

Explicit Convection and Scale-Aware Cumulus Parameterizations: High-Resolution Simulations over Areas of Different Topography in Germany

ANDREAS WAGNER

Institute of Geography, University of Augsburg, Augsburg, Germany

DOMINIKUS HEINZELLER^a

Institute of Meteorology and Climate Research IMK-IFU, Karlsruhe Institute of Technology (KIT-Campus Alpin), Garmisch-Partenkirchen, Germany

SVEN WAGNER^b AND THOMAS RUMMLER

Institute of Geography, University of Augsburg, Augsburg, Germany

HARALD KUNSTMANN

Institute of Meteorology and Climate Research IMK-IFU, Karlsruhe Institute of Technology (KIT-Campus Alpin), Garmisch-Partenkirchen, and University of Augsburg, Institute of Geography, Augsburg, Germany

(Manuscript received 17 August 2017, in final form 20 April 2018)

ABSTRACT

An increase in the spatial resolution of regional climate model simulations improves the representation of land surface characteristics and may allow the explicit calculation of important physical processes such as convection. The present study investigates further potential benefits with respect to precipitation, based on a small ensemble of high-resolution simulations with WRF with grid spacings up to 1 km. The skill of each experiment is evaluated regarding the temporal and spatial performance of the simulation of precipitation of one year over both a mountainous region in southwestern Germany and a mainly flat region in northern Germany. This study allows us to differentiate between the impact of grid spacing, topography, and convection parameterization. Furthermore, the performance of a state-of-the-art convection parameterization scheme in the gray zone of convection is evaluated against an explicit calculation of convection only. Our evaluation demonstrates the following: high-resolution simulations (5 and 1 km) are generally able to represent the diurnal cycle, structure, and intensity distribution of precipitation, when compared to observational datasets such as radar data and interpolated station data. The influence of the improved representation of the topography at higher resolution (1 km) becomes apparent in complex terrain, where the localization of precipitation maxima is more accurate, although these maxima are overestimated. In flat areas, differences in spatial evaluations arise between simulations with parameterized and explicitly calculated convection, whereas smaller grid spacings (1 km vs 5 km) show hardly any impact on precipitation results.

1. Introduction

General circulation models (GCMs), in the context of long-term climate simulations, often use horizontal grid

spacing larger than 100 km (Taylor et al. 2012), whereas nested regional climate models (RCMs) are able to reduce this resolution to the lower kilometer range. In principle, a finer model grid allows a better representation of biogeophysical variabilities such as land use and topography (Prein et al. 2015), which should translate into more accurate simulations. Still, the recent Coordinated Downscaling Experiment-European Domain (EURO-CORDEX) RCM simulations with a 12-km horizontal resolution for Europe only show minor improvements such as a less pronounced southern European warm summer bias compared to the ENSEMBLES simulations (50-km

^a Current affiliation: Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and NOAA/OAR/ESRL/Global Systems Division, Boulder, Colorado.

^b Current affiliation: Fraunhofer IAO, Stuttgart, Germany.

Corresponding author: Andreas Wagner, andreas.wagner@geo.uni-augsburg.de

grid spacing). Yet, the biases inferred from both ensembles of simulations lie in the same range (Jacob et al. 2014; Kotlarski et al. 2014; Smiatek et al. 2016). Current research efforts aim at reducing the grid spacing until processes are resolved at grid scale, as the parameterization of critical subgrid processes is regarded as a major source of uncertainty in climate modeling and weather forecasting (Knight et al. 2007; Sanderson et al. 2008; Prein et al. 2013a).

Large uncertainty is associated with the parameterization of deep convection (e.g., Weisman et al. 1997; Hohenegger et al. 2008). The interaction of convective parameterizations with microphysics, planetary boundary layer (PBL), and radiation schemes at kilometer-scale resolutions is complex (Prein et al. 2015). According to Bryan and Morrison (2012) a grid spacing of 100 m would be necessary to start resolving in-cloud turbulence and particularly the entrainment process, which is a measure for convective overturning. Consequently, only deep convection can be explicitly calculated for grid spacings in the lower kilometer range. However, for deep convection the gap in resolution between explicitly resolved flows and implicitly parameterized flows—called the “gray zone for convection”—has to be taken into account (Wyngaard 2004; Prein et al. 2015). Undesired side-effects in this gray zone may deteriorate simulation results with and without convective parameterizations (Zhou et al. 2014).

Arnault et al. (2016) and Klein et al. (2017) performed their simulations of the monsoon regime in West Africa with the Weather Research and Forecasting (WRF) Model without convection parameterization for all grid spacings below 10 km. They showed that in this region, where mesoscale systems are larger, the explicit calculation of convection improved the representation of model-based land–atmosphere interactions. In the temperate zone, Weisman et al. (1997) have shown that for grid spacings below 4 km, an explicit calculation of deep convection is reasonable. Thus, the quality of the simulations depends on the ability to resolve the dominant processes rather than purely on the grid spacing.

In recent years, a number of high-resolution simulations in the convection permitting scale (CPS; grid spacing ≤ 4 km) were performed that are summarized in Prein et al. (2015). The authors mainly investigate precipitation, temperature (e.g., Prein et al. 2013a; Brisson et al. 2016b), cloud cover (Langhans et al. 2012; Prein et al. 2013a; Fosser et al. 2015; Brisson et al. 2016b), and the improvement of physical parameterizations by, for example, accounting for aerosol effects in explicit simulations of clouds (Heinzeller et al. 2016).

Most attention is paid to the simulation of precipitation, since it aggregates many physical processes including moisture advection, microphysics, and

atmospheric dynamics. If precipitation is correctly simulated on a high spatiotemporal scale, many processes within the model are likely well captured. Early high-resolution simulations in the CPS were presented by Grell et al. (2000), Kunstmann et al. (2004), and Kunstmann and Stadler (2005). They showed significant differences of the location of precipitation patterns and precipitation amounts with and without convection parameterization in complex terrain. In the following years, the benefit of a better representation of topography in complex terrain was studied with respect to improvements regarding the diurnal cycle of precipitation, the frequency of heavy hourly events, the representation of the size and shape of precipitation patterns, and the spatial variability (e.g., Hohenegger et al. 2008; Weusthoff et al. 2010; Prein et al. 2013a). Similar improvements were found both for longer simulations of two to three decades (Kendon et al. 2012; Fosser et al. 2015) and in coastal regions without orographic forcing (Brisson et al. 2016a). However, several shortcomings were observed such as erroneous intensification of heavy precipitation for different regions and orography (Kendon et al. 2012; Langhans et al. 2012; Chan et al. 2013; Prein et al. 2013a; Ban et al. 2014). A general conclusion whether this was an effect of the chosen model, the investigation area, or the analysis approach could not be drawn (Prein et al. 2015). Prein et al. (2013b) showed that in principle an added value is easier to detect in mountainous regions than in a hilly area and that during the winter months, where convection is generally less intensive, the added value of high-resolution simulations is related to the better orographic resolution.

While the benefit of CPS simulations is generally accepted, the question remains whether an even higher resolution is beneficial in terms of the model accuracy. According to Weisman et al. (2008), Mass et al. (2002), and Kain et al. (2008) a more realistic representation of convection was observed with smaller grid spacings. However, no improvements in forecast skill for precipitation were obtained. In contrary, the skills got worse when reducing grid spacings up to 2 km (Kain et al. 2008) and 1 km (Mass et al. 2002). Clark et al. (2009) argued that increasing finescale details seems to be responsible for this effect. In their abstract, Roberts and Lean (2008) summarized the problem as follows: “greater realism does not necessarily mean more accurate precipitation forecasts.” However, Roberts and Lean (2008) showed improvements for 1-km grid spacing compared to 4 km for certain statistical measures such as the fractions skill score. Langhans et al. (2012) found only small differences between simulations at various grid spacings between 4.4 and 0.55 km in their

9-day evaluation of convective systems over the Alps. Uncertainty remains whether added value can be associated with a further increase of grid spacings, and if so, for which characteristics they are observed.

A third option for high-resolution simulations besides the explicit calculation of convection for very small and coarser grid spacings, is applying convection parameterization. To balance the trade-off between an increase in model accuracy and higher computational costs, in particular in the context of long-term climate simulations, new parameterizations of deep convection are developed that can be applied in between the classical convection-permitting (≤ 4 km) and mesoscale (≥ 10 km) regime (Baldauf et al. 2011; Grell and Freitas 2014). Mahoney (2016) evaluated the impact of explicit versus parameterized convection at 4-km grid spacing. Their investigation of convective events revealed that an explicit calculation of convection improves over classic convective parameterizations. However, these studies suggest that certain aspects of a convection parameterization scheme (e.g., shallow cumulus) are still required to account for unresolved processes at this scale. They further showed that a revised scale-aware version of the Kain–Fritsch convection parameterization scheme is able to provide similar results as the explicit calculation.

In this work, a combination of high-resolution WRF Model simulations with different grid spacings, parameterizations, and over different topographies is analyzed. Our simulations make use of the urban canopy model to account for the impact of the urban geometry on surface energy budgets and wind shear calculations. We investigate grid spacings of 1 and 5 km without convection parameterization and also apply the scale-aware Grell–Freitas convection parameterization for 5-km grid spacing. This allows us to evaluate the strength and weakness of each dataset regarding its performance of simulating the temporal, spatial, and intensity distribution of precipitation.

The aim of this study is to investigate (i) whether an increase in resolution from 5 to 1 km is leading to an added value in the simulations, and whether this added value justifies the increase in computational costs; (ii) if convection parameterizations are able to provide similar results, or if they even show advantages for certain aspects; and (iii) how topography influences the performance of high-resolution precipitation simulations.

The simulations are performed for two domains with different topography in the northern and southwestern part of Germany for a time period between 3 months and 1 year. The spatial and temporal evaluations (e.g., correlograms, fractions skill score, probability density functions) are achieved with several validation datasets (radar data, interpolated station data),

mainly on 3 hourly time steps with grid spacings between 5 and 15 km.

2. Experiment design

a. WRF Model setup

The WRF Model (Skamarock et al. 2008) is used in this work as an RCM to downscale global atmospheric reanalysis data. The best setup of physical parameterization schemes in Table 1 is obtained through extensive testing and evaluating the performance regarding precipitation and temperature (2 m), compared to validation datasets such as the first high-resolution gridded dataset of daily climate over Europe (E-OBS; Haylock et al. 2008), Regionalized Precipitation Amount (REGNIE; Rauthe et al. 2013), and station measurements (not shown). The choice of the model configuration is made with focus on high-resolution simulations such as applying a microphysics scheme that accounts for graupel processes, which are important for reliable simulation of convection (Brisson et al. 2016a), a scale-aware convection parameterization scheme that attempts to smooth the transition to cloud-resolving scales (Grell and Freitas 2014), and a PBL scheme that shows weaker influences of undesired effects such as “grid-dependent convection” (Wyngaard 2004; Zhou et al. 2014; Ching et al. 2014) as a result of partly resolving large eddies.

At finer scales, biogeophysical features such as urban areas also become more important. The shape of a city influences momentum, turbulence, and thermal fluxes between the surface and the atmosphere. We use the single-layer urban canopy model (SLUCM; Kusaka et al. 2001; Kusaka and Kimura 2004) in this investigation to parameterize the major processes within urban areas. Hence, the Noah land surface model (LSM) is used to calculate the relevant processes for natural surfaces and the SLUCM for the three-dimensional structure of urban surfaces (Chen et al. 2011).

b. Investigation areas and domains

Two different investigation areas in northern and southwestern Germany with a minimum grid spacing of 1 km are chosen to analyze and compare similar high-resolution simulations over different topography (Fig. 1). The northern innermost domain (D03N) around Hamburg, Germany, is characterized by an almost flat landscape and a maritime influence. The southwestern innermost domain around Freiburg, Germany (D03S), with the Vosges in the west and the Black Forest in the east show a much higher degree of orographically enhanced precipitation and more intense convection.

ERA-Interim reanalysis data (Dee et al. 2011) serves as input for the RCM simulations at a spatiotemporal resolution of 80 km and 6 h. For downscaling the 80-km forcing

TABLE 1. WRF Model setup and settings: RRTM (Mlawer et al. 1997), Dudhia (Dudhia 1989), WSM6 (Hong and Lim 2006), YSU (Hong et al. 2006), GF (Grell and Freitas 2014), and Noah (Tewari et al. 2004).

Simulation name	Longwave radiation	Shortwave radiation	Microphysics	PBL	Convection parameterization	LSM	Duration	Grid spacing	Temporal resolution	Cells D-North	Cells D-South
RIP-15	RRTM	Dudhia	WSM6	YSU	GF	Noah	Jun 2005–May 2006	15 km	3 h	8 × 12	8 × 8
RIP-5	RRTM	Dudhia	WSM6	YSU	GF	Noah	Jun 2005–May 2006	5 km	1 h	24 × 36	24 × 24
RIE-1	RRTM	Dudhia	WSM6	YSU	—	Noah	Jun 2005–May 2006	1 km	5 min	120 × 180	120 × 120
R2E-5	RRTM	Dudhia	WSM6	YSU	—	Noah	Jun–Aug 2005	5 km	1 h	24 × 36	24 × 24

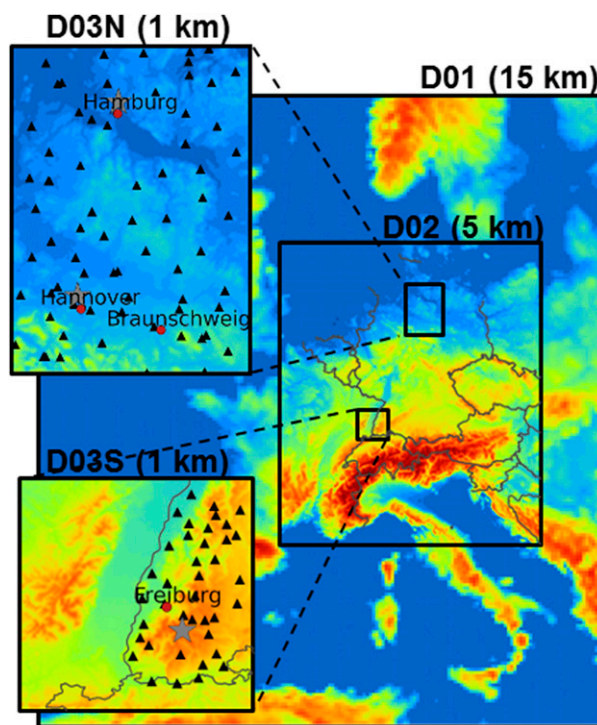


FIG. 1. Overview of the four WRF domains showing the altitude for the respective grid spacing to downscale ERA-Interim data. The map is overplotted by the stations used for the interpolation of STA-5 (black triangle), the radar sites (gray star), and national borders (dark gray). The investigation areas D-North and D-South are based on D03N and D03S discarding the outer 15 km of both domains to minimize numeric artifacts.

data to a fine grid of 1 km, two intermediate steps are used in a one-way-nesting mode: a large parent domain D01 of 15-km grid spacing (181×202 grid columns) is used to downscale the reanalysis data. The intermediate 5-km domain (D02, 241×280 grid columns) serves as input to the two domains at 1-km resolution (Fig. 1). All domains use 42 vertical eta levels in a default (i.e., automatic) distribution by the WRF Model. 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 (Matte et al. 2017). To further reduce deviations from the large-scale forcing patterns, spectral nudging (von Storch et al. 2000) for temperature, horizontal wind, and humidity is applied every 6 h on the outer 15 km domain for features with wavelengths of about 1300 km and upward. This downscaling technique constrains the large atmospheric conditions within an RCM to the global model and the propagation of large-scale atmospheric patterns into the RCM is better realized.

c. Spinup

Soil properties such as moisture content or temperature are changing slowly over time and often need several

weeks or months to spin up. To save computational efforts, a spinup of four months is performed on the outer domain D01 only. After four months, the inner domains are added and adopt subsurface and surface fields such as skin temperature, soil moisture, and soil temperature from D01. One additional month of spinup on all domains is appended to reach satisfying soil conditions for each of the domains.

3. Simulation and validation data

a. Simulation data

The analysis is performed for the areas covered by domains D03N and D03S. For both areas, a 15-km-wide boundary zone is discarded to minimize numeric artifacts in the analysis. The resulting 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 1 shows the resulting grid columns for each area based on the respective domains.

Two different simulation runs are conducted to investigate the differences among high-resolution simulations. For both runs, the same spinup procedure and WRF setup is used, except for the parameterization of convection. The label R1P-15 for instance means “run 1 with cumulus parameterization and 15-km grid spacing.” R2E-5 means “run 2 with explicit calculation of convection and 5-km grid spacing.” Table 1 also gives an overview of all datasets used in this chapter in its original spatial and temporal resolution and the period for which the data are produced. Each dataset is available for both areas D-North and D-South. Overall, four WRF datasets are analyzed for each area. Run 1 simulations are generated for the whole year (June 2005–May 2006), whereas the remaining run 2 simulation is conducted only from June to August 2005 (summer), where the main convective activity can be expected.

b. Validation data

Reliable precipitation data at high spatial and temporal resolution are required for the analysis and are often the biggest challenge to obtain. Interpolated station data usually represent the natural temporal variability of precipitation and its distribution of intensities well at the location of the recording stations. But interpolation techniques are hardly able to represent location and exposure and may lead to smoothing effects and limitations in spatial precipitation patterns, especially in complex terrain and for short-lived convective events. In contrast, radar data show a good representation of the spatial distribution of precipitation echoes, but have limitations regarding the evaluation of temporal aspects of precipitation and absolute

precipitation intensities. Furthermore, the existing radar datasets in Germany from the Radar-Online-Adjustment (RADOLAN) project (Bartels et al. 2004) suffer from limitations such as not being corrected for systematic measuring effects like the vertical profile of reflectivity (VPR), clutter, or compositing effects, all of which may influence precipitation patterns (Fabry and Zawadzki 1995; Vignal et al. 1999; Franco et al. 2006; Krajewski et al. 2011). Consequently, an additional radar dataset, corrected for systematic measuring effects, is applied for the spatial evaluation in this work, while interpolated rain gauge data is used in the temporal evaluation.

1) CORRECTED RADAR DATA

The basis for the radar dataset is the RX product of the German Meteorological Service (DWD) with 256 classes (31.5–95.5 dBZ) and a resolution of 0.5 dB. It is based on terrain-following precipitation scans of up to 16 radar systems at high temporal and spatial resolution (5 min; $1 \times 1 \text{ km}^2$). This product undergoes the usual corrections within the signal processor [e.g., Doppler filtering, clutter correction, speckle remover, and thresholding for noise (LOG) and signal quality (SQI; Seltsmann (1997))]. Additional corrections are applied to correct systematic effects in radar data such as clutter remnants, spokes, and the dependence of radar measurements on altitude. Precipitation intensities are calculated by a three-part Z–R relationship (Bartels et al. 2004), and an adjustment to rain gauge data is performed. These corrections are based on Wagner et al. (2012). In Table 2, mean hourly RMSE values between station data and radar data are shown. The highest RMSE values over Germany are observed in the south, due to clutter and shading effects in the alpine regions. The higher RMSE values in summer (0.5 mm) are caused by higher precipitation intensities and convective patterns than in winter (0.37 mm). The summer precipitation events are more difficult to capture by stations, and the translation from radar reflectivities into rain rates during heavy precipitation shows larger errors. Heavy precipitation intensities are slightly underestimated, whereas smaller and medium precipitation intensities are well captured. In D-North, clutter remnants lead to higher RMSE values (0.48 mm) than in D-South (0.43 mm), albeit the flatter topography. Consequently, the RMSE values in winter are lower, but a higher influence of snow, especially in D-South, impedes both measurements (0.21–0.24 mm). The abbreviation of this dataset is RAD-1.

2) INTERPOLATED STATION DATA

The second dataset is solely based on station data from a dense network of rain gauges in Germany.

TABLE 2. RMSE of hourly precipitation amounts for summer and winter between radar and station data (516 stations) for Germany (BRD), D-North, and D-South.

Area	Summer	Winter
BRD	0.5	0.37
D-North	0.48	0.21
D-South	0.43	0.24

Hourly data of approximately 40–70 quality-controlled stations from DWD for each investigation area (Fig. 1) are interpolated by an external drift-kriging technique based on altitude as drift variable (Deutsch and Journel 1992). The interpolation of the station data is performed on a 5-km grid for both areas D-North and D-South. The characteristics of the topography dominate the interpolated precipitation patterns. Naturally induced variations of precipitation amounts due to windward and leeward effects cannot be adequately considered. Since our observational dataset is limited to the German station network, precipitation data are available for the entire area D-North, but only for the German part of D-South. This dataset is called STA-5 in the following.

3) REGNIE DATA

For analyses on a longer time scale (days or seasons), the REGNIE data are additionally used for means of comparison. This dataset of daily interpolated rain gauge values is available on an approximately 1 km² grid for Germany. About 2000 quality-controlled rain gauges are interpolated using long-term monthly background fields and linear regression techniques. The geographical location as well as the direction and degree of exposure are taken into account. REGNIE data are regarded as a standard dataset for resolutions of one day or longer in Germany. As this dataset is only based on station data without using additional pattern information (from radar or satellites), its accuracy in complex terrain is limited. REGNIE data are based on a much higher density of station data than the STA-5 dataset, but are only available on a daily scale. This dataset is called REG-1 in the following.

4. Evaluation methods

The aim of this investigation is to assess differences between high-resolution simulations with WRF and highlight the potential added value and also drawbacks of each dataset with respect to precipitation. At small scales, precipitation is partly regarded to be non-deterministic and unpredictable (Hohenegger et al. 2008). Small spatial and temporal shifts between modeling results and observations may lead to the “double

penalty” problem: a slight spatial shift first misses the observed precipitation object, and second produces precipitation where no precipitation is observed (Prein et al. 2013a). Common statistical values such as root-mean-square error (RMSE) or correlation are only useful on a longer time scale or averaged over sufficiently large areas. For the analysis of simulations with high temporal and spatial resolution, a statistical approach that is independent of space or time is useful, for example, the comparison of the spatial extent and correlation of precipitation objects, or the goodness of fit of temporally averaged model results and observation. Statistical values or means, such as the distribution of precipitation amounts, provide additional information on mean precipitation characteristics. While the emphasis of this work lays on spatial evaluation techniques, attention is also paid to the temporal component to fully evaluate the performance of our simulation results.

Spatial and temporal evaluation methods

Most evaluations are performed on the same grid resolution to reveal the added value due to a better representation of processes instead of the improved representation of terrain and elevation (Berg et al. 2013). In this study, we use the coarse-grid resolution of 15 km for the bulk of the analysis, and the intermediate-grid resolution of 5 km for the investigation of small-scale precipitation patterns to better represent the spatial variability of precipitation. The entire evaluation period of one year is subdivided into seasons with a focus on summer and winter. These two seasons are characterized by different synoptic-scale forcing and mesoscale processes (Prein et al. 2013a; Fosser et al. 2015), in particular with respect to heavy precipitation. Most convective events occur in summer (June–August), during which the largest added value of high-resolution simulations has been found (e.g., Kendon et al. 2012; Prein et al. 2013a; Ban et al. 2014). The additional simulation R2E-5 is performed for this period only.

For almost all evaluation aspects, the two distinct validation datasets STA-5 and RAD-1 are used. Since REG-1 and STA-5 only provide measurements for the German part of D-South, the evaluations are performed for this part of the domain D-South. Exceptions, as to when the entire domain including the French territory is considered, are explicitly mentioned.

First, the seasonal domain-averaged precipitation amounts are calculated for each model experiment. Second, the main characteristics of precipitation patterns are analyzed on a longer time scale for a subset of the simulations, based on annual precipitation patterns at their respective grid resolution. Additionally, annual precipitation patterns are evaluated on a common

15-km grid based on mean correlation coefficients (cor) and normalized standard deviations (sn) compared to REG-1.

Third, an in-depth investigation of the spatial precipitation patterns at the scale of convective events is performed based on correlograms at 5-km and 3-h resolution. The evaluations reveal the average decrease of the correlation with distance, individually for each dataset. It is performed for a threshold (domain average) $q = 0.01$ and 0.5 mm h^{-1} to represent time steps with light precipitation or more, and such with more intensive precipitation only (approximately 95th to 98th percentile). Therefore, the correlations between each grid cell and all other grid cells are calculated for the respective time steps, averaged and plotted versus distance. This approach is based on the principle that a stronger decrease in correlation with distance corresponds to smaller-scale precipitation patterns in the observations or simulations.

To overcome the problem of double penalty, certain measures exist that consider those displacements [for more information see Roberts and Lean (2008); Wernli et al. (2008); Prein and Gobiet (2011)]. In this study, the fractions skill score (FSS) is applied. The basic idea is that a simulation is useful if the spatial frequency of events is similar in forecast and in observation (Prein 2013). FSS is a neighborhood (or fuzzy) verification measure that directly compares forecasts and observed fractional coverages of precipitation. First, the precipitation fields of the simulations and the observations are transferred into binary fields with the help of a precipitation threshold q (e.g., 0.1, 0.5, 1.0, and 2.0 mm). In the second step, a spatial moving average is applied to the binary fields, based on a squared window of length n and uniform weights. Here, n is the horizontal scale (neighborhood size) with values between $n = 1$ and $n = 2N - 1$ where $N = \max(N_x, N_y)$. The third step is the computation of the $\text{FSS}_{(n)}$, which requires the calculation of the mean square error for each n in advance:

$$\text{FSS}_{(n)} = 1 - \frac{\text{MSE}_{(n)}}{\text{MSE}_{(n)\text{ref}}}. \quad (1)$$

The term $\text{MSE}_{(n)\text{ref}}$ used in Eq. (1) can be regarded as the largest obtainable MSE with the given observed and simulated fractions. $\text{FSS}_{(n)}$ values range between 0 and 1 (perfect fractional coverage). Two further thresholds are defined. The lower threshold is the observed fractional precipitation 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 ($\text{FSS} = 1$) and is defined as the limit where “reasonable skill” is achieved. More detailed

information about the FSS can be found at Roberts and Lean (2008) and Roberts (2008).

This procedure is performed at hourly time steps based on a 5-km grid for the entire areas D-North and D-South to apply this measure on a meaningful amount of grid cells. This excludes the datasets STA-5 because of the coverage of only the German part of the domain as well as R1P-15. RAD-1 serves as observational reference. The seasonal FSS is calculated as the median of all hourly FSS for summer or winter. Varying the thresholds q and the size of the window n , both intensity- and scale-dependent analyses are performed with FSS.

To assess the temporal accuracy of the high-resolution simulations in summer and winter, cor and sn compared to STA-5 (15 km and 3 h) are calculated. Furthermore, the diurnal cycle of precipitation is computed. It is based on a 15-km grid spacing and hourly precipitation values and calculated by spatially and temporally averaging the precipitation amounts for the same hour of a day over the entire domain.

Finally, the distribution of precipitation amounts is analyzed in the form of probability density functions (PDFs).

5. Results

a. Overview of mean values

As a first overview, Fig. 2 displays the seasonal rainfall amounts for all evaluated datasets for the areas D-North and D-South as relative differences to the observed total amount of precipitation in REGNIE. This evaluation is similar to the approach taken in Berg et al. (2013), who compared their ensemble of 9 high-resolution simulations (7-km grid spacing) over Germany to REGNIE data.

With a total of 701 mm, almost half as much precipitation is observed for D-North than for the German part of D-South (1198 mm) between June 2005 and May 2006. The validation datasets STA-5 and RAD-1 vary up to $\pm 13\%$ from REG-1 in all seasons and both areas. In winter, larger variations of up to -26% for RAD-1 in D-South are observed as a consequence of frequent snowfall.

Our simulations overestimate precipitation amounts in summer and winter months between 0.5% and 18% and reveal a maximum deviation from REG-1 of -31% in autumn. These are promising results compared to Berg et al. (2013), who presented seasonal variations between -8% and 83% for the period 1971–2000. Additionally, the mean wet bias of annual precipitation amounts between 14% and 59% presented by Berg et al. (2013) is not observed in our simulations.

R1P-15 differs from the other simulations in both domains with partly opposite tendencies of over- and

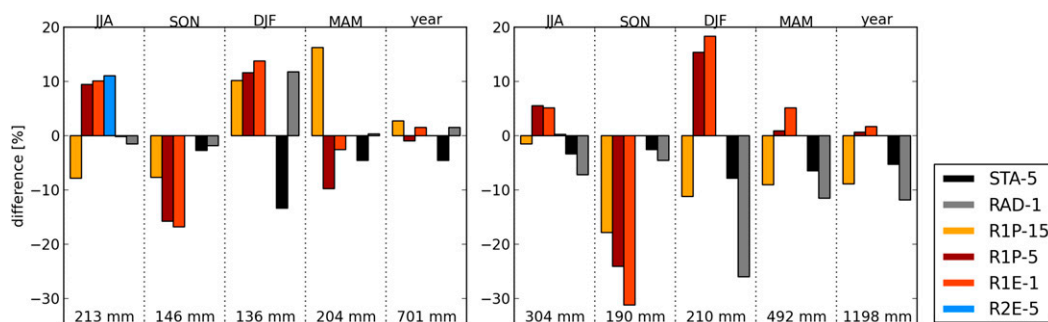


FIG. 2. Seasonal precipitation amounts (June 2005–May 2006) for all datasets for (left) D-North and (right) D-South. The columns display the relative difference between the respective dataset and REGNIE, averaged over the entire area. The simulation R2E-5 is available for the summer months only.

underestimations (e.g., in summer). Only R1P-15 consistently underestimates precipitation in all seasons in D-South, a consequence of the coarse grid spacing in complex terrain dominated by orographically forced precipitation events.

The high-resolution simulations (with 1- or 5-km grid spacing) behave similarly in both domains. The 5-km simulations (R1P-5 and R2E-5) only differ in D-South by 5% while they are almost identical in D-North. For the smaller grid spacing of 1 km, the overestimations and underestimations compared to REG-1 are intensified by up to 7% in almost all seasons except summer and both areas compared to grid spacings of 5 km (R1E-1 versus R1P-5).

In summary, the high-resolution-simulations differ considerably from R1P-15, but they differ only slightly from each other.

b. Spatial evaluation

1) ANNUAL PRECIPITATION PATTERNS IN D-NORTH

REG-1 is used as reference in Figs. 3 and 4 and shows detailed spatial patterns of precipitation amounts. For D-North (Fig. 3a), the precipitation distribution follows a northwest to south/southeast gradient from maximum to minimum precipitation.

Figure 3c reveals some constraints of the precipitation data RAD-1. Two spokes are visible near Hamburg, and clutter remnants are obvious in the southwestern part (Hannover radar). The latter features are responsible for the weak statistical measures ($sn = 1.51$, $r = 0.40$). These patterns are mainly present in winter and only show a minor influence in summer ($sn = 0.90$, $r = 0.81$; not shown).

The simulated precipitation patterns for D-North in Figs. 3d–f show less agreement with the validation datasets. The precipitation maximum in the northwestern

part is underestimated and shifted toward the center; the minimum in the southeastern part is absent in the simulations. The simulated summertime precipitation is mainly responsible for this shift (not shown). Only R1P-15 (Fig. 3d) roughly reproduces the location of the observed maximum precipitation ($r = 0.47$). The main patterns of R1P-5 (Fig. 3e) and R1E-1 (Fig. 3f) are similar on an annual basis.

2) ANNUAL PRECIPITATION PATTERNS IN D-SOUTH

Figure 4a shows the corresponding results of REG-1 for D-South. The precipitation patterns in D-South clearly follow the topography: precipitation maxima are produced in the southern, central, and northern part of the Black Forest (47.6°–49°N, 7.7°–9°E) and in the Vosges in the western part of the domain (47.6°–48.8°N, 6.3°–7.5°E), with significantly higher precipitation amounts than in D-North.

The sn of the interpolated station dataset STA-5 in D-South (0.72) are much lower than REGNIE data on the same 15-km grid with respect to annual precipitation patterns. This supports the visual impression that STA-5 (Fig. 4b) shows similar spatial distributions ($r = 0.91$), albeit with the structure being smoother and the precipitation maxima lower than REGNIE data. According to Fig. 4c, the overall distribution of precipitation patterns is reproduced well by RAD-1, although radar artifacts reduce the agreement. Apart from artificial ringlike structures, a result of the radar compositing algorithm, topographic shading effects occur for RAD-1, which cover the precipitation maxima in the Black Forest and lead to a lower correlation coefficient of 0.40.

Over the complex terrain in D-South (Fig. 4), the simulated precipitation patterns are close to the observations with spatial correlation between 0.79 and 0.85. The locations of minima and maxima and the transitions in between are fit well, even for the finescale structures

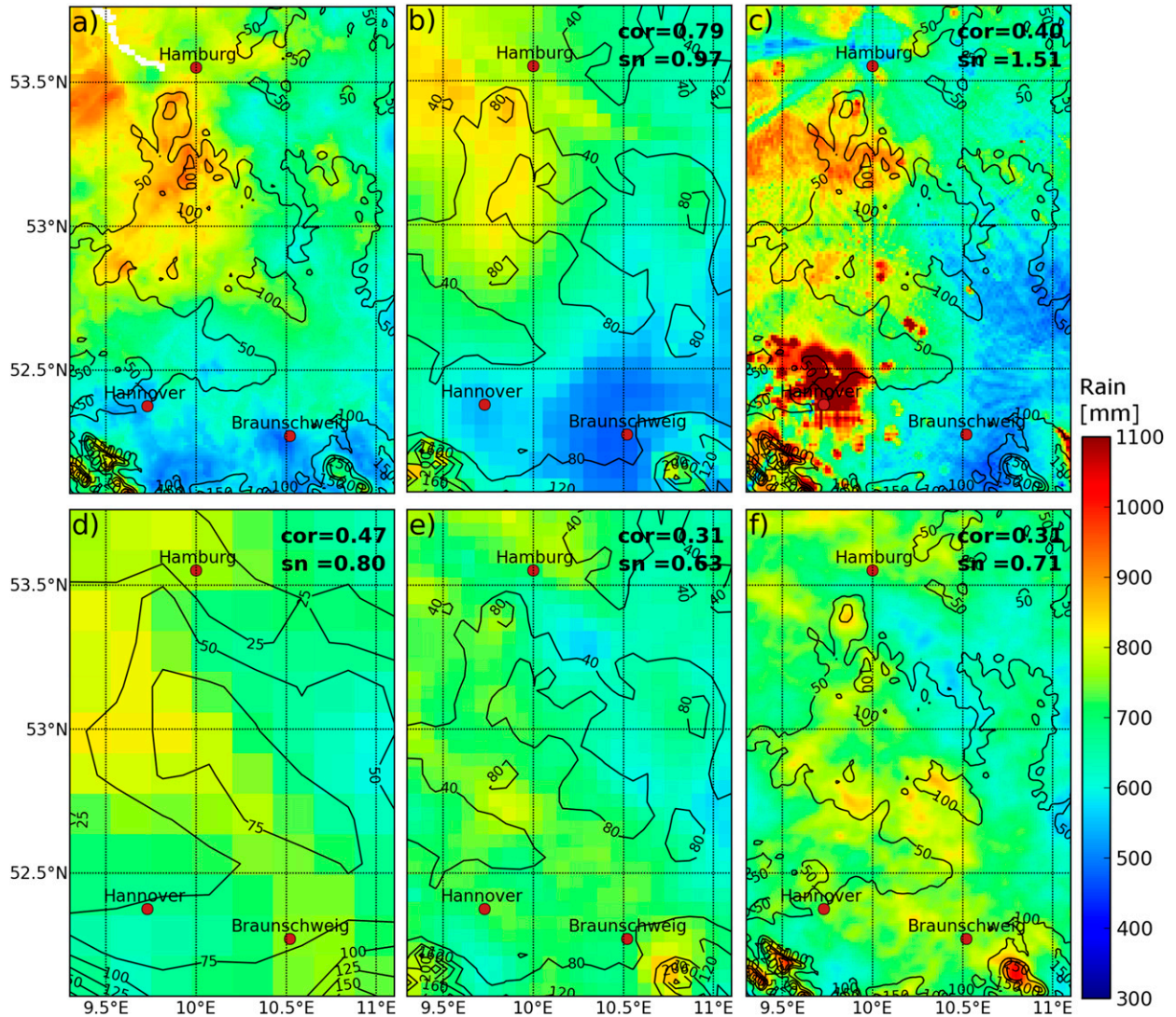


FIG. 3. Distribution of annual precipitation amounts of validation data (a) REG-1, (b) STA-5, and (c) RAD-1 and of simulation data (d) R1P-15, (e) R1P-5, and (f) R1E-1 for D-North at their original resolution, overlaid by contour lines of elevation in meters. In the top-right corner of each plot, the correlation coefficient (cor) and the normalized standard deviation (sn) compared to REG-1 based on a common 15-km grid spacing are shown. The evaluation period is June 2005–May 2006.

in R1E-1 (Fig. 4f). However, in this case the minima and maxima are too pronounced, which is also expressed in an increase in sn for R1P-5 (1.28) and R1E-1 (1.35).

3) SPATIAL SIMILARITIES IN SUMMER

The plots in Fig. 5 show the average decrease of the temporal correlation of 3-h precipitation with distance based on two precipitation thresholds, individually for each dataset. It should be noted that the investigation of 1-h data (excluding R1P-15) leads to comparable results (not shown).

The differences between the observational datasets reveal their respective shortcomings. STA-5 is not able

to adequately reproduce small-scale features in summer and complex terrain, which leads to smaller decreases of correlations with distance (Figs. 5c,d). In fact, radar data are ideal for this purpose and as such are chosen as the validation dataset.

In summer, the differences among the simulations are large due to a high influence of small-scale convective events, which lead to specific variations with topography. In the flat terrain of D-North (Figs. 5a,b), a separation depending on convection treatment is observed with a high similarity of simulations with convection parameterization (R1P-15 and R1P-5) and with an explicit calculation of convection (R1E-1 and R2E-5).

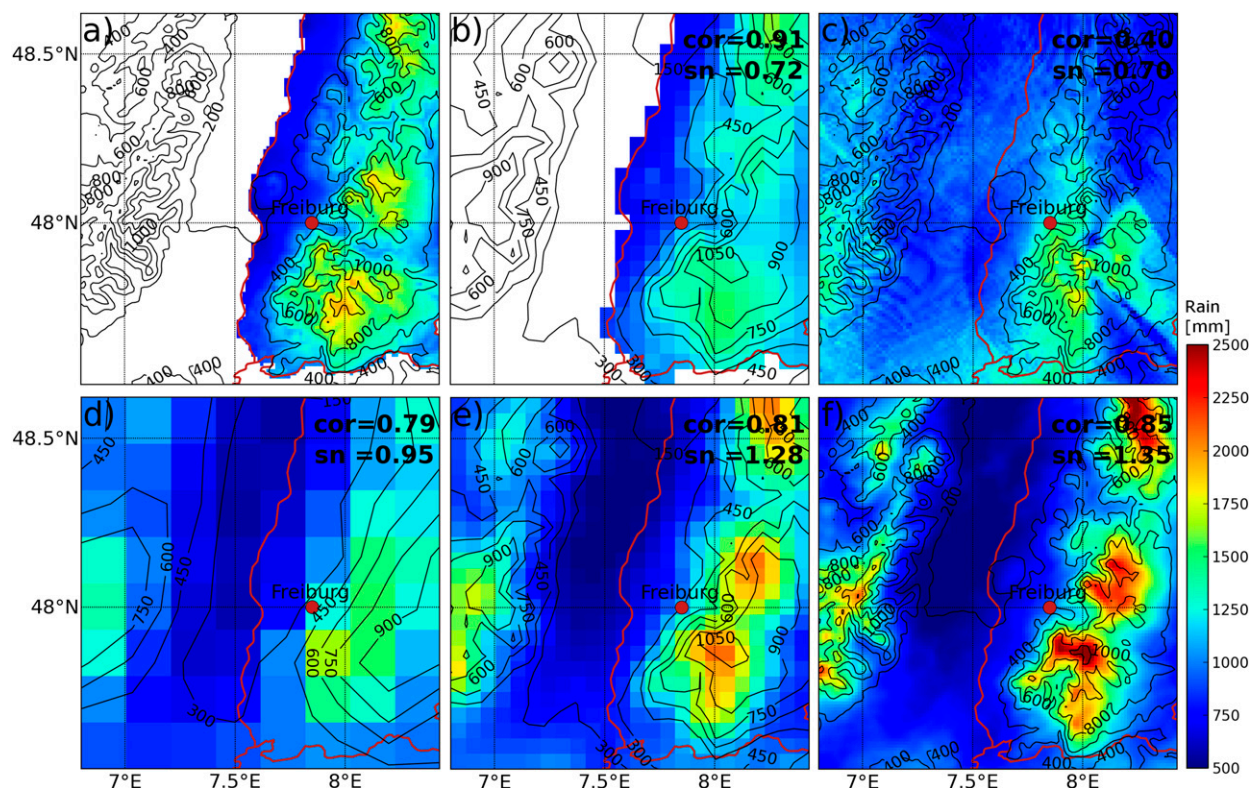


FIG. 4. As in Fig. 3, but for D-South.

Compared to RAD-1, the latter simulations are superior with a very similar decrease of correlation with distance. In D-South, a separation by spatial resolution becomes apparent with a similar decrease of correlations of R1P-5 and R2E-5 with distance and even smaller correlations for R1E-1 with distance. For all precipitation events (Fig. 5c), the simulations with 5 km show a high accordance with RAD-1, and for higher precipitation amounts (Fig. 5d) R1E-1 performs best, especially for distances below 25 km.

In general, the decrease in correlation of the simulation results with smaller grid spacing indicates an increase of the variability of precipitation patterns in space. The high-resolution simulations (1 and 5 km) perform well, whereas the decrease of R1P-15's correlations is too weak for all precipitation intensities and in both domains. This is most pronounced for areas with complex terrain and higher precipitation intensities (Fig. 5d).

4) SPATIAL SIMILARITIES IN WINTER

The limitations in RAD-1 lead to a faster drop in correlation with distance. Therefore, in contrast to summer, the STA-5 dataset is used for the validation for domain D-North in winter (Figs. 6a,b).

In winter, the high-resolution simulations R1P-5 and R1E-1 perform similarly for both domains (Fig. 6) with a stronger decline with distance in D-South compared to validation datasets. The lack of convection activity in the winter is probably responsible for the fact that even in the complex terrain of D-South only minor differences between both simulations exist if the grid spacing is further reduced (1 km vs 5 km).

In fact, R1P-15 performs similarly in flat terrain and winter (Fig. 6a). However, in complex terrain and for higher precipitation intensities, significant differences between R1P-15 and the high-resolution simulations remain (Figs. 6b-d).

5) SKILL SCORE TO COMPARE HIGH-RESOLUTION PRECIPITATION PATTERNS IN SUMMER

The FSS in Fig. 7 is presented for four thresholds in summer for both investigation areas. All FSS values exceed the limit of random skill (dashed line in Fig. 7). The FSS of all simulations generally increases with horizontal scale and decreases with larger precipitation thresholds.

In line with the correlograms, the FSS also reveals a separation according to convection treatment in D-North and according to grid spacing in D-South in summer. This

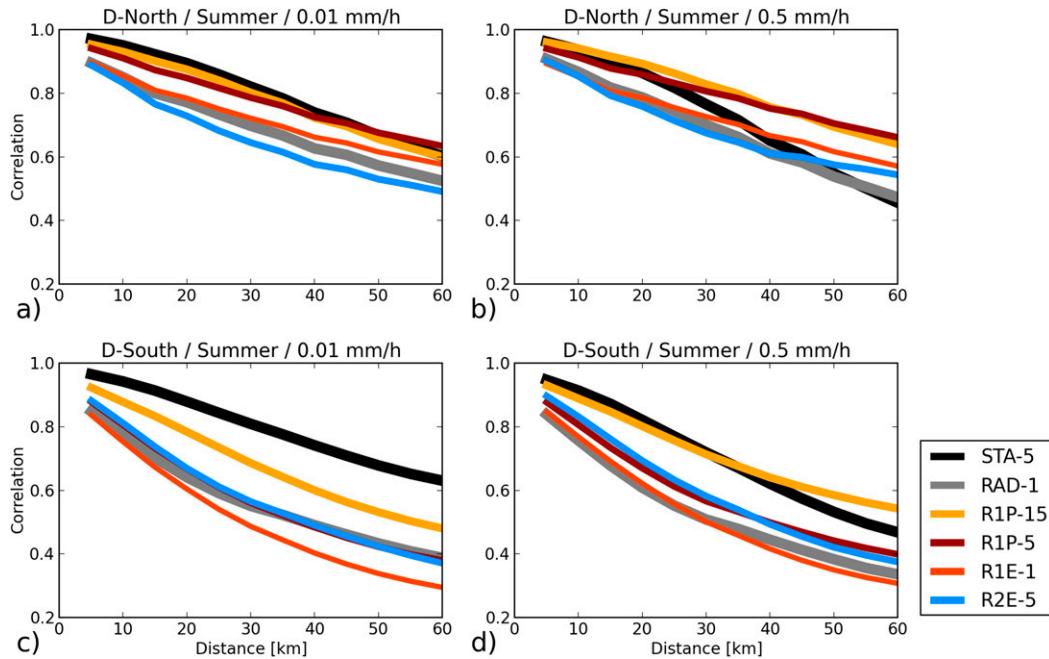


FIG. 5. Correlograms of all available datasets based on a 5-km grid and 3-h values for (a),(b) D-North and (c),(d) D-South for the summer months. Two different thresholds (0.01 and 0.5 mm h^{-1}) represent (left) all precipitation intensities and (right) the higher precipitation intensities. For the summer months, the additional simulation R2E-5 is also available.

separation is only very weak for the lowest threshold. Simulations without convective parameterization show the highest FSS values in D-North. R2E-5 outperforms R1E-1 for the largest thresholds (Figs. 7c,d). In D-South,

R1E-1 performs best, whereas the 5-km simulations (R1P-5 and R2E-5) are similar. Only for the largest threshold (Fig. 7h), is R2E-5 similar to R1E-1. Hence, the FSS in complex terrain is mainly influenced by grid

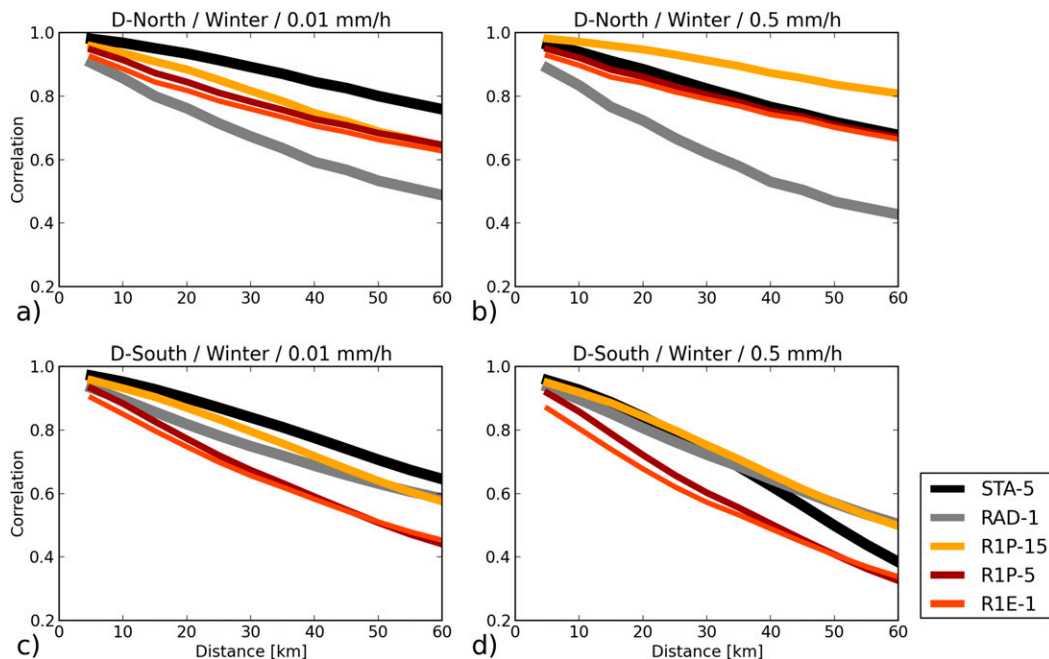


FIG. 6. As in Fig. 5, but for winter.

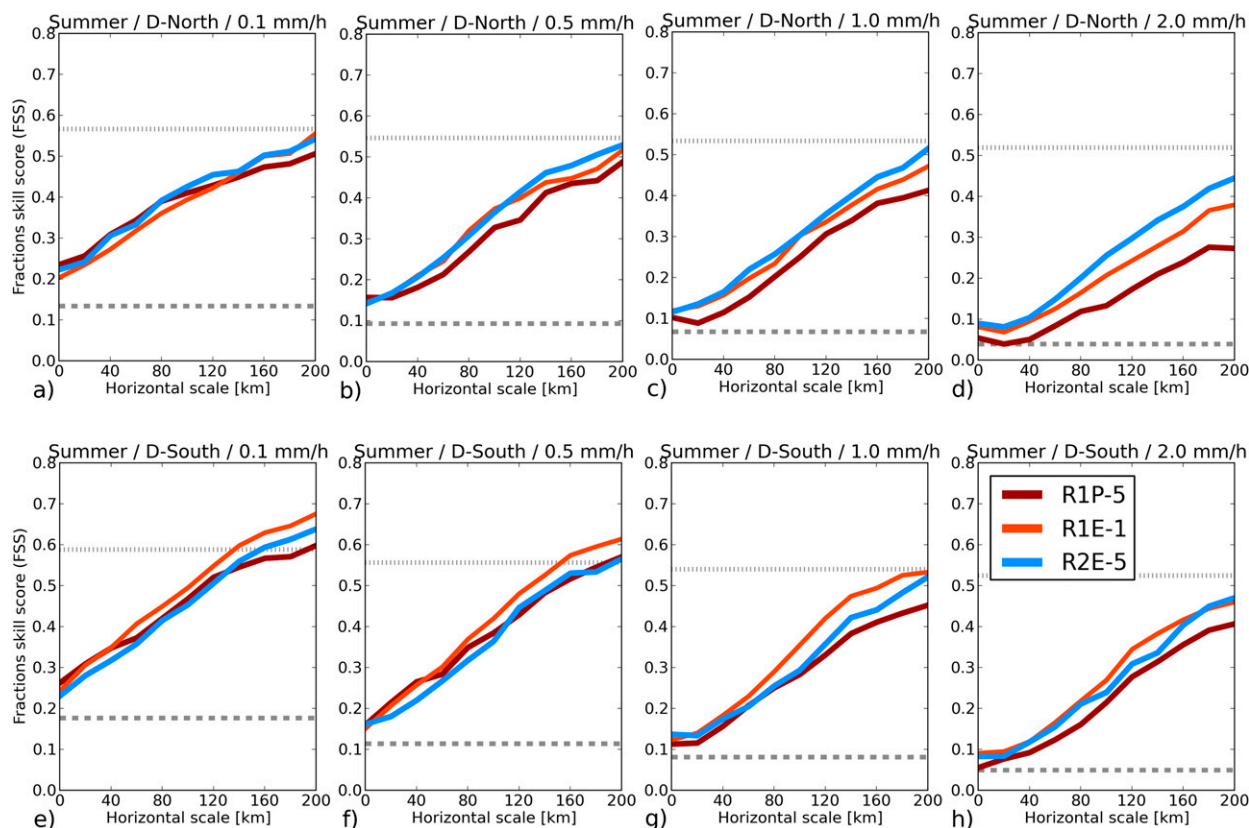


FIG. 7. Fractions skill scores (FSSs) for simulation results of hourly precipitation remapped to 5-km grid spacing in (a)–(d) D-North and (e)–(h) D-South, for different thresholds (0.1, 0.5, 1.0, and 2.0 mm) and different horizontal scales for the summer months. The FSS of a random simulation is plotted as a dashed line (bottom line), whereas reasonable skill is achieved for values larger than the dotted line (top line). RAD-1 is used as a reference. The evaluation period is June 2005–May 2006.

spacing, whereas for higher precipitation intensities, the treatment of convection dominates.

6) SKILL SCORE TO COMPARE HIGH-RESOLUTION PRECIPITATION PATTERNS IN WINTER

In winter, only R1E-1 and R1P-5 are available for comparison. The FSS in Fig. 8 is presented for only three thresholds for both investigation areas because of lower precipitation intensities in winter. Again, all FSS values exceed the limit of random skill (dashed line in Fig. 8), and the FSS decreases with larger precipitation thresholds.

For small thresholds, the FSS is slightly higher than in summer (Figs. 8a and 8d). For thresholds of 0.5 mm h^{-1} and above, a similar behavior to summer is observed: R1E-1 outperforms R1P-5 for all thresholds and horizontal scales.

c. Temporal evaluation

1) MEAN TEMPORAL MEASURES

According to Table 3, RAD-1 agrees well with STA-5, which serves as a reference with cor of approximately

0.8 in summer and winter. The sn is also similar, except for winter in D-North where clutter influence leads to an increased sn of 1.40.

The high-resolution simulations show low cor values between 0.24 and 0.33 in summer in both domains. Despite the overestimation of sn values, R2E-5 performs best in D-North and R1E-1 performs best in D-South, but the differences are weak. The agreement of sn values of both simulations indicate that the explicit calculation of convection is responsible for this similar behavior. The coarser R1P-15 performs better with a slightly higher cor (0.4 and 0.3), but a reduced sn in summer. This is likely due to the already mentioned double-penalty problem. Since the high-resolution simulations can resolve spatial gradients better than course resolution simulations, they get penalized stronger if small shifts in space and time do occur. Additionally, R1P-15 is tied closer to the driving data by spectral nudging than the high-resolution simulations.

In winter, all simulations are very similar irrespective of grid spacing, convection parameterization, and terrain. The cor values (0.52–0.59) are higher than

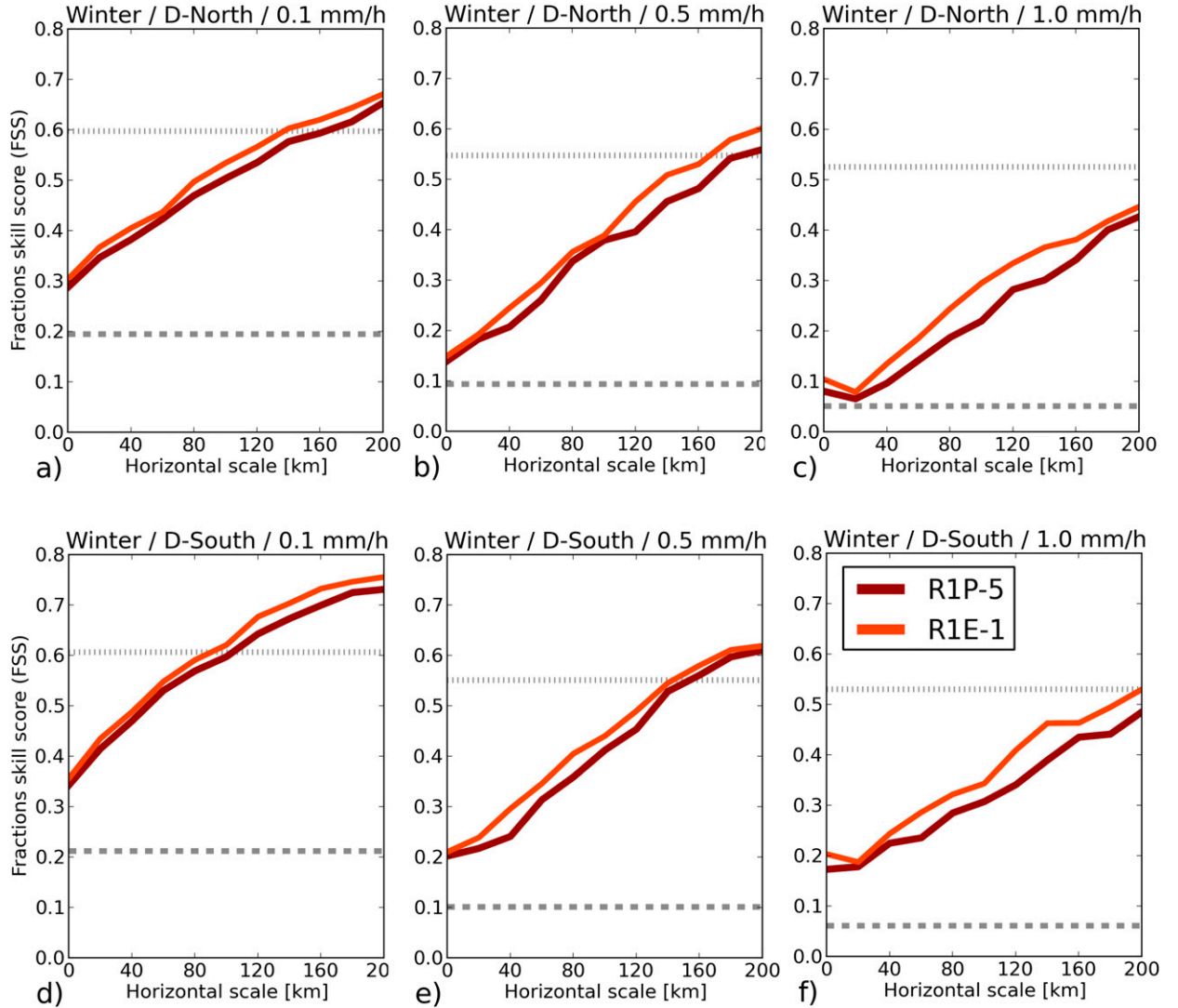


FIG. 8. As in Fig. 7, but for the winter months and only for the thresholds (0.1, 0.5, and 1.0 mm).

in summer, and sn is increased by about 20% compared to STA-5.

This analysis of the temporal aspect of simulated precipitation patterns does not reveal improved results for grid spacings smaller than 15 km.

2) DIURNAL CYCLE OF PRECIPITATION AMOUNTS

Both validation datasets show a similar diurnal cycle of precipitation in summer (Figs. 9a,c). However, a shift of the peak in D-South for RAD-1 appears that cannot be entirely explained within this framework, but is likely caused by the transfer of reflectivities into rain rate and by measuring artifacts.

The high-resolution simulations mimic the daily increase and decrease of precipitation well. The explicit calculation of convection for R2E-5 leads to the best

performance, especially in D-North. In contrast, R1E-1 and R1P-5 show a very similar daily cycle, suggesting that the daily cycle for R1E-1 is mainly adopted from the coarser R1P-5 simulation and not produced by the

TABLE 3. Normalized standard deviation (sn) and correlation coefficient (cor) of precipitation of all available datasets compared to STA-5 based on 3-h and 5-km resolution for summer and winter.

Dataset	Summer (D-North)		Summer (D-South)		Winter (D-North)		Winter (D-South)	
	sn	cor	sn	cor	sn	cor	sn	cor
RAD-1	0.98	0.82	1.03	0.75	1.40	0.78	0.90	0.80
R1P-15	0.90	0.40	0.75	0.30	1.24	0.52	1.18	0.59
R1P-5	1.45	0.29	0.96	0.24	1.17	0.53	1.23	0.58
R1E-1	1.56	0.28	1.19	0.29	1.22	0.54	1.23	0.59
R2E-5	1.57	0.33	1.14	0.26	—	—	—	—

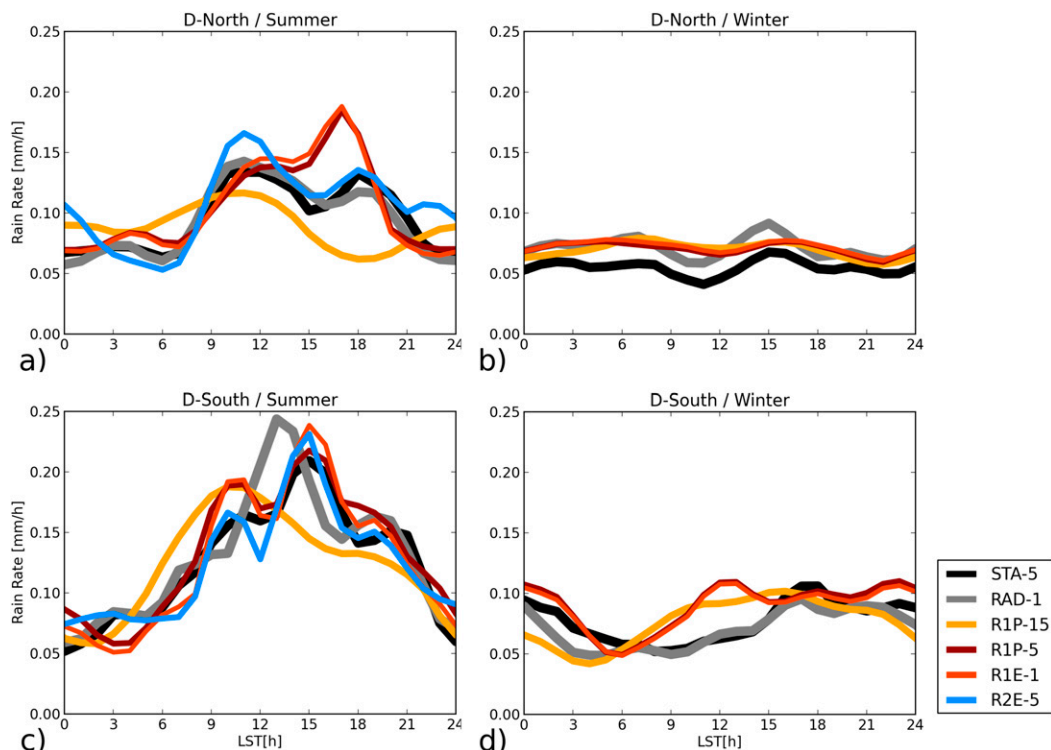


FIG. 9. Mean diurnal cycle of precipitation for the temporal evaluation of simulation results based on 1-h values for (a),(b) D-North and (c),(d) D-South, and for (left) the summer and (right) winter months. A 3-h running mean is applied. The additional simulation R2E-5 is only available for the summer months.

explicit calculation of convection. The domain of R1E-1 seems to be too small to allow an independent temporal evolution comparable to R2E-5. In D-South, the differences between these three simulations are weak but the close accordance of R1E-1 and R1P-5 remains. The coarser R1P-15 simulation shows a much poorer performance with a premature increase and the lack of the late afternoon maximum. This is an issue commonly seen at resolutions where convection is under-resolved at grid scale and instead parameterized (Prein et al. 2015).

In winter (Figs. 9b,d), no clear daily cycle is visible and the differences are difficult to interpret. Convective activity, which is usually the driver for a specific daily precipitation regime, is minor in winter. In this way, strong single events can dominate the daily cycle of precipitation in the winter season.

d. Evaluation of the probability density function of precipitation

Figure 10 shows the probability density function of precipitation and the dry 3-h probabilities. This probability of dry time steps ($<0.1 \text{ mm h}^{-1}$) for each dataset is calculated and added in the legend. For the validation datasets, the dry 3-h probability is approximately 0.85 in both summer and winter. Over the summer months, the

simulations with convection parameterization (R1P-5 and R1P-15) tend to underestimate this probability, whereas R1E-1 and especially R2E-5 tends to overestimate it. This is in line with Berg et al. (2013) and Ban et al. (2014), who report an underestimation of dry day probabilities and trace this result to convection parameterizations.

The PDF curves of both validation datasets are comparable in summer. This is also true for the high-resolution simulation PDFs for intensities below 7 mm (3 h)^{-1} . For higher intensities they tend to cluster by model setup (e.g., R1E-1 and R1P-5 vs R2E-5) in D-North and by handling convection (R1P-5 vs R1E-1 and R2E-5) in D-South. The agreement of the PDFs with the observations is better in D-South, even for higher precipitation intensities. The R1P-15 PDF reveals the lowest probabilities in both areas, different from the validation datasets. Small, but systematic differences become apparent for R1P-5 and R2E-5 with lower probabilities for low intensities and higher probabilities for medium and most high intensities for R2E-5 compared to R1P-5. A detailed PDF-analysis (not shown) reveals the slightly better agreement of R1P-5 with validation datasets more clearly.

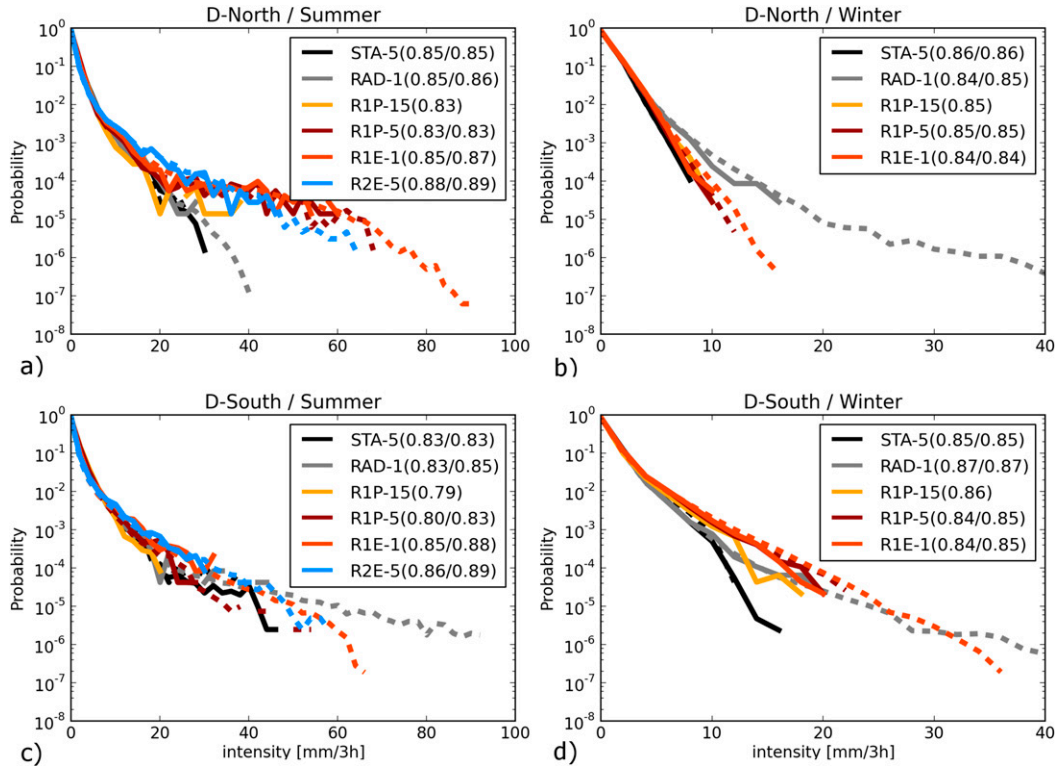


FIG. 10. Probability density functions (PDFs) of all available datasets with 15-km grid spacing and 3-h resolution for (left) summer and (right) winter for (a),(b) D-North and (c),(d) D-South. The dashed lines display the PDFs of all datasets at their original spatial resolution (if differing from 15-km grid spacing). The dry 3-h probability of each dataset is annotated in the legend in brackets (15-km grid spacing/original grid spacing).

In winter (Figs. 10b,d), all simulations exhibit similar dry 3-h probabilities as the validation datasets. The curves of interpolated station data and radar data differ in D-North, probably because of the clutter in the radar data. The PDFs of the three simulations are similar in both domains. Only R1P-15 shows slightly lower probabilities for higher precipitation intensities in D-South as a result of the smoother topography of the Black Forest at 15-km resolution, where usually the highest precipitation amounts are observed.

6. Discussion

The results reflect a combination of influencing factors, including the calculation of convection, grid spacing, and topography. The decisive factor is how well the relevant processes of precipitation formation such as orographic forcing in complex terrain and the timing and spatial distribution of convective precipitation can be resolved by the model. In flat areas in winter and for all precipitation events (i.e., when these factors are of lesser importance), the differences between the simulation results of R1P-15 and R1E-1 are almost negligible (Fig. 6a).

However, in summer and especially for high intensities and in complex terrain (Figs. 5c,d) the variations among the simulations increase and lead to an improved representation of temporal and spatial characteristics of precipitation with smaller grid spacings. This is in line with the findings of, for instance, Kendon et al. (2012) and Prein et al. (2013b), with the difference that their identified added value of smaller grid spacings refers to the comparison of coarser grid spacings (>10 km) and simulations with fine grid spacings (≤ 5 km) mixing the potential added value of the explicit calculation of convection and the better representation of land use and terrain. The emphasis of our work lies in the further separation of reasons and consequences of those variations, focusing on several high-resolution simulations.

a. Impact of the calculation of convection

The strongest argument for the expected added value of high-resolution simulations is the removal of the convection parameterization, so its influence is evaluated first. The simulation results of R1P-5 are almost exclusively closer to R2E-5 and R1E-1 than to R1P-15, even though for both simulations the

GF-parameterization scheme is applied. This parameterization offers a smooth transition to cloud resolving scales. It acts as a typical cumulus parameterization scheme for deep convection with a cloud top at about 300 hPa for 15-km grid spacing. With higher spatial resolutions, the entrainment rates are forced to increase and the heating and drying tendencies decrease. According to Grell and Freitas (2014), the scheme acts as a shallow-cumulus scheme with the cloud top at about 800 hPa for 1-km grid spacing. In this way, the unresolved convective activity and parameterized precipitation amounts decrease significantly when using smaller grid spacings as shown in Langhans et al. (2012) and Grell and Freitas (2014). At a grid spacing of 5 km, the parameterization is rather inactive and the differences between the high-resolution simulations are small for almost all evaluations.

In general, R2E-5 and the corresponding R1P-5, employing the Grell–Freitas scale-aware cumulus parameterization, are of comparable quality. The results confirm the applicability of this convection parameterization for smaller grid spacings even within the gray zone of convection. In this way, the spatial patterns (Figs. 5–8) and the diurnal cycle (Fig. 9) are represented successfully in both domains within this framework. Donner et al. (2011) and Bechtold et al. (2014) also reported improvements in the representation of the diurnal cycle due to a new parameterization scheme.

Nevertheless, differences between R1P-5 and R2E-5 remain. In this work, the main added value of simulations without convection parameterization at 5-km grid spacing (R2E-5) can be observed for the domain with flat terrain of D-North. The spatial patterns (correlograms in Figs. 5a,c and the FSS in Figs. 7a–d) and also the diurnal cycle (Fig. 9a) are slightly better represented by R2E-5 than by R1P-5. In D-South, the improvements are restricted to more intense precipitation (FSS in Fig. 7h), whereas a poorer performance is observed for the probability density function of precipitation (PDFs in Fig. 10a). According to Weisman et al. (1997) and Prein et al. (2015), convective instability at the grid scale is forced toward an unrealistic scale for grid spacings larger than 4 km. This leads to an overestimation of the convective mass flux and the associated precipitation. This effect may explain the overestimation of probabilities for the observed intensities from approximately 10–30 mm (3 h)^{−1} in R2E-5.

In summary, the advantages of the explicit calculation of convection clearly dominate the drawbacks in our evaluations. Therefore, our results suggest that an application of a convective parameterization is not necessary at 5-km grid spacing, independent of the complexity of the terrain.

b. Added value of smaller grid spacings and the influence of topography

Leveraging the advantages of an explicit calculation of convection and a smaller grid spacing, R1E-1 should outperform the other simulations. However, in our study this is only true in the complex terrain of D-South, where the higher spatial resolution shows its advantage especially for evaluations of spatial patterns (correlograms in Fig. 5d and the FSS in Figs. 7e–h), and mainly for higher precipitation amounts. Chan et al. (2013) in the southern United Kingdom and Ban et al. (2014) in Switzerland also showed improvements of daily precipitation amounts only for higher percentiles (90% and 95%). However, for R1E-1, the more pronounced maxima and smaller precipitation cells in complex terrain for smaller precipitation intensities (correlograms in Fig. 6a and sn in Figs. 4c,d) suggest an increased influence of orographic forcing. The reasons for this behavior are not further investigated within this framework, but this feature might also not indicate a shortcoming of the models themselves, but possibly also a shortcoming in the validation datasets because of the spatiotemporal smoothing effects.

Conversely, in D-North, the simulations without convective parameterization R1E-1 and R2E-5 perform very similar for spatial evaluations with no clear improvements when reducing the grid spacing from 5 to 1 km in this region.

These results clearly show the influence of topography even on high-resolution simulation results. In D-North, added value is achieved by the explicit calculation of convection, whereas in D-South, the respective similarities of spatial evaluations reveal a dominating influence of resolving the orography. As the topography induces convection and guides the location where precipitation occurs, the location of precipitation maxima are better captured than in areas without any topographic forcing, as long as the model resolution is high enough to reflect the topographic variation (precipitation patterns in Figs. 3 and 4). As a result, the 15-km runs fail to generate quality precipitation amounts and distribution for D-South, while the 5-km and 1-km runs perform well.

c. Influence of parent domain

The prerequisite for a successful development of precipitation patterns in the model is a sufficiently large domain. Brisson et al. (2016a) analyzed the influence of the domain size in the CPS with 3-km grid spacing. They observed a significant decrease of precipitation amounts for smaller domain sizes than about 160 × 160 to 180 × 180 grid columns and concluded that a minimum distance from the boundary to the domain of interest

of ~ 150 km is necessary for a proper spatial spinup. In our case, the finest domains D03N and D03S have about the same number of grid columns, albeit for a smaller grid spacing. In another study, [Matte et al. \(2017\)](#) concluded that the spinup distance corresponds to a fixed number of grid points in the RCM. Since our 5-km domain ought to be sufficiently large and our results for R1E-1 reproduce the precipitation patterns of R1P-5 well, we conclude that, despite the smaller total domain extent, the number of grid columns in the finest model is large enough to provide meaningful precipitation patterns.

However, since our 1-km domains are too small to encompass entire mesoscale systems, the independent development of large-scale precipitation patterns is limited. Only smaller systems such as convective cells can develop independently within such a domain. Accordingly, the very similar behavior of R1P-5 and R1E-1 regarding the diurnal cycle ([Fig. 9](#)) are hints for a significant influence of the parent domain on the results of the child domain.

The good performance of R2E-5 suggests further experiments with nesting our 1-km simulations in this domain in addition to R1P-5. Slight differences between simulation results are expected because of varying boundary conditions from the parent domain.

7. Conclusions and outlook

According to our results, a grid spacing of 15 km can provide only limited details of precipitation patterns, whereas a 5-km grid spacing captures the spatiotemporal variability of precipitation, adequately. Further improvements for smaller grid spacings (1 km) are observed in complex terrain (D-South) and especially for higher precipitation intensities in summer ([Fig. 5d](#)). Consequently, computational costs can be saved in winter, when convective activity plays a minor role, and in flat areas. The representation of spatial patterns is dominated by the treatment of convection in D-North and by resolving orography in D-South.

The temporal performance (diurnal cycle in [Fig. 9](#)) of R1E-1 by contrast is influenced by the parent domain R1P-5 in both areas. Thus, an independent development of precipitation patterns in time and space is partly limited for R1E-1. Larger domains comprising mesoscale weather systems may reduce this influence, but require extensive computation devices ([Leutwyler et al. 2017](#)).

The WRF simulations produce the largest precipitation amounts at the highest altitudes. For the complex terrain in D-South, a tight correlation of observed precipitation maxima and altitude is found,

whereas in D-North, this is less pronounced in the observations. Consequently, the locations of precipitation maxima are much better captured by our simulations in D-South than in D-North. In D-North it is shown that improvements of small-scale patterns ([Figs. 5a,b](#) and [6a,b](#)) may also cause a deterioration of annual precipitation patterns with smaller grid spacings ([Figs. 3e,f](#)). As indicated before, an improvement may be made by using spectral nudging also on the 5-km grid spacings, rather than on the outermost domain only. This should be investigated in the future.

At the lower end of the gray zone of convection (here: at 5 km), reasonable results are achieved when using the new, scale-aware Grell–Freitas convection parameterization, as well as by relying on an explicit calculation of convection only. Depending on the statistical measure and investigation area, either of the two methods can lead to improvements. Consequently, for simulations at this resolution, it is worth testing both approaches and evaluating their performance with respect to the specific problem to solve. While grid spacings in the gray zone of convection are generally avoided by the modeling community, they are of interest in the context of long-term climate simulations, for which experiments at smaller grid spacings are simply too expensive. Thus, the progress made in the development of scale-aware cumulus parameterizations may be of high importance for the climate modeling community.

With smaller grid spacings, the impact of processes such as orographic forcing seems to increase. A thorough analysis of influencing factors such as wind and radiation, beyond the scope of this work, may help to achieve further indications for this behavior. In addition, new PBL parameterization schemes that are designed for resolutions on the order of 1 km should be tested (e.g., [Shin and Hong 2013](#)), since partly resolved vertical transport in combination with a PBL parameterization scheme that is designed for different grid spacings might also affect precipitation results.

Limitations of the analysis presented here are the relatively small areas of investigation and short evaluation periods, at least in the context of climate modeling, which is mainly due to the computational costs required to perform these simulations. A larger statistical basis, both spatially and temporally, will help to reduce the influence of single events on, in particular, the probability density functions ([Fig. 10](#)) and the fractions skill score analysis ([Figs. 7](#) and [8](#)) presented here.

The presented investigation looks beyond the common comparison of simulations with coarse and fine grid spacings. It reveals aspects where potentials for further improvements for high-resolution simulations are offered and also where computational time can be saved.

Acknowledgments. The research was conducted in the framework of the SYNOPSE project, with the financial support of the BMBF (Federal Ministry of Education and Research, Germany). The authors would like to acknowledge the German Meteorological Service (DWD) for supplying us with radar data, rain gauge data, and REGNIE data; the ECMWF for making their reanalysis dataset ERA-Interim available; and the European Environment Agency for providing the high-resolution land use dataset CORINE. The modeling experiments were conducted on the Linux Cluster of the Leibniz Supercomputing Centre of the Bavarian Academy of Science and Humanities (LRZ) in Garching, Germany, near Munich. Finally, the authors thank three anonymous reviewers for their valuable feedback.

REFERENCES

- Arnault, J., S. Wagner, T. Rummeler, B. Fersch, J. Bliefernicht, S. Andresen, and H. Kunstmann, 2016: Role of runoff-infiltration partitioning and resolved overland flow on land-atmosphere feedbacks: A case study with the WRF-Hydro coupled modeling system for West Africa. *J. Hydrometeorol.*, **17**, 1489–1516, <https://doi.org/10.1175/JHM-D-15-0089.1>.
- Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905, <https://doi.org/10.1175/MWR-D-10-05013.1>.
- Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res. Atmos.*, **119**, 7889–7907, <https://doi.org/10.1002/2014JD021478>.
- Bartels, H., E. Weigl, T. Reich, P. Lang, A. Wagner, O. Kohler, and N. Gerlach, 2004: *Projekt RADOLAN—Routineverfahren zur Online-Aneicherung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen (Ombrometer)* [Routine procedure for online calibration of radar precipitation data using automatic rain gauges (ombrometer)]. Deutscher Wetterdienst, 111 pp.
- Bechtold, P., N. Semane, P. Lopez, J.-P. Chaboureaud, A. Beljaars, and N. Bormann, 2014: Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, **71**, 734–753, <https://doi.org/10.1175/JAS-D-13-0163.1>.
- Berg, P., S. Wagner, H. Kunstmann, and G. Schädler, 2013: High resolution regional climate model simulations for Germany: Part I-Validation. *Climate Dyn.*, **40**, 401–414, <https://doi.org/10.1007/s00382-012-1508-8>.
- Brisson, E., M. Demuzere, and N. P. M. van Lipzig, 2016a: Modelling strategies for performing convective permitting climate simulations. *Meteor. Z.*, **25**, 149–163, <https://doi.org/10.1127/metz/2015/0598>.
- , K. Van Weverberg, M. Demuzere, A. Devis, S. Saeed, M. Stengel, and N. P. M. van Lipzig, 2016b: How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics? *Climate Dyn.*, **47**, 3043–3061, <https://doi.org/10.1007/s00382-016-3012-z>.
- Bryan, G., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140**, 202–225, <https://doi.org/10.1175/MWR-D-11-00046.1>.
- Chan, S., E. Kendon, H. Fowler, S. Blenkinsop, C. A. Ferro, and D. B. Stephenson, 2013: Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate Dyn.*, **41**, 1475–1495, <https://doi.org/10.1007/s00382-012-1568-9>.
- Chen, F., and Coauthors, 2011: The integrated WRF/urban modelling system: Development, evaluation, and applications to urban environmental problems. *Int. J. Climatol.*, **31**, 273–288, <https://doi.org/10.1002/joc.2158>.
- Ching, J., R. Rotunno, M. A. LeMone, A. Martilli, B. Kosovic, P. A. Jimenez, and J. Dudhia, 2014: Convectively induced secondary circulations in fine-grid mesoscale numerical weather prediction models. *Mon. Wea. Rev.*, **142**, 3284–3302, <https://doi.org/10.1175/MWR-D-13-00318.1>.
- Clark, A. J., W. A. Gallus, M. Xue, and F. Kong, 2009: A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting*, **24**, 1121–1140, <https://doi.org/10.1175/2009WAF2222222.1>.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **137**, 553–597, <https://doi.org/10.1002/qj.828>.
- Deutsch, C. V., and A. G. Journel, 1992: *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, 340 pp.
- Donner, L. J., and Coauthors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484–3519, <https://doi.org/10.1175/2011JCLI3955.1>.
- Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, [https://doi.org/10.1175/1520-0469\(1989\)046<3077:NSOCOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2).
- Fabry, F., and I. Zawadzki, 1995: Long-term radar observations of the melting layer of precipitation and their interpretation. *J. Atmos. Sci.*, **52**, 838–851, [https://doi.org/10.1175/1520-0469\(1995\)052<0838:LTROOT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1995)052<0838:LTROOT>2.0.CO;2).
- Fosser, G., S. Khodayar, and P. Berg, 2015: Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Climate Dyn.*, **44**, 45–60, <https://doi.org/10.1007/s00382-014-2242-1>.
- Franco, M., R. Sanchez-Diezma, and D. Sempere-Torres, 2006: Improvements in weather radar rain rate estimates using a method for identifying the vertical profile of reflectivity from volume radar scans. *Meteor. Z.*, **15**, 521–536, <https://doi.org/10.1127/0941-2948/2006/0154>.
- Grell, G. A., and S. R. Freitas, 2014: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. *Atmos. Chem. Phys.*, **14**, 5233–5250, <https://doi.org/10.5194/acp-14-5233-2014>.
- , L. Schade, R. Knoche, A. Pfeiffer, and J. Egger, 2000: Non-hydrostatic climate simulations of precipitation over complex terrain. *J. Geophys. Res.*, **105**, 29 595–29 608, <https://doi.org/10.1029/2000JD900445>.
- Haylock, M., N. Hofstra, A. Klein Tank, E. Klok, P. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *J. Geophys. Res.*, **113**, D20119, <https://doi.org/10.1029/2008JD010201>.
- Heinzeller, D., W. Junkermann, and H. Kunstmann, 2016: Anthropogenic aerosol emissions and rainfall decline in

- southwestern Australia: Coincidence or causality? *J. Climate*, **29**, 8471–8493, <https://doi.org/10.1175/JCLI-D-16-0082.1>.
- Hohenegger, C., P. Brockhaus, and C. Schär, 2008: Towards climate simulations at cloud-resolving scales. *Meteor. Z.*, **17**, 383–394, <https://doi.org/10.1127/0941-2948/2008/0303>.
- Hong, S., and J. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- , Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, <https://doi.org/10.1175/MWR3199.1>.
- Jacob, D., and Coauthors, 2014: Euro-CORDEX: new high-resolution climate change projections for European impact research. *Reg. Environ. Change*, **14**, 563–578, <https://doi.org/10.1007/s10113-013-0499-2>.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Kendon, E., N. Roberts, C. Senior, and M. Roberts, 2012: Realism of rainfall in a very high-resolution regional climate model. *J. Climate*, **25**, 5791–5806, <https://doi.org/10.1175/JCLI-D-11-00562.1>.
- Klein, C., J. Bliefernicht, D. Heinzeller, U. Gessner, I. Klein, and H. Kunstmann, 2017: Feedback of observed interannual vegetation change: A regional climate model analysis for the West African monsoon. *Climate Dyn.*, **48**, 2837–2858, <https://doi.org/10.1007/s00382-016-3237-x>.
- Knight, C. G., and Coauthors, 2007: Association of parameter, software, and hardware variation with large-scale behavior across 57,000 climate models. *Proc. Natl. Acad. Sci. USA*, **104**, 12 259–12 264, <https://doi.org/10.1073/pnas.0608144104>.
- Kotlarski, S., and Coauthors, 2014: Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model Dev.*, **7**, 1297–1333, <https://doi.org/10.5194/gmd-7-1297-2014>.
- Krajewski, W. F., B. Vignal, B.-C. Seo, and G. Villarini, 2011: Statistical model of the range-dependent error in radar-rainfall estimates due to the vertical profile of reflectivity. *J. Hydrol.*, **402**, 306–316, <https://doi.org/10.1016/j.jhydrol.2011.03.024>.
- Kunstmann, H., and C. Stadler, 2005: High resolution distributed atmospheric-hydrological modelling for alpine catchments. *J. Hydrol.*, **314**, 105–124, <https://doi.org/10.1016/j.jhydrol.2005.03.033>.
- , K. Schneider, R. Forkel, and R. Knoche, 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, <https://doi.org/10.5194/hess-8-1031-2004>.
- Kusaka, H., and F. Kimura, 2004: Coupling a single-layer urban canopy model with a simple atmospheric model: Impact on urban heat island simulation for an idealized case. *J. Meteor. Soc. Japan*, **82**, 67–80, <https://doi.org/10.2151/jmsj.82.67>.
- , H. Kondo, Y. Kikigawa, and F. Kimura, 2001: A simple single-layer urban canopy model for atmospheric models: Comparison with multi-layer and slab models. *Bound.-Layer Meteor.*, **101**, 329–358, <https://doi.org/10.1023/A:1019207923078>.
- Langhans, W., J. Schmidli, and C. Schär, 2012: Bulk convergence of cloud-resolving simulations of moist convection over complex terrain. *J. Atmos. Sci.*, **69**, 2207–2228, <https://doi.org/10.1175/JAS-D-11-0252.1>.
- Leutwyler, D., D. Lüthi, N. Ban, O. Fuhrer, and C. Schär, 2017: Evaluation of the convection-resolving climate modeling approach on continental scales. *J. Geophys. Res. Atmos.*, **122**, 5237–5258, <https://doi.org/10.1002/2016JD026013>.
- Mahoney, K. M., 2016: The representation of cumulus convection in high-resolution simulations of the 2013 Colorado Front Range flood. *Mon. Wea. Rev.*, **144**, 4265–4278, <https://doi.org/10.1175/MWR-D-16-0211.1>.
- Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430, [https://doi.org/10.1175/1520-0477\(2002\)083<0407:DIHRPM>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2).
- Matte, D., R. Laprise, J. Thériault, and P. Lucas-Picher, 2017: Spatial spin-up of fine scales in a regional climate model simulation driven by low-resolution boundary conditions. *Climate Dyn.*, **49**, 563–574, <https://doi.org/10.1007/s00382-016-3358-2>.
- Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated k model for the long-wave. *J. Geophys. Res.*, **102**, 16 663–16 682, <https://doi.org/10.1029/97JD00237>.
- Prein, A., 2013: Added value of convection permitting climate simulations. Scientific Rep. 53–2013, Wegener Center Verlag, University of Graz, Graz, Austria, 168 pp., <http://wegcwww.uni-graz.at/publ/wegcreports/2013/WCV-SciRep-No53-APrein-Jul2013.pdf>.
- , and A. Gobiet, 2011: NHCM-1: Non-hydrostatic climate modelling. Part I: Defining and detecting added value in cloud-resolving climate simulations. Wegener Center Tech. Rep. 39–2011, Wegener Center Verlag, University of Graz, Graz, Austria, 74 pp., <http://wegcwww.uni-graz.at/publ/wegcreports/2011/WCV-SciRep-No39-APreinAGobiet-Feb2011.pdf>.
- , —, M. Suklitsch, H. Truhetz, N. Awan, K. Keuler, and G. Georgievski, 2013a: Added value of convection permitting seasonal simulations. *Climate Dyn.*, **41**, 2655–2677, <https://doi.org/10.1007/s00382-013-1744-6>.
- , G. Holland, R. Rasmussen, J. Done, K. Ikeda, M. Clark, and C. Liu, 2013b: Importance of regional climate model grid spacing for the simulation of heavy precipitation in the Colorado headwaters. *J. Climate*, **26**, 4848–4857, <https://doi.org/10.1175/JCLI-D-12-00727.1>.
- , and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 323–361, <https://doi.org/10.1002/2014RG000475>.
- Rauthe, M., H. Steiner, U. Riediger, A. Mazurkiewicz, and A. Gratzki, 2013: A Central European precipitation climatology Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteor. Z.*, **22**, 235–256, <https://doi.org/10.1127/0941-2948/2013/0436>.
- Roberts, N., 2008: Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteor. Appl.*, **15**, 163–169, <https://doi.org/10.1002/met.57>.
- , and H. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Sanderson, B., C. Piani, W. Ingram, D. Stone, and M. Allen, 2008: Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations. *Climate Dyn.*, **30**, 175–190, <https://doi.org/10.1007/s00382-007-0280-7>.
- Selmann, J., 1997: Radarforschung im DWD: Vom Scan zum Produkt. *Promet*, **26**, 32–42.

- Shin, H. H., and S.-Y. Hong, 2013: Analysis of resolved and parameterized vertical transports in convective boundary layers at gray-zone resolutions. *J. Atmos. Sci.*, **70**, 3248–3261, <https://doi.org/10.1175/JAS-D-12-0290.1>.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. Tech. Note NCAR/TN-475+STR, 113 pp., <http://dx.doi.org/10.5065/D68S4MVH>.
- Smiatek, G., H. Kunstmann, and A. Senatore, 2016: EURO-CORDEX regional climate model analysis for the Greater Alpine Region: Performance and expected future change. *J. Geophys. Res. Atmos.*, **121**, 7710–7728, <https://doi.org/10.1002/2015JD024727>.
- Taylor, K., R. Stouffer, and G. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Tewari, M., and Coauthors, 2004: Implementation and verification of the unified Noah land surface model in the WRF model. *20th Conf. on Weather Analysis and Forecasting/16th Conf. on Numerical Weather Prediction*, Seattle, WA, Amer. Meteor. Soc., 14.2a, https://ams.confex.com/ams/84Annual/techprogram/paper_69061.htm.
- Vignal, B., H. Andrieu, and H. Creutin, 1999: Identification of vertical profiles of reflectivity from volume scan radar data. *J. Appl. Meteor.*, **38**, 1214–1228, [https://doi.org/10.1175/1520-0450\(1999\)038<1214:IOVPOR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1999)038<1214:IOVPOR>2.0.CO;2).
- von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673, [https://doi.org/10.1175/1520-0493\(2000\)128<3664:ASNTFD>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<3664:ASNTFD>2.0.CO;2).
- Wagner, A., J. Seltmann, and H. Kunstmann, 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, <https://doi.org/10.5194/hess-16-4101-2012>.
- Weisman, M., W. Skamarock, and J. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125**, 527–548, [https://doi.org/10.1175/1520-0493\(1997\)125<0527:TRDOEM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2).
- , C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW model. *Wea. Forecasting*, **23**, 407–437, <https://doi.org/10.1175/2007WAF2007005.1>.
- Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, **136**, 4470–4487, <https://doi.org/10.1175/2008MWR2415.1>.
- Weusthoff, T., F. Ament, M. Arpagaus, and M. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. *Mon. Wea. Rev.*, **138**, 3418–3433, <https://doi.org/10.1175/2010MWR3380.1>.
- Wyngaard, J. C., 2004: Toward numerical modeling in the terra incognita. *J. Atmos. Sci.*, **61**, 1816–1826, [https://doi.org/10.1175/1520-0469\(2004\)061<1816:TNMITT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2).
- Zhou, B., J. Simon, and F. Chow, 2014: The convective boundary layer in the terra incognita. *J. Atmos. Sci.*, **71**, 2545–2563, <https://doi.org/10.1175/JAS-D-13-0356.1>.