal spin up

WRF uses input data from reanalysis or GCMs to describe the initial state of the atmosphere. For the ground surface, table values are additionally applied that often significantly varies from true conditions. Quickly changing conditions such as atmospheric patterns are unproblematic in this respect but soil properties such as moisture content or temperature, often need several weeks or months to spin up. To exclude undesired effects due to inadequate soil conditions a temporal spin up is used even though these spin-up simulations are time consuming. The higher the resolution, the more problematic a long spin-up becomes. Here, a temporal spin-up of four months (January to April 2005) is performed for the coarsest domain $D01$. Then the other domains are nested and take over some surface fields such as skin temperature, soil moisture and temperature. This is realized by the option *fine_input_stream*. Then, another one month temporal spin-up is performed to achieve reliable domain specific conditions.

5.3 Urban modeling with WRF

With finer scales bio-geophysical features such as urban areas become more important. The shape of a city influences momentum, turbulence and thermal fluxes between the surface and the atmosphere. To explicitly resolve the processes occurring in the very heterogenic urban environment resolutions in the lower meter-scale both vertically and horizontally would be required. Two ways to parametrize the major processes within urban areas seem feasible. The first one is to modify the parameters in the surface heat balance equation such as albedo, roughness, thermal conductivity or soil moisture (bulk parametrization). The second one is to link an Urban Canopy Model (UCM) to an atmospheric model to refine the simulation of those processes that influence the urban climate most [Taha, 1999].

Bulk parametrization

The NOAH-LSM already contains a bulk parametrization of urban land use [Chen et al., 2011, Liu et al., 2006, Tewari et al., 2004]: the roughness length is increased from 0.5 m to 0.8 m to account for turbulence effects produced by buildings and the resulting drag, the surface albedo is reduced from 0.18 to 0.15, a larger volumetric heat capacity value (3.0 $Jm^{-3}K^{-1}$) for urban surfaces (walls, roofs and roads) is assigned, the soil thermal conductivity is increased to represent the heat storage of urban building materials (3.24 $Wm^{-1}K^{-1}$) and the green vegetation fraction is reduced to limit evaporation. According to the results of Liu et al. [2006], this simple representation of urban climate is able to mimic the bulk effects of the boundary layer in urban areas. Urban features cannot be represented below the grid size.

Urban Canopy Models

For higher complexities, Urban Canopy Models (UCM) have to be applied to consider the urban geometry in its surface energy budgets and calculations of wind shear. Three different urban classes are distinguished within WRF using USGS classification. Class 31 represents areas with built-up structures and shares of 20 to 70 % of vegetation ('low-intensity residential'). The percentages of vegetation decreases below 20 % for 'high-intensity residential' in class 32. Residual highly developed areas and infrastructures are referred to as 'industrial/commercial' (class 33) [European Environment Agency, 2007]. For each class a parameter set of vegetation parameters and land use properties has to be defined.

Two UCMs with different degrees of complexity that are linked to the LSM within WRF-ARW are tested: the Single-Layer Urban Canopy Model (SLUCM) and the Multi-Layer Urban Canopy Models BEP (Building Environment Parametrization). Figure 5.2 illustrates both models. The coupling between LSM and UCM is realized through the parameter urban fraction (F_{urb}) representing the portion of impervious surfaces in WRF subgrid scale. Temperature and surface fluxes for the part of natural (vegetated) surfaces are provided by the LSM; those for the anthropogenically influenced areas are calculated by the UCM and are subsequently combined [Chen et al., 2011].

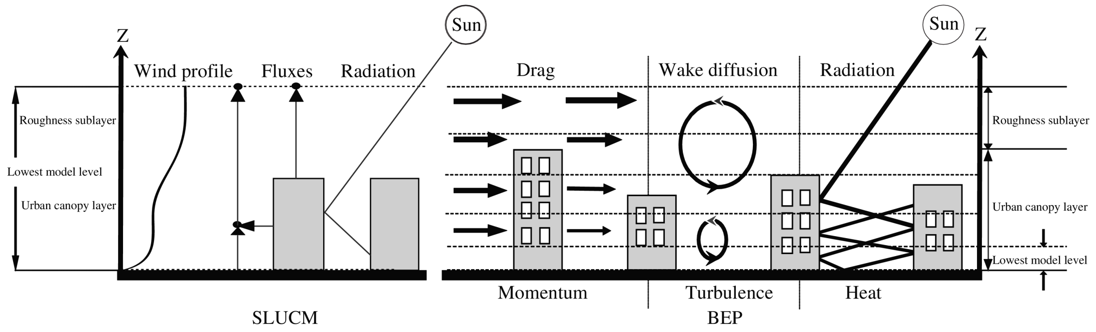

Figure 5.2: A schematic illustration of the SLUCM (left) and the multi-layer BEP models (right). Adapted from Chen et al. [2011].

Single-layer Urban Canopy Model (SLUCM)

The SLUCM [Kusaka and Kimura, 2004, Kusaka et al., 2001] uses infinitely-long two-dimensional parametrized street canyons to characterize the urban geometry. Reflections, shadowing and trapping of radiation is taken into account in each canyon and an exponential wind profile is derived (see Fig. 5.2left). The model accounts for the three-dimensional structure of urban surfaces including diurnal variation of solar azimuth angle and canyon orientation. Temperature profiles within roof, wall, and road layers are calculated based on the thermal conduction equation; surface skin temperatures at the roof, wall, and roads are computed using the surface energy equation. The surface-sensible heat fluxes are derived on the basis of the Monin-Obukhov similarity theory and the Jurges formula. A similarity stability function for momentum is applied to calculate the canyon drag coefficient and friction velocity. The WRF boundary-layer schemes receive the total friction velocity from urban and natural surfaces. Anthropogenic heating and its diurnal cycle enhances the sensible heat flux of the urban canopy layer. SLUCM has 20 parameters characterizing the surface parameters and canyon dimensions of the respective study area [Chen et al., 2011]. The canopy model is driven by the present state of the boundary condition within WRF. The results of the canopy model gained by variables and constants that characterize urban surfaces are passed back to update the current state of the atmospheric model. For a further differentiation of urban areas, the above mentioned USGS land use classes 31 to 33 are used.

Multi-layer Urban Canopy Models (BEP)

In contrast to the SLUCM, the BEP directly interacts with the PBL [Martilli et al., 2002]. Buildings and urban surfaces produce sources and sinks of moisture, momentum and heat, both horizontally (streets and roofs) and vertically (walls) and affect the whole urban canopy layer (see Fig. 5.2right). The three-dimensional approach of the BEP model accounts for changes in the lower part of the urban boundary layer such as the thermodynamic nature of the urban roughness. Effects on Turbulent Kinetic Energy (TKE), momentum and potential temperature are considered as well as radiation effects of roads and walls through trapping of radiation in street canyons, shadowing and reflections. The application of BEP requires a high vertical resolution near the ground to receive more than one level within the urban canopy. Currently, only the BouLac and the MYJ turbulence schemes provide the required source term in the TKE equation and the modified turbulent length scales due to the presence of buildings and other urban surfaces. BEP is able to simulate the urban heat island effect and the elevated inversion layer above the city in principle but at the expense of high computation rates [Chen et al., 2011]. The same three urban classes that are used for the SLUCM are utilized here to reconstruct the features of each city. For each class, a certain canyon width is set with a distribution of different building heights with a certain probability. The horizontal grid size and the length of the street canyon are equal. The grid of the mesoscale model and the grid of the urban structure vary to achieve the utmost flexibility to consider sub-scale processes [Fallmann, 2014].

Further information about the Urban Canopy Models linked to WRF can be found in Chen et al. [2011]. A more detailed overview for SLUCM and BEP is given by Fallmann [2014].

Application of Urban Canopy Models

Before applying UCMs in WRF some additional work is necessary to enable the coupling of the UCM with the NOAH-LSM.

The standard USGS land use data with 24 classes is commonly used in WRF. This dataset only includes one urban class which would limit the full capacity of UCMs. Therefore, the CORINE dataset [European Environment Agency, 2007] is used to refine the distribution of urban pixels into three classes. The reclassification used here is based on Fallmann [2014] and Pineda et al. [2004]. There is no defined procedure available on how to achieve this aim. Fallmann [2014] has developed a scheme based on ArcGIS. The procedure that is developed in this work is presented in the appendix B and is briefly summarized in the following: The CORINE data has to be reclassified according to the USGS land use classification but with 33 classes including the three new urban classes. Then, this reclassified CORINE data has to be transformed to the same binary format as the USGS dataset. As a last step, the CORINE data replaces fully or partly (only urban classes) the USGS data when creating the land use grids for each domain. Here, only the urban classes are adopted from CORINE data. The more recent CORINE data of 2006 leads to a significant difference of urban pixels in both domains compared to the USGS data base that was created before 2000. The percentage of urban pixels increased for the northern domain (southern domain) from 3.3 % (1.5 %) to 9 % (7.8 %).

The distribution of urban pixels is as follows:

- Class 31 (low intensity residential) represents 83 % (82.5 %)
- Class 32 (high intensity residential) 5.3 % (2.8 %)
- Class 33 (industrial/commercial/transportation) 11.7 % (14.7 %)

Additional information is necessary to account for the new land use classes. Both, the LSM and the UCM require information about soil types and properties, soil temperatures, green vegetation fraction and albedo. These values are adopted from the default settings [Skamarock et al., 2008] and are archived in the VEGPARM.TBL and LANDUSE.TBL. Both tables have to be updated to contain 33 classes. The actual parametrizations of roofs, walls and surfaces to be used by the UCMs can be found in the URBPARM.TBL. Two versions of the URBPARM.TBL exist by default. The original version (URBPARM.TBL) - developed for American cities - was modified in 2012 by Loridan and Grimmon for the UK (URBPARM_UZE.TBL). The entire exploitation of capabilities of UCMs actually requires the definition of those parameters exclusively for the specific city. High resolution elevation data and maps as well as an extensive investigation of those datasets are necessary to achieve a specific parametrization for the regarded city. Fallmann [2014] generated such a new URBPARM.TBL for the city of Stuttgart. In this work, the three available datasets of the URBPARM.TBL are tested regarding the representation of precipitation. The best results (not shown here) are achieved with the parametrization generated by Fallmann [2014].

5.4 Key findings during WRF setup process

The testing procedure is a protracted process to achieve the best possible model setup. Here, the most relevant results are briefly summarized:

(i) The extension of the **relaxation zone** from 5 to 10 pixels to achieve a smoother transition between the lateral boundary conditions [Prein et al., 2015] does not add value to the simulation results.

(ii) The **convection parametrization** scheme Kain-Fritsch (KF) leads to results that are similar in quality compared to the chosen Grell-Freitas scheme for the 5 km grid. But for $D03N$ and $D03S$ an enormous dry bias resulted with KF. These results are in accordance with the results from Brisson et al. [2015], who also observe a significant dry-bias when simulating in the gray zone of convection but with the Tiedtke scheme and the COSMO-CLM model.

(iii) WRF offers a special type of one-way nesting called **ndown**. The initial and lateral boundary conditions for the two fine grids (1 km) are obtained for every hour from the results of a coarse grid simulation (5 km). The separation of coarse grid simulations and fine grid simulation can save time, especially for two domains. This setup works well for nearly six months, until some wave structures in the Western part of the domain become apparent. These wave structures are visible in several variables such as 2 m temperature or wind speed and produce enormous rain amounts. It turns out that this process is gradual, reproducible (but only with long-term simulations) and no errors or numerical instabilities are observed. The exact reason for this behavior cannot be identified so far. It seems that the simulation in the gray zone of PBL and the rare update of boundary conditions each hour causes high deviations between the scales. Additional variations may be caused by the UCM that was first used for the 1 km resolutions only. As a consequence for the final setup, the three domains are simulated simultaneously using the new WRF version 3.7.1 instead of version 3.6.1 and applying the UCM to all domains (which is a prerequisite for version 3.7.1). In this way, reliable high resolution simulations for one year in $D03N$ and $D03S$ are achieved with this setup. The problems with ndown illustrate the huge effort and problems that may be linked to high resolution simulations.

(iv) The influence of **urban** areas simulated by different models (UCMs) or parametrizations is also evaluated. The comparison of the bulk parametrization and the SLUCM will be presented in Chapter 6. For the application of the most sophisticated UCM, the BEP, the eta levels near the surface have to be reduced to capture the influence of buildings at different heights. The two possible PBL schemes BouLac and MYJ applicable for the BEP are tested. The simulations with this model take significantly longer and show no added value regarding the performance of precipitation amounts and distribution. Additionally, the BouLac scheme sometimes simulates pixelated signatures that do not seem to be reliable. The MYJ scheme is widely used for high resolution simulations (e.g. Clark et al. [2007] and Fallmann [2014]), but the obtained results show a cold bias compared to the YSU scheme (e.g. -0.3 °C for May 2005 as a mean for Germany based on the 5 km domain). Those results do not justify the preference of the BEP model instead of the SLUCM.

6 High resolution simulations with WRF

6.1 Simulation and validation data

Simulation data

The investigation of simulation results are performed for the areas of domain $D03N$ and $D03S$. Both areas are additionally cut by at least 15 km at each sideline to minimize boundary effects. The two new areas (domains) are called $D-North$ and $D-South$ in the following. The simulation results are then compared for all domains including $D01$ (15 km, 3 h) and $D02$ (5 km, 1 h). Table 6.1 shows the resulting pixels for each area based on the respective domains. For $D01$ only 8 x 12 pixels for $D-North$ and 8 x 8 pixels for $D-South$ remain.

Grid size	15 km	5 km	1 km
$D-North$	8 x 12	24 x 36	120 x 180
$D-South$	8 x 8	24 x 24	120 x 120

Table 6.1: Amount of grid cells for $D-North$ and $D-South$ for the different grid sizes.

Three different simulation runs are performed to investigate the added value of high resolution simulations including certain aspects such as the influence of UCMs (Urban Canopy Model) or the convective parametrization. The abbreviation for first simulation runs is URB[1] to indicate the additional application of an Urban Canopy Model (UCM). These main simulation runs are performed for one year. The second high-resolution simulation run named $NURB$[2] is accomplished without UCMs to evaluate their influence. After a temporal spin-up that is similar to the URB simulations, the three summer months (June to August 2005) are simulated based on the same setup and data basis except the UCM. The same procedure is also applied to the third simulation run named $NCUP$[3]. Besides not applying an UCM (similar to 'NURB'), the cumulus parametrization is turned off for the 5 km resolution to investigate its influence on simulation results. No high resolution run (1 km) has been performed for this setup. The evaluation period is again June to August 2005.

Table 6.2 gives an overview of all datasets used in this Chapter in its original spatial and

---

[1] URB means URBan Canopy Model applied.

[2] $NURB$ means No URBan Canopy Model applied.

[3] $NCUP$ means No CUmulus Parametrization applied.

	$D01$	$D02$	$D03N$	$D03S$	period
URB	URB-15	URB-5	URB-1	URB-1	June 05 - May 06
NURB	(NURB-15)	NURB-5	NURB-1	NURB-1	June 05 - August 05
NCUP	(NCUP-15)	NCUP-5			June 05 - August 05

Table 6.2: Overview of simulations regarding applied grid sizes and periods. The abbreviations show the grid size and the setup of the respective datasets. Those configurations which are performed but not used in the following investigation are in brackets. Missing abbreviations reveal configurations that are not performed.

temporal resolution and the period for which the data is produced. Each dataset is available for both areas. The simulation names in this table are combinations of an abbreviation that describes the simulation run and the spatial resolution in kilometers. E.g. 'URB-5' describes a simulation run with UCM for the 5 km resolution.

Overall, six WRF (Weather Research and Forecasting Model) datasets with different temporal and spatial resolutions are analyzed for each area. The simulations with UCM are generated for the whole year whereas the remaining simulations only last from June to August 2005 (summer).

Validation data

Reliable, high temporal and spatial precipitation data is rare, even in Germany. Within RADOLAN[4], a variety of adjusted radar data products on an hourly data basis exist. The freely available RW-product and the more sophisticated RU-product, for instance, use hourly station data to adjust radar data or finally merge both datasets. See Bartels et al. [2004] for more details on the products. But both datasets are not corrected for systematic measuring effects which may influence precipitation patterns. The advantages and disadvantages of radar data and rain gauge data are merged, there. Instead of using one validation dataset based on two different measuring techniques, two validation datasets each based on only one measuring technique (or only adjusted on average) is applied here. In this way, the possible variation of spatially distributed precipitation data shall be better represented.

The first one is the radar dataset presented in Chapter 4. The corresponding correction scheme only uses relative radar data values to downscale below daily resolutions, but systematic measuring effects are additionally corrected. The original resolution of this dataset is 5 min and 1 km grid size. Its abbreviation is $RAD-1$.

The second dataset is solely based on station data, due to the fact that a dense network of rain gauges exists in Germany. Hourly data of those stations used for comparison in Chapter 4 are interpolated by an external drift-kriging technique based on altitude as drift variable analogous to the interpolation of monthly rain gauge data (see A 1 in the appendix). The relationship of precipitation data on altitude is less pronounced for a high temporal resolution but the bulk altitude effect could be better represented for a longer period or for missing stations at higher altitudes. The interpolation is performed on a 5 km grid for both domain

---

[4] Radar Online Adjustment; DWD project.

areas. The verification of this new dataset is only performed with REGNIE[5] data. The patterns are similar even on a longer temporal scale but the amounts are slightly underestimated (see Fig. 6.2 and Fig. 6.3). Both datasets are based on the German station network and provide precipitation values for the whole of $D-North$, but only for the German part of $D-South$. The abbreviation for this dataset is $STA-5$.

For analysis on a longer temporal scale (day or season) the REGNIE data is also used for means of comparison [REGNIE, 2017]. REGNIE data is regarded as a standard dataset for resolutions of one day or longer periods. As this dataset is only based on station data without using additional pattern information (from radar or satellite) its performance in complex terrain is also uncertain. The abbreviation for this dataset is $REG-1$.

6.2 Evaluation methods for high resolution simulations

How to evaluate high resolution simulations?

The aim of this investigation is to highlight the added value of higher resolution simulations with WRF with respect to precipitation. At small scales precipitation is partly regarded to be non-deterministic and unpredictable [Hohenegger et al., 2008]. Small temporal and spatial shifts between modeling results and observations can lead to the “double penalty” problem: A slight spatial shift firstly misses the observed precipitation object and secondly produces rain where no rain is observed [Prein et al., 2013a]. Common statistical values such as Root-Mean-Square-Error (RMSE) or correlation has therefore to be interpreted with care or at least not as the only measuring value. To evaluate the performance of simulation results for different grid sizes both a spatial and a temporal component have to be considered. Comparisons are useful on a longer temporal scale and for coarse grids or even for the whole domain. For the analysis of simulations with high temporal and spatial resolution a statistical approach that is independent from space or time is feasible. The comparison of the spatial extent of precipitation objects and the spatial dependences for instance, or if the mean temporal behavior of simulations and observations match. Statistical values or means that further investigate the precipitation structure (e.g. the distribution of precipitation amounts) are also useful. The emphasis of this work is placed on spatial evaluations.

General aspects

Most evaluations are performed on the same grid size to reveal the added value due to a better representation of processes instead of the better representation of terrain and elevation. Also Berg et al. [2013] claim: *’If the finer resolutions bring added value besides that of higher spatial resolution, this should be obvious also on a coarser resolution.’*. Therefore, the finer grids are remapped to the coarser grid (usually 15 km) with conservative techniques to maintain total rain amounts. The remapping feature ’remapcon’ within the Climate Data Operators [CDO, 2015] is used for that purpose. For the investigation of small scale precipitation patterns a 5 km grid size is applied to better represent the spatial variability of precipitation. The total evaluation period of one year is subdivided into seasons with a focus on winter and especially

---

[5] Gridded daily precipitation dataset of DWD that is based on rain gauge data.

summer. These two seasons show typically different synoptic-scale forcing and mesoscale processes regarding in particular heavy precipitation. Spring and autumn are regarded as a mixture of both types [Fosser et al., 2015, Prein et al., 2013a] and are widely disregarded in the following evaluations. Most convective events occur in the summer months (June to August) where the largest added value of high resolution simulations is expected. The two additional simulations are also only performed for this period. Consequently, the key emphasis is placed on the results of precipitation simulations for the summer months.

For almost all evaluation aspects at least two validation datasets are used, all of them revealing specific limitations. Interpolated station data (STA-5) represents the natural temporal variability of precipitation and its distribution of intensities well, at least at station's location. The areal interpolation leads to a smoothing effect that limits the reliability of spatial precipitation patterns based on station data especially in complex terrain and for small-scale rain patterns such as in convective cases during summer. Additionally, the influence of interpolation on the distribution of rain intensities and their temporal performance cannot be evaluated here. The applied radar dataset is not ideal regarding the evaluation of temporal aspects of precipitation and rain structure due to the uncertainties in calculating precipitation based on a Z/R relationship and due to the adjustment on a daily basis. The main advantage of radar data is the spatial distribution of rain echoes. But the influence of remaining systematic deficiencies such as clutter or composting effects is unknown. A disturbance in spatial evaluations of precipitation is also possible. The influence of these effects is reduced in summer (see Chapter 4). In general, interpolated rain gauge data is preferred for temporal evaluations and radar data for spatial ones.

Due to the fact that REGNIE data and interpolated station data only provide measurements for the German part of $D-South$, the following evaluations are also only performed for the German part of that domain. In case the whole domain is considered, it is explicitly mentioned.

The methods and measures are described first, and then the results are presented and discussed.

Seasonal precipitation amounts

The seasonal precipitation amounts are calculated for a first overview (see Fig. 6.1). An average domain value of precipitation amounts is derived for each dataset and compared to REGNIE data that serves as the validation dataset. For each season and the entire year from June 2005 to May 2006 the difference in percent is computed and opposed for a rough comparison.

Annual precipitation patterns

Figures of annual rain amounts serve as a first overview to get an impression about the influence of different spatial resolutions and the overall performance of different simulations (see Fig. 6.2 and Fig. 6.3). Therefore, the three different validation datasets that are used to reveal the range of probable distributions of precipitation amounts and the three URB simulations lasting one year are shown with their respective original grid sizes.

Spatial correlation and variability

The performance of the different simulation runs and resolutions regarding the spatial correlation, RMSE and variability is evaluated on the basis of Taylor plots (see Fig. 6.4). These three measures are mathematically linked and can be plotted jointly. The advantage of this figure is that all three of them can be regarded independently, but show a common performance at the same time. The variation of a validation dataset is usually used to normalize all datasets which are investigated. The three measures are calculated for each dataset and plotted in the diagram (one cross per dataset). The dataset with the shortest distance to the validation dataset shows the best performance [Taylor, 2001].

First, all available datasets are remapped and aggregated to the lowest common resolution (15 km, 3 h). Then the Pearson's correlation coefficient, the standard deviation and the Mean Square Error (MSE) are calculated and compared to REGNIE data (REG-1) for each pixel and finally averaged for the whole domain. The MSE is additionally converted to the RMSE. This is done separately for the summer months and the winter months.

Spatial similarities

Correlograms are used to investigate the spatial organization of precipitation patterns (see Fig. 6.5 and Fig. 6.6). They show the dependence of correlation on the spatial scale. The 3-hour precipitation amount for all available datasets remapped on a 5 km grid size is analyzed, therefore. First, distance classes are defined (here: each 5 km up to 60 km). Then for each pixel the correlation is calculated between the regarded pixel and all pixels within the respective distance class. The mean correlation values for each distance class are derived and plotted versus distance (class). Two thresholds are used to analyze rain patterns of all rain events as well as those time steps with higher rain amounts. Therefore, those time steps are chosen, where a 3-hour threshold of 0.03 mm (0.01 mm/h) as a mean value of all pixels based on interpolated station data is exceeded. For the analysis of heavy precipitation events or time steps a threshold of 1.5 mm (0.5 mm/h) is chosen which represents approx. the 95 to 98 percentile depending on region and season. Consequently, this measure is calculated for both thresholds, for summer and winter and for both areas.

Measures to compare high resolution precipitation features

To overcome the problem of 'double penalty' certain measures exist that consider those displacements. Different methods are shown and applied by Wernli et al. [2008] and Prein et al. [2013a] and an overview is given by Prein and Gobiet [2011]. In this study the Fractions Skill Score (FSS) is applied [Roberts, 2008, Roberts and Lean, 2008]. The basic idea is that a simulation is useful if the spatial frequency of events is similar in forecast and in the observation [Prein, 2013]. FSS is a neighborhood (or fuzzy) verification measure that directly compares forecast and observed fractional coverage of precipitation. Here, all simulations except URB-15 with 15 km resolution are used and remapped again to the 5 km resolution (if necessary) on an hourly temporal basis to obtain a reliable amount of grid cells in each investigation area. The whole area of $D-South$ is used for that evaluation, which excludes the application of interpolated station data. The radar dataset with 5 km grid size serves as validation data. First, the precipitation fields of simulations (Xr) and observations (Or) are transferred into binary fields I_o and I_x with the help of a precipitation threshold q. According to Eq. 6.1, all grid cells with values greater or equal to the regarded threshold are set to one and the residuals to zero.

$$I_o = \begin{cases} 1 & o \geq q \\ 0 & o < q \end{cases} \quad and \quad I_x = \begin{cases} 1 & x \geq q \\ 0 & x < q \end{cases} \tag{6.1}$$

In the second step a spatial moving average is used on the binary fields based on a squared window of side length n and uniform weights.

$$O_{(n)}(i,j) = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_O \left[i + k - 1 - \frac{(n-1)}{2}, j + l - 1 - \frac{(n-1)}{2} \right] \tag{6.2}$$

$$X_{(n)}(i,j) = \frac{1}{n^2} \sum_{k=1}^{n} \sum_{l=1}^{n} I_X \left[i + k - 1 - \frac{(n-1)}{2}, j + l - 1 - \frac{(n-1)}{2} \right] \tag{6.3}$$

$O_{(n)}(i,j)$ in Eq. 6.2 represents the field of observed fractions for a squared window with length n that is calculated from the binary field I_O. $X_{(n)}(i,j)$ is the corresponding field of simulated fractions based on I_X (Eq. 6.3). In both equations (6.2 and 6.3) i ranges from 1 to N_x in longitude direction and j from 1 to N_y in latitude direction. n is the horizontal scale (neighborhood size) with values between $n = 1$ and $n = 2N - 1$, where N = max(N_x,N_y). All grid points of the moving window which lie outside the investigation area are set to 0.

The third step is the computation of the $FSS_{(n)}$. Therefore, the MSE has to be calculated in advance for each n (see Eq. 6.4). The $MSE_{(n)ref}$ used in Eq. 6.6 can be regarded as the largest obtainable MSE with the given observed and simulated fractions (see Eq. 6.5).

$$MSE_{(n)} = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left[O_{(n)i,j} - X_{(n)i,j} \right]^2 \tag{6.4}$$

$$MSE_{(n)ref} = \frac{1}{N_x N_y} \left[\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} O^2_{(n)i,j} + \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} X^2_{(n)i,j} \right] \tag{6.5}$$

Then, the $FSS_{(n)}$ is calculated from the MSE as follows:

$$FSS_{(n)} = \frac{MSE_{(n)} - MSE_{(n)ref}}{MSE_{(n)perfect} - MSE_{(n)ref}} = 1 - \frac{MSE_{(n)}}{MSE_{(n)ref}} \tag{6.6}$$

$FSS_{(n)}$ values range between 0 and 1 (perfect fractional coverage). Two further thresholds are defined. The lower threshold is the observed fractional rainfall coverage (wet-area ratio) and represents the FSS value of a random simulation. The upper threshold is the value halfway between the random limit and a perfect match (FSS = 1) and is defined as the limit where

’reasonable skill’ is achieved. Consequently, reasonable skill is desireable, but all FSS values larger than the random limit have ’skill’ and can be evaluated. More detailed information about the FSS can be found in Roberts and Lean [2008].

This procedure is performed on all hourly time steps of one season where either the simulation data or the observation data provide values above the regarded precipitation threshold. The seasonal FSS is then calculated as the median of all hourly FSS. Varying the thresholds q and the size of the window n, both intensity and scale-dependent analysis are performed with the FSS. The following precipitation thresholds are applied in this work: q=0.1 mm/h, 0.5 mm/h, 1.0 mm/h and 2.0 mm/h. The window side length n varies between 1 and 41 grid cells (steps: 1, 5, 9, 13, 17, 21, 25, 29, 33, 37, 41). Finally, the seasonal FSS values for the summer and winter months are plotted against window side length n (see Fig. 6.7 and Fig. 6.8). The MET (Model Evaluation Tools) software is used to calculate the FSS [Brown et al., 2016]. One constraint of the application of the FSS in the presented work is the limited area. According to Mittermaier and Roberts [2010] the domain should ideally be large enough to comprise a typical mesoscale rain area which is in the magnitude of several hundred kilometers. Nevertheless, for convective rain events and also for a comparison of different simulations this evaluation is concluded to be useful.

Temporal correlation and variability

For the analysis of the temporal correlation and variability Taylor plots are used again (see Fig. 6.9). 3-hour precipitation amounts of all available datasets remapped onto the 15 km grid are analyzed for the summer and the winter season, therefore. It is similar to the preparation of spatially related Taylor plots. STA-5 serves as a reference.

Diurnal Cycle

The site-specific diurnal cycle of precipitation is the result of diurnal varying radiation, water vapor and large scale circulations [Li and Gao, 2012, Wallace, 1975]. Atmospheric circulation changes are responsible for intra-annual variabilities with more or less influence on convective processes [Trenberth et al., 2003]. Especially the simulation of the timing of the daily precipitation maximum is a challenge for mesoscale models. Here, the diurnal cycle is used as a measure how well simulations can reproduce a mean temporal distribution of precipitation amounts on an hourly scale for summer months and winter months. It is calculated by averaging the precipitation amounts for the same hours of a day temporally and spatially for the whole domain. Finally the 24 resulting values are plotted versus hours of a day. This is done for all available simulations and validation datasets. The 15 km simulation ($D01$) only offers a data resolution of three hours. Hence, these eight values per day are linearly interpolated to obtain values for each hour per day. This procedure leads to lower maximum values but the mean cycle should be well represented in this way. Additionally, a running mean of 3 hours is used for all datasets to make them even more comparable (see Fig. 6.10) .

Distribution of precipitation amounts

The distribution of rain amounts (Probability Density Function, PDF) is independent from time and serves as a statistical measure that describes mean precipitation characteristics.

First, all precipitation values of all grid cells are sorted according to size. Then, the range of precipitation values is separated into classes and the quantity of values per class is counted. As a final step, those quantities are normalized by the total amount of values and plotted versus precipitation class. Additionally, the probability of dry 3-hour time steps for each dataset is calculated and added in the legend. A threshold of 0.1 mm/h is used to differentiate between dry and wet. A 15 km grid size and 3-hour precipitation amounts are used as a basis for comparison. The original resolution of spatially higher resolved simulations is analyzed and plotted as a dashed line, too. This analysis aims at investigating the influence of originally different spatial resolutions of simulations on the distribution of rainfall intensities.

The PDF based on hourly data and 5 km grid sizes of all simulations except URB-15 are compared to evaluate the influence of the UCM, convection parametrization and grid size on rain structure for the summer months. Only data for three months are available for each season. Although all grid cells are used, the time serie is short and it cannot be precluded that single rain events influence the results. But this investigation may serve as additional valuable information for the comparison of different simulations.

The results are presented in Fig. 6.11 and Fig. 6.12.

6.3 Results of the analysis of simulations of precipitation

Figure 6.1 serves as a first rough overview of precipitation amounts. A total rain amount of 701 mm for the regarded annual period (June 2005 to May 2006) based on REGNIE data is shown in $D-North$ in Fig. 6.1(top). 30 % of total rain amount is produced in summer and in spring. Autumn and winter contribute each 20 % of total rain amount. The annual variations of rain amounts of all datasets compared to REGNIE are less than +/- 5 % in $D-North$. The seasonal variations are higher but do not exceed +/- 15 %. The high resolution WRF simulations with UCM (URB-5 and URB-1) show a similar performance: rain amount is overestimated in summer and winter as well as underestimated in spring and autumn. The results of URB-15 are slightly different, revealing overestimations in winter and spring as well as underestimations in summer and autumn. The additional simulations performed for the summer months show even smaller deviations than the other simulations in summer. The validation datasets reveal similar results as REGNIE data. Only for the winter months, interpolated rain gauge data overestimates REGNIE rain amounts whereas radar data show underestimations of the same magnitude.

With 1198 mm almost twice the amount of precipitation is observed only for the German part of $D-South$ according to Fig. 6.1(bottom). Most of the precipitation occurs in spring (41 %) and summer (25 %), lower amounts in autumn (16 %) and winter (18 %). Annual rain amounts only slightly differ from REGNIE data for all datasets except URB-15 and STA-5, both showing underestimations of approximately 10 %. The maximum deviations of simulation results from REGNIE data are higher in $D-South$ than in $D-North$, ranging between - 30 % and + 20 %. Again URB-5 and URB-1 agree well and show a similar behavior. Rain amounts are slightly overestimated in summer, winter and also in spring. Significant underestimations occur in autumn. Simulation results for URB-15 reveal an underestimation of rain amounts for all seasons. The additional simulations in summer show similar rain amounts compared to other simulations. The simulation run without cumulus parametrization for the 5 km grid size shows a slight underestimation of rain amounts, whereas the other simulations of grid sizes 5 km and below overestimate up to 8 %. The residual validation datasets reveal slight underestimations of rain amounts for all seasons. This underestimation becomes significant for interpolated station data in winter.

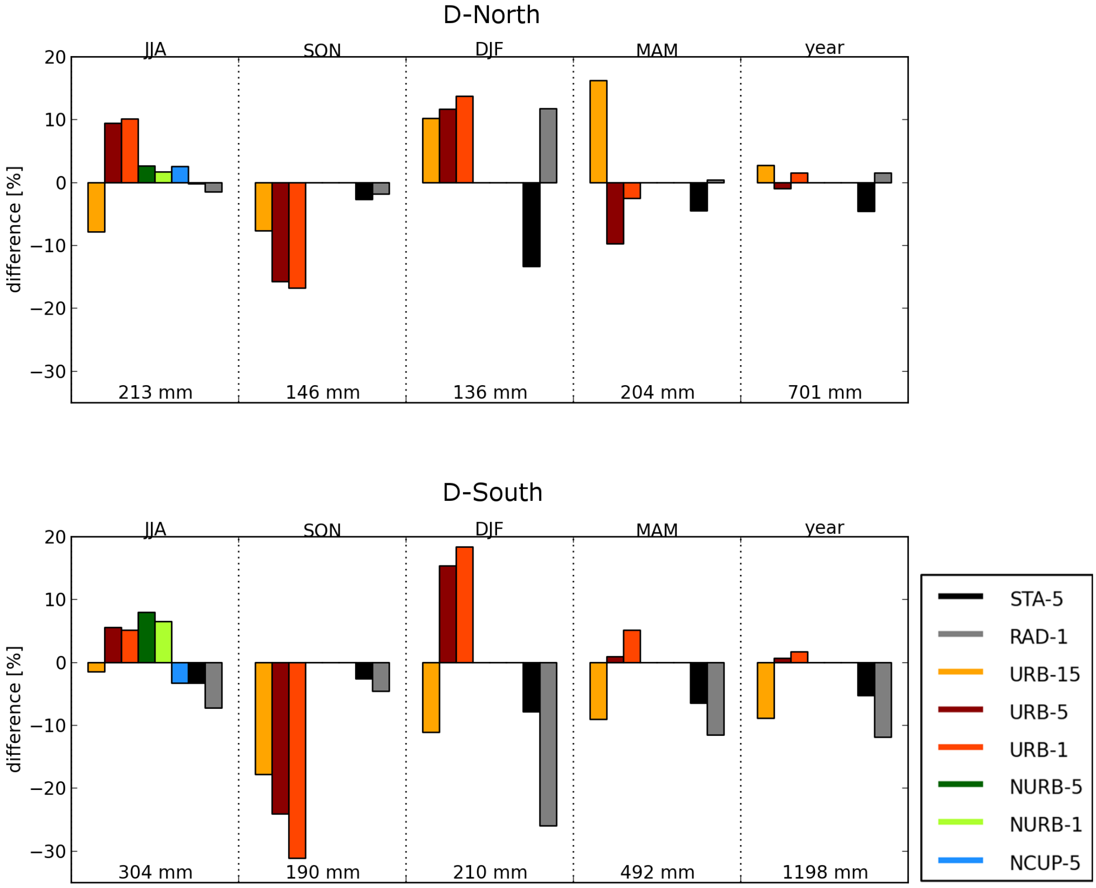

Figure 6.1: Seasonal and annual rain amounts (June 2005 to May 2006) for all datasets and $D-North$ (top) and $D-South$ (bottom). The columns mark the relative difference between the respective dataset and REGNIE for the whole domain. The additional simulations (NURB-1, NURB-5 and NCUP-5) are only available for the summer months. The total rain amount for the domain based on REGNIE data is shown at the bottom of each diagram for each season and year.

6.3.1 Spatial evaluation

Annual precipitation patterns

The following evaluations highlight different aspects of the analysis of precipitation patterns.

An overview of annual rain amounts for each investigation area and each dataset in its original grid size is given in Fig. 6.2 for $D-North$ and in Fig. 6.3 for $D-South$.

REGNIE data with a spatial resolution of one kilometer shows detailed spatial patterns of rain amounts even in complex terrain. For $D-North$ (Fig. 6.2a) a distinct rain maximum becomes obvious in the north eastern part of the domain and a minimum in the southern and south-eastern part. Figure 6.3 shows the corresponding results for $D-South$. The precipitation patterns in $D-South$ are closely linked to the terrain and its elevations. Precipitation maxima are produced in the southern, middle and northern part of the Black Forest and in the Vosges in the western part of the domain with significant higher rain amounts than in $D-North$. The precipitation minima can be found in the upper Rhine valley between the Black Forest and the Vosges.

The interpolated station data (Fig. 6.2c and Fig. 6.3c) show similar results but the maxima are slightly reduced and the patterns are smoother than REGNIE on a 5 km grid (not shown here) in both areas.

The spatial patterns of the previous datasets and radar data (Fig. 6.2e and Fig. 6.3e) match very well for both domains, but some clutter remnants and compositing effects become apparent. In the northern part of $D-North$ the Hamburg radar with two spokes is still visible, whereas significant clutter influence in the south western part appears. In $D-South$ significant influence of shaded areas in the north-eastern and south-eastern part of the domain cover the precipitation maxima in the Black Forest. Additionally, the conspicuous ring structure is a side-effect of the compositing algorithm that cannot be corrected by the correction schemes presented in Chapter 4.

The simulated precipitation patterns on the right hand side of Fig. 6.2 only partly match with those of the validation datasets. The precipitation maximum in the north-western part is underestimated and the minimum in the south-eastern part is absent in the simulations. In addition, the precipitation maximum calculated by WRF is lower and shifted to the south-east. Especially, the precipitation patterns for the summer months are responsible for that shift (not shown here). Only URB-15 (Fig. 6.2b) roughly reproduces the location of the correct precipitation maximum. The main patterns of URB-5 (Fig. 6.2d) and URB-1 (Fig. 6.2f) visually match very well on an annual basis but with a varying accentuation of small scale patterns.

Despite the complex terrain in $D-South$ (Fig. 6.3), the simulated precipitation patterns by WRF are similar compared to those datasets. The locations of minima and maxima and also their transition are well simulated, even for fine-scale structures in URB-1 (Fig. 6.3f). But minima and maxima are too pronounced for the high resolution simulation URB-1. Although URB-5 (Fig. 6.3d) shows similar patterns, significant lower maxima and minima are produced there than for the high resolution simulation remapped to a 5 km grid (not shown here). This is also true for the comparison of precipitation patterns from URB-15 (Fig. 6.3b) and URB-5 on a common 15 km grid (not shown here).

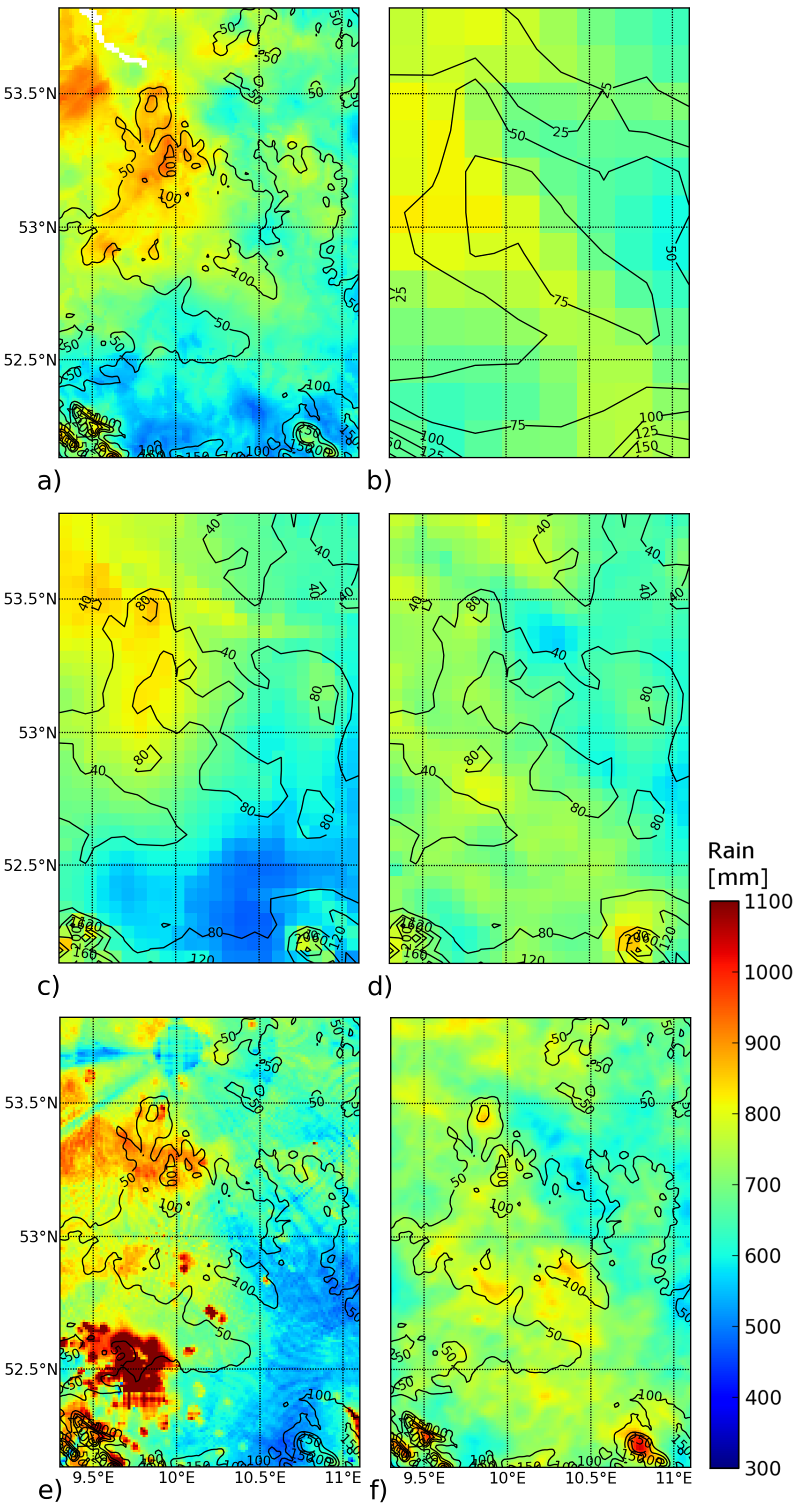

Figure 6.2: Distribution of annual rain amounts (June 2005 to May 2006) of validation data ((a) REGNIE, (c) interpolated station data, (e) radar data) and of simulation data with UCM ((b) URB-15, (d) URB-5, (f) URB-1) for $D-North$ in its original resolution overplotted by contour lines of elevation.

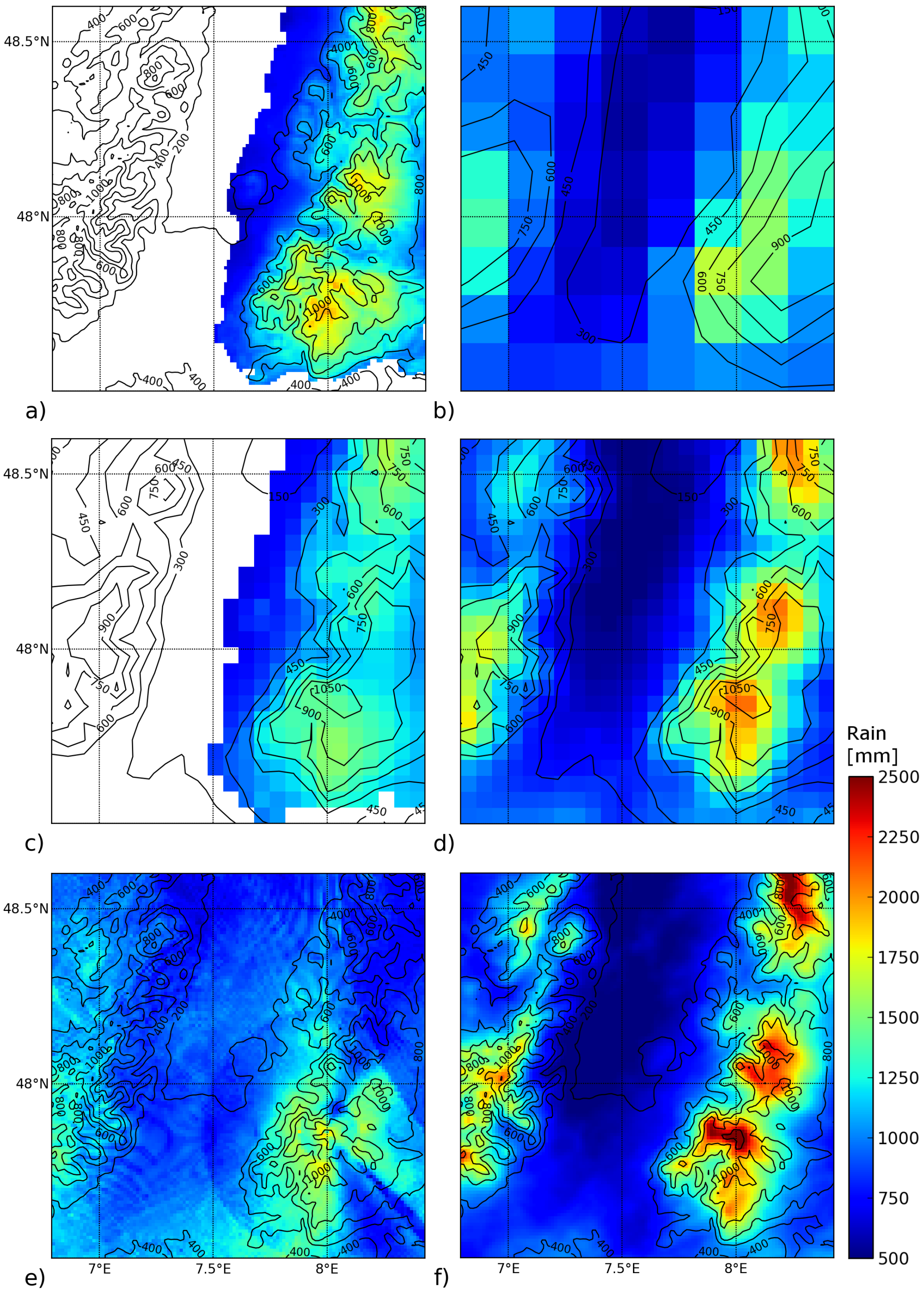

Figure 6.3: Same as Fig. 6.2 but for $D - South$.

Spatial correlation and variability

Taylor plots in Fig. 6.4 reveal mean statistical values of spatial patterns of all datasets for summer (left) and winter months (right) for both areas and shall help to support the visual derived findings above.

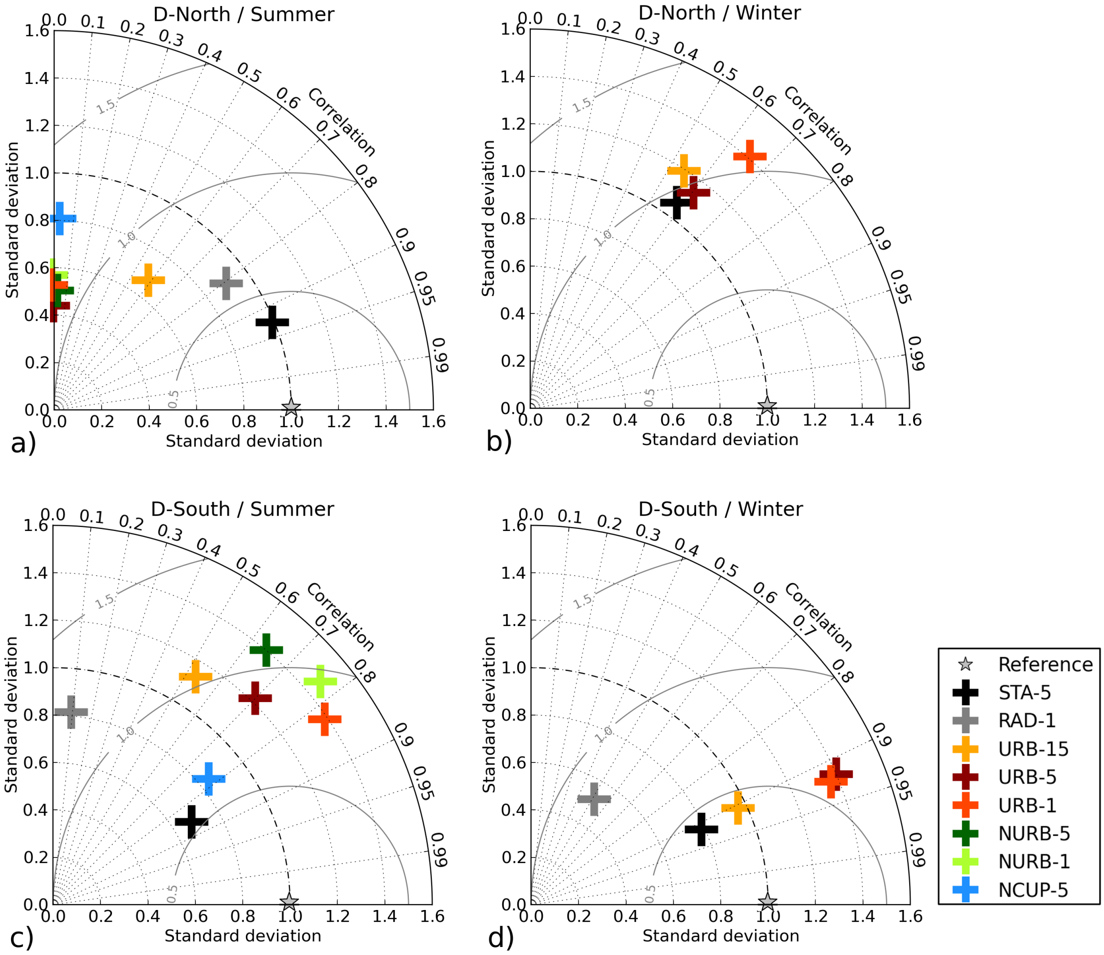

Figure 6.4: Taylor plots for the spatial evaluation of all available datasets based on a 15 km grid and seasonal precipitation values for $D-North$ (top) and $D-South$ (bottom) and for the summer (left) and winter months (right) are compared to REGNIE data. The additional simulations (NURB-1, NURB-5 and NCUP-5) are only available for the summer months. In (b) the radar dataset is not shown due to its high variability.

The high correlation coefficients of interpolated station data and REGNIE data (Fig. 6.4) confirm their good spatial accordance. The variability in $D-North$ is similar for both datasets and seasons; smaller variabilities for STA-5 in $D-South$ are observed.

Regarding radar data, a high variability of approx. 3.3 and a small correlation coefficient of 0.3 is calculated in $D-North$ in winter. Due to this high variability, radar data is not included in Fig. 6.4(b). The clutter influence in $D-North$, especially in winter, seems to be responsible for this feature. The correlation coefficients are also small in $D-South$ and the variability is reduced. This is probably due to remaining shading effects of radar data. Even if the radar image is not perfect small-scale precipitation patterns are probably better represented there than by interpolation techniques.

The visually weak spatial accordance in Fig. 6.2(d) and Fig. 6.2(f) is confirmed by correlation coefficients around 0 in summer. URB-15 (Fig. 6.2b) performs better regarding the large scale structures and reveals a correlation coefficient of approx. 0.6. The variability of precipitation patterns based on simulation results in the summer months are significantly smaller than in the observations. NCUP-5 shows a slightly improved performance. In winter, the correlation coefficients of all simulations range between approx. 0.55 (URB-15) and 0.65 (URB-1). As the variability of precipitation patterns increases with initially higher spatial resolutions, observations' variability is overrated.

The good accordance of simulated precipitation patterns in $D-South$ by WRF and the validation datasets is shown by correlation coefficients of around 0.7 in summer and over 0.9 in winter (see Fig. 6.4c, d). A significant improvement becomes obvious regarding the correlation coefficients with higher spatial resolutions. But the variability is simultaneously increased. This is also true for additional simulations without UCMs. Correlation coefficients of simulations without UCM are slightly smaller. Only NCUP-5 performs best according to the correlation coefficient. But compared to $D-North$ its variability of precipitation patterns is reduced in $D-South$.

Spatial similarities

An investigation of spatial rain patterns especially of convective cells requires datasets with high spatial and temporal resolutions. A 3-hour resolution on a 5 km grid size is used as a compromise to meet those conditions and to compare all available datasets. The investigation of hourly data (excluding URB-15) leads to comparable results (not shown here).

The correlograms of all datasets are shown for the threshold 0.01 mm/h (a, b) and 0.5 mm/h (c, d), for summer (left) and winter (right) and for $D-North$ in Fig. 6.5 and for $D-South$ in Fig. 6.6. The differences between the curves of validation datasets reveal the shortcomings of both datasets. Interpolated station data are often not able to reproduce small scale features especially in complex terrain adequately. This leads to an overestimation of spatial correlations with distance as already mentioned above. In principle, radar data are ideal for this investigation, but remaining clutter pixels and other systematic effects may deteriorate correlations in space. This can be seen in $D-North$ in winter where the influence of clutter increases due to a lower magnitude of precipitation intensities. As a consequence, the radar dataset is to favor in principle, but its validity is limited during winter in $D-North$, whereas interpolated station data may sufficiently represent precipitation patterns, there.

For the summer months and all rain events (threshold = 0.01 mm/h) NURB-5 and URB-1 show the best accordance with radar data up to a distance of 30 km in $D-North$ (Fig. 6.5a). Then both curves approach the higher correlations of interpolated station data (STA-5). URB-15 and URB-5 reveal similar correlations as STA-5 for all distance classes. NCUP-5 and NURB-1 have comparable correlations but both are considerably lower than the other datasets. For higher rain amounts (threshold = 0.5 mm/h) a separation of correlation curves due to the

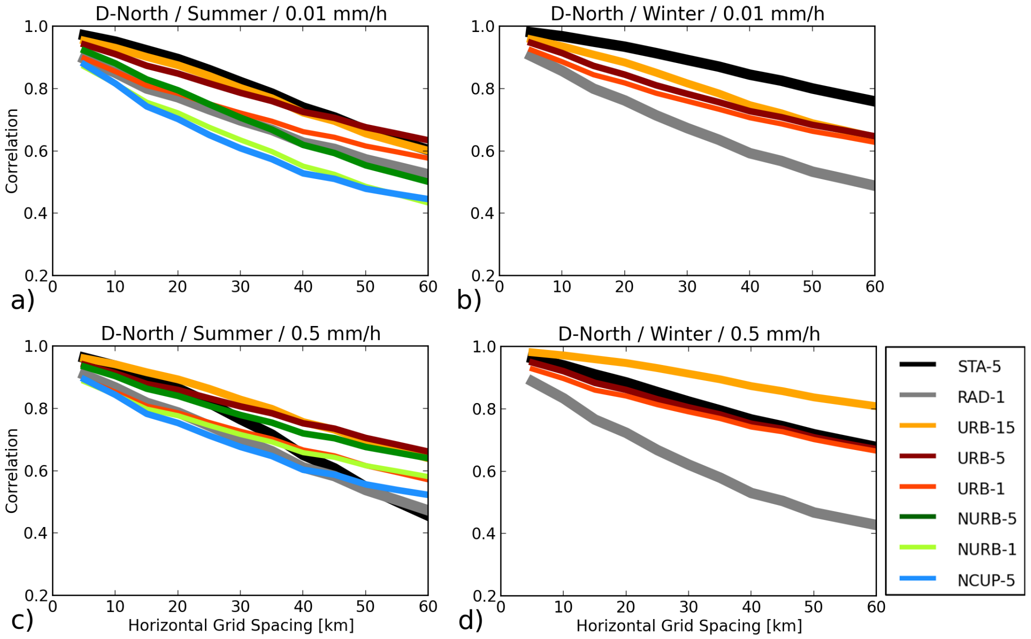

Figure 6.5: Correlograms of all available datasets based on a 5 km grid and 3-hour values for $D-North$ and for the summer (left) and winter months (right). Two different thresholds (0.01 mm/h and 0.5 mm/h) represent all rain intensities (a, b) and the higher rain intensities (c, d). The additional simulations (NURB-1, NURB-5 and NCUP-5) are only available for the summer months.

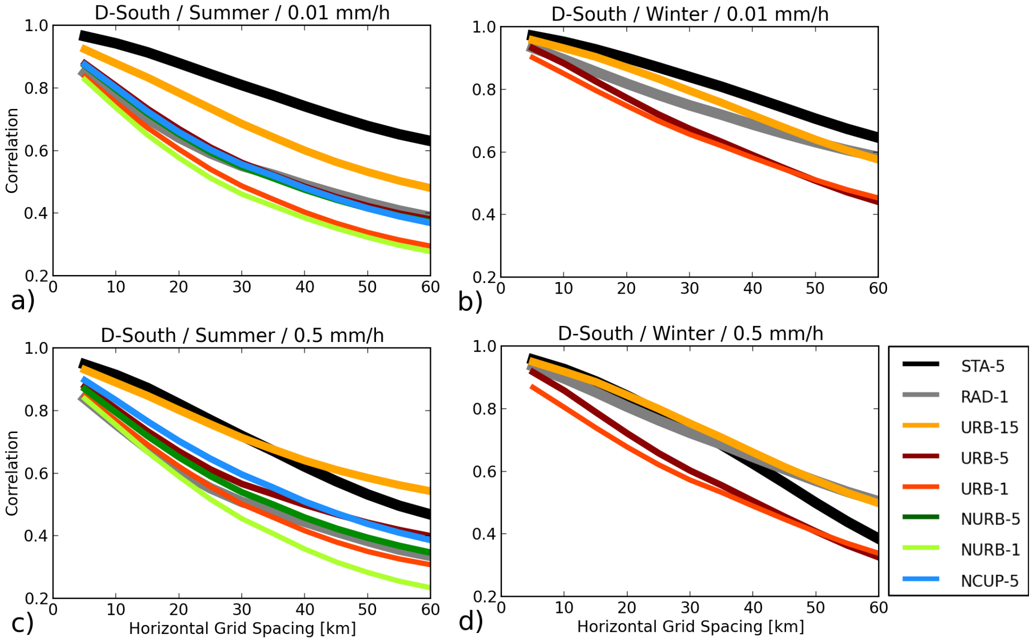

Figure 6.6: As Fig. 6.5, but for $D-South$.

original spatial resolution becomes apparent. URB-15 overestimates spatial correlations most. At a distance of 40 km the correlations of URB-15, URB-5 and NURB-5 are comparable but significantly too high. URB-1 and NURB-1 show the best accordance with radar data up to a distance of 30 km. NCUP-5 reveals even smaller spatial correlations. For distances above 50 km all simulations show overestimations of the spatial correlation.

In $D-South$, the separation according to the original spatial resolution is already apparent for all rain events (Fig. 6.6a). URB-15 significantly overestimates the spatial correlation of radar data for all distance classes, whereas URB-1 and NURB-1 underestimate it with the same magnitude. URB-5, NURB-5 and NCUP-5 almost perfectly match the correlation coefficients of radar data. Those relations remain even for higher rain intensities (Fig. 6.6c), but an influence of the UCM is noticeable. URB-15, URB-5 and NCUP-5 now slightly overestimate the spatial correlation of radar data and NURB-1 still underestimates it. URB-1 and NURB-5 show the best accordance with radar data, URB-1 up to a distance of 35 km and NURB-5 for distances greater than 35 km. In winter, the accordance of URB-1 and URB-5 is very high with slightly lower correlations with the URB-1 dataset. URB-15 reveals higher spatial correlations. In $D-North$, all simulations' correlations show small differences compared to observational datasets (Fig. 6.5b). For higher rain amounts (Fig. 6.5d) URB-15 overestimates spatial correlations whereas URB-1 and URB-5 match the correlations of interpolated station data well. In $D-South$, spatial correlations of URB-15 differ again from those of URB-1 and URB-5 but match very well with those of radar data. Results for URB-1 and URB-5 underestimate the spatial correlations of observational datasets. This is true for all rain events and higher intensities.

Skill score to compare high resolution precipitation patterns

The Fractions Skill Score (FSS) in Fig. 6.7 and Fig. 6.8 is presented for four thresholds in summer and three in winter for horizontal scales up to 200 km and for both investigation areas. Almost all FSS values exceed the limit of random skill (dashed line in Fig. 6.7 and in Fig. 6.8). These limits are very low and indicate the high degree of time steps with only very few pixels exceeding the regarded threshold. The FSS of all simulations typically increases with horizontal scale and it decreases the higher the precipitation thresholds are. This is true for all seasons and areas in general.

For small thresholds an indifferent behavior of the FSS for all simulations becomes obvious for $D-North$ in summer. With higher precipitation thresholds, the high resolution simulations URB-1 and NURB-1 perform better than the respective simulations with 5 km grid size. The URB simulations (with UCM) reveal slightly better results than the corresponding NURB simulations (without UCM). The NCUP-5 shows the highest FSS values for thresholds of at least 1 mm/h. Consequently, the separation of simulations with and without convection parametrization becomes obvious for higher rain amounts and almost all horizontal scales. Without convection parametrization slightly better FSS values are achieved.

In winter, only URB-1 and URB-5 are available for this investigation. The magnitude of FSS values for small thresholds is slightly higher than in summer in $D-North$. For thresholds of 0.5 mm/h and above a similar performance regarding the FSS is observed. URB-1 outperform the URB-5 simulations regarding FSS values for all thresholds and horizontal scales.

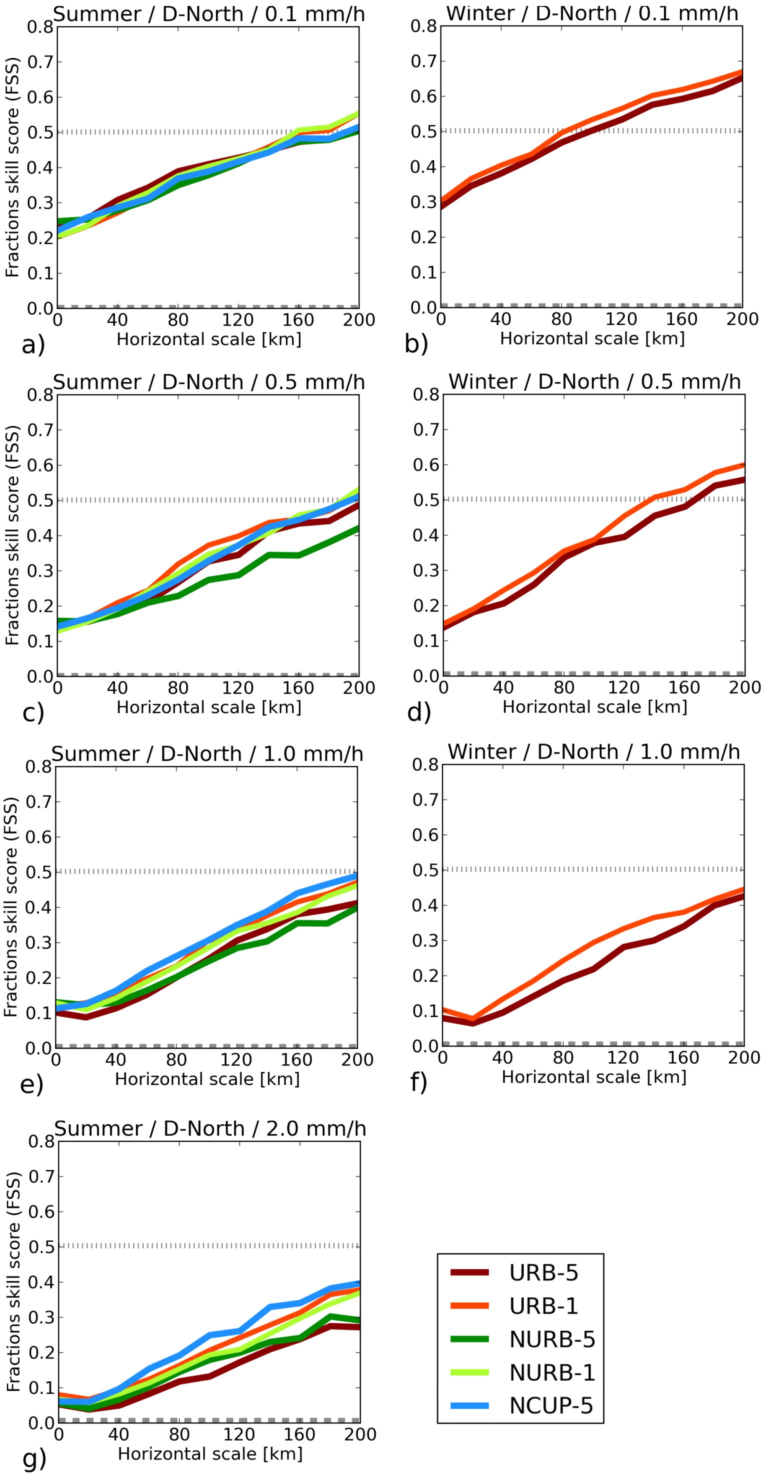

Figure 6.7: Fractions Skill Scores (FSS) for simulation results of hourly precipitation remapped to 5 km grid size for summer (left) and winter months (right) and for different thresholds (top to bottom: 0.1 mm, 0.5 mm, 1.0 mm, 2.0 mm) and different horizontal scales in $D-North$. The radar dataset serves as validation data. The FSS of a random simulation is plotted as a dashed line (near 0) whereas reasonable skill is achieved for values larger than the dotted line (near 0.5).

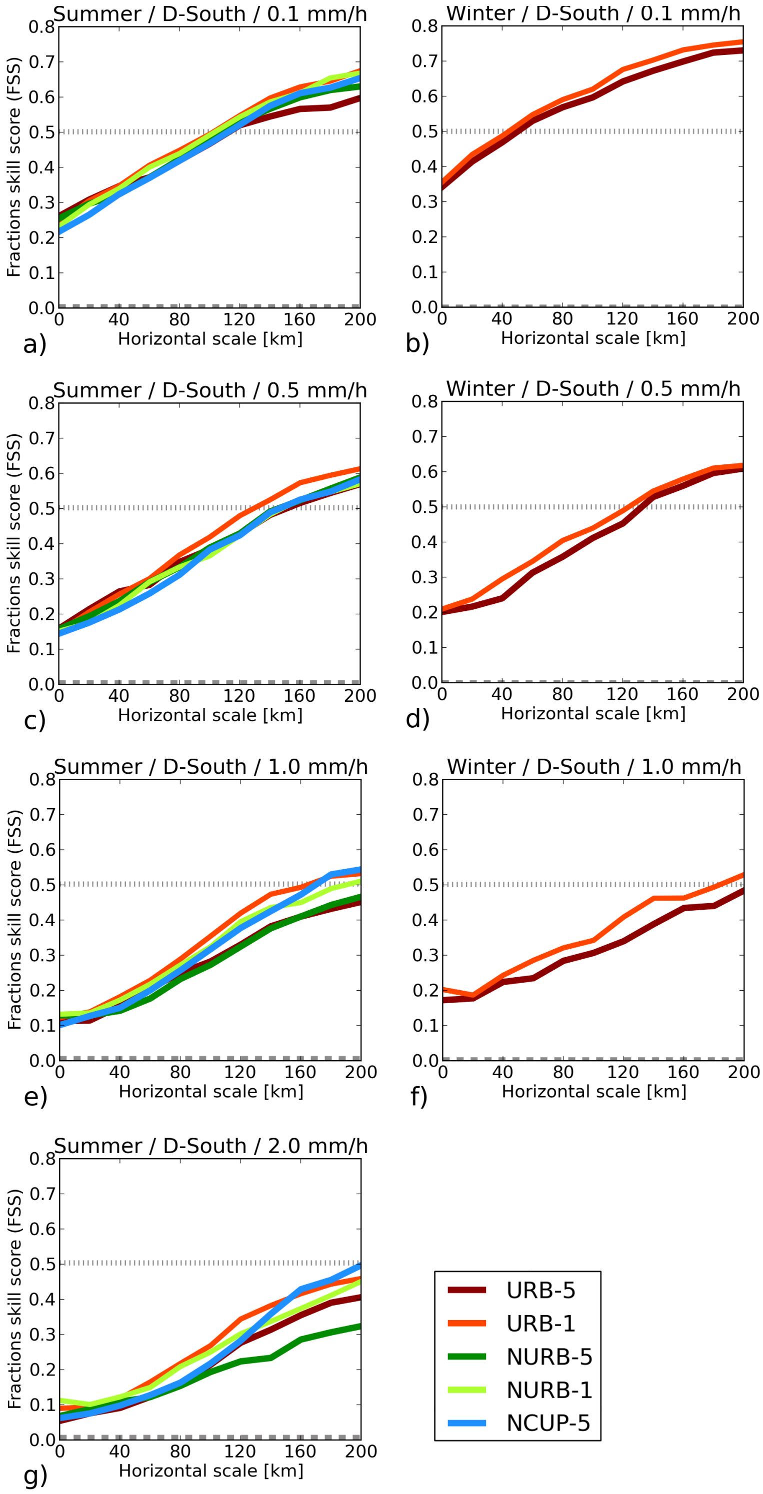

Figure 6.8: Same as Fig. 6.7 but for $D - South$.

In $D-South$, URB-1 performs best for all thresholds. NURB-1 and URB-1 show similar FSS values for a threshold of 0.1 mm/h as well as for horizontal scales below 80 to 100 km and higher thresholds. URB-5 and NURB-5 reveal again smaller FSS values for almost all thresholds and horizontal scales. The behavior of NCUP-5 is different. For horizontal scales below 80 km it is quite similar to the 5 km grid size simulations whereas it matches the FSS values of the 1 km grid size simulations for larger horizontal scales.

In winter and $D-South$, the same characteristics are observed than in $D-North$. The magnitude of FSS values for the lowest threshold of 0.1 mm/h is higher in winter than in summer, but not for the residual thresholds. Again, in general, URB-1 performs better than URB-5.

6.3.2 Temporal evaluation

Temporal correlation and variability

In Fig. 6.9 spatial statistical measures are presented to evaluate the overall temporal performance of precipitation simulations for winter and summer months for both investigation areas. Radar data reveal a correlation coefficient of 0.8 for both domains and seasons. The respective standard deviations in summer (left) match those of interpolated rain gauge data (reference) and significantly underestimate them in winter (right).

In summer, the correlation coefficients for all simulations slightly vary around 0.3 in $D-North$ and around 0.2 in $D-South$. Nearly the same correlation coefficients are calculated for simulations with the same setup but different spatial resolutions in both areas. Simulations without the UCM perform slightly better in $D-North$ but not in $D-South$. The RMSE shows a minor increase and the standard deviation a significant increase exceeding the observation's standard deviation with higher spatial resolutions.

Higher correlation coefficients can be observed for the winter months ranging between 0.35 and 0.6 with higher values in $D-South$. URB-1 and URB-5 reveal similar results. A significant improvement of both simulations compared to URB-15 regarding correlation coefficient, especially in winter, and RMSE becomes obvious for both investigation areas. Only their standard deviations are similar but slightly higher than the deviation of observed data.

Diurnal cycle of precipitation amounts

Figure 6.10 shows the diurnal cycle of all available datasets for $D-North$ (top) and $D-South$ (bottom) for summer (left) and winter (right) precipitation amounts.

Both validation datasets show a similar diurnal cycle of precipitation in summer (see Fig. 6.10a, c). In the summer months all simulations with a spatial resolution of 1 or 5 km show a similar performance. The timing of the increase and the decrease of precipitation amounts is almost identical, only maxima are accentuated slightly different. Even the accordance of the simulated diurnal cycle with the validation datasets is very good except a small delay of one or two hours when precipitation amounts increase in the morning. Minor differences between simulations with and without UCMs exist but no preference regarding the diurnal cycle can be identified. URB-15 also represents the main characteristics of the diurnal cycle but its performance shows significant higher deficiencies than the other simulations due to

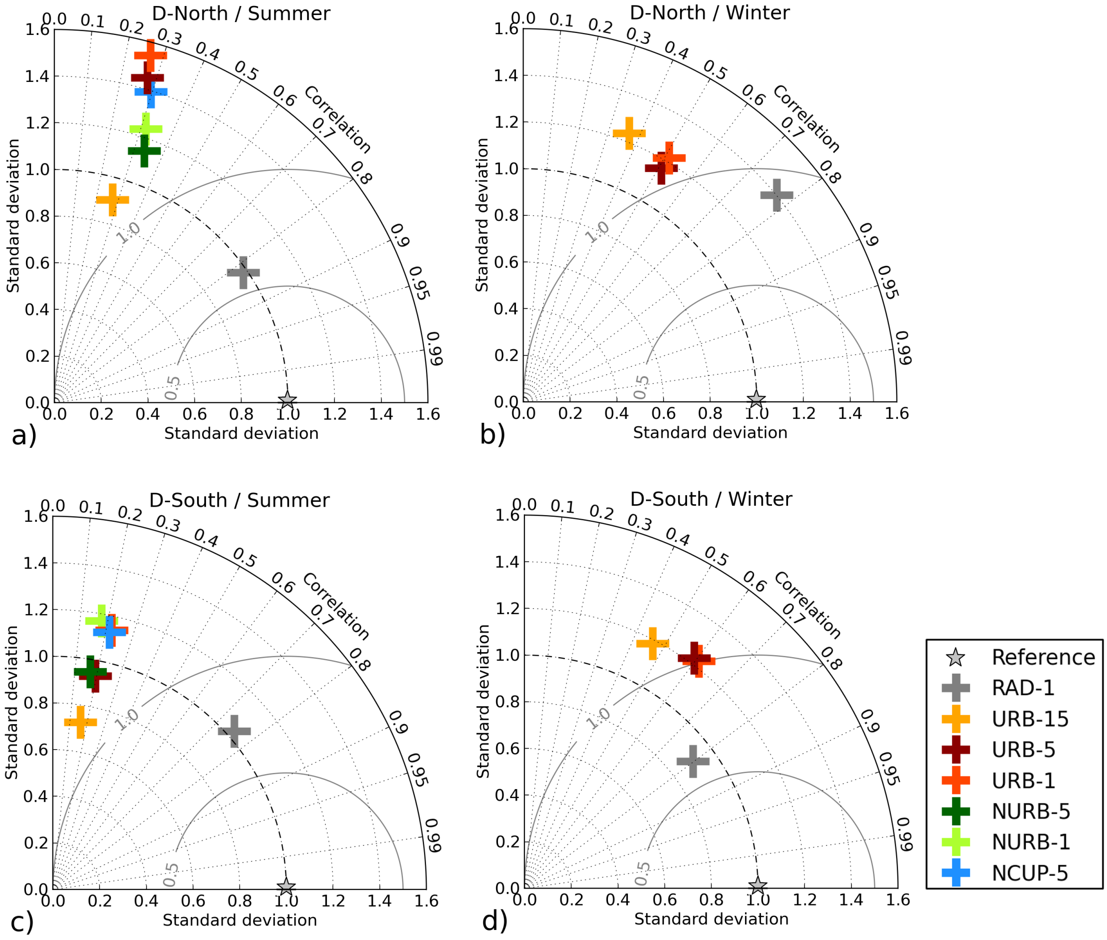

Figure 6.9: Taylor plots for the temporal evaluation of all available datasets based on a 15 km grid and 3-hour precipitation values for the northern domain (top) and the southern domain (bottom) and for the summer (left) and winter months (right) compared to interpolated station data. The additional simulations (NURB-1, NURB-5 and NCUP-5) are only available for the summer months.

the less pronounced increase of precipitation amounts and the early decrease.

The diurnal cycle in the winter months is less pronounced due to the lower influence of convection (see Fig. 6.10b, d). URB-1 and URB-5 show a good temporal accordance with validation data for $D-North$. The increase of precipitation amounts is again slightly delayed in the morning. By contrast, URB-15 shows a poor performance with anticyclical minima and maxima. In $D-South$, the accordance of the diurnal cycle between all simulations and validation datasets is very good.

In general, the high resolution simulations with 1 km and 5 km grid sizes show a better temporal accordance with validation data than the results of the 15 km grid size simulations. Between 1 and 5 km simulations, no major differences regarding the temporal performance can be identified.

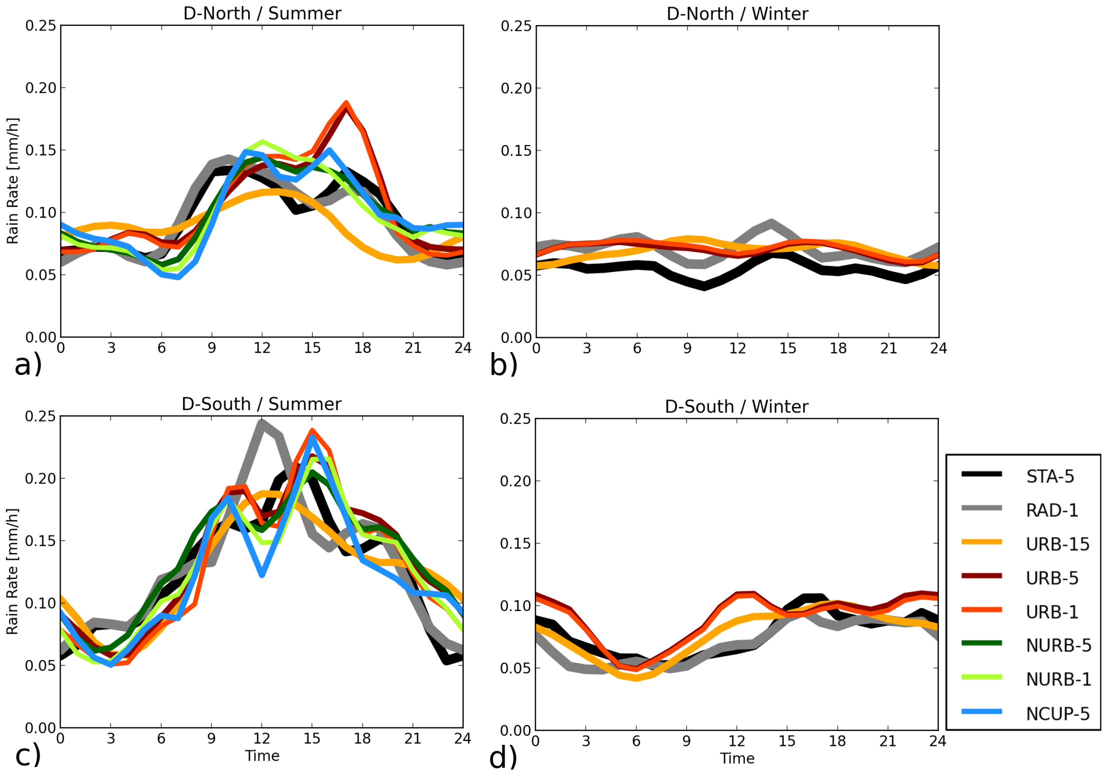

Figure 6.10: Mean diurnal cycle of precipitation for the temporal evaluation of simulation results based on 1-hour values for $D-North$ (top) and $D-South$ (bottom) and for the summer (left) and winter months (right). A 3-hour running mean is applied. The additional simulations (NURB-1, NURB-5 and NCUP-5) are only available for the summer months.

6.3.3 Evaluation of the distribution of precipitation amounts

The Probability Density Function (PDF) generated from 3-hour rain amounts for the URB-simulations and validation datasets are shown in Fig. 6.11. These PDF curves are based on a common 15 km grid size as well as in their original spatial resolution (dashed). The dry period probability are shown for summer (left) and winter (right) and $D-North$ (top) and $D-South$ (bottom).

The dry 3-hour probability of the validation datasets of approx. 0.85 is similar in summer and winter. In the summer months URB-5 and URB-15 show nearly the same values in both domains and slightly underestimate the observational dry period probability. By contrast, URB-1 shows a slight tendency of overestimation. In winter only URB-15 reveals similar probabilities as the validation datasets. URB-1 and URB-5 show similar values but underestimate the dry period probability. With higher spatial resolutions the dry period probability slightly increases.

The PDF curves of both validation datasets for the summer months are comparable. The URB-15 curve reveals the lowest probabilities for intensities above 3 mm, also differing from the validation datasets. URB-1 and URB-5 show a similar performance with slightly higher probabilities at medium and higher rain rates for URB-1. Both datasets overestimate the

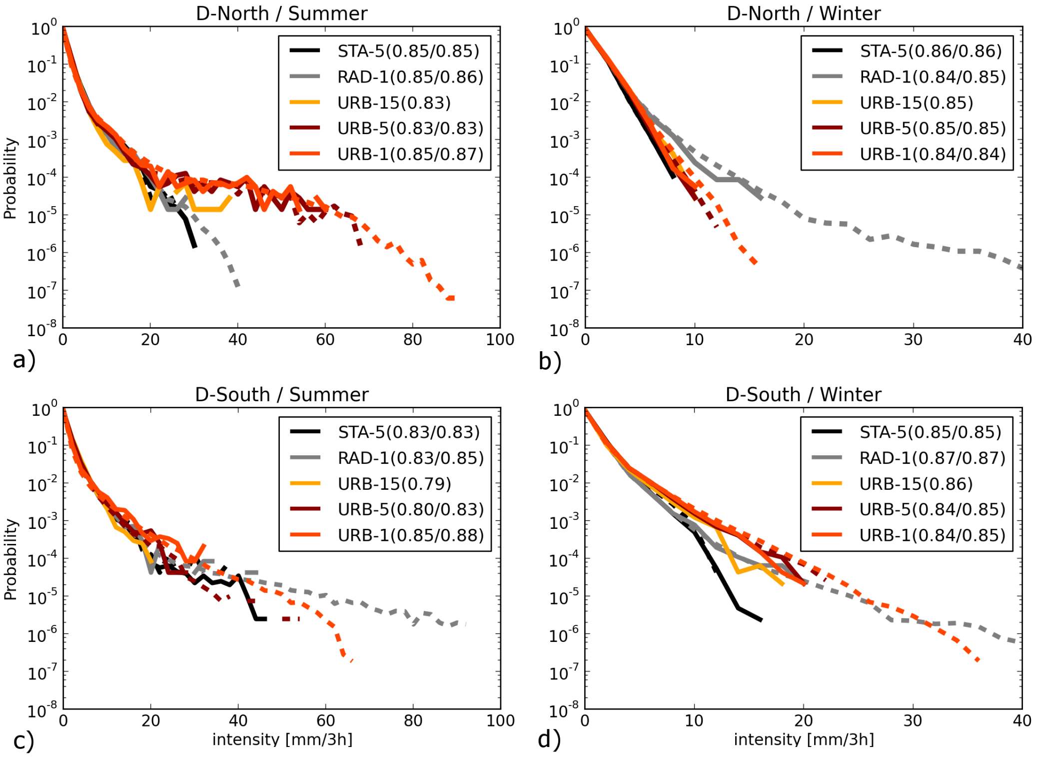

Figure 6.11: Probability Density Functions (PDF) of URB-simulations and validation datasets with 15 km grid size and 3-hour resolution for summer (left) and winter months (right) for $D-North$ (top) and $D-South$ (bottom). As a dashed line the PDF curves of all datasets are added with their original spatial resolution (if differing from 15 km grid size). The dry 3-hour probability of each dataset is added to the legend in brackets (15 km grid size / original grid size).

observational probabilities for intensities above 7 mm in $D-North$ and display a longer tailing even for high rain intensities. The accordance of probabilities in $D-South$ is better even for higher rain intensities.

In winter, the curves of interpolated station data and radar data vary probably due to the influence of remaining clutter effects in radar data (see Chapter 4). The curves of the three simulations are similar in both domains, only URB-15 shows slightly lower probabilities for higher rain intensities in $D-South$. This is probably the result of the coarse resolution of elevations in the Black Forest where usually the highest rain amounts are produced. In general, the curves of the same simulation with a different resolution are almost identical. Only the tailing varies due to smoothing of precipitation maxima for larger grid sizes.

In general, the findings above agree with those for hourly data on a 5 km grid size in Fig. 6.12. The dry hour probability is similar for simulations with and without an UCM showing small underestimations for URB-5 and NURB-5 and small overestimations for URB-1 and NURB-1. Even higher values are achieved without cumulus parametrization for the 5 km grid size (NCUP-5). The PDF curves for all simulations in $D-North$ show higher probabilities for

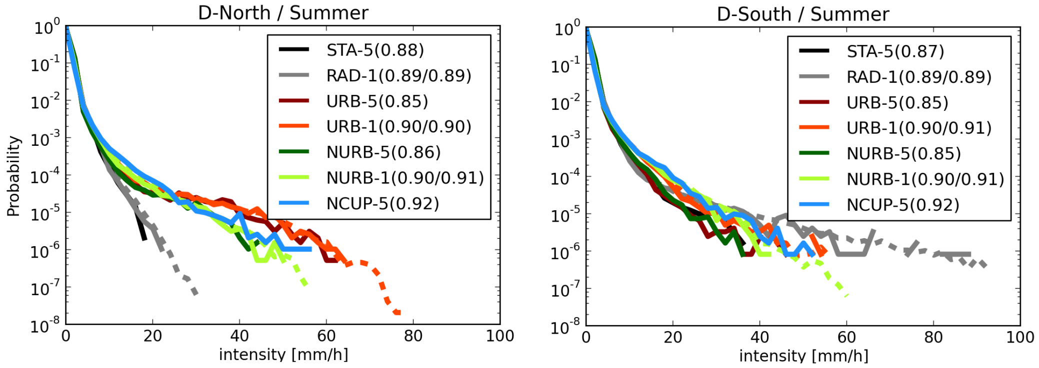

Figure 6.12: Probability Density Functions (PDF) of all available simulations and validation datasets with 5 km grid size and hourly resolution in summer for $D-North$ (left) and $D-South$ (right). As a dashed line the PDF curves of all datasets are added with their original spatial resolution (if differing from 5 km grid size). The dry hour probability of each dataset is added to the legend in brackets (5 km grid size / original grid size).

rain amounts over 3 mm. URB-5 and NURB-5 as well as URB-1 and NURB-1 behave similarly again for intensities below approx. 12 mm. NCUP-5 shows the highest probabilities for this rain intensity range. For higher amounts, the influence of the UCM dominates and leads to a comparable behavior of URB-simulations on the one hand and of NURB-simulations and NCUP on the other hand. In $D-South$ the similar behavior of URB and NURB simulations with the same original grid size counts for all rain intensities. NCUP-5 shows comparable PDF curves as the high resolution simulations (URB-1 and NURB-1) do.

6.4 Discussion

The discussion of simulation results is arranged according to the type of simulation (setup) referring simultaneously to the research questions that are formulated in Chapter 1.

Differences of WRF precipitation results of simulations with varying spatial resolutions

The evaluation of the added value of smaller grid sizes and the influence of the topography on simulation results is mainly based on the simulations with UCM (URB) lasting for one year.

The main differences among the simulations' results become obvious between the 15 km resolution and the higher resolutions (5 km and 1 km). Land use and especially the topography are only very roughly represented with 15 km grid size. The underestimation of total rain amounts in all seasons in $D-South$ may be attributed to the smoothing effect of elevations in this complex terrain with a reduced orographic forcing (Fig. 6.1). Smoothing effects are responsible for most limitations of URB-15 in the presented investigations. They are dominant in temporally and spatially highly variable rain fields. As a consequence, the spatial correlation with distance is too high and the probability of intense precipitation within the PDFs is underestimated (Fig. 6.5, Fig. 6.6 and Fig. 6.11). Ban et al. [2014] also show that their simulations with 12 km grid size underestimate the frequency of heavy hourly events, whereas the 2.2 km simulation performs very good. The temporal correlation is only slightly reduced compared to URB-5 and URB-1, but the diurnal cycle shows the typical premature maximum and a reduced accordance with observational datasets [Brisson et al., 2016]. The best performance of URB-15 can accordingly be observed in winter (Fig. 6.9 and Fig. 6.10).

Another major reason that is responsible for the variations among the different simulations results is the usage of spectral nudging only for URB-15. The better accordance of precipitation patterns in the summer months in $D-North$ can be explained in this way, whereas the further downscaling with WRF on the basis of boundary conditions deteriorates these patterns for this season (Fig. 6.2).

URB-5 and URB-1 show a similar performance for spatial and temporal analysis. The temporal performance is almost identical and even the diurnal cycle of both datasets agree well (Fig. 6.9 and Fig. 6.10). The latter one is often argued to be one of the main constraint of convection parametrizations. Comparable improvements regarding the diurnal cycle are achieved by Fosser et al. [2015], Kendon et al. [2012], Prein et al. [2013a] by omitting convection parametrization.

Despite the fact that URB-5 and URB-1 show the same tendencies, slightly differences between the results of simulated rain patterns and rain intensities remain. URB-1 usually shows a further improvement of precipitation simulations in summer and winter. The dry period probabilities (3h and 1h) in summer are the same for URB-1 and validation datasets, whereas the respective probabilities for URB-5 are reduced (Fig. 6.12). This is in line with Berg et al. [2013] who show underestimation of dry day probabilities and trace this result to convection parametrizations. But this is only true for the summer season. Furthermore, URB-1 reveals slightly higher correlation coefficients of spatial patterns and a higher variability than URB-5 (Fig. 6.4). The Fractions Skill Scores based on URB-1 are higher for all horizontal scales and intensities (Fig. 6.7 and Fig. 6.8). The largest differences between both simulations and the main added value are observed for higher rain intensities. With higher intensities also the correlograms and the PDF curve show that the size and the structure of rain cells as well as the corresponding rain intensities are better captured by URB-1 (Fig. 6.12, Fig. 6.5 and Fig. 6.6). Kendon et al. [2012] and Prein et al. [2013b] draw the same conclusions based on their evaluations for Convection Permitting Scales (CPS) simulations.

For lower rain intensities, the results are domain-specific. In $D-North$, an improvement of the above mentioned rain structure is still valid even for lower rain rates. In $D-South$ by contrast, the spatial correlations are underestimated and the probabilities for small and medium intensities are overestimated compared to validation datasets (Fig. 6.6 and Fig. 6.11). This also explains the overestimation of maxima and the underestimation of minima in the annual rain amounts of $D-South$ based on URB-1 (Fig. 6.3f). A similar behavior can be derived from the simulation results of Chan et al. [2013], Ban et al. [2014] and Brisson et al. [2016]. Chan et al. [2013] and Ban et al. [2014] perform their simulations in mountainous regions and observe improvements of daily rain amounts only for higher percentiles (90 % and 95 %). The domain of Brisson et al. [2016] is in a more flat area; they show a general improvement with smaller grid sizes. Of course, different model setups are applied but the variability of rain patterns seems to be overestimated due to a higher influence of orographic forcing. The reasons for this behavior cannot be completely identified within this framework.

A more thorough analysis of processes or variables such as wind and radiation that influence orographic forcing have to be performed to achieve further indications for this behavior.

Despite more demanding simulations in complex terrain in $D-South$, the location of precipitation maxima are better captured there than in areas without any topographic particularities (Fig. 6.3). The model tends to produce highest rain amounts in those areas with the highest altitudes. This usually matches well in complex terrain, but in slightly hilly areas precipitation maxima might be dislocated.

Impact of omitting convection parametrization at 5 km grid size

A 5 km grid size is in the gray zone of convection modeling. New parametrization schemes are developed for that purpose but one should not reject an explicit modeling of convection even though convection processes can only partly be resolved. Weisman et al. [2008], Schwartz et al. [2009] and Prein et al. [2013a] show improvements for such simulations based on grid sizes between 4 and 5 km.

The simulation run NCUP-5 should serve as a reference dataset to evaluate the performance of convection parametrization at 5 km grid sizes (URB-5 and NURB-5). The main added value of NCUP-5 can be observed for main spatial distribution of precipitation including the total amounts (Fig. 6.4). But it also reveals deficiencies in the exact location of maxima (not shown here). The dry period probability is improved compared to NURB-5 (Fig. 6.12). The evaluation of the FSS reveals a significant improvement for $D-North$ by omitting convection parametrization. In $D-South$ only improvements for higher precipitation thresholds and horizontal scales beyond 80 km are observed (Fig. 6.7 and Fig. 6.8). A poorer performance is shown for the distribution of rain amounts within the PDF curves (Fig. 6.12). According to Weisman et al. [1997] and Prein et al. [2015], convective instability is forced onto an unrealistic scale for grid sizes larger than 4 km. There it leads to an overestimation of the convective mass flux and precipitation. This effect can explain the overestimation of probabilities for the observed intensities from approx. 3 to 12 mm based on NCUP-5.

In general, the simulation results of NCUP-5 are not outperformed by those of NURB-5. NURB-5 and URB-5 are based on a convection parametrization which is developed for smaller grid sizes even within the gray zone of convection. In this way the diurnal cycle and the spatial patterns are represented successfully (Fig. 6.10 and Fig. 6.4). Donner et al. [2011] and Bechtold et al. [2014] also report improvements in the representation of the diurnal cycle due to a new parametrization scheme. Slightly better results are achieved for NCUP-5 regarding measures of the spatial distribution and for location of precipitation maxima in summer (not shown here). However, especially in complex terrain overestimations due to unrealistic convective forcing is also observed.

An additional benefit of this further simulation is that certain conclusions can be drawn from the comparison of all datasets: NCUP-5 behaves similarly as URB-1 and NURB-1 for precipitation probabilities (PDF) and dry period probabilities as well as for main temporal and spatial performance (Fig. 6.12, Fig. 6.4) and Fig. 6.9). This suggests that the parametrization or rather the explicit simulation of convection is responsible for the different behavior shown in these evaluations between URB-1 (or NURB-1) and URB-5 (or NURB-5) and not the grid size.

This is also true for the correlograms in $D-North$. But in $D-South$ the results of NCUP-5 are closer to URB-5 and NURB-5 (Fig. 6.5 and Fig. 6.6). The grid size seems to dominate and suggests that the higher spatial resolution in complex terrain intensifies processes such as orographic forcing which results in this scale-specific behavior.

Influence of an Urban Canopy Model (UCM) on simulation results

The application of an UCM usually aims at improving the simulated weather explicitly in urban areas. Here, the overall effect on the complete domain is analyzed without any focus on the urban areas.

The total rain amounts for the whole domain in summer already differs slightly between simulations with (URB-1 and URB-5) and without UCM (NURB-1 and NURB-5) (Fig. 6.1). Some evaluations presented above show slightly better results for NURB in $D-North$ and for URB in $D-South$. These investigations are based on common statistical measures which focus on the main temporal and spatial accordance such as the total rain amount for summer months and the temporal and spatial accordance based on Taylor diagrams (Fig. 6.9 and Fig. 6.4). No differences between URB and NURB can be observed for the dry period probability and also the accordance of PDF curves is similar for wide ranges. Only the probabilities for the highest rain rates are overestimated by URB in $D-North$ and the medium rain rates are slightly overrated by NURB in $D-South$ (Fig. 6.12). Regarding the diurnal cycle no preferred simulation setup can be identified. The increase and the decrease of mean precipitation amounts are similar but the shapes of the peaks are specific for simulations with the same setup (Fig. 6.10). URB-simulations show a better accordance with validation datasets for correlograms. Nevertheless, the variations between correlograms from simulations based on 1 km and those based on 5 km are higher (Fig. 6.5 and Fig. 6.6). The results of the FSS reveal a better performance for URB-simulations for nearly all intensity thresholds and horizontal scales (Fig. 6.7 and Fig. 6.8).

In general, an influence of the usage of an UCM can be observed in precipitation results. Its influence is widely subordinate compared to different grid sizes or differences due to the usage of a convection parametrization. The URB-simulations are advantageous in $D-South$ and less superior in $D-North$. The main added value of the UCM is observed for evaluations regarding the spatial distribution of precipitation.

7 Joint discussion of weather radar data and WRF simulations

Reliability of observational data

Prein et al. [2015] emphasized the importance of reliable observational datasets. Both, interpolated rain gauge data and radar data may serve as possible validation datasets to analyze the temporal and spatial performance of high resolution simulations. Frei and Schär [1998] and Isotta et al. [2014] presented the constraints of interpolated rain-gauge data in complex terrain. Limitations of radar data are discussed in Chapter 3. The usage of at least two validation datasets may indicate possible variations of observational datasets for the respective analysis. If both datasets agree well for certain evaluations, it can be regarded as a proof of reliability of observational results. If the results of these datasets differ, the limitations of one dataset may be identified in this way.

The main limitations of radar data and of interpolated station data are already obvious in the patterns of total annual precipitation amounts in Fig. 6.2 and Fig. 6.3. Mainly, the correlograms in Fig. 6.5 and Fig. 6.6 are influenced by these effects.

The spatial correlations within radar data in $D-North$ are reduced (Fig. 6.4) due to a higher influence of remaining clutter such as in winter and for higher rain intensities compared to the residual datasets. Clutter is also responsible for the higher rain intensities observed there, according to the PDF (Fig. 6.11) in winter. Partly beam-filling and overshooting effects in winter may further decrease correlations in space. In $D-South$, shading effects in radar data are apparent, but they do not influence spatial correlations.

The spatial smoothing effect of interpolated station data also affects correlograms. This dataset shows significant higher spatial correlation coefficients for distances below 50 km, especially in complex terrain and in summer (Fig. 6.6). Regarding the temporal performance of validation datasets, an influence of the effects mentioned above cannot be identified (Fig. 6.9).

Comparison of weather radar results and simulation results with WRF

The high spatiotemporal variability of precipitation places high demands on measurement and simulation of this parameter. Weather radar data measurements and dynamical downscaling are selected as promising tools to capture these variabilities especially in space.

Long temporal scale

On a long temporal scale, systematic effects in radar data influence the quality of precipitation data. Both, gradual effects such as the altitude dependence of radar measurements as well as local, abrupt limitations of precipitation patterns such as clutter effects are observed. The new developed correction schemes are able to improve radar data significantly on a long temporal scale. The accumulated corrected 5-minute measurements also show improvements compared to uncorrected radar data. In winter, the presented long-term analysis and correction of radar data reveals limits of radar measurements that are hardly correctable. These effects are largely caused by shallow precipitation echoes that cannot be detected by radar measurements at far ranges from the radar site. The higher vertical extent of rain echoes at mean and measurements of largely pure rain in summer reveal that radar measurements are usually of high quality in summer and may show some constraints in winter.

Accumulated precipitation measurements of WRF (Weather Research and Forecasting Model) also show a high potential in this respect. The total rain amounts are comparable to validation datasets and also the patterns in complex terrain are well represented (Fig. 6.3). Only in flat or gently-hilly terrain the locations of precipitation maxima are not well captured in summer (Fig. 6.2). The higher variability and heterogeneity of precipitation in summer is more difficult to simulate. In winter, even the simulation with the coarse grid size of 15 km reveals reasonable results.

Short temporal scale

On a short temporal scale, corrections of individual radar images may lead to a satisfying data quality. Radar reflectivity has to be transferred to rain intensities, but this can reasonably be achieved by a well-defined Z/R relationship such as the three-part Z/R relationship used in this work (see Tab. 3.1). Adjustment or additional post-corrections for systematic effects may further improve radar based precipitation amounts. Areal patterns of precipitation are well represented – after correction – but the transfer of reflectivity into rain rate or the adjustment lead to uncertainties especially for higher rain amounts [Morin et al., 2003].

The simulations show the best performance for high resolution simulations. The topography, land use and also the dynamical processes can be better represented with smaller grid sizes. The diurnal cycle is well represented for small grid sizes (Fig. 6.10). This is a further indication that the dynamics are well captured in current Regional Climate Models (RCM). The temporal comparison between simulation results and interpolated station data (see Fig. 6.9) shows a reasonable accordance especially in winter. Nevertheless, the simulations are outperformed by radar measurements.

Large spatial scale

On a large spatial scale the gradual processes such as the altitude dependence of radar measurements become important. Differences of precipitation amounts at close and at far ranges from the radar site are apparent for longer periods. In complex terrain, the radar beam has to be raised to prevent blockage and shading effects. This enhances the areal variations of rain amounts. The presented post-correction scheme shows its main impact, there. The quality of radar measurements decreases with distance from the radar site [Joss and Lee, 1995], as already mentioned. Only the measurements close to the radar site may be deteriorated by side-lobes of the radar measurement itself. Consequently, nationwide radar composites reveal varying data quality in space.

The performed WRF simulations aim at improving the representation of the dynamic of those processes that dominate small scale patterns such as convection. The large scale patterns are transferred from General Circulation Models (GCM) or reanalysis data such as ERA-Interim only at the boundaries of the domain. The simulation of small-scale processes can superimpose the large scale patterns and may deteriorate simulation results. This is observed in the northern domain in summer for high resolution simulations on a long temporal scale and for large spatial scales. Berg et al. [2013] showed that biases in the GCM are not only transferred to the RCM but also enhanced by the RCM itself.

Small spatial scale

The representation of precipitation patterns at small spatial scales is the strength of radar data. All small-scale convective events are usually detected which is often a shortcoming of precipitation datasets. The gradual effects such as the altitude dependence only show minor influence at short distances. Shortcomings of single pixels or small areas such as clutter effects locally affect the spatial patterns. Corrections for individual radar images are important in this regard as many of these corruptions are temporary and cannot be identified by a long term analysis.

The simulation results on small spatial scales are dominated by the ability of the model to reproduce the topography and the dynamical processes of precipitation. The high resolution simulations with 5 km and 1 km grid size reveals a much better representation of the small-scale spatial rain patterns and rain structure than the 15 km grid size results do.

Comparison of radar data and simulations on high temporal and spatial resolution

Weather radar data and the simulations with the highest resolutions provide almost the same resolution of precipitation data with 1 km grid size and 5-minute time step. To compare the performance of both datasets on these scales, the Probability Density Function (PDF) is calculated for the simulations URB-1, NURB-1 as well as for corrected radar data RAD-1 and uncorrected radar data RADU-1 for $D-North$ and the entire domain of $D-South$. The latter dataset has already been applied in Chapter 3 and Chapter 4. In Fig. 4.17(a) the uncorrected radar data reveal some constraints for small rain intensities compared to rain gauges but show a very good accordance for higher rain intensities. The corrected radar data minimizes the shortcomings for small rain intensities but on the expense of a weaker performance for higher rain intensities. The QQ-Plot (Fig. 4.17a) clearly shows an underestimation of high rain intensities for this dataset. These findings are valid for Lower Saxony and therefore $D-North$. Figure 7.1 shows the PDF curves for both simulations and both radar datasets based on 1 km grid size and 5-minute time step in both domains for summer 2005 including the dry 5-minute probability.

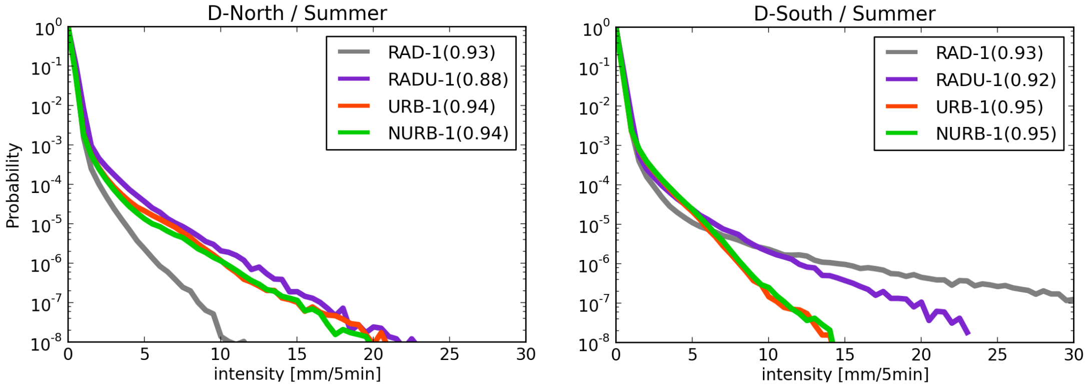

Figure 7.1: Probability Density Functions (PDF) for the simulations URB-1, NURB-1, corrected radar data RAD-1 and uncorrected radar data RADU-1 based on 1 km grid size and 5-minute time step for the summer months in $D-North$ and $D-South$. The dry 5-minute probability of each dataset is added to the legend in brackets.

The dry period probabilities of simulation results and corrected radar data match very well in both domains. Slightly more dry time steps are simulated than observed by radar data. RADU-1 shows less dry time steps in $D-North$. This may partly be explained by clutter influence for dry time steps. URB-1 and NURB-1 show the similar curves in both domains. Slightly higher probabilities for higher rain intensities are observed for URB-1 in $D-North$. The simulations reveal an almost identical behavior for lower rain intensities as RAD-1. RADU-1 reveals higher probabilities, there. In $D-North$, the PDF curves for higher rain intensities of both simulations and RADU-1 match well. Transferring the results from the QQ-Plots in Fig. 4.17(a) to this figure implies that the simulation results are able to represent the rain intensity distribution very well. Comparable results become apparent for $D-South$. Only the higher rain intensities of simulation data are lower, but the reliability of intense precipitation measurements based on radar data is also uncertain in complex terrain.

8 Conclusions and outlook

8.1 Conclusions

The specific research questions for this investigation were (Chapter 1.3):

(i) Which types of systematic effects in radar data do exist and what is their influence on data quality?

(ii) How can systematic effects in radar data be corrected and which quality can be achieved in this way?

(iii) What is the added value of high resolution modeling?

(iv) What is the impact of key land surface properties such as topography or a more complex representation of urban areas on the simulation results?

(v) Are weather radar measurements and high resolution Regional Climate Models (RCMs) able to represent the spatiotemporal variability of precipitation in a suitable way?

Measurements – weather radar data

(i) The question of quality is closely linked to the limitations of possible datasets. For precipitation measurements the ’circumstances’ of the measurements are important, as already mentioned in Chapter 1. Systematic effects in datasets are often not noticeable in single radar images. But they can lead to major sources of errors on a longer temporal scale and also affect single radar images. Based on accumulated radar images, a variety of shortcomings (dependence of reflectivities on altitude, shading effects, clutter, etc.) in single radar data and composite radar data have been identified. It is possible to describe the dependence of systematic effects on time and on reflectivity level. Radar measurements’ dependence on altitude is the most conspicuous feature in accumulated radar images. It can be regarded as the result of radar measurements at different altitudes associated with the natural variability of reflectivity with height. The latter one is due to the continuous process of melting, coagulation, evaporation and the high dependence of radar reflectivity on the aggregation state of the hydrometeor. In Bartels et al. [2004] the importance and the need for such a correction of the Vertical Profile of Reflectivity (VPR) for the German Met. Service (DWD) has already been stated. But no correction scheme has been established so that these effects still remain in

DWD's radar products.

(ii) The net effect of these limitations can be corrected by a statistical correction that is based on the findings mentioned above. At the current stage, the presented correction scheme is only based on statistical corrections. It should not replace common single radar image corrections. It rather offers an additional possibility to correct systematic variations. The algorithm shows one possible option to realize this. Even when only the variations due to the different beam elevations of one radar site are equalized and a mean (e.g. monthly) altitude correction factor is applied, radar data quality also of individual radar images can be improved. A combination with corrections for individual radar images regarding VPR and attenuation effects [Krämer and Verworn, 2008] is promising. The development and application of this correction algorithm show the limits of such a correction and whether this mean statistical correction is also applicable on a short temporal scale.

The presented correction scheme provides reasonable results for long term data but shows some constraints for very high temporal resolutions and heterogenic rain events. For weather radar data, the most critical part is the correction of the variation of rain amounts with altitude. The presented correction scheme based on a linear regression is not able to correct all effects (e.g. bright band effect) and it is only an approximation to reality. An additional feedback mechanism is useful to adjust static correction factors to actual conditions. The deduction of this feedback correction based on real patterns seems promising and allows the further temporal disaggregation of static correction factors in principle. But it only works properly if a sufficient coverage with measurements exists.

The question may arise, how far it is reasonable to downscale such a correction. The best answer might be: 'A reasonable correction is ensured, as long as the correction scheme is able to capture the variability of the respective systematic effect.' If the variation of the effect is small compared to the magnitude of this effect, a mean correction will usually improve data quality, otherwise a deterioration of data quality is possible. In the present case, the plain application of the annual altitude correction factor for instance, will probably provoke a significant overestimation of rain rates at far ranges from the radar site in the summer months and is therefore not advisable. The presented modified correction scheme in Chapter 4.3.1 has reached its limits, when the feedback corrections are not reliable anymore.

Furthermore, it is inevitable to take into account the pure limits of the measuring technique besides the systematic effects such as measurements at higher altitudes in winter. A statistical correction of radar measurements is only meaningful as long as the underlying precipitation patterns are represented adequately. Or, seen from another perspective: the investigation of accumulated radar images and the correction on a short temporal scale have the potential to identify shortcomings in radar data that cannot adequately be corrected. On a longer temporal scale most effects can be corrected as shown in Chapter 4. But when single radar images are statistically corrected, missing measurements at far ranges due to overshooting in winter may lead to underestimations. In this way, the correction algorithm is also able to identify uncorrectable limitations in radar data.

Simulations with WRF

Weather radar data is already accepted as one possible option to serve as input to hydrological models or as a validation dataset. In comparison to radar data, model simulations have to prove their ability to represent high resolution precipitation data.

(iii) The grid size of a model is a central issue that can significantly influence simulation results. It determines which processes can explicitly be resolved and which ones have to be parametrized. The quality of the representation of topography and land use is also dominated by the spatial resolution. The presented results clearly show that a grid size of 15 km is usually not sufficient to produce reliable precipitation patterns. Temporal and spatial deficiencies are observed especially for higher temporal resolutions and in complex terrain.

Simulations with both 5 km and 1 km grid size should be regarded as high resolution simulations and are superior to the simulation with 15 km grid size regarding the precipitation results. They show a similar level of quality and reveal a good accordance with observational datasets for temporal and spatial statistical measures.

The added value of the 1 km grid size compared to the 5 km grid size becomes apparent for the dry period probability. An excellent accordance with observational datasets is observed. But the minor deviations of this measure based on simulations with 5 km grid size and cumulus parametrization can be reduced by omitting convection parametrization. The Fractions Skill Score evaluates the ability of simulations to represent the spatial patterns of validation datasets. It also shows an improvement for all rain intensities and horizontal scales for the 1 km simulations. Especially for higher rain intensities, simulations on smaller grid sizes (1 km) are superior to coarser ones (5 km).

Furthermore, it depends on the investigation area and the prevailing rain type which grid size should be used at least. The simulation of convective events in complex terrain, for example, requires very small grid sizes. The added value of the 1 km grid size which is presented in this study may only be important to specific applications. For many applications less computational expensive simulations (e.g. with a grid size of 5 km) may provide a sufficient data quality.

Reasonable results in the gray zone of convection (here: 5 km) are achieved by a new convection parametrization scheme as well as by the explicit calculation of convection. Spatial patterns and even the diurnal cycle are well represented by both simulations. The benefit of simulations with convection parametrization is the representation of the distribution of rain intensities (PDF). The advantages of simulations with explicit modeling of convection are observed for the dry period probability and slightly for spatial patterns. Consequently, for simulations in this gray zone of convection it is worth to test both approaches and to evaluate their performance regarding e.g. the representation of the diurnal cycle, the dry period probability and the spatial correlations. These grid sizes are still interesting (see Chapter 1) especially for climate simulations over several years.

(iv) WRF (Weather Research and Forecasting Model) simulations seem to produce highest rain amounts at the highest altitudes. For the complex terrain in $D-South$ a high consistency of precipitation maxima and altitude is observed whereas in $D-North$ this accordance is less pronounced in the observations. Consequently, the locations of precipitation maxima are much better captured in $D-South$ than in $D-North$ by the WRF simulations.

Even though, the locations of precipitation maxima are very well captured in $D-South$, the simulation results of 1 km grid size reveal smaller values in the correlograms than observed. Additionally, the minima and maxima in complex terrain are too pronounced. Orographic forcing effects might be responsible for this behavior and a further examination of this effect in future would be worthwhile. The PBL (Planetary Boundary Layer) scheme might influence these results. The 1 km grid size is in the gray zone of unresolved vertical transport. Partly resolved vertical transport in combination with a parametrization scheme that is designed for larger grid sizes might also affect precipitation results.

A further option to improve simulation results is to apply additional modules that represent the land use and the specific meteorological conditions more exactly, such as Urban Canopy Models (UCM). The application of an UCM is able to influence precipitation patterns in the whole domain, although the overall performance of simulations with and without UCMs is almost similar according to the presented results. Minor improvements in simulations with UCM become apparent only regarding the size and the correlation of spatial patterns.

A larger added value for urban grid cells is very likely but this is not part of this investigation. The usage of an UCM for grid sizes of 5 km and above is not common. The parametrization of many important processes such as convection or turbulent movements by PBL schemes on one hand and the application of a sophisticated UCM on the other hand may sound strange. But especially for domains with a high degree of urban grid cells, an UCM may be superior compared to the bulk parametrization of these urban grid cells and should not be excluded from the very beginning.

A huge effort can be made in creating a city-specific 'URBPARM.tbl' with parameters representing mean road widths or building heights as input for an UCM (e.g. Fallmann [2014]). The more sophisticated BEP model can be applied that significantly slows down simulations. For the investigation of specific effects in cities, this effort seems to be justified. Here, the precipitation patterns in the whole domain occupy center stage and tests with the BEP do not show improvements. So, the Single-Layer Urban Canopy Model (SLUCM) is applied here with the city-specific URBPARM.tbl which is adopted from Stuttgart [Fallmann, 2014]. The CORINE land use data [European Environment Agency, 2007] is applied for urban pixels. For the residual pixels the USGS dataset remains unmodified to provide a comparable setup and database as the long-term reanalysis within the SYNOPSE project (see Chapter 1). But more recent land use data may significantly influence simulations results and are further possibilities to improve model results [Tölle et al., 2014].

Representing the spatiotemporal variability of precipitation

(v) Both, weather radar and dynamical downscaling with WRF show a high potential to represent the spatiotemporal variability of precipitation, even for very high spatial and temporal resolutions. Nonetheless, each dataset has its respective advantages and disadvantages.

Weather radar measurements show a good performance on a short temporal scale, in summer and for small scale precipitation patterns. In winter and on a large temporal and spatial scale, systematic measuring effects as well as clutter or shading effects may add up and lead to inconsistencies of data quality in space. The calculation of rain rate from radar reflectivity is reasonable for light and moderate rain, at least when adjusted to rain gauges [Morin et al., 2003]. Whereas the accuracy of radar derived extreme precipitation values is rather weak.

Current RCMs such as WRF are also able to represent the general temporal and spatial characteristics of precipitation well. But the simulation of the exact temporal accordance with rain gauge measurements and the location of precipitation maxima should be improved. A large improvement can already be observed for high resolution simulations compared to coarser grid sizes.

Regarding these aspects, radar data outperforms dynamic modeling results. A combination of both systems can be achieved by the assimilation of radar data [Juanzhen and Wang, 2013, Sokol, 2011] to adopt precipitation patterns from measurements. But this reduces the applicability of model simulations. The simulation of precipitation is possible for any region and period worldwide without any measurement failures and measuring effects, as long as climate data or reanalysis data is available.

Even if disadvantages or limitations in precipitation datasets exist, this would not restrain their application, as long as these effects can be corrected. Radar data has to be corrected in advance as mentioned in Chapter 1 and Chapter 4. The presented simulation results are uncorrected, so data quality may be further improved. A possible systematic underestimation or overestimation of rain amounts is not desirable and a hint for limitations in the model structure, in the model setup or in the input data, but this can usually be corrected in a reasonable way [Cannon et al., 2015]. For many applications this correction provides reliable results, even though it is an interference into the model continuum. Limitations of precipitation patterns or the diurnal cycle are more difficult as they can hardly be corrected. In this respect, the presented results of simulations with high resolutions are very promising.

8.2 Outlook

How can weather radar data be improved?

The presented investigation of radar data on a long temporal scale shows the importance and the necessity to consider systematic effects in radar data. For the usage of radar data in future, a simple accumulation of radar images can serve as a first hint for systematic effects. DWD's RW-product[1], for instance, is widely used for quantitative analysis or validation purposes (e.g. Kühnlein et al. [2014]) and provides a good data quality for most areas of Germany. In Fig. 8.1, the annual rain amount of accumulated radar images of the RW-product (right) is opposed to REGNIE rain amounts (left) for the northern investigation area of WRF simulations $D-North$ from June 2005 until May 2006. A significant overestimation of rain amounts near the radar Hamburg is observed, probably caused by clutter effects. This figure shows, that the RW-product is not suitable to serve as a validation dataset in this area. Some of these effects may exist only temporary or periodically, so that in general, the respective investigation area

---

[1] Radar rain rates after adjustment with the weighted mean from two standard procedures [DWD, 2016b].

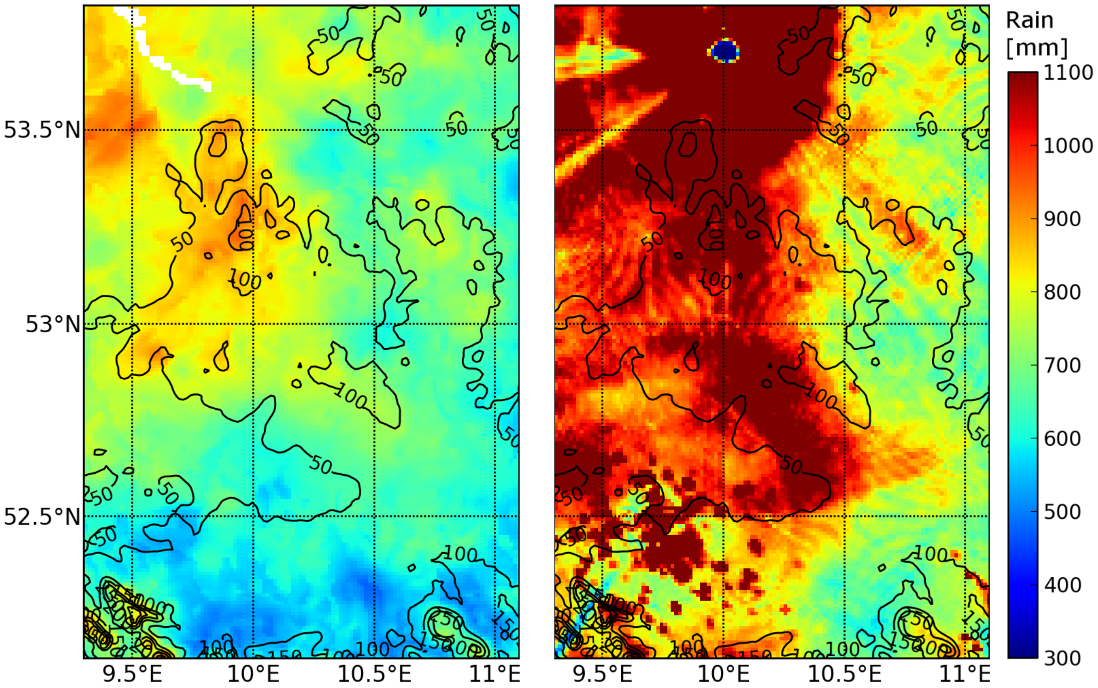

Figure 8.1: Distribution of annual rain amounts (June 2005 to May 2006) of REGNIE (left; same as Fig. 6.2(a) and the RW-product (right) for $D-North$ with 1 km grid size overplotted by contour lines of elevation.

and the regarded time span should be analyzed before using a radar dataset.

Additional corrections of radar data are useful, especially from autumn till spring, due to the dependence of radar based rain amounts on altitude. A mean annual correction is not able to capture the high variability of this dependence. The presented approach to analyze and use current measured precipitation patterns to derive correction factors is promising, although improvements regarding robustness or a higher variability (correction factors for sub-daily time spans) are desirable. Alternatively, VPR correction schemes on radar images of individual time steps have to be developed and applied.

Furthermore, there are effects in radar data, that are difficult to correct solely relying on radar data such as the simultaneous measurement of rain and snow, 'overhanging precipitation' that never reaches the ground or dislocated precipitation due to wind influence. A combination of different measuring techniques to merge the respective advantages and also that serves as a plausibility check, is preferable [Bartels et al., 2004, Vogl et al., 2012]. It is conceivable to extend one of the radar merging products within RADOLAN[2] by an additional module that corrects systematic effects. Important corrections should be applied as early as possible in the processing chain. E.g. a skilled adjustment and merging of radar data and rain gauge data within RADOLAN may reduce the systematic differences between areas close to and far from the radar site. But a correction to minimize those systematic effects in advance is preferable.

---

[2] Radar Online Adjustment; DWD project.

Besides the correction of data, progress in producing data with higher qualities are desirable. The development of new techniques such as polarization for weather radar measurements proceeds and offers a variety of possible improvements of radar data quality. New parameters such as K_{DP} and Z_{DR} (see Chapter 2) are able to improve the detection of clutter, the identification of the melting layer (bright band) and the calculation of rain rate from radar measurements, but only minor improvements can be expected regarding the dependence of radar measurements on altitude.

Technical progress and promising approaches to enhance the quality of precipitation simulations

Technical progress is also of high importance for model simulations. The enhancement of computer hardware and storage capacities allow high resolution simulations or dynamical downscaling in the kilometer range and below, aiming at resolving explicitly critical dynamical processes such as convection and representing complex terrain and land use in detail. These are opportunities one should exploit.

For this study, the grid sizes of 1 km and of 5 km have been prerequisites within the SYNOPSE project to downscale long time series (see Chapter 1). These grid sizes are important to analyze the added value of two high resolution simulations. If the main emphasis is laid only on the high resolution simulations, a grid size of 10 to 15 km for domain 1 and 2 km or 3 km for domain 2 would avoid the gray zones of convection. In addition, it would save computational costs and probably lead to similar results.

The circumvention of gray zones (convection or vertical fluxes) seems to be a feasible way to minimize unwanted side effects. Nevertheless, the further development of parametrization schemes that are also designed for these gray zones is of utmost importance. On the one hand, a compromise between calculation time and spatial resolution for longer simulations often leads to grid sizes within the gray zone of convection. On the other hand, even below 4 km an explicit simulation of convection is not able to resolve the respective processes entirely [Prein et al., 2015]. This is also true for PBL schemes in a similar way. For future investigations new parametrization schemes that are designed even for gray zones should be tested (e.g. Shin and Hong [2013]).

Regarding the limited accordance of simulated precipitation patterns with observational datasets especially in the summer season, an improvement may be achieved by spectral nudging even for the finer grid sizes. The adoption of patterns of large scale processes for the small domains of high resolution simulations may help to improve data quality. This is not performed within this framework, but it seems to be an interesting and promising approach.

Another aspect that requires further investigation, is the overestimation of rain amounts and the simultaneous underestimation of spatial correlations of simulations with 1 km grid size in complex terrain. A search for reasons for this behavior including the analysis of radiation terms and also tests with new PBL schemes are necessary to really gain from the higher spatial resolutions. Brisson et al. [2016] showed some shortcomings of cloud cover in their high resolution simulation results with COSMO-CLM. It is conceivable, that this might also influence the presented results and should be taken into account in future simulations.

What is the cost of quality?

Apart from the specific results and conclusions presented in this framework, some general aspects of precipitation datasets shall be addressed:

There is no perfect precipitation dataset especially when spatial patterns are paramount. So, working with precipitation data means accepting compromises. One central question of the application of precipitation datasets is: which quality of precipitation data is necessary? The answer to this question is closely linked to the intended purpose of the respective analysis one aims at. Are for example the exact rain amounts at certain places or the precipitation patterns most important? For which season and area will precipitation data be used and for how long? Consequently, the user of precipitation data should be aware which precipitation characteristics are important and which limitations are still acceptable for the respective work. In addition, the characteristics of different precipitation datasets should be addressed to choose the appropriate dataset. In this context, several validation datasets are used in Chapter 6 to reveal advantages of certain datasets despite of apparent deficiencies and to identify hidden limitations.

Summarized, the intended purpose of the application of precipitation data determines the dataset and the expense of simulations or additional corrections of measurements.

The presented study may help to make users of precipitation data aware of the possibilities and the constraints of those datasets. The section of precipitation simulations reveals the status quo of high resolution simulations by dynamical downscaling. The analysis of the influence of various grid sizes, certain modules and parametrizations in areas with different terrain on the simulation results may serve as a guideline for prospective investigations.

APPENDICES

A Preparation of a spatially high resolution precipitation dataset

Monthly rain gauge data for the time span 2005 to 2009 are quality checked, aggregated to mean annual values and finally interpolated to the same grid as the composite radar data. An external drift-kriging technique based on altitude as drift variable is applied. The GSLIB libraries [Deutsch and Journel, 1992] are used for this task. Rainfall data usually have a high variation in time and space, but on a longer temporal scale, rain amounts depends primarily on altitude. So, annual values of gauge data are interpolated. The comparison with interpolated monthly data leads to similar results. The dependence of rain amounts on absolute values of the altitude tends to overestimate the rain amounts at the summit locations. Better results can be achieved by using the square root of the altitude as drift-variable [Beck, 2013]. The rain gauge data are converted to logarithmic scale before being interpolated to get at least log-normally distributed values. This is not a mandatory precondition, but the quality and the reliability of the results of the interpolation is improved. Additionally, the estimation of the kriging variance and the calculation of the confidence interval are deducible if the data is normally distributed. The quality of the interpolation results is closely related to the quality of the rain gauge data and to the density of the monitoring network. The patterns of the topography dominate the interpolated rain patterns especially if the network's density is low. Naturally induced variations of rain amount like windward and leeward effects cannot be adequately considered. Therefore, spatial patterns of interpolated data usually have limitations.

To validate the results of interpolation, the RMSE of the 516 independent rain gauges used for validation are opposed to the corresponding pixels (9-pixel-value) of the interpolated annual rain field. Figure A.1 shows the scatterplots for each year of the investigation period 2005 to 2009 and the respective RMSE value.

According to Fig. A.1 a close relationship between measurements and interpolated data become apparent. The data basis for interpolation improves continuously from 255 rain gauges in 2005 to 668 in 2009. Despite these differences, the interpolation results are quite similar ranging from 78 mm to 105 mm for the RMSE. The highest RMSE (105 mm) is calculated for the wettest year (2007). The annual rain amounts seem to have more effect on the RMSE value than the number of contributing rain gauges. Due to these results, the quality of interpolated rain amounts is almost similar for each year and it is sufficient to serve as a validation dataset.

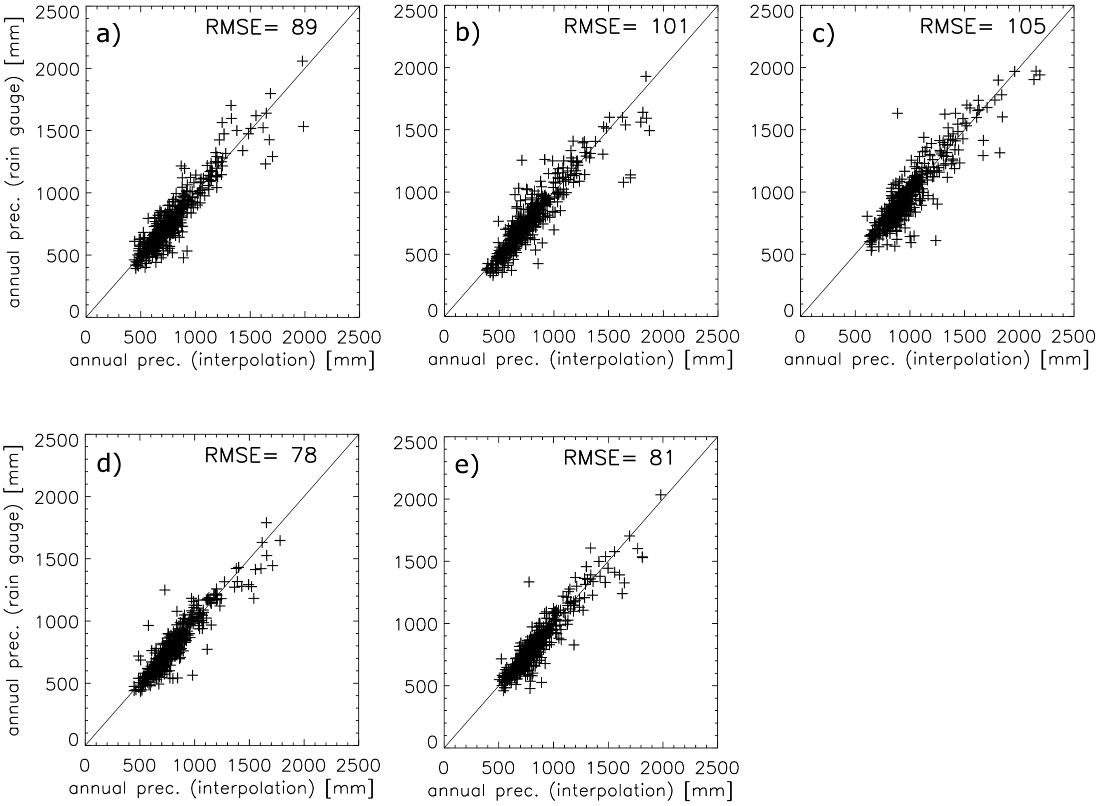

Figure A.1: Scatterplots of annual rain amounts for the 516 independent rain gauges used for validation and the corresponding pixels (9-pixel-value) of the interpolated rain field for (a) 2005, (b) 2006, (c) 2007, (d) 2008 and (e) 2009.

B Procedure to adopt CORINE data as land use information

The following procedure describes one possible option to transfer CORINE land use data [European Environment Agency, 2007] to the binary files in WRF (GEOGRID):

- Re-project CORINE-data to tif-format with GDALWARP
- Reclassify the urban CORINE-data to the three urban classes of the UCMs (Urban Canopy Models) with PYTHON
- Change resolution with GDALWARP, if desired.
- Transform tif-format to the binary format used by GEOGRID (in WPS pre-processing):

  For example: ConvLANDUSE[1] or Convert_geotiff[2]

  The resulting files are in binary code plus one additional index textfile.
- These files are transferred to the '../WPS/geog'-folder and the GEOGRID.TBL is modified to link to that data base, if all classes should be adopted.
- If only the urban classes have to be adopted from CORINE data:
  - The index textfile is modified in the following way:
    'category_min =31' and 'category_max =33'
  - Then these files are transferred to the '../WPS/geog'-folder.
  - The GEOGRID.TBL has to be modified: an additional data block 'name_LANDUSEF' has to be augmented with the respective link to the new dataset and a 'priority=2'.
- Finally, the pre-processing GEOGRID has to be performed in the common way.
- A final visual check of the two- and three-dimensional land use data is advisable.

---

[1] http://dl.dropbox.com/u/58758735/HighResgeogrid.tar.gz. Accessed 19 January 2017

[2] https://github.com/openwfm/convert_geotiff. Accessed 19 January 2017

Bibliography

Atlas, D. (1990). *Radar in Meteorology - Battan Memorial and 40th Anniversary Radar Meteorology Conference*, page 806. Amer. Meteor. Soc.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T. (2011). Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Weather Rev.*, 139(12):3887–3905.

BALTRAD (2014). *http://baltrad.eu.* [Online; accessed 19 January 2017].

Ban, N., Schmidli, J., and Schär, C. (2014). Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res. Atmos.*, 119:7889–7907.

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

Battan, L. J. (1973). *Radar Observation of the Atmosphere*, page 323. University of Chicago Press.

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J.-P., Beljaars, A., and Bormann, N. (2014). Representing equilibrium and nonequilibrium convection in large-scale models. *J. Atmos. Sci.*, 71(2):734–753.

Beck, C., Grieser, J., Rudolf, B., and Schneider, U. (2005). Globale Niederschlagsanalysen des Weltzentrums für Niederschlagsklimatologie für den Zeitraum ab 1951 (Analysis Products of the Global Precipitation Climatology Centre for the Period Since 1951). *Promet 31*, 2–4:191–196.

Beck, F. (2013). Generation of spatially correlated synthetic rainfall time series in high temporal resolution: a data driven approach. *Mitteilungen / Institut für Wasser- und Umweltsystemmodellierung, Universität Stuttgart.*

Berg, P., Wagner, S., Kunstmann, H., and Schädler, G. (2013). High resolution regional climate model simulations for Germany: Part I-Validation. *Clim. Dyn.*, 40(1–2):401–414.

Berne, A., Delrieu, G., Creutin, J.-D., and Obled, C. (2004). Temporal and spatial resolution of rainfall measurements required for urban hydrology. *J. Hydrol.*, 299:166–179.

Beven, K. J. (2011). *Rainfall-runoff modelling: the primer*, page 457. John Wiley & Sons.

Bougeault, P. and Lacarrère, P. (1989). Parameterization of orography-induced turbulence in a mesobeta-scale model. *Mon. Weather Rev.*, 117:1872–1890.

Bárdossy, A. and Pegram, G. (2013). Interpolation of precipitation under topographic influence at different time scales. *Water Resour. Res.*, 49:4545–4565.

Bringi, V. N. and Chandrasekar, V. (2001). *Polarimetric Doppler Weather Radar: Principles and Applications*, page 636. Cambridge University Press.

Brisson, E., Demuzere, M., and van Lipzig, N. P. M. (2015). Modelling strategies for performing convective permitting climate simulations. *Meteorol. Z.*

Brisson, E., Van Weverberg, K., Demuzere, M., Devis, A., Saeed, S., Stengel, M., and van Lipzig, N. P. M. (2016). How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics? *Clim. Dyn.*, 47(9):3043–3061.

Brown, B., Bullock, R., Fowler, T., Halley Gotway, J., Newman, K., and Jensen, T. (2016). Model Evaluation Tools version 5.2 (METv5.2). *User's guide 5.2. Developmental Testbed Center Rep.*, page 328. Available at: http://www.dtcenter.org/met/users/docs/users_guide/MET_Users_Guide_v5.2.pdf. [Online; accessed 19 January 2017].

Cannon, A. J., Sobie, S. R., and Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: How well do methods preserve changes in quantiles and extremes? *J. Clim.*, 28:6938–6959.

Carbone, R. E. and Tuttle, J., D. (2008). Rainfall occurrence in the U.S. warm season: the diurnal cycle. *J. Clim.*, 21:4132–4146.

CDO (2015). *Climate Data Operators. http://www.mpimet.mpg.de/cdo.* [Online; accessed 19 January 2017].

Chan, S., Kendon, E., Fowler, H., Blenkinsop, S., Ferro, C. A., and Stephenson, D. B. (2013). Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Clim. Dyn.*, 41(5–6):1475–1495.

Chen, F. and Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part II: Preliminary model validation. *Mon. Weather Rev.*, 129:587–604.

Chen, F., Kusaka, H., Bornstein, R., et al. (2011). The integrated WRF/urban modelling system: development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.*, 31:273–288.

Ching, J., Rotunno, R., LeMone, M. A., Martilli, A., Kosovic, B., Jimenez, P. A., and Dudhia, J. (2014). Convectively induced secondary circulations in fine-grid mesoscale numerical weather prediction models. *Mon. Weather Rev.*, 142:3284–3302.

Chwala, C., Gmeiner, A., Qiu, W., Hipp, S., Nienaber, D., Siart, U., Eibert, T., Pohl, M., Seltmann, J., Fritz, J., and Kunstmann, H. (2012). Precipitation observation using microwave backhaul links in the alpine and pre-alpine region of southern Germany. *Hydrol. Earth Syst. Sci.*, 16(8):2647–2661.

Clark, A. J., Gallus, W. A., and Chen, T. C. (2007). Comparison of the diurnal precipitation cycle in convection-resolving and non-convection-resolving mesoscale models. *Mon. Weather Rev.*, 135:3456–3473.

Collier, C. G. (1989). *Applications of Weather Radar Systems/: a Guide to Uses of Radar Data in Meteorology and Hydrology*, page 406. Ellis Horwood Limited, Chichester, England.

Collier, C. G. (1992). *International Weather Radar Networking. Final Seminar of the COST Project 73*, page 332. Kluwer Academic Publishers, Dordrecht.

COST (2017). *COST program, www.cost.eu/COST_Actions/.* [Online; accessed 19 January 2017].

de Ela, R., Laprise, R., and Denis, B. (2002). Forecasting skill limits of nested, limited-area models: A perfect-model approach. *Mon. Weather Rev.*, 130(8):2006–2023.

de Rooy, W. C., Bechtold, P., Fröhlich, K., Hohenegger, C., Jonker, H., Mironov, D., Pier Siebesma, A., Teixeira, J., and Yano, J.-I. (2013). Entrainment and detrainment in cumulus convection: An overview. *Quart. J. Roy. Meteor. Soc.*, 139(670):1–19.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The era-interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, 137:553–597.

Deutsch, C. V. and Journel, A. G. (1992). *GSLIB: Geostatistical Software Library and User's Guide*, page 340. Oxford University Press, New York, NY.

Donner, L. J. et al. (2011). The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Clim.*, 24(13):3484–3519.

DWD (2016a). *Radar products DWD, http://www.dwd.de/DE/leistungen/radarprodukte/radarprodukte.html.* [Online; accessed 19 January 2017].

DWD (2016b). *RADOLAN products, http://www.dwd.de/DE/leistungen/radolan/produktuebersicht/radolan_produktuebersicht.html.* [Online; accessed 19 January 2017].

DWD (2017). *RADOLAN changes, http://www.dwd.de/DE/leistungen/radolan/radolan_info/radolan_change_management_pdf.* [Online; accessed 19 January 2017].

Dyck, S. and Peschke, G. (1995). *Grundlagen der Hydrologie*, page 386. VEB Verlag für Bauwesen, Berlin.

Einfalt, T., Hatzfeld, F., Wagner, A., Seltmann, J., Castro, D., and Frerichs, S. (2009). URBAS: forecasting and management of flash floods in urban areas. *Urban Water Journal*, 6:5:369–374.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D. (2003). Implementation of Noah land surface model advancements in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, 108(D22):8851.

European Environment Agency (2007). CLC2006 technical guidelines. *Technical report No 17/2007. Available at: http://www.eea.europa.eu/publications/technical_report_2007_17/at_download/file*, 19. [Online; accessed 19 January 2017].

Fabry, F., Cazenave, Q., and Basivi, R. (2013). Echo climatology, impact of cities, and initial convection studies: New horizons opened using 17 years of Conterminous US radar composites. *36th Int. Conf. on Radar Meteor., Breckenridge, CO, AMS.*

Fabry, F. and Zawadzki, I. (1995). Long-term radar observations of the melting layer of precipitation and their interpretation. *J. Atmos. Sci.*, 52:838–851.

Fairman, J. G., Schultz, D. M., Kirshbaum, D. J., Gray, S. L., and Barrett, A. I. (2015). A radar-based rainfall climatology of Great Britain and Ireland. *Weather*, 70:153–158.

Fallmann, J. (2014). *Numerical simulations to assess the effect of urban heat island mitigation strategies on regional air quality.* PhD thesis. University of Cologne. Available at: http://kups.ub.uni-koeln.de/id/eprint/5913. [Online; accessed 19 January 2017].

Faures, J. M., Goodrich, D. C., Woolhiser, D. A., and Sorooshian, S. (1995). Impact of small-scale spatial variability on runoff modeling. *J. Hydrol.*, 173:309–326.

Feser, F. (2006). Enhanced detectability of added value in limited-area model results separated into different spatial scales. *Mon. Weather Rev.*, 134(8):2180–2190.

Fosser, G., Khodayar, S., and Berg, P. (2015). Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Clim. Dyn.*, 44:45–60.

Franco, M., Sanchez-Diezma, R., and Sempere-Torres, D. (2006). Improvements in weather radar rain rate estimates using a method for identifying the vertical profile of reflectivity from volume radar scans. *Meteorol. Z.*, 15(5):521–536.

Frech, M. (2013). Monitoring the data quality of the new polarimetric weather radar network of the German Meteorological Service. *36th Int. Conf. on Radar Meteor., Breckenridge, CO, AMS.*

Frei, C. and Schär, C. (1998). A precipitation climatology of the alps from high-resolution rain-gauge observations. *Int. J. Climatol.*, 18(8):873–900.

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

Gabella, M. and Notarpietro, R. (2002). Ground clutter characterization and elimination in mountainous terrain. *In Use of radar observations in hydrological and NWP models*, pages 305–311.

Goovaerts, P. (1998). Ordinary cokriging revisited. *Mathematical Geology*, 30(1):21–42.

Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *J. Hydrol.*, 228(1):113–129.

GPM (2013). *Precipitation measurement missions.* Global Precipitation Measurement Program, NASA, http://gpm.nasa.gov/. [Online; accessed 19 January 2017].

Grell, G., Schade, L., Knoche, R., Pfeiffer, A., and Egger, J. (2000). Nonhydrostatic climate simulations of precipitation over complex terrain. *J. Geophys. Res.*, 105:29595–29608.

Grell, G. A. and Freitas, S. R. (2014). A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, 14:5233–5250.

Hand, W. H. (1996). An object-oriented technique for nowcasting heavy showers and thunderstorms. *Meteor. Appl.*, 3:31–41.

Haylock, M., Hofstra, N., Klein Tank, A., Klok, E., Jones, P., and New, M. (2008). A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950-2006. *J. Geophys. Res.*, 113(D20119).

Heistermann, M., Jacobi, S., and Pfaff, T. (2013). Technical Note: An open source library for processing weather radar data (wradlib). *Hydrol. Earth Syst. Sci.*, 17:863–871.

Hohenegger, C., Brockhaus, P., and Schär, C. (2008). Towards climate simulations at cloud-resolving scales. *Meteorol. Z.*, 17(4):383–394.

Holleman, I. (2007). Bias adjustment and long-term verification of radar-based precipitation estimates. *Meteor. Appl.*, 14:195–203.

Holleman, I., Huuskonen, A., Kurri, M., and Beekhuis, H. (2010). Operational Monitoring of Weather Radar Receiving Chain Using the Sun. *J. Atmos. Ocean. Technol.*, 27:159–166.

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. *Bull. Amer. Meteorol. Soc.*, 95:701–722.

Isotta, F. A. et al. (2014). The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int. J. Climatol.*, 34(5):1657–1675.

Janjic, Z. I. (1994). The step-mountain eta coordinate: further development of the convection, viscous sublayer, and turbulent closure schemes. *Mon. Weather Rev.*, 122:927–945.

Joss, J. and Lee, R. (1995). The application of radar-gauge comparisons to operational precipitation profile corrections. *J. Appl. Meteorol.*, 34:2612–2630.

Juanzhen, S. and Wang, H. (2013). Wrf-arw variational storm–scale data assimilation: current capabilities and future developments. *Advances in Meteorology*, 2013(815910):13.

Keefer, T. O. (2003). *Precipitation Simulation Models*, pages 729–733. Encyclopedia of Water Science.

Kendon, E. J., Roberts, N. M., Senior, C. A., and Roberts, M. J. (2012). Realism of rainfall in a very high-resolution regional climate model. *J. Clim.*, 25(17):5791–5806.

Kühnlein, M., Appelhans, T., Thies, B., and Nauss, T. (2014). Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—a random forests-based approach applied to msg seviri. *Remote Sens. Environ.*, 141:129–143.

Kitchen, M., Brown, R., and Davies, A. G. (1994). Real-time correction of weather radar data for the effects of bright band, range and orographic growth in widespread precipitation. *Quart. J. Roy. Meteor. Soc.*, 120:1231–1254.

Knight, C. G. et al. (2007). Association of parameter, software, and hardware variation with large-scale behavior across 57,000 climate models. *Proc. Natl. Acad. Sci.*, 104(30):12259–12264.

Koistinen, J. (1991). Operational correction of radar rainfall errors due to vertical reflectivity profile. *Proc. 25th Conf. on Radar Met.*

Koistinen, J., Pohjola, H., and Hohti, H. (2003). Vertical reflectivity profile classification and correction in radar composites in Finland. *in Preprints AMS 31st Int. Conf. on Radar Met.*

Krajewski, W. F., Vignal, B., Seo, B.-C., and Villarini, G. (2010). Statistical model of the range dependent error in radar-rainfall estimates due to the vertical profile of reflectivity. *J. Hydrol.*, 402:306–316.

Krämer, S., Massmann, S., Schmidt, N., Fuchs, L., Schröder, K., Sympher, K.-J., Rohde, S., Grosskopf, K.-I., Kuchenbecker, A., Callau, A., Müller, H., Haberlandt, U., Mosthaf, T., Müller, T., Lorenz, M., Wagner, A., and Wagner, S. (2016). Synthetische Niederschlagszeitreihen – Alternative zu Niederschlagsmessungen – Untersuchung an großstädtischen Kanalnetzen. *AQUA & GAS*, 10/2016:38–42.

Krämer, S. and Verworn, H. R. (2008). Improved C-band radar data processing for real time control of urban drainage systems. *11th International Conference on Urban Drainage, Edinburgh, Scotland, UK.*

Krige, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. of the Chem., Metal. and Mining Soc. of South Africa*, 52(6):119–139.

Kunstmann, H., Schneider, K., Forkel, R., and Knoche, R. (2004). Impact analysis of climate change for an alpine catchment using high resolution dynamic downscaling of echam4 time slices. *Hydrol. Earth Syst. Sci.*, 8:1031–1045.

Kunstmann, H. and Stadler, C. (2005). High resolution distributed atmospheric-hydrological modelling for alpine catchments. *J. Hydrol.*, 314:105–124.

Kusaka, H. and Kimura, F. (2004). Coupling a single-layer urban canopy model with a simple atmospheric model: impact on urban heat island simulation for an idealized case. *Journal of the Meteorological Society of Japan*, 82:67–80.

Kusaka, H., Kondo, H., Kikegawa, Y., and Kimura, F. (2001). A simple single-layer urban canopy model for atmospheric models: comparison with multi-layer and slab models. *Boundary-Layer Meteorology*, 101:329–358.

Lang, P. (2001). Cell tracking and warning indicators derived from operational radar products. *30th Int. Conf. Radar Met., AMS, Munich, Germany.*

Langhans, W., Schmidli, J., and Schär, C. (2012). Bulk convergence of cloud-resolving simulations of moist convection over complex terrain. *J. Atmos. Sci.*, 69(7):2207–2228.

Leduc, M. and Laprise, R. (2009). Regional climate model sensitivity to domain size. *Clim. Dyn.*, 32(6):833–854.

Li, X. F. and Gao, S. T. (2012). *Precipitation Modeling and Quantitative Analysis*, page 240. Springer, Dordrecht.

Liu, Y., Chen, F., Warner, T., and Basara, J. (2006). Verification of a mesoscale data-assimilation and forecasting system for the Oklahoma city area during the Joint Urban 2003 Field Project. *J. Appl. Meteorol.*, 45:912–929.

Lorenz, E. N. (1969). The predictability of a flow which possesses many scales of motion. *Tellus*, 21(3):289–307.

Martilli, A., Clappier, A., and Rotach, M. (2002). An urban surface exchange parameterization for mesoscale models. *Boundary-Layer Meteorology*, 104:261–304.

Meinke, I., Geyer, B., Feser, F., and von Storch, H. (2006). The impact of spectral nudging on cloud simulation with a regional atmospheric model. *J. Atmos. Ocean. Technol.*, 23(6):815–824.

Meischner, P. (2004). *Weather radar: principles and advanced applications*, page 337. Springer, Berlin.

Mishra, A. K. (2013). Effect of rain gauge density over the accuracy of rainfall: A case study over Bangalore, India. *Springer Plus*, 2:1–7.

Mittermaier, M. and Roberts, N. (2010). Intercomparison of Spatial Forecast Verification Methods: Identifying Skillful Spatial Scales Using the Fractions Skill Score. *Wea. Forecasting*, 25:343–354.

Morin, E., Krajewski, W. F., Goodrich, D. C., Gao, X. G., and Sorooshian, S. (2003). Estimating rainfall intensities from weather radar data: the scale-dependency problem. *J. Hydrometeorol.*, 4:782–797.

NAWAM-INIS (2014). *INIS program, https://nawam-inis.de/en/inis-projects/synopse.* [Online; accessed 19 January 2017].

OPERA (2017). *EUMETNET OPERA, http://www.eumetnet.eu/opera.* [Online; accessed 19 January 2017].

Overeem, A., Buishand, T. A., Holleman, I., and Uijlenhoet, R. (2010). Extreme-value modeling of areal rainfall from weather radar. *Water Resour. Res.*, 46.

Overeem, A., Holleman, I., and Buishand, T. A. (2009). Derivation of a 10-year radar-based climatology of rainfall. *J. Appl. Meteorol.*, 48(7):1448–1463.

Pedersen, L., Jensen, N., and Madsen, E. (2008). *Extreme rainfall statistics based on rain gauges and radar measurements.* WRaH Grenoble, France, http://radar.dhigroup.com/publications/ Extreme rainfall statistics based on rain gauges and radar.pdf. [Online; accessed 19 January 2017].

Pineda, N., Jorba, O., Jorge, J., and Baldasano, J. M. (2004). Using NOAA AVHRR and SPOT VGT data to estimate surface parameters: application to a mesoscale meteorological model. *Int. J. Remote Sens.*, 25:129–143.

Prein, A. (2013). *Added Value of Convection Permitting Climate Simulations*, pages 168. Available at: http://wegcwww.uni–graz.at/publ/wegcreports/2013/WCV–SciRep–No53–APrein–Jul2013.pdf. Scientific Report No. 53-2013, Wegener Center Verlag, Graz, Austria. [Online; accessed 19 January 2017].

Prein, A. and Gobiet, A. (2011). *Non-hydrostatic climate modelling. Part I: Defining and detecting added value in cloud resolving climate simulations*, page 74. Wegcenter Rep. 39, Tech. Rep., Wegener Center Verlag, Graz, Austria.

Prein, A., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N., Keuler, K., and Georgievski, G. (2013a). Added value of convection permitting seasonal simulations. *Clim. Dyn.*, 41(9–10):2655–2677.

Prein, A. F., Holland, G. J., Rasmussen, R. M., Done, J., Ikeda, K., Clark, M. P., and Liu, C. H. (2013b). Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters. *J. Clim.*, 26(13):4848–4857.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., et al. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, 53:323–361.

REGNIE (2017). *REGNIE documentation, ftp://ftp-cdc.dwd.de/pub/CDC/grids_germany/daily/regnie/REGNIE_Beschreibung_20170304.pdf.* [Online; accessed 25 April 2017].

Rinehart, R. E. (2004). *Radar for Meteorologists*, page 482. 4th ed., Rinehart Publications.

Roberts, N. M. (2008). Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorological Applications*, 15:163–169.

Roberts, N. M. and Lean, H. W. (2008). Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.*, 136(1):78–97.

Rudolph, J., Friedrich, K., and Germann, U. (2011). Relationship between radar-estimated precipitation and synoptic weather patterns in the European Alps. *J. Appl. Meteorol. Climatol.*, 50:944–957.

Rummukainen, M. (2010). State-of-the-art with regional climate models. *WIREs Clim. Change*, 1(1):82–96.

Sanderson, B. M., Piani, C., Ingram, W., Stone, D., and Allen, M. (2008). Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Clim. Dyn.*, 30(2–3):175–190.

Sauvageot, H. (1992). *Radar Meteorology*, page 384. Artech House.

Schwartz, C. S., Kain, J. S., Weiss, S. J., Xue, M., Bright, D. R., Kong, F., Thomas, K. W., Levit, J. J., and Coniglio, M. C. (2009). Next-day convection-allowing WRF model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Weather Rev.*, 137(10):3351–3372.

Seltmann, J. (1997). Radarforschung im DWD: Vom Scan zum Produkt. *promet*, 26:32–42.

Seo, B.-C., Cunha, L. K., and Krajewski, W. F. (2013). Uncertainty in radar-rainfall composite and its impact on hydrologic prediction for the eastern Iowa flood of 2008. *Water Resour. Res.*, 49:2747–2764.

Shin, H. H. and Hong, S.-Y. (2013). Analysis of resolved and parameterized vertical transports in convective boundary layers at gray-zone resolutions. *J. Atmos. Sci.*, 70:3248–3261.

Skamarock, W. C., Klemp, J. B., Dudhia, J., et al. (2008). *A Description of the Advanced Research WRF Version 3, Tech. Note, NCAR/TN-475+STR*, page 113. National Center for Atmospheric Research, Boulder.

Sokol, Z. (2011). Assimilation of extrapolated radar reflectivity into a NWP model and its impact on a precipitation forecast at high resolution. *Atmos. Res.*, 100:201–212.

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

Steinert, J. (2014). Hydrometeor classification for the DWD weather radar network: First verification results. *Proc. 8th Europ. Conf. On Radar in Meteor. and Hydrol., Garmisch-Partenkirchen, Germany.*

Tabary, P., Dupuy, P., L'Henaff, G., et al. (2012). A 10-year (1997–2006) reanalysis of quantitative precipitation estimation over France: methodology and first results. *in Weather Radar and Hydrology (Proceedings of a Symposium held in Exeter, UK, April 2011), Publication 351. IAHS: Wallingford, UK*, pages 255–260.

Taha, H. (1999). Modifying a mesoscale meteorological model to better incorporate urban heat storage: a bulk-parameterization approach. *J. Appl. Meteorol.*, 38:466–473.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106:7183–7192.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.*, 93(4):485–498.

Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M. A., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R. H. (2004). Implementation and verication of the unied NOAH land surface model in the WRF model. *20th conference on weather analysis and forecasting16th conference on numerical weather prediction*, pages 11–15.

Tölle, M. H., Gutjahr, O., Busch, G., and Thiele, J. C. (2014). Increasing bioenergy production on arable land: Does the regional and local climate respond? Germany as a case study. *J. Geophys. Res. Atmos.*, 119:2711–2724.

Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B. (2003). The changing character of precipitation. *Bull. Am. Meteorol. Soc.*, 84:1205–1217.

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

Vannitsem, S. and Chomé, F. (2005). One-way nested regional climate simulations and domain size. *J. Clim.*, 18(1):229–233.

Vignal, B., Andrieu, H., and Creutin, H. (1999). Identification of vertical profiles of reflectivity from volume scan radar data. *J. Appl. Meteorol.*, 38:1214–1228.

Vogl, S., Laux, P., Qiu, W., Mao, G., and Kunstmann, H. (2012). Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. *Hydrol. Earth Syst. Sci.*, 16:2311–2328.

von Storch, H. (2005). Conceptual basis and applications of regional climate modeling. In *Extended Abstracts "High-resolution Climate Modelling: Assessment, Added Value and Applications"*, volume 29, page 26–27. WMO/WCRP-sponsored regional-scale climate modelling Workshop, Lund, Sweden.

von Storch, H., Langenberg, H., and Feser, F. (2000). A spectral nudging technique for dynamical downscaling purposes. *Mon. Weather Rev.*, 128(10):3664–3673.

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

Wagner, A., Seltmann, J., and Lang, P. (2006). URBAS_Radar - a statistical approach to radar climatology. *Abstr. 4th European Conference on Radar in Meteorology and Hydrology (ERAD4), Sept. 18-22, 2006, Barcelona, Spain, ERAD Publ. Series*, 3:61.

Wagner, S. and Kunstmann, H. (2016). High resolution precipitation fields for the planning of urban drainage systems using WRF. *Internal annual report. Steinbuch Centre for Computing (SCC). Karlsruhe Institute of Technology (KIT)*, page 11.

Wallace, J. M. (1975). Diurnal variations in precipitation and thunderstorm frequency over the conterminous United States. *Mon. Weather Rev.*, 103:406–419.

Wang, W., Bruyère, C., Duda, M., Dudhia, J., Gill, D., Kavulich, M., Keene, K., Lin, H.-C., Michalakes, J., Rizvi, S., Zhang, X., J., B., and Fossell, K. (2015). *ARW User Guide*, page 408. Available at: http://www2.mmm.ucar.edu/wrf/users/docs/user_guide_V3/ ARWUsersGuideV3.pdf. [Online; accessed 19 January 2017].

Warner, T. T., Peterson, R. A., and Treadon, R. E. (1997). A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bull. Am. Meteorol. Soc.*, 78(11):2599–2617.

Weigl, E. and Winterrath, T. (2010). Radar based precipitation analysis and forecasting (RADOLAN, RADVOR-OP). *promet*, 35:78–86.

Weisman, M. L., Davis, C., Wang, W., Manning, K. W., and Klemp, J. B. (2008). Experiences with 0-36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, 23(3):407–437.

Weisman, M. L., Skamarock, W. C., and Klemp, J. B. (1997). The resolution dependence of explicitly modeled convective systems. *Mon. Weather Rev.*, 125(4):527–548.

Werner, M. (2014). A New Radar Data Post-Processing Quality Control Workflow for the DWD Weather Radar Network. *Proc. 8th Europ. Conf. On Radar in Meteor. and Hydrol., Garmisch-Partenkirchen, Germany.*

Wernli, H., Paulat, M., Hagen, M., and Frei, C. (2008). Sal-A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Weather Rev.*, 136(11):4470–4487.

Weusthoff, T., Ament, F., Arpagaus, M., and Rotach, M. (2010). Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Mon. Weather Rev.*, 138(9):3418–3433.

Whiton, R. C., Smith, P. L., Bigler, S. G., Wilk, K. E., and Harbuck, A. C. (1998). History of operational use of weather radar by U.S. Weather Services. Part 1: the pre-NEXRAD era. *Wea. Forecasting*, 13:219–243.

WMO (2014). *WMO GUIDE TO METEOROLOGICAL INSTRUMENTS AND METHODS OF OBSERVATION WMO-No.* 8, http://www.wmo.int/pages/prog/www/IMOP/CIMO-Guide.html. [Online; accessed 19 January 2017].

Wouters, H., Ridder, K. D., Demuzere, M., Lauwaet, D., and van Lipzig, N. (2013). The diurnal evolution of the urban heat island of Paris: A model-based case study during summer 2006. *Atmos. Chem. Phys.*, 13(17):8525–8541.

Wyngaard, J. C. (2004). Toward numerical modeling in the “terra incognita". *J. Atmos. Sci.*, 61:1816–1826.

Yoon, S.-S., Phuong, A. T., and Bae, D.-H. (2012). Quantitative comparison of the spatial distribution of radar and gauge rainfall data. *J. Hydrometeor.*, 13(6):1939–1953.

Zhang, F., Snyder, C., and Rotunno, R. (2003). Effects of moist convection on mesoscale predictability. *J. Atmos. Sci.*, 60(9):1173–1185.

Zhou, B., Simon, J. S., and Chow, F. K. (2014). The convective boundary layer in the Terra Incognita. *J. Atmos. Sci.*, 71:2545–2563.

Zoccatelli, D., Borga, M., Zanon, F., Antonescu, B., and Stancalie, G. (2010). Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. *J. Hydrol.*, 394:148–161.

Zrnić, D. S. and Ryzhkov, A. V. (1999). Polarimetry for weather surveillance radars. *Bull. Amer. Meteorol. Soc.*, 80:389–406.

Danksagung

Mein Dank gilt:

- *Prof. Dr. Harald Kunstmann* für die Möglichkeit meine Radarstudien weiterzuführen, meinen wissenschaftlichen Horizont zu erweitern und zu promovieren. Vielen Dank für deine Betreuung und das in mich gesetzte Vertrauen.
- *PD Dr. Christoph Beck* für die Übernahme des Zweitgutachtens und die dafür aufgewendete Zeit und Arbeit.
- *Dr. Jörg Seltmann* für die gemeinsame Arbeit in zwei Projekten am Observatorium Hohenpeißenberg und die Einführung in die Wissenschafts- und Radarwelt. Ich danke ihm für die kontinuierliche Unterstützung meiner Studien auf dem Radarsektor, für die unzähligen Diskussionen und für die Zeit, die er investiert hat.
- *Sven Wagner* für die gemeinsame Projektarbeit und die Unterstützung bei den WRF-Simulationen.
- *Thomas Rummler* für viele Fragen und Diskussionen rund um WRF, Programmierung und Computerprobleme aller Art.
- meinen Kollegen *Jan Bliefernicht, Barbara Haese, Manuel Lorenz* und *Dominikus Heinzeller* für das angenehme wissenschaftliche und soziale Umfeld.
- dem *Deutschen Wetterdienst* für die Bereitstellung von Radardaten.
- meiner *Schwester Claudia* für hilfreiche Anmerkungen zu dieser Arbeit.
- besonders meiner *Frau Silke* für die Zeit und Nerven, die auch sie in diese Arbeit gesteckt hat.

Entschuldigen möchte ich mich an dieser Stelle bei meinen beiden Töchtern *Maia* und *Lene*, dass ich viel zu wenig Zeit für sie hatte ...