Spatial-Scale Characteristics of Precipitation Simulated by Regional Climate Models and the Implications for Hydrological Modeling

S. H. RASMUSSEN, J. H. CHRISTENSEN, AND M. DREWS*
Danish Meteorological Institute, Copenhagen, Denmark

D. J. GOCHIS
National Center for Atmospheric Research, Boulder, Colorado

J. C. REFSGAARD
Geological Survey of Denmark and Greenland (GEUS), Copenhagen, Denmark

(Manuscript received 5 January 2012, in final form 24 April 2012)

ABSTRACT

Precipitation simulated by regional climate models (RCMs) is generally biased with respect to observations, especially at the local scale of a few tens of kilometers. This study investigates how well two different RCMs are able to reproduce the spatial correlation patterns of observed summer precipitation for the central United States. On local scales, gridded precipitation observations and simulated precipitation are compared for the period of the 1987 First International Satellite Land Surface Climatological Project Field Experiment (FIFE) campaign. The results show that spatial correlation length scales on the order of 130 km are found in both observed data and RCM simulations. When simulations and observations are aggregated to different grid sizes, the pattern correlation significantly decreases when the aggregation length is less than roughly 100 km. Furthermore, the intermodel standard deviation between simulations with different domains or resolutions increases for aggregation lengths below ~130 km. Below this length scale there is a high level of randomness in the precise location of precipitation events. Conversely, spatial correlation values increase above this length scale, reflecting larger predictive certainty of the RCMs at larger scales. The findings on aggregated grid scales are shown to be largely independent of the underlying RCMs grid resolutions but not of the overall size of RCM domain. With regard to hydrological modeling applications, these findings indicate that precipitation extracted from the present RCM simulations at a catchment scale below the intermodel standard deviation length cannot be expected to accurately match observations.

1. Introduction

Global warming is expected to lead to considerable hydrological change (Huntington 2006). A large number of hydrological modeling studies have been carried out in order to quantify the magnitude of the future hydrological change (e.g., van Roosmalen et al. 2007; Andreasson et al. 2004; Lettenmaier et al. 1999). Typically, the forcing data for hydrological model simulations under future climate conditions are derived from simulations with general circulation models (GCMs). However, the resolution of nearly all GCMs is presently too coarse to adequately represent regional and local climate conditions. Downscaling of GCM results for especially precipitation is therefore required for many purposes. Statistical downscaling relies on a combination of climate model output and present-day climatologies derived from observations. In contrast, dynamical downscaling by regional climate models (RCMs) relies solely on model output and attempts to present a more physically consistent, although possibly biased, simulation of the future climate (Fowler et al. 2007; Maraun et al. 2010). In fact, most RCM results are known to be biased with respect to observed precipitation (Christensen et al. 2008; Jacob

* Current affiliation: DTU Climate Centre, Roskilde, Denmark.

Corresponding author address: S. H. Rasmussen, Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, Denmark.
E-mail: sohr@DMI.dk

DOI: 10.1175/JHM-D-12-07.1

© 2012 American Meteorological Society
et al. 2007). Hence, some form of bias correction is often applied before the results are used in hydrological models (van Roosmalen et al. 2011; Graham et al. 2007). While some recent results have shown marked progress in the reduction of RCM precipitation biases in some regions (e.g., Ikeda et al. 2010; Rasmussen et al. 2011), many precipitation regimes, particularly warm-season-dominated regimes, present significant challenges.

In the international model intercomparison project Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects (PRUDENCE) (Christensen et al. 2007), the ensemble mean for the summer precipitation in Europe (1961–90) is found to have a bias of approximately ±50% within selected regions (Jacob et al. 2007). Similar results are found within the most recent European model intercomparison project Ensemble-Based Predictions of Climate Changes and their Impacts (ENSEMBLES) (van der Linden and Mitchell 2009), where Christensen et al. (2008) find that the ENSEMBLES models generally seem to be too dry in the wettest month, although in some regions models seem to be too wet in all months. In the U.S. intercomparison project North American Regional Climate Change Assessment Program (NARCCAP) (Mearns et al. 2009, 2012), model simulation of precipitation are also biased. For example, Wang et al. (2009) found the annual fluctuations to be too strong in a study over the intermountain region in the west. Sobolowski and Pavelsky (2012) analyzed the ensemble mean and found bias in summer precipitation (1971–2000) in the range ~5%–30%, which is significant in 63% of the study area in the southeast United States.

Jacob et al. (2007) report a small general increase in the magnitude of the simulated precipitation with increasing model resolution based on two of the PRUDENCE RCMs that were run with different resolutions. In a series of nested high-resolution RCM experiments, carried out in the southwest part of South America, Rojas (2006) finds large improvements in going from a 135- to 45-km horizontal resolution and smaller improvements when going from 45 to 5 km. Similarly, in simulating western U.S. winter precipitation, Ikeda et al. (2010) and Rasmussen et al. (2011) showed marked reductions in both station and regional averaged biases as model resolution increased from 36 to 18 to 6 km with minor further improvements at 2 km. These studies suggest that the magnitude of the precipitation bias is related to model resolution. Miguez-Macho et al. (2004) downscaled NCEP–NCAR reanalysis over North America at 50-km resolution in several simulations with shifted location of domain and with and without use of spectral nudging (interior relaxation of long waves) (Waldron et al. 1996). Among the simulations with use of spectral nudging, the variation in location of precipitation was eliminated. They show how their proposed spectral nudging method constrains the RCM to follow the synoptic scale of the driving GCM and allows the small-scale dynamic to develop. In RCMs that only rely on lateral boundary forcing (i.e., that have no interior nudging), the synoptic-scale circulation may drift from the driving field and thereby create location uncertainty in the RCM simulation (Miguez-Macho et al. 2004). When analyzing model performance, information is often aggregated to subregions or shown as maps in order to get a broader view of the quality of the simulations. However, it remains an open question to define at which scale models will be able to give meaningful information about precipitation in the sense that it contains similar spatial-temporal patterns as compared to observations to a sufficient degree for use in the desired hydrological assessment. Here we will explore this in some detail for the case of warm-season precipitation over the central United States.

Hydrological models generally rely on observations of precipitation and streamflow for calibration and model validation. Using precipitation data that differ significantly from the observed calibration data in hydrological models can cause misleading or incorrect results (Fu et al. 2011; Obled et al. 1994). The effects of uncertainties in the precipitation input to hydrological models in terms of spatial variability or representation issues have previously been investigated (van de Beek et al. 2011; Fu et al. 2011; Moulin et al. 2009; Vischel and Lebel 2007; Bell and Moore 2000; Koren et al. 1999; Obled et al. 1994). For large catchments spanning several 10 000 km², the precipitation pattern is potentially important for the temporal variability of the runoff. For small catchments where the response time from rainfall to runoff is smaller, the precipitation pattern has been shown to play a significant role (He et al. 2011; Obled et al. 1994). Fu et al. (2011) found the importance of the precipitation pattern to be small for catchments above 250 km² and negligible at 1000 km².

The typical horizontal grid resolution of a state-of-the-art RCM is 25–50 km with high-resolution, long-term, continuous experiments routinely being run at ~10 km (Mearns et al. 2009; Christensen and Christensen 2007; van der Linden and Mitchell 2009). Consequently, a medium-sized hydrological catchment area is generally only covered by a few RCM grid cells. To represent subcatchment variability, even higher model resolution is needed to resolve precipitation patterns. RCM results are typically validated on a local scale by direct comparison to station data (e.g., Rojas 2006) or via diagnosis of general bias patterns over the model domain (e.g.,
Kjellström et al. 2010; van Roosmalen et al. 2010; Jacob et al. 2007). Owing to data limitations or lack of RCM resolution, precipitation patterns simulated by RCMs generally have not been validated rigorously at local scales, such as for small catchments of less than few 100 km².

Within the last decade various research teams have been working on coupling hydrological and atmospheric models at different levels of complexity (Maxwell et al. 2011; Jiang et al. 2009; Anyah et al. 2008; Yuan et al. 2008; Maxwell et al. 2007; Miguez-Macho et al. 2007; Seuffert et al. 2002; York et al. 2002; Walko et al. 2000). When coupling the model codes dynamically, bias correction of precipitation during runtime is particularly problematic because it creates violations of the fundamental equations of conservation of water and energy. Presently, it is not clear whether RCM biases will be reduced or amplified, or whether the ability of atmospheric models to simulate correct spatial precipitation patterns will be improved in a coupled modeling code. From a hydrological modeling perspective it is particularly important to know at what scale the precipitation can be adequately simulated by RCMs for different precipitation regimes—something that is not clear from a theoretical or an empirical point of view. In this study we 1) analyze the potential capability of two different RCMs to represent precipitation at different spatial scales, 2) investigate how domain size and grid size influence the simulation of precipitation by the RCMs, and 3) discuss the implications of the RCMs’ capability of realistically capturing precipitation events when used as input for hydrological modeling.

2. Methodology

a. Datasets

Several observational datasets of precipitation were included in the analysis.

1) FIFE

During the First International Satellite Land Surface Climatological Project Field Experiment (FIFE) (Sellers et al. 1992), an area of 15 × 15 km² was monitored in 1987 during the period 26 May–16 October. FIFE is located in the Konza Prairie Reserve near Manhattan, Kansas. Precipitation was recorded half-hourly at 10 meteorological stations. Data from the meteorological stations have been quality checked and an area mean was calculated by Betts and Ball (1998). The precipitation data are not corrected for wind or turbulence effects on gauge measurements, which typically results in some degree of underestimation or “undercatch” of the rain gauge. Nevertheless, given the high density of stations deployed during the experiment, the FIFE data were assumed to provide a good estimate of precipitation occurrence, intensity, and total amount in a local area corresponding to an RCM cell.

2) UWAS

University of Washington (UWAS) is a gridded dataset of daily total precipitation covering the United States from the National Oceanic and Atmospheric Administration (NOAA) Cooperative Observer (Co-op) stations (Maurer et al. 2002). The average density of stations is approximately one station for each 700 km² (~26 km). The resolution of the gridded data is 0.125° (~13 km). Interpolation is done with a search radius between 50 and 100 km depending on the number of stations; some spatial correlation can thereby be expected within the range. The data are scaled to the long-term mean of the Parameter-Elevation Regressions on Independent Slopes Model (PRISM) precipitation climatology (Daly et al. 1994, 1997), the impact of which is largest in regions of complex terrain in the western United States. The UWAS dataset is used for analysis and model validation because of its high resolution.

3) UDEL

University of Delaware (UDEL) collects global monthly total precipitation from rain gauges from 1950 to 2008, and interpolates them to a grid of 0.5° × 0.5° resolution (~55 km) (Willmott and Matsuura 1995). UDEL builds on data from the Global Historical Climate Network (Peterson et al. 1998) and the archives of Legates and Willmott (1990).

4) GPCC

Global Precipitation Climatology Centre (GPCC) is a gridded dataset of monthly totals from rain gauges at 0.5° × 0.5° resolution. The data coverage varies from less than 10 000 to more than 40 000 stations (Rudolf and Schneider 2005). Information about the number of stations used within each grid cell is also available. The average density of stations within our analysis area is approximately 2.5 stations per grid cell (one each ~25 km).

5) CRU

Climatic Research Unit, University of East Anglia (CRU TS 2.0) is a time series of global monthly gridded precipitation data at 0.5° × 0.5° resolution (Mitchell et al. 2004). The number of stations varies temporally and spatially. Some of the highest station densities are found in the United States. The number of stations was 14 579 in 1981 and 8237 in 1995 (New et al. 2000).
b. Regional climate models

In this study two different climate models have been applied: namely, the Weather Research and Forecasting Model (WRF) and HIRHAM.

1) HIRHAM, VERSION 5

The HIRHAM regional climate model was originally developed by Christensen and van Meijgaard (1992) and has been updated by Christensen et al. (1996, 2006). It consists of the dynamical core of the synoptic-scale weather forecast model High-Resolution Limited-Area Model, version 7 (HIRLAM7) (Eerola 2006; Undén et al. 2002) and the physics from the global circulation model ECHAM5 (Roeckner et al. 2003). HIRHAM is a hydrostatic model with a semi-implicit time-differencing and a semi-Lagrangian advection scheme. The prognostic variables are the horizontal wind components, temperature, specific humidity, cloud water, turbulent kinetic energy, and surface pressure. The horizontal discretization uses an Arakawa C grid; vertically, hybrid terrain-following coordinates are used near the surface, while pressure coordinates are used at the top of the atmosphere. The inherent physics is implemented vertically in columns (Christensen et al. 2006, 1996).

The subgrid cloud cover distribution is derived by a statistical–dynamical method, where the cloud cover only depends on fluctuations in total water content (Tompkins 2002). The cloud microphysical scheme has three classes, where the liquid or ice phase depends on the temperature (Lohmann and Roeckner 1996). The convective parameterization is a mass-flux scheme (Nordeng 1994; Tiedtke 1989). The condensation scheme with subgrid-scale cloud formation is by Sundqvist (1978). The radiation scheme is taken from ECMWF model cycle 36 (Morcrette 1989). The planetary boundary layer scheme is based on the eddy diffusivity concept (K-model), building on Monin and Obukhov (1954). The parameterization of gravity wave drag follows Miller et al. (1989) with the modifications proposed by Palmer et al. (1986). The land surface model has five layers for calculation of soil temperature, and the water budget is formulated for four reservoirs: snow intercepted by canopy, snow at surface, rainwater intercepted by canopy, and soil water (Roeckner et al. 2003).

2) WRF, VERSION 3.2.1

WRF is a community model currently maintained by the National Center for Atmospheric Research (NCAR), Mesoscale and Microscale Meteorology Division. It is a unified flexible system ranging from global circulation to large eddy simulation scales. The WRF Advanced Research and Weather (ARW) dynamic core is fully compressible, Eulerian, and nonhydrostatic. It uses explicit time integration by the third-order Runge–Kutta scheme. Prognostic variables are the horizontal and vertical velocity wind components, perturbation potential temperature, perturbation geopotential, and perturbation surface pressure of dry air. Horizontal discretization is by an Arakawa C grid and in the vertical by hybrid terrain-following coordinates (Skamarock et al. 2008). WRF offers a number of different parameterization options. For this study we use the Thompson six-class microphysical scheme, which includes representation of graupel (Thompson et al. 2008). The cumulus scheme is the Grell 3D multiclosure multiparameter ensembles method, which is an improved version of Grell and Devenyi (2002) suitable for high-resolution modeling. The short- and longwave radiation scheme is the Rapid Radiative Transfer Model (Mlawer et al. 1997) including a random cloud overlap (Pincus et al. 2003). The surface layer scheme is based on Monin and Obukhov (1954) with viscous sublayer following Carlson and Boland (1978). The planetary boundary layer scheme is by Yonsei University with nonlocal K (Hong et al. 2006). The land surface scheme is the Noah model with four soil layers for heat and moisture (Chen and Dudhia 2001; Ek et al. 2003).

c. Experimental setup

The two RCMs were set up and forced as identically as possible. Both models were forced with ERA-40 at their lateral boundaries. Simulations were made from 1 January to 31 August 1987. The analysis period is the summer months of June, July, and August (JJA), corresponding largely to the duration of the FIFE experiment. The analysis area is a subset of the “small” model domain and is illustrated in Fig. 1. To avoid placing the western boundary in the mountainous region of Colorado, the “small” domain is not completely centered over FIFE.

As noted above there are some structural differences between the two models, which complicate the identical setup (e.g., the two models do not use the same
parameters to define the domains). Specifications of the different setups are given in Table 1. The locations of the model domains and analysis area are shown in Fig. 1. All simulations are made on a rotated longitude–latitude grid with origin located at the FIFE area (location approximately longitude –96.5°, latitude 39°). The number of grid cells in the boundary relaxing zone is 10 for HIRHAM and 5 for WRF.

The two models have different complexity of their land surface models and therefore different periods of spinup. HIRHAM uses a bucket model, whereas the Noah land surface model used in WRF is more complex with a Richard’s equation–based soil moisture accounting scheme that has several vertical layers. Effectively, the Noah model requires more time to spin up than the HIRHAM bucket model. An offline spinup period of 1 year was applied to Noah before it was run as a part of WRF. In the offline spinup period Noah was forced by the High-Resolution Land Data Assimilation System (HRLDAS) (Chen et al. 2007) using ERA-40 data for 1986. HIRHAM and WRF were then run for 5 months (January–May) before the period of analysis. In total Noah was spun up for 1 year and 5 months, while HIRHAM was effectively spun up for 5 months, which was assumed sufficient for our area of interest (Christensen 1999). Soil moisture maps for all simulations at 1 June are seen at Fig. 2 together with ERA-40 and the topography. All simulations show comparable patterns of soil moisture at the beginning of the summer.

d. Spatial scales

In the following we describe the precipitation patterns by three different spatial scales. The first one is of daily precipitation and the next two of total summer precipitation at aggregation lengths. Aggregation lengths are the size of the grids where the data are regirded by aggregation under the analysis, for both observations and simulations. All fitting of parameters were done by the Interactive Data Language (IDL) algorithm “curve-fit,” which is a gradient-expansion algorithm for estimating a nonlinear least squares fit (Research Systems, Inc. 2007).

1) Spatial correlation length

For each cell in the analysis area the cross correlation coefficients of the daily time series to all other cells are calculated. For each pair of cells only the days with precipitation amounts over a fixed threshold of, for example, 1 mm day⁻¹ are included in the calculation. An exponential function is subsequently fitted to a scatterplot of the correlation coefficients between cells against distance between cells:

\[
S(d) = e^{-\frac{d}{s}}
\]

where S is the correlation, d distance, and s is the spatial correlation length. This spatial correlation length is a measure of the distance within which precipitation is expected to occur at the same time and at similar strength.

Table 1. Model setups.

<table>
<thead>
<tr>
<th>Setup</th>
<th>Model</th>
<th>Domain</th>
<th>Resolution (°)</th>
<th>Size (No. cells in x and y direction)</th>
<th>Time step (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL25</td>
<td>HIRHAM</td>
<td>Large</td>
<td>0.25</td>
<td>182 × 101</td>
<td>5</td>
</tr>
<tr>
<td>HS25</td>
<td>HIRHAM</td>
<td>Small</td>
<td>0.25</td>
<td>62 × 62</td>
<td>5</td>
</tr>
<tr>
<td>HS12</td>
<td>HIRHAM</td>
<td>Small</td>
<td>0.125</td>
<td>122 × 122</td>
<td>2.5</td>
</tr>
<tr>
<td>HS05</td>
<td>HIRHAM</td>
<td>Small</td>
<td>0.05</td>
<td>302 × 302</td>
<td>2</td>
</tr>
<tr>
<td>WL25</td>
<td>WRF</td>
<td>Large</td>
<td>0.247</td>
<td>182 × 101</td>
<td>3</td>
</tr>
<tr>
<td>WS25</td>
<td>WRF</td>
<td>Small</td>
<td>0.25</td>
<td>62 × 62</td>
<td>3</td>
</tr>
<tr>
<td>WS05</td>
<td>WRF</td>
<td>Small</td>
<td>0.05</td>
<td>302 × 302</td>
<td>1</td>
</tr>
</tbody>
</table>
2) PATTERN CORRELATION LENGTH

To calculate the pattern correlation between gridded maps of simulated and observed total summer precipitation, the grid cells of each of the two maps are collected in one row, column by column, starting in the southwest corner. The pattern correlation was inferred from these two rows. The pattern correlation is calculated at different resolutions (i.e., by aggregating grid cells) based on the criteria that the total coverage of the analysis area must be at least 7.5 and that only full grid cells are counted (the analysis area itself covers an area of 8×8). Aggregation grid sizes are 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, and 4. An exponential function is subsequently fitted to an aggregation size correlation plot, assuming that the pattern correlation increases with the resolution of the aggregation:

\[ P(a) = x - xe^{-ap} , \]

where \( P \) is the pattern correlation, \( a \) the aggregation size, \( p \) the pattern correlation length, and \( x \) an estimate of the maximum correlation. The pattern correlation length is a measure of the scale at which patterns are similar, while below the pattern is largely random.

Significant levels are found by bootstrapping. The map of observed precipitation is divided in blocks greater than the spatial correlation length (2°). From these blocks 100 000 random maps are made by shuffling the blocks and pattern correlation are calculated to each of the simulation. The 5th and 95th percentiles are found as the lower and upper limit for pattern correlation to a random precipitation map. Significant levels are found for grid aggregation size of 0.25, 0.5, 1, and 2°.

3) INTERMODEL STANDARD DEVIATION LENGTH

For each of the RCMs the standard deviation (square root of mean squared deviation) between simulations of total summer precipitation carried out using different model grids/model domains is calculated at different grid resolutions by aggregating grid cells, while using the same coverage criterion as described above. An exponential function is then fitted to a scatterplot of the
resulting aggregation size versus the median of the standard deviation in each cell:

\[ I(a) = ye^{-a/i} + z \]

Here \( I \) is the median of standard deviation, \( a \) the aggregation size, \( i \) the intermodel standard deviation length, \( y \) an estimate of the median of the intermodel standard deviation at an aggregation size of zero, and \( z \) an estimate of the lowest median of the intermodel standard deviation. The intermodel standard deviation length is a measure of the scale at which the precipitation patterns simulated by the different RCMs agree internally—the scale of internal spatial variability.

3. Local bias in simulated precipitation

Typically, RCM data used for hydrological modeling are extracted based on the geographical location of the catchment. In this way input from the seven RCM runs is extracted for the FIFE area based on the overlapping RCM grid cells (Fig. 3). Compared to observations spanning the whole FIFE period, the simulations exhibit a bias \([ (P_{\text{model}} - P_{\text{obs}})/P_{\text{obs}} ] \) between \(-21\%\) and \(144\%\) in total precipitation amounts. June thirtieth is one example of a day with large differences between the simulations. To illustrate some of the causes of the spread in the pattern of precipitation amounts, the period of 29 June–1 July is depicted in Fig. 4 based on the UWAS observations and several different model realizations. FIFE is located within the star on the maps. In the observations we track a frontal precipitation system moving across the map in the southern half of the domain. On 29 June FIFE is located in a zone without precipitation and on 30 June the system has moved south of FIFE. In the simulations on 29 June the precipitation system is a band across the map from the Midwest to the northeast corner of the map. Only simulations HS25 and HS05 show the precipitation window in the system near FIFE. On 30 June all simulations show a precipitation system moving across the southern half of the map with high intensity. In simulations HS12 and WL25 there is no precipitation over FIFE or in the observations. There is some precipitation over FIFE in HL25 and HS05. In HS25, WS25, and WS05 there is heavy precipitation at FIFE. In all simulations it is only a matter of a few grid cells whether the heavy precipitation is located over FIFE or not. On 1 July the system has moved farther west. There is no significant topography in the area to influence the release of precipitation at a specific location. All simulations agree on the band of precipitation on 30 June, but disagree on the strength and width.

4. Characteristics of observed precipitation

The upper four panels of Fig. 5 show the total summer precipitation (JJA) in 1987 as found from four different gridded observation datasets. While the overall pattern in general is the same, the four datasets do not agree on a gridcell by gridcell basis. The large differences on a gridcell scale are noteworthy, since parts of the underlying observations are overlapping. The total precipitation at the FIFE area in JJA 1987 was 288 mm (Betts 1994). The root-mean-square error (RMSE) between the datasets on that location is \(7\%\)–\(102\%\) of the observed precipitation at FIFE. These observed differences could be attributed to differences among the rain gauges used in the analyses or to quality checking and interpolation methods. Out of the four gridded datasets only UWAS has a resolution that is comparable to the resolution in our RCM simulations. Bearing in mind the listed spread in the observations, we compare our model output primarily to the UWAS data and to the FIFE gauge data.

Examining the observations closely, the daily mean precipitation data have a peak at the upper-middle part
of the analysis area, north–northeast of the FIFE site (cf. Fig. 5). A peak of mean intensity and the maximum intensity are also found in the same area. The pattern correlations between the daily mean to the daily mean intensity and to the daily maximum intensity are 0.80 and 0.72, respectively. The correlation between the daily mean precipitation and the number of wet days is 0.40. Combined, this indicates that the peak in the precipitation amounts is often caused by intense precipitation events rather than by the number of precipitation days. The mean spatial correlation length for the analysis area in the UWAS data is 130 km (Fig. 6, upper-left map). At the FIFE location (indicated by the star symbol) the spatial correlation length is $\sim 215$ km. These spatial correlation lengths are around the typical scale of mesoscale convective systems. This is consistent with expectations, as mesoscale convective systems are found to account for 30%–70% of the precipitation in the warm season and even more during JJA in the region from the Rocky Mountains to the Mississippi River (Fritsch et al. 1986).

As shown in Fig. 7, the probability of precipitation within the analysis area in the summer of 1987 is quite similar to the climatological statistics for 1961–90, based on UWAS data. When considering the FIFE dataset and the UWAS grid cell over the FIFE area, it should be noted that these datasets only include 92 days of observations, while the full analysis area counts 92 days of observations for each of the 5676 grid cells. As a result, probabilities below 0.01 cannot be statistically resolved at the FIFE area. From Fig. 7, a number of noteworthy differences are seen between the local-scale FIFE data and UWAS data for the FIFE grid. For example, the FIFE data seem to have a fairly low probability of precipitation in the range 3–6 mm day$^{-1}$ compared to UWAS. The FIFE data have a higher probability of precipitation at 17 mm day$^{-1}$, because two events were recorded compared to only one in UWAS data. In general, the local-scale observations would be expected to exhibit much larger variations than the gridded data, where the local signal is smoothed during interpolation to larger scales. As such, Fig. 7 illustrates that observations at the FIFE area during 1987 show some statistical differences compared with long-term values from its surroundings.

Figure 8 shows different box-and-whiskers plots of the daily precipitation for the nearest grid points around FIFE. The top panel compares the precipitation at FIFE with the surrounding 100 grid cells of UWAS. The area shown is centered at FIFE and spans approximately $1.25^\circ \times 1.25^\circ \ [\sim (130 \times 130 \ km^2)]$, corresponding to the mean spatial correlation length derived from the gridded observations. As depicted in the figure, the spatial
variability of the observations even within the mean spatial correlation length is large—up to \(50 \text{ mm day}^{-1}\) on single days. However, with mesoscale convective system dominating, the observed high spatial variability is to be expected both in terms of the scale and the spatial pattern. The high spatial variability of precipitation could thus easily be explained by the large differences between the gridded datasets discussed above. A different choice of which rain gauges to include in the interpolations may also lead to large discrepancies, especially when there is large variability between the gauges such as when gauges are widely spaced or precipitation is very localized.

5. Model analysis

Maps of the spatial correlation lengths derived from the different simulations using WRF and HIRHAM are seen in Fig. 6, along with UWAS. In general, the simulations using the large model domain (HL25 and WL25) yield a longer mean spatial correlation length than found from the observations. The HIRHAM simulations using the small model domain (HS25, HS12, and HS05) on the other hand show a similar or slightly smaller mean spatial correlation length than UWAS, while the corresponding WRF simulations (WS25 and WS05) show a longer mean spatial correlation length, though still smaller than those using the larger model domain. Likewise, the pattern of the spatial correlation length also shows notable differences between the simulated precipitation and the gridded observations. For simulations using the larger model domain this is expected, as the RCMs then have a high degree of freedom in terms of localizing the precipitation systems. In the case of the smaller model domain, the pattern is more strongly constrained by the lateral boundary conditions (i.e., the reanalysis). On the other hand, stronger constraints on the boundary could limit the small-scale circulation, but this is not seen among the simulations. All maps of spatial correlation length by the simulations show no pattern correlation to the map of spatial correlation length by UWAS. The patterns are very different compared to the similarity of the soil moisture maps (Fig. 2). In general, the resolution only has a minor effect on the length of the spatial correlation length.

We have evaluated how the choice of wet day threshold affects the pattern of the spatial correlation length (not shown). Increasing the wet day threshold implies a reduction of the number of available wet days for the calculation. We found that the correlation lengths did not change up to a threshold of \(8 \text{ mm day}^{-1}\) in WRF and \(3 \text{ mm day}^{-1}\) in HIRHAM, while a further increase in threshold results in decreased spatial correlation lengths.

The pattern correlation between UWAS and the simulated total summer precipitation at different grid scales are shown in Fig. 9. For each of the simulations the pattern correlation is stronger with increased length of aggregation. All simulations, except for WS25, show pattern correlations close to either the upper and lower 5% significant levels. The pattern correlation lengths are between 0.35° and 1.26°. Simulation WS25 has a pattern correlation length of 4.49°, but the pattern of this simulation is not significantly different from a random pattern. The total summer precipitation does not show patterns like soil moisture (Fig. 2). Except for HS12, no precipitation patterns show significant correlations to soil moisture. In general, simulations with
small domains and high resolution have a higher absolute value for the maximum of pattern correlation to precipitation by UWAS. At coarser grid scales fewer data points are available. At a grid scale of $4^\circ$ there is only four model–observation pairs on the map. For experiments HS12, HS05, and WL25, the pattern correlation at $4^\circ$ looks like an outlier when compared to the pattern correlation at $2.5^\circ$, and therefore the data point at $4^\circ$ is not used to fit the exponential curve. WS05 and WS25 show a negative pattern correlation. Assuming the pattern has a wave structure where the phase of the wave is shifted somewhat, or if there is only a slight spatial offset between corresponding observed and modeled precipitation features, such shifts could explain the negative correlation. In such cases, more robust spatial verification metrics that can account for such changes should be employed.

Figure 10 shows the intermodel standard deviation for each grid cell and at different grid scales between the simulations of HIRHAM and WRF. The range and median of the standard deviation is seen to decrease with coarser grid scale. The regression function of intermodel standard deviation has been fitted with or without the data point at $4^\circ$ grid resolution. The intermodel standard deviation lengths are found to be 1.02$^\circ$ and 1.26$^\circ$ for the two models. Using all aggregation sizes.

Fig. 6. Maps of spatial correlation length for summer precipitation; the mean length ($M$) over the analysis area is shown at the top of each map, along with pattern correlation to UWAS ($P$). Observations are from UWAS (Maurer et al. 2002). See Table 1 for names of simulations.
the WRF intermodel standard deviation length is 4.9° and 1.26° for HIRHAM. The minimum intermodel standard deviation is similar for both RCMs. If the simulations are grouped depending on domain size or resolution (not shown), and not on models as shown in Fig. 10, a comparable standard deviation length is found.

Revisiting Fig. 8, the lower panels show different box-and-whiskers plots of the daily precipitation for the nearest grid points around FIFE. At 0.25° resolution the FIFE area is located in the central grid cell and the surrounding 5 × 5 cells are used. At higher resolutions all the grid cells within the same area are used. The correlation between the FIFE data and the time series of the median from UWAS is found to be moderate (0.49) (cf. Fig. 8). The RCM simulations also have a moderate correlation to FIFE. In general, the RCM simulations using the small model domain have slightly larger correlation to FIFE than the simulation using the large domain.

6. Discussion

A comparison of the different gridded observational datasets for the total summer precipitation shows a large spread (Fig. 5). For the summer precipitation in the analysis area, the spatial correlation length on wet days is found to be approximately 130 km on average (Fig. 6). Within a region corresponding to the mean spatial correlation length, the spread in precipitation amount on a single day can be up to 50 mm day$^{-1}$ between the grid cell of highest and lowest daily precipitation in the UWAS dataset (Fig. 8).

The time series for the median amount of UWAS in the area around FIFE shows a moderate correlation with the FIFE data. Furthermore, for two-thirds of the FIFE wet days, the amount of precipitation falls within the spatial variability range of the minimum and maximum of the gridded UWAS data; one-third falls within the 25%–75% quartiles. Therefore, when comparing the simulations to the FIFE dataset, one should not expect exceptional correspondence, because data from the 1987 FIFE experiment only partly represent (in a statistical sense) the precipitation within a larger area corresponding to the spatial correlation length scale. When comparing the simulations to FIFE in the same way, the simulations are found to match FIFE to nearly the same extent at UWAS to FIFE (Fig. 8).

Examining the single precipitation event over the FIFE area on 30 June all simulations reconstruct the same kind of precipitation system across the domain. The location, however, is different (Fig. 4). At that day the difference between nearly no precipitation and heavy precipitation is only the matter of a few grid cells. Generally, when compared to reanalysis or observations, climate models should be expected to reproduce...
Fig. 8. Box-and-whiskers plot of daily precipitation for nearest model grid points around FIFE; the area shown is 1.25° × 1.25°. (top) Observed data and (lower panels) model simulations. The simulations are bias corrected by scaling to the same mean over the period as the observations (i.e., UWAS). “Corr. FIFE” designates the correlation between FIFE and the median of the data. “Corr. UWAS” is correlation between median of simulation data and median of UWAS top panel. “Dry” is the number of days with less precipitation than 1 mm day⁻¹. “Min-Max” is the number and percent of wet days, where FIFE observation is between min and max of the data. “Q25-Q75” is the number and percent of wet days, where FIFE observations are between the 25% and 75% quartiles of the data. Note that ideally Q25-Q75 would be 50% and Min-Max should be 100%.
weather statistics to a reasonable degree. As the boundary forcing is provided by an observations-based reanalysis, the simulated weather on a particular day in the simulation would also be expected to look reasonably similar to observations. However, RCMs forced exclusively at the lateral boundaries can “freely” develop their own weather with only moderate constraint from the boundaries—a behavior that is influenced by the simulation domain size (Warner 2011; Miguez-Macho et al. 2004). Small inaccuracies in the forcing will tend to drift the simulation away from the observed weather (Lorenz 1963a,b). By use of spectral interior nudging, the synoptic-scale circulation can be constrained to follow the forcing data, while small-scale circulation at the surface can freely develop (Miguez-Macho et al. 2004; von Storch et al. 2000). In this way, location uncertainty due to shifts in storm tracks—the synoptic-scale circulation—is minimized and uncertainty is then mainly due to model physics parameterizations. With the use of spectral nudging it is implicitly assumed that the large-scale circulation of the driving GCM is correct and that the RCM does not add any value to the simulation of large-scale circulation (von Storch et al. 2000). In relation to parameterization error and the total summer precipitation, all simulations at the small domain are found to be too wet while the large domain shows comparable amounts compared to the observational data from UWAS (Fig. 6). The mean spatial correlation lengths did not prove to be a reasonable length. Below this length scale randomness in the simulated weather will therefore disturb the pattern with the result that the pattern correlation will decrease. The potential capability of RCMs to realistically reproduce precipitation at different spatial scales has been assessed by comparing the simulations at different grid scale with the observational data from UWAS. The three different measures of scales are of a comparable size and comparable to spatial correlation length found by the observational dataset UWAS 130 km (Fig. 6). The spatial correlation lengths within the simulations are 116–189 km (Fig. 6) and the pattern correlation length is 0.35°–1.26° [~(40–140 km)] (Fig. 9). The intermodel standard deviation is calculated for both of the models, with a spatial length scale of 1.02°–1.26° [~(40–140 km)] (Fig. 10). It should be emphasized that the absolute values of the length scales are rather uncertain. The pattern correlation and the intermodel standard deviation length are calculated from relatively few data points, while the spatial correlation length are calculated from many data points showing a considerable scatter. Similar scales were identified by van de Beek et al. (2011) for the Netherlands.

b. Influence of domain size and grid resolution

The domain size has been found to affect the spatial correlation length and the highest level of the pattern correlation (Fig. 9). For the small model domain, the pattern correlations with observations are stronger than for the large domain. Similar results were reported by Leduc and Laprise (2009), who performed so-called “big brother–little brother” experiments and found the pattern correlation to improve for small domains. This can be explained by the smaller distance to the lateral boundaries, which means that the regional climate model has less space and time to drift away from the forcing. On the other hand, small domain sizes are found to underestimate small-scale features where the larger domain has more space to develop disaggregated convection (Leduc et al. 2011). Anyway, mean spatial correlation length is found to be shorter for the small domains than the large domains, which could indicate that underestimation of small features due to too-small domain size is not an issue (Fig. 6). The negative correlation on the small domain indicates a shift in the pattern—perhaps a shift in storm tracks. The simulations on the small domain show a better timing of precipitation at FIFE (Fig. 8).

The mean precipitation amounts produced by simulations in the area around FIFE are clearly higher when using the small domain than when using a larger domain (Figs. 3 and 8). To understand this behavior we compared the vertically integrated atmospheric moisture in the western boundary of the small domain (i.e., the direct moisture input from the forcing) and the moisture crossing the same line in the simulations using the large domain (not shown). This quick analysis showed that more moisture was entering the small domain by the forcing than by the models with larger domains.

The mean spatial correlation lengths did not prove to depend strongly on grid resolution (Fig. 6). Conversely, pattern correlations seem to get stronger at higher resolution, but show no clear trend with increasing resolution (Fig. 9). A stronger pattern correlation could indicate better resolved important processes at higher resolution. Jacob et al. (2007) have previously reported that HIRHAM tends to produce more precipitation at higher resolution (over Europe); however, we were not able to find a clear dependence on the resolution in the present study. The intermodel standard deviations for the two models are similar at coarse resolution, while at fine resolution HIRHAM shows more spatial variability than WRF (Fig. 10). Compared to HIRHAM,
Fig. 9. (first column) Pattern correlation of total summer precipitation between UWAS and simulations at different aggregation sizes: $p$ and $x$ are the fitted parameters of exponential function, and dashed lines are 5th and 95th percentiles. Total summer precipitation aggregated at (second column) 0.25° grid, (third column) 1°, and (fourth column) 2°; the simulations are scaled to the same mean as UWAS for easier comparing of the patterns. (first row) Observations by UWAS (Maurer et al. 2002). (remaining rows) Simulation results; see Table 1 for names.
WRF exhibits less spatial variability, a longer mean spatial correlation length, and a smaller intermodel standard deviation, and it has a smoother precipitation signal. Altogether, this may indicate that WRF needs higher grid resolution in order to resolve the same spatial features as HIRHAM, or that other modeling decisions in WRF, such as in its numerics, may be influencing this response.

In this study only one type of forcing data is used. Other reanalysis products might have resulted in different simulations of precipitation patterns as the synoptic circulations between different reanalysis products are likely to differ. Furthermore, this study relies on only one parameterization suite for each RCM. Use of different parameterization schemes does have an effect on the precipitation pattern (Warner 2011; Miguez-Macho et al. 2005) as well as on regional circulation, moisture convergence patterns, and water budgets (e.g., Gochis et al. 2003). Attribution of parameterization uncertainty could have been investigated by use of spectral nudging for constraining the RCM to follow the synoptic-scale circulation and thereby eliminate the uncertainty due to shift in storm tracks (Miguez-Macho et al. 2004; von Storch et al. 2000). When using existing downscaled reanalysis products (e.g., ENSEMBLES), the precipitation location uncertainty herein would be a mix of shift in synoptic circulation and dynamical and parameterization uncertainty—similar to this study.

c. Implications for hydrological modeling

The spatial scale at which precipitation can be simulated correctly by RCMs has implications for hydrological modeling. For small catchments below the intermodel standard deviation length, the precipitation pattern as simulated by an RCM should not be expected to adequately resemble reality. For large catchments above the intermodel standard deviation length, the RCM may in principle be expected to simulate the amount and pattern of the precipitation. For flood predictions, timing in precipitation events is important, but in a climate context it is the probability distribution over a longer period (e.g., 30 years) that is more important (Mearns et al. 2012; Salzmann and Mearns 2012). The uncertainty of the precipitation simulation can be accounted for in the hydrological modeling with ensemble modeling using various realizations of precipitation based on different RCM perturbations. Or, given the location uncertainty a neighboring grid cell of the RCM simulation may provide at least as likely representation of the precipitation for small catchments (e.g., when extracting data from ENSEMBLES).

This inability to represent small scales will also have implications when using coupled climate–hydrological modeling (e.g., Maxwell et al. 2007). An interesting question in this respect is whether a fully coupled model will be able to reduce this length scale. This has not been addressed in previous studies of coupled modeling, and our analysis has not addressed this issue either, but it deserves considerable attention in future studies.

The findings of the present study suggest that the scale of predictive capability of a hydrological model forced by RCM output is greater than 100 km for this regime with warm-season convective precipitation. However,
RM simulations of different precipitation regimes (e.g., cold-season frontal systems or large organized tropical storms) may exhibit very different behaviors if the model more faithfully simulates the scaling characteristics of precipitation in those regimes.

7. Conclusions

The present study has analyzed the potential capability of RCMs to predict precipitation for the 15 km × 15 km FIFE area (central United States) during a 3-month, summertime analysis period. The results show that different versions of the two RCMs with different domain and grid sizes have similar statistical characteristics but exhibit quite different precipitation results.

Three different length scales have been calculated to characterize the results. First, the spatial correlation length within observed and simulated precipitation fields was estimated to be about 130 km for the observed dataset (UWAS) while it varied between 119 and 189 km for the different versions of the two RCMs. Second, the pattern correlation between simulated and observed total summer precipitation was found to increase as the area over which the data are aggregated increases, with pattern correlation length scales ranging between 40 and 140 km. Third, the intermodel standard deviation length decreased for increasing areas included in the analyses with length scales in the order of 100 km. These findings imply that the RCM simulations of precipitation in regions dominated by unforced convection show less random behavior and stronger correlation structures with observations when aggregated over larger areas, or, in other words, that the RCM predictions have larger predictive certainty at larger scale than at small scale. At scales below approximately 100 km, the simulations of unforced convective precipitation in this study cannot be expected to match the observations well. This does not imply that models need not run at resolutions higher than 100 km; instead, it suggests that the statistical patterns of rainfall from the models are similar to observations at this length scale and greater.

Our analysis of how the domain size and grid size influence the simulations of precipitation in RCMs shows that the pattern correlation is larger for the small model domain. The model resolution did not significantly affect the mean spatial correlation length. The two models exhibited different spatial variability at the finest scales. The HIRHAM RCM had a larger spread in intermodel standard deviation than WRF. We have found indications that both models seem to resolve the important processes better at higher resolution, but that WRF requires higher resolution than HIRHAM for resolving the same processes.

The inability to represent small-scale precipitation does have implications for hydrological modeling at small catchments and in coupled climate–hydrological modeling. However, in a climate context, only the long-term statistics are important; the effect in coupled models is not addressed here or in previous studies.

Our analysis has been done only for one summer and for one particular climate region. Hence, reservations about the general validity and the transferability of the results to other regions must be made, especially with regard to the specific numbers for the various length scales. To analyze the general validity of the results, the analysis needs to be extended to include all seasons because of the seasonality in convective precipitation.

To be seen in the full context of climate simulation, our findings need to be tested over a full climate period (e.g., several years). Furthermore, analysis of spatial correlation length in other regions with different topography, land use, or groundwater–land surface–atmosphere interaction is also warranted.

Acknowledgments. The present study was funded by a grant from the Danish Strategic Research Council for the project Hydrological Modelling for Assessing Climate Change Impacts at Different Scales (HYACINTS; www.hyacints.dk) under Contract DSF-EnMi 2104-07-0008. The authors would like to acknowledge Philippe Lucas-Picher (previously at the Danish Meteorological Institute) for helping to set up and run HIRHAM; Wei Yu [NCAR Research Application Laboratory (RAL)], Beau Paisley, Cindy Bruyere, and Jimy Dudhia (NCAR, Mesoscale and Microscale Meteorology) for help compiling, setting up, and running WRF; and Michael Barlage (NCAR, RAL) for processing ERA-40 data to use in HRLDAS and instruction to modification of HRLDAS to use ERA-40 data. The comments of three anonymous reviewers are acknowledged.

REFERENCES


   [Available online at http://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=88.]


