AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2014JD022748

Key Points:
- RCMs reproduce overall spatial dependence yet with considerable discrepancies
- These discrepancies only moderately affect area-aggregated heavy precipitation
- Pair-copula constructions can model dependence in mean and heavy precipitation

Supporting Information:
- Readme
- Tables S1–S2 and Figures S1–S30

Correspondence to:
I. Hobæk Haff,
ingrid@nr.no

Citation:
Hobæk Haff, I., A. Frigessi, and D. Maraun (2015), How well do regional climate models simulate the spatial dependence of precipitation? An application of pair-copula constructions, *J. Geophys. Res. Atmos.*, *120*, 2624–2646, doi:10.1002/2014JD022748.

Received 24 OCT 2014
Accepted 20 FEB 2015
Accepted article online 27 FEB 2015
Published online 3 APR 2015

How well do regional climate models simulate the spatial dependence of precipitation? An application of pair-copula constructions

Ingrid Hobæk Haff[1], Arnoldo Frigessi[1,2], and Douglas Maraun[3]

[1]Statistics for Innovation, Norwegian Computing Center, Oslo, Norway, [2]Oslo Centre for Biostatistics and Epidemiology, Department of Biostatistics, University of Oslo, Oslo, Norway, [3]Ocean Circulation and Climate Dynamics, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

Abstract We investigate how well a suite of regional climate models (RCMs) from the ENSEMBLES project represents the residual spatial dependence of daily precipitation. The study area we consider is a 200 km × 200 km region in south central Norway, with RCMs driven by ERA-40 boundary conditions at a horizontal resolution of approximately 25 km × 25 km. We model the residual spatial dependence with pair-copula constructions, which allows us to assess both the overall and tail dependence in precipitation, including uncertainty estimates. The selected RCMs reproduce the overall dependence rather well, though the discrepancies compared to observations are substantial. All models overestimate the overall dependence in the west-east direction. They also overestimate the upper tail dependence in the north-south direction during winter, and in the west-east direction during summer, whereas they tend to underestimate this dependence in the north-south direction in summer. Moreover, many of the climate models do not simulate the small-scale dependence patterns caused by the pronounced orography well. However, the misrepresented residual spatial dependence does not seem to affect estimates of high quantiles of extreme precipitation aggregated over a few grid boxes. The underestimation of the area-aggregated extreme precipitation is due mainly to the well-known underestimation of the univariate margins for individual grid boxes, suggesting that the correction of RCM biases in precipitation might be feasible.

1. Introduction

Potential regional impacts of future climate change are commonly assessed by impact models. These are often driven with regional climate change scenarios, downscaled either statistically or dynamically from global climate model simulations [*Fowler et al.*, 2007; *Maraun et al.*, 2010; *Rummukainen*, 2010]. For many advanced impact models, it is indispensable to provide input fields that not only correctly represent relevant meteorological variables at a given location but also their dependence structure in space. For instance, it is important for distributed hydrological models that the amount of precipitation aggregated across a catchment is correctly simulated. Temperature and precipitation are generally considered the most relevant input variables for hydrological models [*Bronstert et al.*, 2007]. Precipitation is considerably more difficult to simulate, mostly owing to the high spatial-temporal variability of precipitation occurrence and intensity [*Maraun et al.*, 2010]. Here we apply pair-copula constructions, a new statistical approach, to validate the spatial dependence of daily precipitation amounts, i.e., the precipitation totals (unit mm) falling within a day, simulated by regional climate models, with a focus on extreme events.

The overall performance of downscaling methods to simulate precipitation has been assessed by many intercomparison projects such as STARDEX (Statistical and Regional dynamical Downscaling of Extremes for European regions) [*Haylock et al.*, 2006; *Goodess et al.*, 2010], PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects) [*Christensen and Christensen*, 2007], ENSEMBLES [*van der Linden and Mitchell*, 2009], and NARCCAP (North American Regional Climate Change Assessment Program) [*Mearns et al.*, 2009]. Other initiatives such as CORDEX (Coordinated Downscaling Experiment - European Domain) [*Giorgi et al.*, 2009] and VALUE [*Maraun et al.*, 2015] have just started; EURO-CORDEX, the European branch of CORDEX, has recently published its first regional climate model (RCM) validation results [*Kotlarski et al.*, 2014], and VALUE (Validating and Integrating Downscaling Methods for Climate Change Research) is expected to deliver validation results over the coming years.

It is generally acknowledged that regional climate models (RCMs) add value to the simulation of precipitation, in particular of high intensity events and over complex terrain [*Frei et al.*, 2006; *Deser et al.*, 2011]. RCMs from the PRUDENCE project tend to simulate too wet (dry) winter conditions in regions with warm (cold) biases. For summer this relationship is reversed. Many of the validation results for earlier model ensembles still hold for the recent EURO-CORDEX ensemble [*Kotlarski et al.*, 2014]. Interannual variability in precipitation is in general well represented [*Jacob et al.*, 2007; *Kotlarski et al.*, 2014]. *Frei et al.* [2006] found that biases in extreme precipitation over the Alps are in general comparable or even smaller than biases in mean intensities. *Nikulin et al.* [2011] report that the Swedish RCA model tends to overestimate precipitation extremes over northern Europe and to underestimate them in southern Europe. In summer, extreme precipitation simulated by RCA is overestimated over mountains and underestimated over the surrounding slopes. Detailed validation studies of extreme precipitation simulated by ENSEMBLES RCMs considered, e.g., the dependence on airflow [*Maraun et al.*, 2012] and seasonality [*Schindler et al.*, 2007].

These studies mostly considered the marginal distribution of precipitation at individual sites, i.e., the unconditional climatological distribution at a certain location, and its systematic spatial variation. Yet, in several situations, a validation of residual spatial dependence is also required. By residual spatial dependence we mean joint variations at different locations about the (climatological) systematic spatial variations. Consider, e.g., two nearby rain gauges along the wind-facing slope of a mountain chain. If a deep cyclone passes by, heavy rain will likely occur at both gauges. Thus, if rain is heavy at one of the gauges, the probability of heavy rain is high at the second gauge as well.

Although it is known that RCMs tend to overestimate the spatial extent of precipitation events [e.g., *Maraun et al.*, 2010], the validation of spatial dependence has been limited in the context of climate modeling. In principle two strategies are conceivable. In numerical weather prediction, skill scores such as the intensity skill score [*Casati et al.*, 2004], the fractions skill score [*Roberts and Lean*, 2007], and the Scale-Aspect-Location (SAL) skill score [*Wernli et al.*, 2008] have been developed to verify the forecast of spatial fields. For an intercomparison of spatial verification methods, see *Gilleland et al.* [2009] and *Gilleland et al.* [2010]. These methods require a temporal correspondence between observational and forecast fields and thus cannot be directly used with a free-running climate model, whose simulated weather is uncorrelated with real weather. Their use is therefore limited to simulations driven with observed (so-called perfect) boundary conditions [*Frei et al.*, 2003]. As a result, spatial verification measures have only recently been applied in a climate context to assess the added value of convection resolving simulations [*Prein et al.*, 2013].

Alternatively, spatial statistical measures can be derived, such as correlation lengths, variogram ranges, or length scales of the tail dependence coefficient [*Coles*, 2001]. The former two describe the overall spatial dependence, i.e., the dependence over the full range of the distribution without explicitly accounting for extremes; the latter explicitly measures the spatial dependence of extreme events (loosely speaking, the probability that an extreme occurs at one location, given that an extreme has occurred at another location). The application of such measures for climate model validation is still limited; see *Rasmussen et al.* [2012] for a rare example.

Spatial dependence measures can either be calculated empirically from the data or by assuming statistical models, e.g., from geostatistics [e.g., *Diggle and Ribeiro*, 2007] or extreme value theory [e.g., *Davison et al.*, 2012]. Empirical estimates are simple but have severe disadvantages: they suffer from large uncertainties without providing proper strategies to quantify these uncertainties. In particular empirical estimates of tail dependence are either biased or very uncertain due to the rareness of extremes. Spatial statistical models attempt to overcome this problem by assuming parametric models for the dependence structure. These models first of all constrain the possible dependence structure and thereby greatly reduce uncertainties; furthermore, they can be utilized to derive confidence intervals. Classical geostatistical models assume Gaussian processes. But these can describe neither the heavy tails of the marginal distributions nor the spatial extremes of precipitation at daily timescales. Some authors have used transformed Gaussian processes to successfully model daily precipitation fields [*Sansó and Guenni*, 2000; *Ailliot et al.*, 2009; *Kleiber et al.*, 2012]. Still, such approaches are not able to capture the spatial dependence of extreme precipitation. Max stable spatial processes have been developed to model spatial extremes but do not describe nonextreme events [*Davison et al.*, 2012].

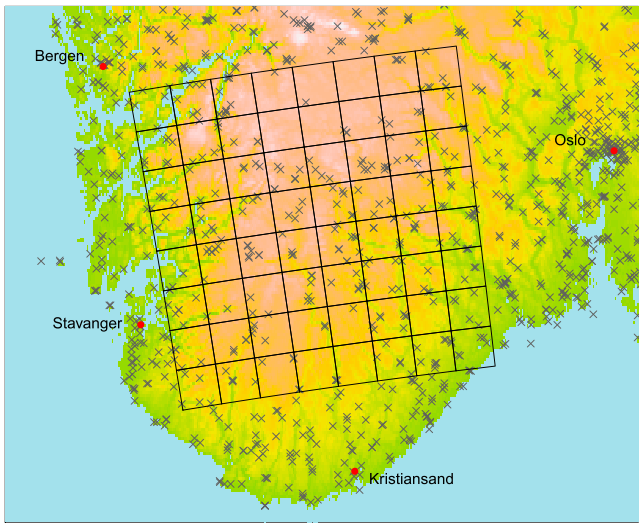

Figure 1. Study area in southern Norway. Crosses indicate the locations of the gauges used in the gridded observational data set.

As an alternative approach, copulas [*Sklar*, 1959] have become popular over recent years; see *Schölzel and Friederichs* [2008] for an introduction with a focus on climate research. The key idea of copulas is to model the dependence structure of a multivariate—e.g., spatially dependent—random variable separately from its marginal distribution. The copula is then the parametric function (or density) that models the dependence structure. Various different families of copulas have been formulated that differ in their detailed shape, symmetries, and tail dependence. A major drawback of many copulas is that they are not easily extended to more than three dimensions, in our context three boxes. Normal copulas have been applied to model spatial precipitation fields [*Wilks*, 1998; *Serinaldi*, 2009; *Yang et al.*, 2005; *Kleiber et al.*, 2012], but they cannot capture tail dependence or nonlinear dependence.

Pair-copula constructions (PCCs) [*Joe*, 1996] are a recent approach to overcome these problems, that–to our knowledge–has not yet been applied in climate science. The underlying idea is to decompose the multivariate dependence structure into layers of bivariate dependencies and model these individually with bivariate copulas. This approach offers high flexibility and is in principle applicable to any dimension—of course with the usual limitations to model selection and parameter uncertainties in the light of limited data.

Here we apply PCCs to validate how well a suite of RCMs from the ENSEMBLES project [*van der Linden and Mitchell*, 2009] represents the residual spatial dependence of daily precipitation amounts. We restrict ourselves to modeling the amounts process, i.e., all dependencies are conditional on having precipitation in the involved grid boxes. As a case study, we consider a region covering south central Norway. The use of PCCs allows us to separately model marginal distributions and dependence structure, and also to consider the overall spatial dependence of precipitation separately from tail dependence. The parametric statistical model provides robust estimates and a straightforward quantification of uncertainties.

In section 2, we lay out the data and models used for the case study. The PCC approach is explained in depth in section 3, with additional detail in the Appendix A. The model derived for our case study is presented in section 4, and the actual results are given in section 5.

2. Data and Models

As an example region, we consider a square of 200 × 200 km^2 covering the major part of southern Norway (Figure 1). The northwest corner is located at 60°06′N, 5°54′E close to the city of Bergen, the southeast corner at 58°48′N, 9°18′E. The topography is characterized by mountains (>1500m) and fjords in the northwest and a hilly landscape in the southeast. The central north is covered by the Hardangervidda, Europe's largest high plateau (>1000m).

We evaluate 10 RCMs available from the ENSEMBLES project (see Table 1). To isolate the RCM performance from errors inherited by the driving global climate model, we chose a so-called perfect boundary setting [e.g., *Frei et al.*, 2006]: simulations are driven with ERA-40 boundary conditions [*Uppala et al.*, 2005] over the period 1 January 1961 to 31 December 2000. All models provide daily data on the same rotated grid with a resolution of 0.22° × 0.22° (approximately 25 km × 25 km). The example region covers a square of 8 × 8 model grid boxes.

The observational reference data have been provided by the Norwegian Meteorological Office. These data were originally gridded from gauge observations to a 1 km × 1 km grid based on triangular interpolation

Table 1. RCMs Considered in the Study

RCM	Institute	Reference
HIRHAM5	DMI	*Christensen et al.* [1996]
HIRHAM	METNO	*Haugen and Haakenstad* [2006]
RCA	SMHI	*Kjellström et al.* [2005]
RCA3	C4I	*Kjellström et al.* [2005]
CLM	ETHZ	*Böhm et al.* [2006]
RACMO2	KNMI	*van Meijgaard et al.* [2008]
HadRM3Q0	UKMO	*Collins et al.* [2006]
HadRM3Q16	UKMO	*Collins et al.* [2006]
HadRM3Q3	UKMO	*Collins et al.* [2006]
REMO	MPI-M	*Jacob* [2001]
RPN_GEMLAM	EC	

[*Jansson et al.*, 2007] and then averaged to the required 25 km × 25 km grid. Undercatch has been accounted for by empirical relationships. The density of stations is in general high but strongly varies across space (see Figure 1). In the western and eastern parts of our study region it is very high and allows for a reasonable estimation of residual spatial dependence. Across the central Hardangervidda plateau, virtually no stations are available. Here the inferred spatial dependence is dominated by the interpolation method but provides only little insight into the actual precipitation dependence. The network has changed during the period we consider, but at any point in time about 100 stations have been operating over the selected study area. These changes are less important in our context, as we consider a time-independent statistical model.

3. Pair-Copula Constructions

Pair-copula constructions (PCCs) have been proposed by *Joe* [1996] and later studied by *Bedford and Cooke* [2001, 2002] and *Aas et al.* [2009]. These models enable one to formulate—at least in principle—multivariate dependence structures in any dimension. We start with a brief review of copulas and then discuss the concept of PCCs in detail.

3.1. Copulas

The purpose of copulas is to study the dependence structure of some variables of interest, $(X_1, \dots, X_d)$, by factoring out each variable's marginal behavior. To this end, all variables are transformed to have uniform margins, $U(0, 1)$, using the respective marginal cumulative distribution functions (CDFs) $F_i(x_i)$. Sklar's theorem [*Sklar*, 1959] states that any d-variate CDF $F_{1\dots d}(x_1, \dots, x_d)$ may be expressed as

$$F_{1\dots d}(x_1, \dots, x_d) = C_{1\dots d}(F_1(x_1), \dots, F_d(x_d)),$$

where $C_{1\dots d}$ is a copula. That is, a copula is simply the joint cumulative distribution function (CDF) of marginally uniform variables $U(0, 1)$. In our application the joint CDF can be assumed to be absolutely continuous, with strictly increasing marginal CDFs. The copula is then unique and has a density $c_{1\dots d}$. For an introduction to copulas, the interested reader may consult *Nelsen* [1999] or *Joe* [1997]. An introduction with a climatological and meteorological focus can be found in *Schölzel and Friederichs* [2008].

3.2. Measures of Dependence

To quantify the dependence between variables, different measures have been defined, addressing different aspects of dependence. Many important measures of dependence are functions merely of the copulas and not of the univariate margins.

A well-known example of the overall dependence is Spearman's rank correlation ρ, defined as

$$\begin{aligned}\rho_S(X_1, X_2) &= 3\left(\mathrm{P}\left(\left(X_1 - \tilde{X}_1\right)\left(X_2 - X_2'\right) > 0\right) - \mathrm{P}\left(\left(X_1 - \tilde{X}_1\right)\left(X_2 - X_2'\right) < 0\right)\right) \\ &= 12\int_0^1\int_0^1 C_{12}(u_1, u_2)\mathrm{d}u_1\mathrm{d}u_2 - 3 = \mathrm{Cor}(F_1(X_1), F_2(X_2)).\end{aligned} \tag{1}$$

Here $\tilde{X}_1$ and X_2' are independent copies of X_1 and X_2, respectively. Compared to the standard linear correlation coefficient, Spearman's rank correlation has the additional advantage of also capturing nonlinear dependence.

Dependence of rare events cannot be measured by overall correlations. In fact, even if two variables are completely uncorrelated, there can be a significant probability of a concurrent extreme event in the two,

i.e., they can still be tail dependent. A measure of upper tail dependence that is a function of just the copula is given by

$$\lambda_U(X_1, X_2) = \lim_{u \nearrow 1} \mathrm{P}\left(X_2 > F_2^{-1}(u) | X_1 > F_1^{-1}(u)\right) = \lim_{u \nearrow 1} \frac{1 - 2u + C_{12}(u,u)}{1-u}. \quad (2)$$

This is the conditional probability that both variables are very "large" simultaneously, given that one of them has a high value. Likewise, the lower tail dependence coefficient is defined as

$$\lambda_L(X_1, X_2) = \lim_{u \searrow 0} \mathrm{P}\left(X_2 < F_2^{-1}(u) | X_1 < F_1^{-1}(u)\right) = \lim_{u \searrow 0} \frac{C_{12}(u,u)}{u}. \quad (3)$$

The copula C_{12} is said to have upper tail dependence if $\lambda_U(X_1, X_2) > 0$ and lower tail dependence if $\lambda_L(X_1, X_2) > 0$. See Figure S1 in the supporting information for simulations from four classical bivariate, i.e., pair-copulas; the Gaussian, which is symmetric and has no tail dependence; the Student's t, also symmetric but with upper and lower tail dependence; the Clayton copula, which is asymmetric with only lower tail dependence; and the Gumbel, with only upper tail dependence. Independent variables can formally be "connected" by the independence copula

$$C_{12}(u_1, u_2) = u_1 u_2.$$

3.3. Pair-Copula Constructions

Whereas the list of bivariate copulas is long and varied, the choice is much more limited in dimensions 3 and higher. Some of the bivariate copulas have multivariate extensions, among others the Gaussian or Archimedean copulas. However, these typically have very confined parameter spaces as the dimension increases, which restricts the degree of dependence they are able to portray. Moreover, they assume that all the bivariate dependencies are of the same type, for instance, that they are either tail dependent or not. The key idea of PCCs is to decompose multivariate copulas into a product of bivariate components. These bivariate components can then be modeled individually by appropriate pair-copulas.

To introduce the concept of PCCs, we start with a three-dimensional example. Let f_{123} be the joint probability density function (pdf) of (X_1, X_2, X_3), and f_i, $i = 1, 2, 3$ the corresponding marginal pdfs. Further, let $f_{2|3}$ and $f_{1|23}$ be the conditional pdfs of X_2 given $X_3 = x_3$ and X_1 given $(X_2, X_3) = (x_2, x_3)$, respectively, defined by

$$f_{2|3}(x_2|x_3) = \frac{f_{23}(x_2, x_3)}{f_3(x_3)}$$
$$f_{1|23}(x_1|x_2, x_3) = \frac{f_{123}(x_1, x_2, x_3)}{f_{23}(x_2, x_3)}.$$

Let us begin with the factorization

$$f_{123}(x_1, x_2, x_3) = f_{1|23}(x_1|x_2, x_3) f_{2|3}(x_2|x_3) f_3(x_3). \quad (4)$$

A corresponding PCC is obtained by expressing (4) as a product of f_i and three bivariate copula densities.

Sklar's theorem, written in terms of densities, states that, e.g., for x_1 and x_2,

$$f_{12}(x_1, x_2) = c_{12}(F_1(x_1),\ F_2(x_2)) f_1(x_1) f_2(x_2). \quad (5)$$

This means that the conditional density $f_{1|2}$ of X_1, given $X_2 = x_2$, can be expressed as

$$f_{1|2}(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1).$$

As mentioned earlier, copulas allow us to isolate the dependence, i.e., the joint behavior of the variables, from the variables' individual behavior and are naturally linked to many important measures of dependence. The purpose of PCCs is to extend (5) to more than two variables, while keeping all components at most bivariate. Now let $c_{12|3}$ be the copula density corresponding to the conditional distribution of (X_1, X_2), given $X_3 = x_3$, determined by

$$f_{12|3}(x_1, x_2|x_3) = c_{12|3}(F_{1|3}(x_1|x_3),\ F_{2|3}(x_2|x_3)) f_{1|3}(x_1|x_3) f_{2|3}(x_2|x_3).$$

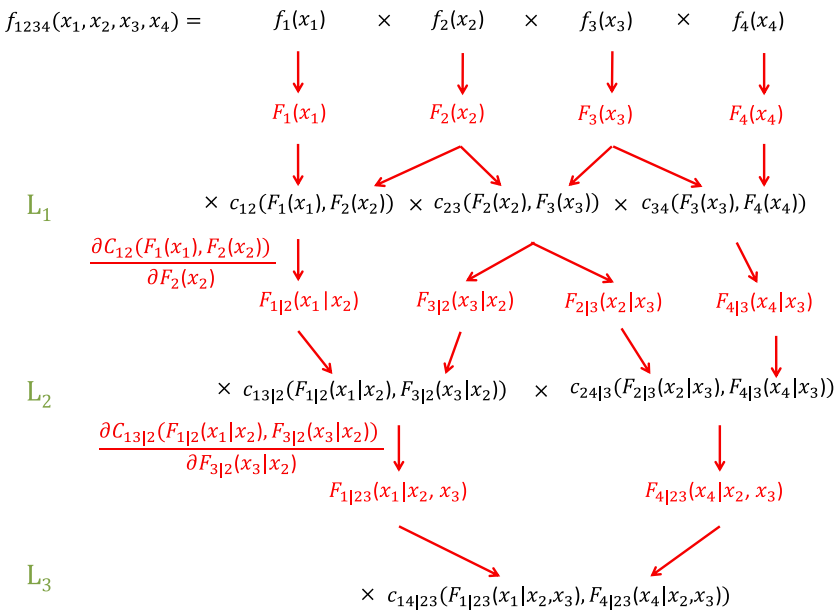

Figure 2. Components of a four-dimensional pair-copula construction. The components of the PCC are organized in three levels L_1 (the ground level), L_2, and L_3.

Note that this copula is a function of conditional CDFs. The factor $f_{1|23}$ from (4) may then be replaced by

$$\begin{aligned} f_{1|23}(x_1|x_2, x_3) &= \frac{f_{12|3}(x_1, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) f_{1|3}(x_1|x_3) \\ &= c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) c_{13}(F_1(x_1), F_3(x_3)) f_1(x_1). \end{aligned}$$

Finally, inserting the above expressions in (4), we obtain the PCC

$$\begin{aligned} f_{123}(x_1, x_2, x_3) = f_1(x_1)\, f_2(x_2)\, f_3(x_3) \\ c_{13}(F_1(x_1), F_3(x_3)) c_{23}(F_2(x_2), F_3(x_3)) \\ c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)). \end{aligned} \tag{6}$$

A key point is that all copulas involved in this decomposition are pair-copulas. In equation (6), the two arguments of the pair-copula in the last line are conditional CDFs. In higher dimensional PCCs, this is the case for all pair-copulas, except the ones at the lowest level of the dependence structure, the ground level, on line 2 of equation (6).

Moreover, the above decomposition is not unique. There are two alternative ways of decomposing f_{123} into a PCC, which are obtained by permuting x_3 with either x_1 or x_2.

Under the given conditions, any d-dimensional absolutely continuous pdf can be decomposed, as described above, into a product of the d marginal pdfs and $d(d-1)/2$ pair-copulas organized in $d-1$ levels. Each of these copulas is a function of two CDFs, more specifically unconditional marginal CDFs at the ground level and conditional CDFs on subsequent levels. The conditioning sets of these conditional CDFs, i.e., the sets of variables that one conditions on, increase in size by 1 with each level. A key result from *Joe* [1996] states that the pair-copulas and the CDFs are linked through

$$F_{1|3}(x_1|x_3) = \frac{\partial C_{13}(F_1(x_1), F_3(x_3))}{\partial F_3(x_3)}$$

for one conditioning variable, and in general,

$$F_{x|\boldsymbol{v}}(x|\boldsymbol{v}) = \frac{\partial C_{xv_j|\boldsymbol{v}_{-j}}(F_{x|\boldsymbol{v}_{-j}}, F_{v_j|\boldsymbol{v}_{-j}})}{\partial F_{v_j|\boldsymbol{v}_{-j}}}, \ \forall j,$$

where $\boldsymbol{V}$ is a random vector, not containing X, and $\boldsymbol{V}_{-j}$ is the same vector reduced by variable V_j.

When using a PCC in practice, one must be able to compute the conditional distributions, which constitute the pair-copula arguments, in a straightforward manner. This is done using the above formula recursively, as demonstrated in Figure 2. This figure shows how a four-dimensional pdf is decomposed into a PCC. The components of the PCC are shown in black, while the copula arguments, which must be computed based on other components of the PCC, are shown in red. For instance, one obtains $F_{1|2}(x_1|x_2)$ by applying the formula to the copula C_{12}, i.e., by taking the derivative of C_{12} with respect to its second argument and substituting $F_1(x_1)$ and $F_2(x_2)$ for u_1 and u_2, respectively. Further, $F_{1|23}(x_1|x_2, x_3)$ results from applying the same formula to $C_{13|2}$, and so on. This requires that the necessary pair-copulas $C_{xv_j|\mathbf{v}_{-j}}$ are available when needed, hence that they are present in the preceding levels. This is not always the case. Consider, for example, the four-variate distribution with pdf

$$\begin{aligned} f_{1234} = f_1 f_2 f_3 f_4 \\ & c_{12}(F_1, F_2) c_{13}(F_1, F_3) c_{14}(F_1, F_4) \\ & c_{23|1}(F_{2|1}, F_{3|1}) c_{34|1}(F_{3|1}, F_{4|1}) \\ & c_{34|12}(F_{3|12}, F_{4|12}). \end{aligned}$$

The conditional CDF $F_{3|12}$, given by

$$F_{3|12} = \frac{\partial C_{31|2}(F_{3|2}, F_{1|2})}{\partial F_{2|1}},$$

is then needed. Yet $C_{31|2}$ is not used in the previous level and thus is not directly available. It may be recovered via integration. However, as d increases, so does the dimension of the required integrals. Therefore, it is important to develop a mechanism to construct PCCs, guaranteeing that all the necessary elements are available when needed. We call such a PCC self-contained.

3.4. Building PCCs and Parameter Estimation

Self-contained PCCs can be built by regular (R-) vines [*Bedford and Cooke*, 2001, 2002], see Appendix A. All the pair-copulas can be selected completely freely, and they are available when needed downstream in the construction. The resulting structure is guaranteed to be valid. This makes PCCs extremely flexible and able to portray a wide range of complex dependencies. Inference on these constructions requires (i) the choice of structure, i.e., which pairs of variables with corresponding conditioning sets should be connected on the different levels, (ii) the choice of each parametric pair-copula, and (iii) the estimation of the copula parameters. In principle, these three steps should be performed simultaneously. Of course it is not feasible in practice and has to be done stepwise, which is suboptimal but still satisfactory.

Strategies for the structure selection step (i) are described in Appendix A. We use the algorithm of *Dißmann et al.* [2013], which is the state-of-the-art for structure selection in R-vines. It has the advantage of using Kendall's τ coefficients, which can also capture nonlinear dependence. As mentioned earlier, there is a long list of bivariate copulas from which to choose in step (ii). A catalogue of pair-copula families can be found in, for instance, *Joe* [1997] or *Nelsen* [1999]. For smaller problems, the pair-copulas may be chosen by goodness-of-fit tests, such as the ones studied in *Genest et al.* [2009]; *Berg* [2009]. As the p values of those tests are computed by bootstrap, this is not feasible for problems of the size treated in this paper. Instead, we use Akaike's information criterion (AIC), which is not justified by likelihood arguments in this setting [*Grønneberg*, 2011] but has been shown to work rather well in practice [*Dißmann et al.*, 2013]. Finally, the model parameters are estimated based on the stepwise semiparametric (SSP) estimator; see Appendix A.

4. Statistical Model

We describe the spatial dependence of daily precipitation amounts by PCCs and compare the results for the RCMs with those from observations. To avoid modeling the seasonal cycle, we divide the year into a winter season from October to March and a summer season from April to September. Both seasons are treated independently.

Commonly, precipitation occurrence $I_{i,t}$ and amount $X_{i,t}$ on a day t in grid box i are modeled separately as independent random variables [*Wilks*, 1998; *Chandler and Wheater*, 2002]. As our focus is on high quantiles of precipitation, we only consider the amount process. We assume $X_{i,t}$ to be independent in time (see Appendix B for a justification) but not in space. As the wet day threshold, we choose 0.1 mm.

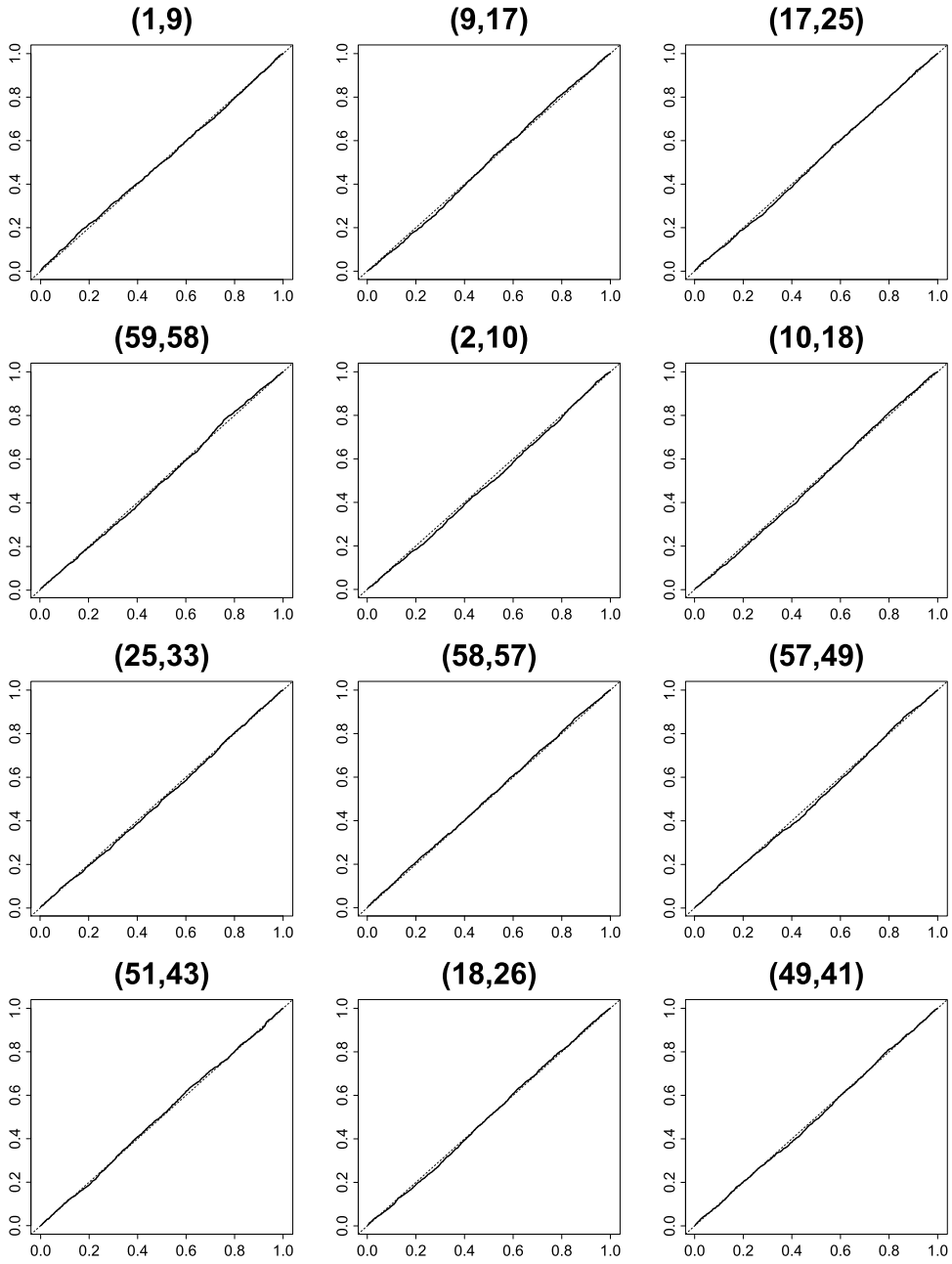

Figure 3. Empirical K function for 12 pairs of grid boxes from the ground level of the PCC for the observations for summer, plotted against the corresponding parametric K function.

The marginal amount distributions vary strongly with orography and are in general rather different from location to location. Therefore, we model the marginal distribution of each grid box independently with the best fitting gamma distribution. To join these 64 univariate margins, we use a PCC, more specifically a regular vine. We follow the three steps described in section 3, i.e., (i) selecting the structure, (ii) choosing the pair-copula types, and (iii) estimating the copula parameters.

In step (i), we choose to select a common structure for all RCMs and the observational data; the copula types and parameter estimates are of course individual. This means that the connected pairs and corresponding conditioning sets on all levels are the same for the observed data and the RCMs. For instance, all the regular vines contain the copula $C_{14|23}$ in the third level. However, $C_{14|23}$ may be Gumbel in the R-vine of the observed data and Student's t in the R-vine of one of the RCMs. The advantage of a common structure is that the resulting statistical models are much easier to compare. The measures we compare, e.g., rank

Table 2. Percentage of Copulas of Each Family Chosen at the 10 First Levels of the Structure for the Observations and the RCMs for Winter

	Independent	Gaussian	Student's *t*	Clayton	Gumbel
Observations	25.5	17.4	42.9	5.1	9.1
METNO-HIRHAM	9.6	10.1	67.4	3.4	9.6
C4IRCA3	12.1	7.5	69.4	2.4	8.5
DMI-HIRHAM5	13.7	10.4	59	4.6	12.3
ETHZ-CLM	14.4	9.4	60.9	5.1	10.3
KNMI-RACMO2	15.7	7.9	64.1	1.9	10.4
METO-HC_HadRM3Q0	19	21.4	39.8	5.6	14.2
METO-HC_HadRM3Q16	18.1	22.6	39	6.2	14.2
METO-HC_HadRM3Q3	15.6	15.7	47.2	5.5	16.1
MPI-M-REMO	13.7	7	65.5	5.3	8.5
RPN_GEMLAM	21.2	12.3	62.2	2.6	1.7
SMHIRCA	19.5	6.5	70.3	2.4	1.4

correlations, are defined for all the copula families considered in the paper. Comparison between the statistical models is therefore straightforward also when the selected copula families are different from model to model. Since the copula types can be chosen freely, they are still very flexible, and the structure is guaranteed to be a valid copula density, so that the estimates will be consistent. We stress that we do not intend to use the statistical models for prediction, in which case it would be more important to choose the best fitting model for each data set.

The idea is to select the structure based on all data sets, so that they all influence the resulting structure more or less. If the structure does not fit one of the data sets very well, the parameter estimates may have larger uncertainty. This will be reflected in broader confidence intervals for the parameters of interest. Selecting the structure based on all data sets is rather time consuming. Moreover, it may be argued that the RCMs are biased and that it would be better to use only the observed data in the selection. In addition, a preliminary study indicated that apparently very different structures produced very similar results, and that the structure selected based only on the observed data is essentially the same as the one selected based on all data series. Hence, the statistical models seem to compensate through the choice of copula types, as well as the parameter estimates.

In the estimation of the marginal gamma distributions, we use all values above the wet day threshold in each location for the given season. Likewise, each bivariate copula in the PCC is fitted based on all observations that are above the threshold in the variables involved in the given copula simultaneously. For instance, the parameters of copula $C_{14|23}$ are estimated using the observations where X_1, X_2, X_3, and X_4 are all above the wet day threshold. For most of the series and both season, this corresponds to between 1700 and 6000 observations, depending on the copula, and at least 1000.

In order to assess how well the PCC fits the data, we have computed the parametric K function

$$K(w) = \mathrm{P}(C_{12}(U_1, U_2) \leq w)$$

Table 3. Same as Table 2 but for Summer

	Independent	Gaussian	Student's *t*	Clayton	Gumbel
Observations	29.2	15.2	38.3	8.7	8.5
METNO-HIRHAM	16.6	10.4	60.7	4.8	7.5
C4IRCA3	17.1	12	58.6	4.6	7.7
DMI-HIRHAM5	15.6	13	57.8	4.1	9.6
ETHZ-CLM	16.6	10.8	57.9	4.8	9.9
KNMI-RACMO2	22.9	6.7	59.7	3.4	7.4
METO-HC_HadRM3Q0	22.1	15.4	41.4	6.5	14.7
METO-HC_HadRM3Q16	25.5	15.7	38.3	4.8	15.7
METO-HC_HadRM3Q3	22.7	12.3	45	6	14
MPI-M-REMO	20.9	5.1	63.4	3.8	6.8
RPN_GEMLAM	37.4	6.8	49.9	3.1	2.7
SMHIRCA	24.3	8.7	61.4	3.9	1.7

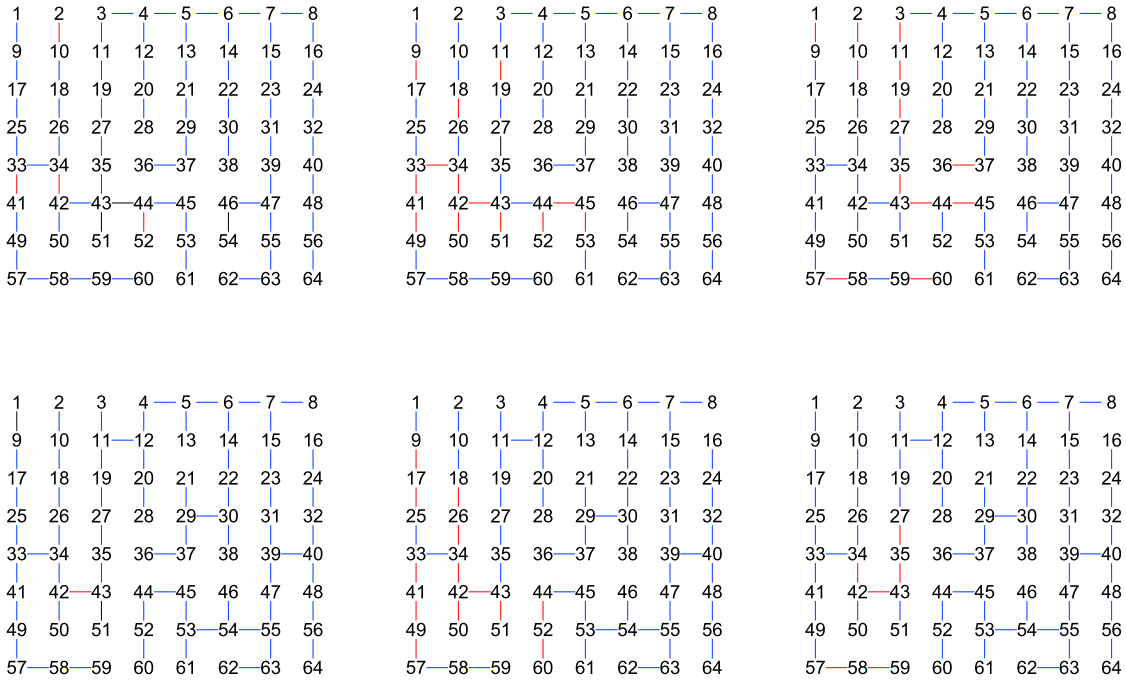

Figure 4. Ground-level pairs selected for (left) the observations, (middle) METNO-HIRHAM, and (right) C4IRCA3 from the (top) winter and (bottom) summer seasons. The color of the lines represents the corresponding copula family (red = Student's *t*, blue = Gumbel, black = Gaussian, green = Clayton; note that the Clayton copula is never chosen on the ground level.)

for each of the pair-copulas of the ground level, as well as the empirical equivalent based on the empirical copula

$$C_n(u,v) = \frac{1}{n}\sum_{j=1}^{n} I(\tilde{u}_{1j} \le u, \tilde{u}_{2j} \le v),$$

where $\tilde{u}_{kj} = \frac{1}{n+1}\sum_{l=1}^{n} I(x_{kl} \le x_{kj}), k = 1, 2$. These K functions are suggested as diagnostic tools for bivariate copulas in *Genest and Favre* [2007]. Figure 3 shows the empirical function for selected pairs plotted against the parametric for summer observations. The resulting lines follow the main diagonal closely, indicating a good fit of the model. A similarly good fit for the remaining pairs, as well as for winter and the METNO-HIRHAM model (not shown) indicates the general suitability of the PCC model.

5. Results

We calibrate the spatial statistical model against both the different RCM simulations and the observational reference data set and compare the results separately for winter and summer. Specifically, we compare the selected model structures, the overall spatial dependence, the tail dependence, and spatially aggregated amounts. Summary statistics are shown for all models and individual maps only for METNO-HIRHAM (one of the best performing) and C4IRCA3 (a representative model). Corresponding maps for the remaining models are shown in Figures S2–S4 in the supporting information.

5.1. Model Structure

As explained above, all 11 data sets from a given season share the same structure, but the associated copula families and parameter estimates may be different. A comparison of the selected copulas and the direction, in which the ground level copulas connect the different grid boxes, already provides insight both into the dependence structure itself, and into the performance of the RCMs to represent it.

Tables 2 and 3 show the percentage of each copula chosen for the observations and RCMs in the first 10 levels for winter and summer, respectively; corresponding tables for all levels are to be found in the supporting information (Tables S1 and S2). The proportion of independence and Gaussian copulas is higher in observed than in simulated precipitation, especially in the first levels. Vice versa, more Student's *t*

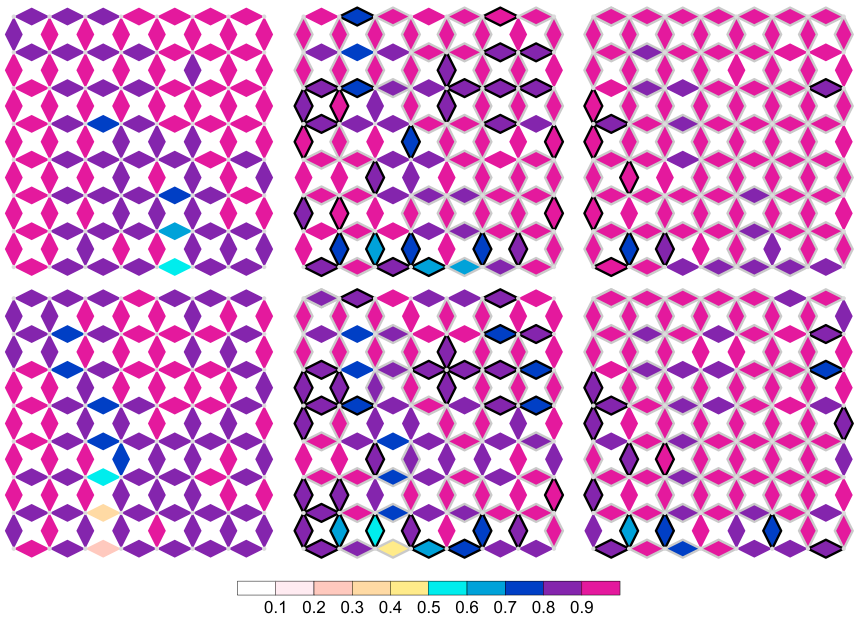

Figure 5. Rank correlation for neighboring grid box pairs. Shown for (left) observations, (middle) METNO-HIRHAM, and (right) C4IRCA3 for (top) winter and (bottom) summer seasons. Grey dots: positions of underlying grid box centers; color: strength of correlation. Black (grey) diamond borders for RCMs indicate a correlation significantly lower (higher) than for observations.

copulas have been chosen for RCM-simulated precipitation than for the observations. These findings indicate that RCM-simulated precipitation exhibits stronger tail dependence than observations. It is, however, worth noting that above the first level, the PCCs model conditional dependencies. Thus, there may be conditional dependencies that are larger in the tails for simulated precipitation.

The ground level is in general dominated by the Gumbel copula, for observations and all RCMs but RPN_GEMLAM and SMHIRCA (see Figure 4 and supporting information). Thus, neighboring pairs tend to exhibit upper, but not lower, tail dependence. Moreover, the selected pairs at this level are mostly in the north-south direction, indicating a stronger dependence in this direction rather than from west to east.

5.2. Overall Dependence

As a measure for the overall spatial dependence, we choose Spearman's rank correlation ρ, equation (1).

Figure 5 shows rank correlations between the amounts of precipitation in neighboring grid boxes for the observations, METNO-HIRHAM, and C4IRCA. The differences between RCMs and observations could in principle stem from parameter uncertainty. We have therefore assessed whether the differences are significant at the 95% level. The critical values for the distribution under the null hypothesis were obtained with parametric bootstrap. More specifically, we have simulated 1000 samples of the same size as the data series in question, estimated the parameters of the statistical model for each of the samples, computed pairwise rank correlations for each set of parameters, and finally determined the empirical 2.5 and 97.5% quantiles of the resulting differences in rank correlations.

Most neighboring grid boxes show a correlation close to 1, i.e., strong dependence. Recall that the values for observed data in the central north should be interpreted carefully due to the low station density on the Hardangervidda plateau. A band showing a tendency of weaker dependence crosses the study area from the northwest to the central south that is particularly pronounced during summer (the very low correlation in the southern part of this band should not be overinterpreted, as the station density is low in that region). This band corresponds to major divides, namely, the Sørfjord in the north and the Setesdal in the south. The area means of the rank correlation and their biases are shown in Table 4. All in all, most models tend to slightly—though often significantly—overestimate the rank correlation, i.e., precipitation is simulated slightly too smooth in space.

Table 4. Area Means of the Dependence Measures and Their Biases

	Spearman Correlation				Upper Tail Dependence			
	Winter		Summer		Winter		Summer	
	Value	Bias	Value	Bias	Value	Bias	Value	Bias
Observations	0.894	N.A.	0.871	N.A.	0.716	N.A.	0.695	N.A.
METNO-HIRHAM	0.898	0.003	0.867	−0.004	0.774	0.058	0.763	0.068
C4IRCA3	0.937	0.042	0.912	0.041	0.807	0.09	0.767	0.073
DMI-HIRHAM5	0.89	−0.004	0.894	0.022	0.765	0.049	0.775	0.081
ETHZ-CLM	0.894	0.000	0.881	0.01	0.766	0.05	0.778	0.084
KNMI-RACMO2	0.949	0.055	0.931	0.059	0.848	0.132	0.789	0.095
METO-HC_HadRM3Q0	0.871	−0.023	0.884	0.013	0.743	0.027	0.755	0.06
METO-HC_HadRM3Q16	0.881	−0.013	0.868	−0.003	0.77	0.053	0.756	0.061
METO-HC_HadRM3Q3	0.905	0.011	0.885	0.014	0.775	0.058	0.75	0.055
MPI-M-REMO	0.903	0.009	0.86	−0.011	0.755	0.038	0.733	0.038
RPN_GEMLAM	0.934	0.04	0.902	0.03	0.775	0.059	0.793	0.098
SMHIRCA	0.956	0.062	0.943	0.071	0.825	0.109	0.839	0.144
Multimodel mean	0.911	0.017	0.894	0.022	0.782	0.066	0.773	0.078

To investigate how well the spatial patterns—apart from mean biases shown in Table 4–of the rank correlation are simulated, we calculated pattern correlations and standard deviations, summarized as Taylor diagrams [*Taylor*, 2001] (see Figure 6). The pattern correlations scatter strongly across models; interestingly, they are slightly higher during summer than winter. Pattern standard deviations are also widely spread. In winter they are distributed around the observed value, whereas in summer most model patterns are too smooth.

The minimum scale at which climate models skillfully simulate climate is considerably larger than an individual grid box [*Grotch and MacCracken*, 1991]; hence, it is recommended to validate RCMs at aggregated scales [*Maraun et al.*, 2010]. Also, the dependence at scales of typical larger catchments is relevant for hydrological impact studies [*Maraun et al.*, 2010]. Therefore, we have computed the rank correlation as a function of distance in the north-south and the west-east directions (Figure 7). The values were obtained by

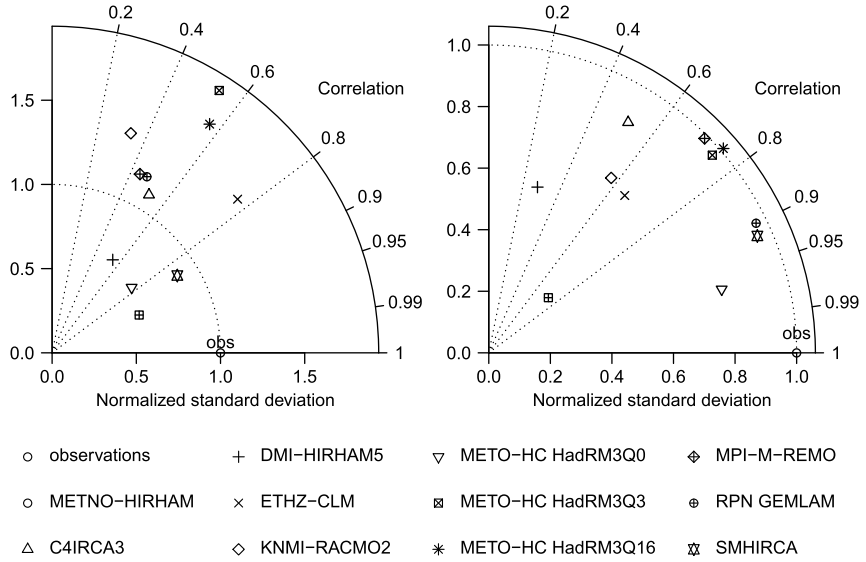

Figure 6. Taylor diagram of rank correlation over all neighboring pairs for (left) winter and (right) summer. Angle: correlation between the spatial patterns of simulated and observed rank correlations; radius: standard deviation of simulated rank correlations, normalized with the standard deviation of the observed rank correlations; distance from (1,1): root-mean-square error between the simulated and observed rank correlation deviations from the corresponding spatial means.

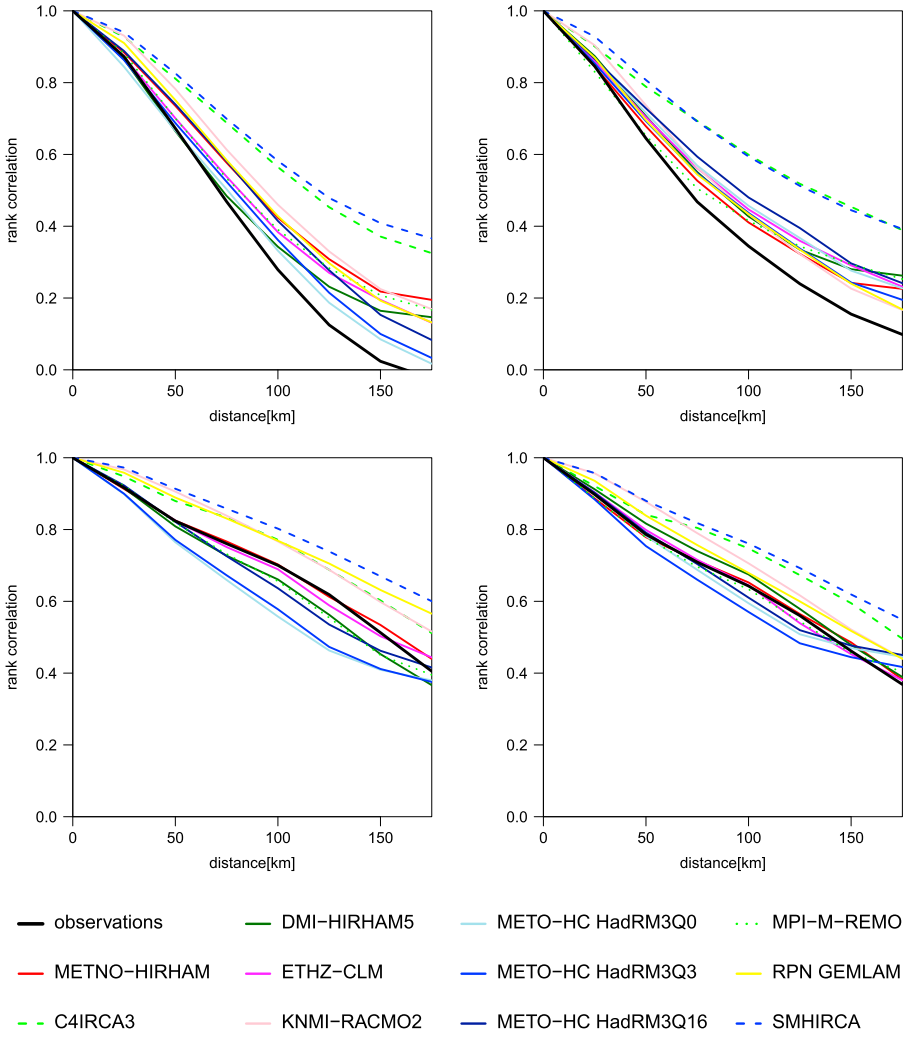

Figure 7. Rank correlation as a function of distance (for definition, see text). Shown for the (first row) west-east and (second row) north-south directions, for (left column) winter and (right column) summer.

averaging the rank correlations between all pairs of grid boxes with the given distance between their center points in the direction in question. This averaging also reduces the influence of the inhomogeneous station coverage. The values decrease with distance, as the spatial dependence is strongest between locations that are close to each other.

As indicated by the selected PCC (section 5.1), the dependence structure is highly anisotropic: it is higher in the north-south than in the west-east direction, and moreover it decays much slower in the north-south direction. The anisotropy reflects the prevailing weather situation: the westerly winds are lifted across the mountain chain, causing a spatially coherent north-south rain band. Most of the RCMs overestimate the spatial dependence, especially for the longer distances in the west-east direction. The orography-related north-south dependence is relatively well simulated.

Tables 5 and 6 show typical decay lengths of the rank correlation, defined as the distance at which the average rank correlation as shown in Figure 7 is reduced to 0.25 for winter and summer, respectively. We derived the decay lengths from third-order polynomials fitted to the decay curves; corresponding 95% confidence intervals are obtained by a parametric bootstrap. In general, the decay lengths for RCMs are longer than for the observed data. This overestimation is significant for all of the RCMs except METNO-HIRHAM and DMI-HIRHAM5.

Table 5. Decay Length of Rank Correlation in Kilometers (for Definition, See Text) for Winter[a]

	West-East	North-South
Observations	108 (99,123)	176 (151,187)
METNO-HIRHAM	153 (133,190)	195 (169,216)
C4IRCA3	216 (180,273)	227 (197,259)
DMI-HIRHAM5	130 (117,163)	194 (187,205)
ETHZ-CLM	141 (126,166)	250 (234,302)
KNMI-RACMO2	146 (132,178)	261 (236,296)
METO-HC_HadRM3Q0	121 (112,136)	297 (243,406)
METO-HC_HadRM3Q16	123 (112,138)	264 (226,351)
METO-HC_HadRM3Q3	137 (124,154)	281 (242,369)
MPI-M-REMO	152 (132,190)	230 (212,267)
RPN_GEMLAM	140 (128,169)	248 (165,305)
SMHIRCA	247 (203,346)	223 (158,292)

[a]Corresponding 95% confidence intervals are shown in parentheses.

5.3. Tail Dependence

The spatial dependence of extreme events is quantified by the coefficients of upper tail dependence, equation (2). Figure 8 shows the upper tail dependence between the amounts of precipitation in neighboring grid boxes for the observations, METNO-HIRHAM and C4IRCA. Area average values and the corresponding model biases are listed in Table 4. Again, results for observed data in the central north should be interpreted carefully.

In the observations, many grid boxes exhibit a rather high upper tail dependence. There is a marked difference between the high dependence in the north-south direction and the comparably low dependence in the southeast direction. The divide between west and east, marked by a band of low dependence values (again, note the relatively low station coverage in the southern part of that band), is strongest for dependence in the north-south direction and spreads over a broad region toward the south. It is much more pronounced than the corresponding pattern for rank correlations.

The spatial pattern of upper tail dependence is reasonably well simulated by both RCMs for winter, whereas both models basically fail during summer. METNO-HIRHAM tends to underestimate the west-east dependence and overestimate the north-south dependence for both seasons; C4IRCA4 generally overestimates the dependence for both seasons.

Figure 9 displays Taylor diagrams for the upper tail dependence. Pattern correlations are in general lower than for the overall dependence; pattern standard deviations are underestimated by almost all models for both seasons.

The decay of the upper tail dependence as a function of distance is shown in Figure 10; the corresponding decay lengths (reduction in upper tail dependence to 0.25) are shown in Tables 7 and 8 for winter and summer, respectively. Like the rank correlation, the upper tail dependence is strongest and decays slower in the north-south direction. Further, about half the RCMs underestimate the upper tail dependence, while the remaining overestimate it. The decay in the north-south direction levels off for long distances at quite high dependencies, indicating that the strongest events occur in large-scale weather systems.

5.4. Spatially Aggregated Amounts

The results indicate that the spatial dependence in the RCMs is considerably misrepresented. To assess how this model deficiency affects the simulation of area-aggregated amounts, we computed upper quantiles of precipitation sums in two neighboring grid boxes. As we do not model the occurrence process, these quantiles are conditional on having a wet day in both considered grid boxes. Given precipitation amounts X_1 and X_2 recorded in two grid boxes on a given day, the aggregate quantiles $q_{\alpha,12}$, given by

$$P(X_1 + X_2 \leq q_{\alpha,12} | X_1 > 0, X_2 > 0) = \alpha,$$

represent the amount that the total precipitation in the two grid boxes exceeds with a certain (small) probability. As

Table 6. Same as Table 5 but for Summer

	West-East	North-South
Observations	131 (118,150)	188 (178,201)
METNO-HIRHAM	181 (154,222)	179 (152,192)
C4IRCA3	250 (221,319)	134 (40,169)
DMI-HIRHAM5	195 (159,258)	209 (193,227)
ETHZ-CLM	184 (155,221)	225 (205,250)
KNMI-RACMO2	152 (135,179)	241 (221,269)
METO-HC_HadRM3Q0	178 (155,213)	443 (336,605)
METO-HC_HadRM3Q16	171 (144,189)	389 (291,517)
METO-HC_HadRM3Q3	176 (158,2169	435 (340,636)
MPI-M-REMO	204 (169,279)	238 (210,274)
RPN_GEMLAM	156 (141,184)	222 (206,257)
SMHIRCA	264 (219,339)	207 (143,264)

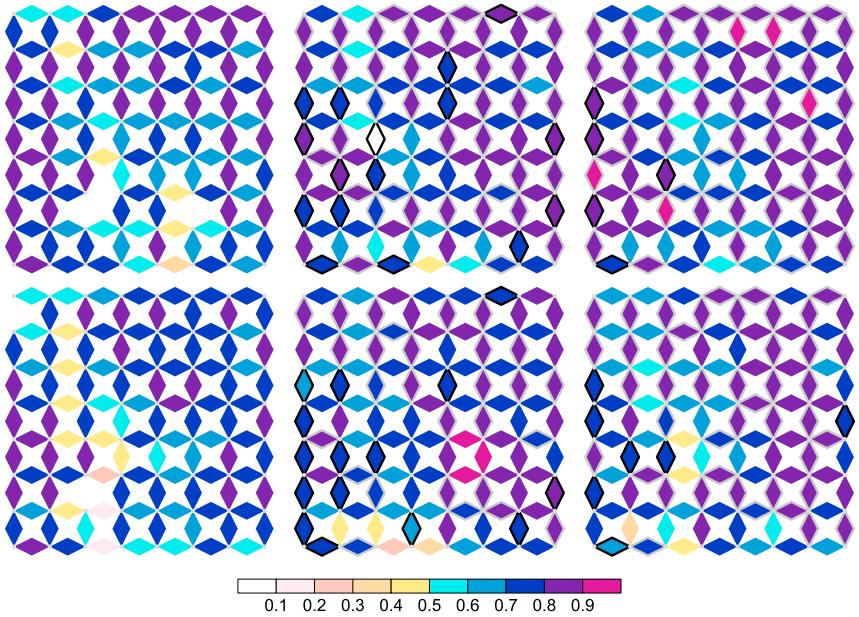

Figure 8. Upper tail dependence for neighboring grid box pairs. Shown for (left) observations, (middle) METNO-HIRHAM, and (right) C4IRCA3 for (top) winter and (bottom) summer. Grey dots: positions of underlying grid box centers; color: strength of tail dependence. Black (grey) diamond borders for RCMs indicate a tail dependence significantly lower (higher) than for observations.

opposed to rank correlations and coefficients of tail dependence, which are functions only of the dependence structure, the quantiles also depend on the univariate margins, i.e., the unconditional amount distribution at a given location. Differences between RCM-simulated and observed quantiles could thus stem from deficiencies in the representation of both spatial dependence and univariate marginal distributions. Here we assess the relative importance of both contributions. Note also that the full, unconditional distribution of spatially aggregated precipitation relies on the occurrence process.

The aggregate 95% quantiles are shown in the top row of Figures 11 (winter) and 12 (summer) for the observations, METNO-HIRHAM and C4IRCA3. Again, results for observed precipitation in the central north

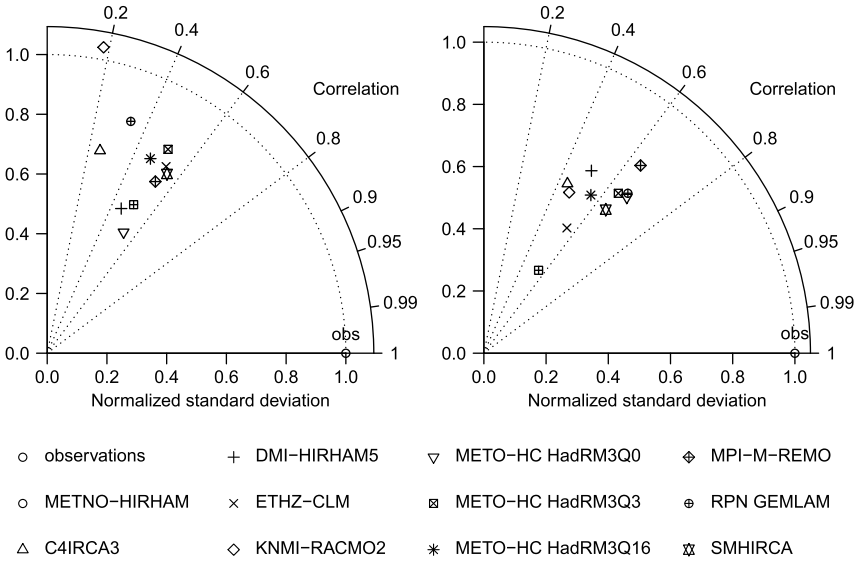

Figure 9. Taylor diagram of upper tail dependence over all neighboring pairs for (left) winter and (right) summer. Angle: pattern correlation; radius: pattern standard deviation; distance from (1,1): pattern root-mean-square error.

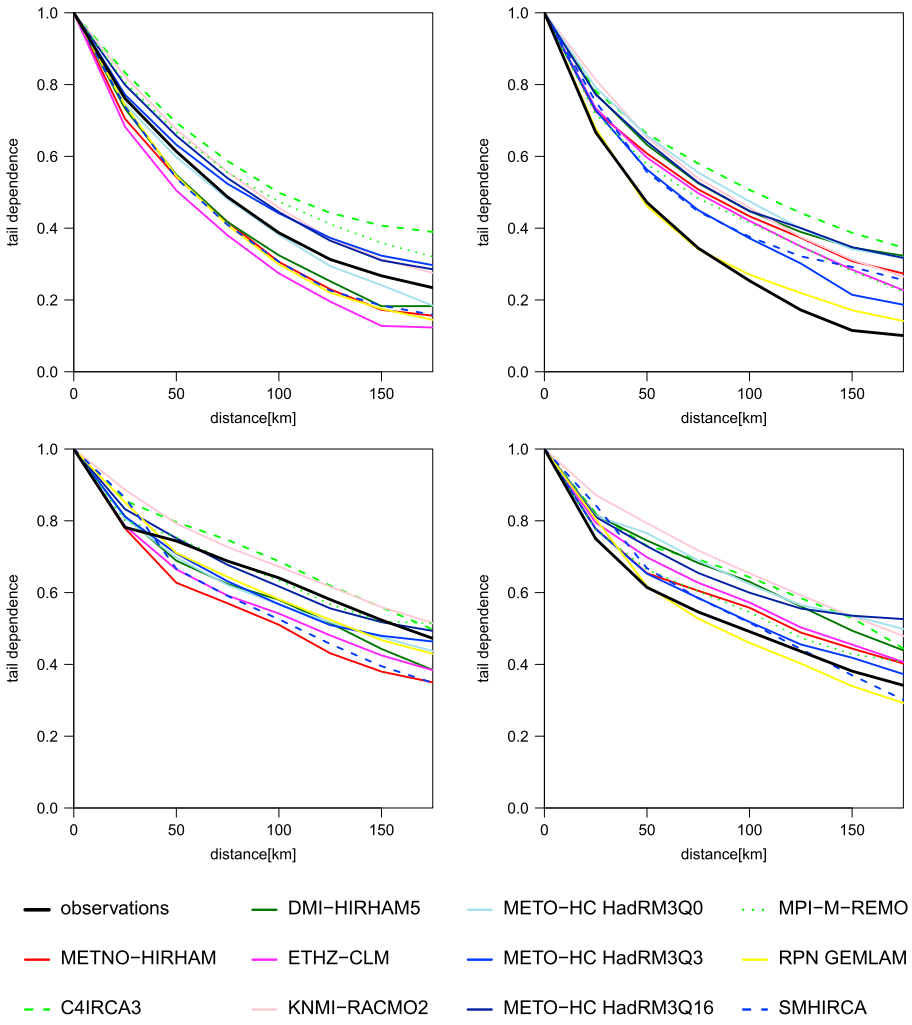

Figure 10. Upper tail dependence as a function of distance (for definition, see text). Shown for the (first row) west-east and (second row) north-south directions, for (left column) winter and (right column) summer.

Table 7. Decay Length of Rank Correlation in Kilometers (for Definition, See Text) for Winter[a]

RCM	West-East	North-South
Observations	205 (99,324)	232 (125,432)
METNO-HIRHAM	194 (109,331)	228 (133,391)
C4IRCA3	322 (136,598)	218 (77,491)
DMI-HIRHAM5	157 (89,275)	232 (133,299)
ETHZ-CLM	169 (58,709)	355 (143,498)
KNMI-RACMO2	195 (121,505)	288 (105,517)
METO-HC_HadRM3Q0	170 (106,395)	377 (166,795)
METO-HC_HadRM3Q16	242 (113,427)	371 (161,866)
METO-HC_HadRM3Q3	310 (138,547)	376 (165,806)
MPI-M-REMO	208 (119,324)	228 (104,440)
RPN_GEMLAM	217 (95,260)	250 (163,437)
SMHIRCA	174 (117,418)	239 (165,396)

[a]Corresponding 95% confidence intervals are shown in parentheses.

should be interpreted carefully. All RCMs appear to underestimate these quantiles, and for almost all of them the underestimation is significant. Thus, area-aggregated extreme precipitation is too low in RCMs.

As expected for RCMs [*Maraun et al.*, 2010], the 95% quantiles of the univariate margins are underestimated (see Figures S25 to S28 in the supporting information), indicating a major contribution to the misrepresentation of area-aggregated precipitation.

To disentangle the effects of misrepresented spatial dependence and univariate marginal distributions on the underestimation, we have reestimated the aggregate quantiles for each RCM: first using each RCM's own dependence

Table 8. Same as Table 7 but for Summer[a]

RCM	West-East	North-South
Observations	193 (99,331)	203 (97,333)
METNO-HIRHAM	247 (123,416)	177 (121,302)
C4IRCA3	255 (146,580)	182 (75,325)
DMI-HIRHAM5	226 (120,479)	211 (91,341)
ETHZ-CLM	225 (118,310)	320 (145,380)
KNMI-RACMO2	266 (126,404)	274 (143,541)
METO-HC_HadRM3Q0	199 (117,378)	379 (147,871)
METO-HC_HadRM3Q16	155 (120,362)	355 (147,742)
METO-HC_HadRM3Q3	194 (124,434)	410 (160,944)
MPI-M-REMO	189 (116,319)	256 (128,428)
RPN_GEMLAM	192 (96,266)	219 (154,332)
SMHIRCA	239 (121,481)	251 (157,345)

[a]Corresponding 95% confidence intervals are shown in parentheses.

structure but the (correct) margins of the observations and second the other way around, i.e., using the RCM's own margins but the (correct) observed dependence structure.

The misspecified RCM spatial dependence together with the correct margins represents the aggregate quantiles fairly well (Figures 11 and 12, bottom row), whereas the misspecified RCM margins with the correct spatial dependence strongly underestimates the aggregate quantiles (see Figures S17 and S18 in the supporting information). In fact, the latter result basically resembles the pure RCM results, where both dependence and margins are taken from the RCMs. Thus, the spatial dependence is misrepresented in RCMs, but the crucial contribution to the misrepresentation of spatially aggregated amounts is the wrong univariate marginal distribution.

In principle, problems in the representation of RCM-simulated marginal distributions can be mitigated by bias correction [*Maraun et al.*, 2010]. But—even though the systematic misplacement of precipitation is to some extent correctable [*Widmann et al.*, 2003]—the spatial extent of actual events can hardly be bias corrected. The case shown in the bottom row of Figures 11 and 12, where marginal distributions are taken from observations and spatial dependence from the RCM, thus corresponds to a bias-corrected RCM.

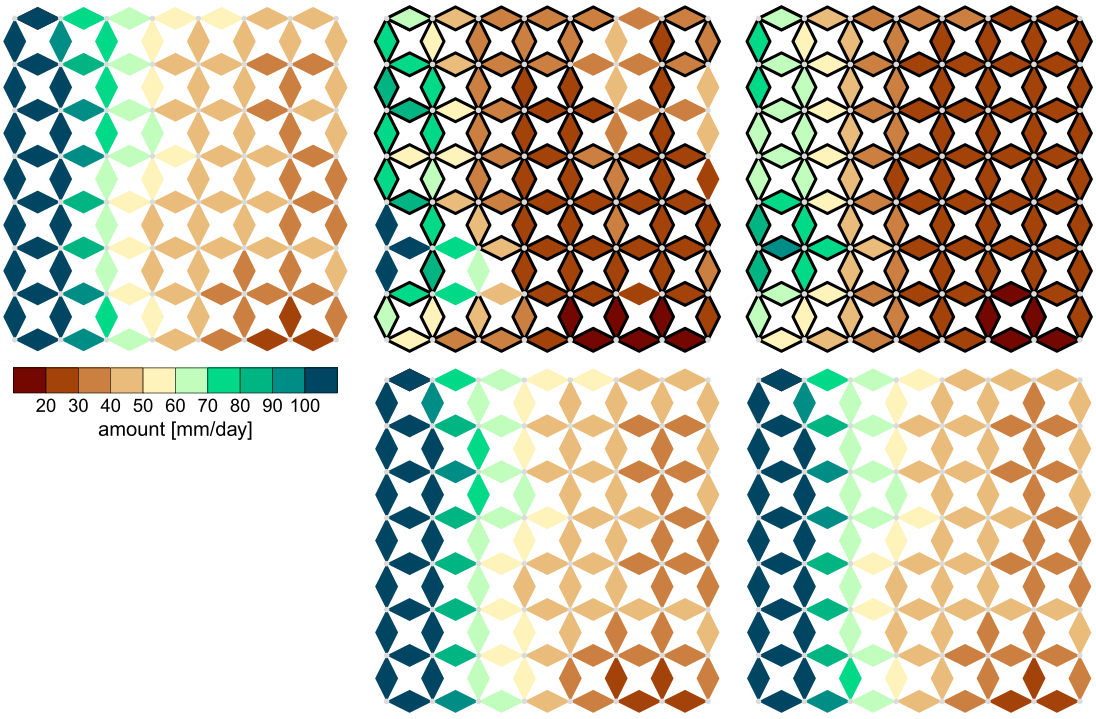

Figure 11. Area-aggregated 95% quantiles of daily winter precipitation. Shown are (left) observations, (middle) METNO-HIRHAM, and (right) C4IRCA3. For models: (top row) RCM margins and dependence (black (grey) diamonds: significantly lower (higher) than observations); bottom row: observed margins and RCM dependence (see text for details).

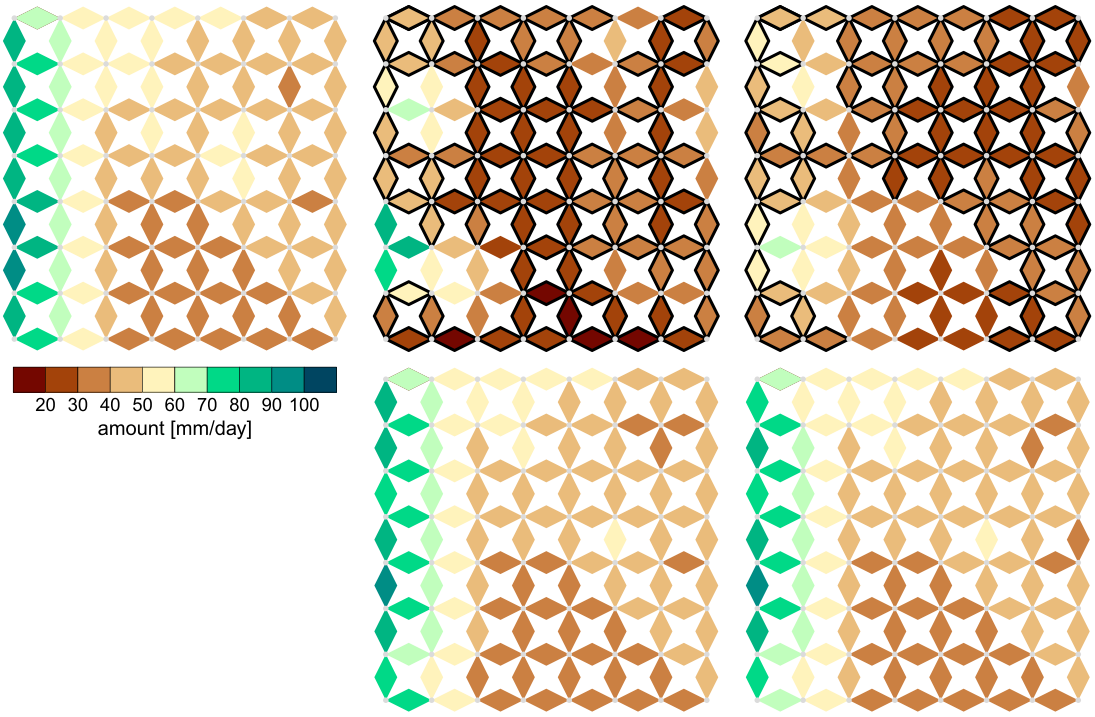

Figure 12. As in Figure 11 but for summer.

6. Conclusions

We have investigated how well a suite of regional climate models from the ENSEMBLES project [*van der Linden and Mitchell*, 2009] represents the residual spatial dependence of daily precipitation amounts. In the current work, we only considered the amount process of precipitation. To fully evaluate the performance of RCMs to simulate spatial precipitation fields, one additionally has to account for the occurrence process. As study area, we considered a 200 km × 200 km region in south central Norway. The RCMs were driven by ERA-40 boundary conditions at a horizontal resolution of approximately 25 km × 25 km.

We quantified the residual spatial dependence by spatial rank correlations for mean precipitation and tail dependence coefficients for extreme precipitation. To model the residual spatial dependence, we employed pair-copula constructions. These are capable of modeling the dependence structure, independently of the marginal distributions, but additionally allow one to easily model extended spatial fields and potential anisotropies in both mean and extreme precipitation. In contrast to purely empirical estimates, in particular of tail dependence, PCCs reduce uncertainties as well as biases and directly provide uncertainty estimates. As we did not consider the occurrence of precipitation, all results for a given copula of the PCC are based on days where all involved grid boxes showed precipitation.

The study area is dominated by fjords and high mountains to the west, a high plateau in the north, and hills toward the southeast. The main mountain ridges and valleys all run in the north-south direction. The topography strongly controls the residual spatial dependence of observed daily precipitation: first of all, deep valleys and fjords tend to separate precipitation events, in particular extremes, in space. Mean precipitation exhibits a strong and highly anisotropic spatial dependence; it decays slowly and almost linearly in the north-south direction, indicating an important influence of orographic lifting of the westerly flow on precipitation. For the same reason, the west-east decay is much faster—precipitation in the west of the domain is independent of that in the east. The north-south decay is similar in both winter and summer, yet the west-east decay is faster during winter than summer. Reasons for this slightly unexpected behavior are (1) that precipitation in southern Norway even in summer is not dominated by small-scale convective events and (2) that the westerly flow is not so dominant during summer, reducing the fraction of orographic precipitation and allowing for extended rainfall approaching from the south and east. The tail dependence is similar, with some interesting differences: the decay in the west-east direction in winter is slower than in

summer—in summer it is very similar to the mean behavior. This result indicates that convective events may be important at least for extreme events during summer. The decay of tail dependence in the north-south direction is overall very similar to the mean behavior.

The selected RCMs reproduce the overall dependence realistically although with considerable discrepancies compared to observations: for mean precipitation, all models overestimate the dependence in the west-east direction, whereas the RCMs scatter evenly around the observed dependence in the north-south direction. For the tail, RCM performance depends on season and direction: during winter, the RCMs scatter evenly around the observed decay in the west-east direction but consistently overestimate the dependence in the north-south direction. During summer, the RCMs overestimate the dependence in the west-east direction but tend to underestimate the dependence in the north-south direction. The shape of the decay is in all cases correctly simulated.

The small-scale dependence patterns caused by the pronounced orography, however, are not very well simulated by many climate models. The spread in performance for mean precipitation is large: pattern correlations range from about 0.3 to over 0.9 (with slightly better performance during summer), pattern standard deviations range from 50% to almost 200% of the observed one during winter and from 20% to 100% during summer. For the representation of tail dependence patterns, the spread is lower, but the performance is also weaker: pattern correlations range from 0.2 to 0.6 during winter and from 0.4 to 0.7 during summer. Pattern standard deviations range from around 30% to 80% of the observed one during both seasons.

A limitation of this—and any—validation study is the limited observational data quality. Overall, the station network underlying the gridded data set is very dense, but in remote mountainous regions only few if any stations are operated. The gridded precipitation data for such regions are thus dominated by the interpolation of distant stations and thus do not provide local area average information, in particular, extreme precipitation. In our study, the results for the observational data on the central Hardangervidda plateau have to be interpreted very carefully.

The misrepresented residual spatial dependence might have important consequences for assessing hydrological impacts of climate change, in particular at small scales. Yet if interpreted at scales of a few grid boxes, the problem is of minor importance. The major contribution to the underestimation of high quantiles of area-aggregated extreme precipitation stems from the well-known underestimation of the univariate margins for individual grid boxes. Problems in the representation of RCM-simulated marginal distributions, however, can in principle be mitigated by bias correction [*Maraun et al.*, 2010]. Note, however, that bias correction should be applied only after demonstrating the skill of the RCM at the scale of interest [*Eden et al.*, 2012, 2014]. If the aim is to downscale to subgrid scales, stochastic bias correction should be applied [*Maraun*, 2013; *Wong et al.*, 2014]. After a proper bias correction, the remaining differences in area-aggregated quantiles are relatively small.

We have demonstrated the usefulness of PCCs in spatially modeling the precipitation amount process. We believe that the flexibility of PCCs in accounting for anisotropy and asymmetric tail behavior makes them a valuable tool in many fields of climatology such as statistical downscaling or the modeling of compound events.

Appendix A: Regular Vines and Inference

Regular (R-) vines [*Bedford and Cooke*, 2001, 2002] are graphical tools to build self-contained PCCs. An R-vine representation of a d-dimensional density is a collection of $d - 1$ trees or levels. The first tree has the d variables as nodes, and its edges represent the pairs it connects in the ground level. The nodes in the next tree are the edges from the first tree, and in general, the edges at level j become the nodes at level $j + 1$. Two nodes in tree $j + 1$ may be connected by an edge only if these nodes, seen as edges in tree j, share a common node. This is called the proximity condition. The edges are indexed with two sets, called the conditioned set and the conditioning set. The former includes the variables which belong to only one of the nodes, and the latter includes the ones that are common for both nodes, i.e., the difference and union of the variables, respectively. See Figure A1 for an example in six dimensions. T_1 is the first tree, connecting the original variables. Its edges denote the pair-copulas from the ground level of the decomposition, i.e., c_{12}, c_{23}, c_{25}, c_{45}, and c_{56}. In the next level, T_2, the five edges from T_1 are the nodes,

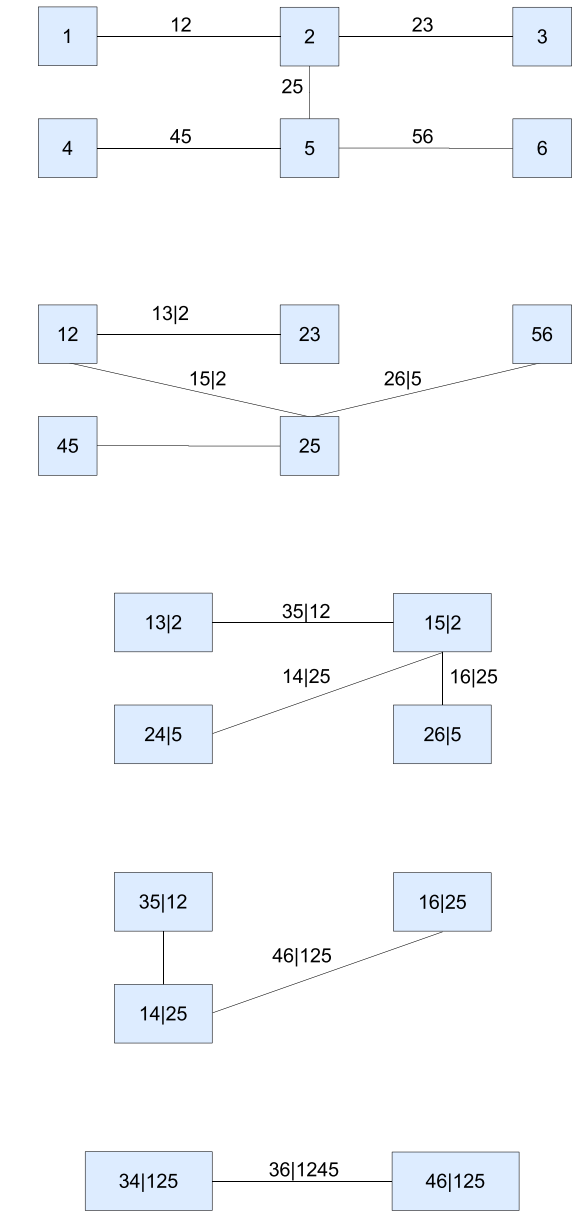

Figure A1. Levels of a regular vine for six variables.

and the new edges represent the second-level copulas $c_{13|2}$, $c_{15|2}$, $c_{24|5}$, $c_{26|5}$, etc. The two key results of *Bedford and Cooke* [2001, 2002] are that the collection of edges identifies all the pair-copulas needed for a legal PCC and that the construction is self-contained; that is, when looking higher up in the hierarchy one finds exactly one pair-copula to differentiate in order to obtain the conditional distribution one needs. For example, in Figure A1, the conditional distributions $F_{1|2}$ and $F_{3|2}$ are necessary for the copula $c_{13|2}$. These are computed by differentiating c_{12} and c_{23}, respectively, from the preceding level.

There are two special classes of R-vines, called canonical (C-) and drawable (D-) vines. These have been given most attention in applications. If there is a variable that governs all the others, the former is useful, whereas the latter is appropriate when there is a natural linear ordering of the variables. This is not the case for the application in this paper. We therefore focus on the more general class of R-vines.

The selection of a globally optimal R-vine is essentially an unsolved problem, but there are several useful strategies. The state-of-the-art is to select the structure that captures the most dependence in the lower levels. The reason for this is that the lower levels are the most influential for the total dependence and that the statistical model uncertainty increases through the structure. *Dißmann et al.* [2013] have proposed such an algorithm, based on Kendall's τ coefficients. This algorithm proceeds level by level, searching among the possible spanning trees for the one with the maximum sum of absolute values of τs. At the ground, all spanning trees are allowed. In the following levels, the possible trees are the ones that satisfy the proximity condition. Kendall's τ coefficients may be substituted by other measures of dependence, such as Spearman's ρ. The properties of Spearman's ρ are somewhat different from the ones of Kendall's τ, but the two essentially measure the same type of dependence. Other dependence measures include the tail dependence coefficients. If focus is on simultaneous extreme events, it would be a good idea to use these measures in the selection algorithm. Unfortunately, they are very difficult to estimate empirically [*Frahm et al.*, 2005], as opposed to Kendall's τ and Spearman's ρ. A different selection approach is to build the structure top down, minimizing the dependence in the upper levels. *Kurowicka* [2011] proposed such a procedure, based on partial correlation coefficients.

The algorithm of *Dißmann et al.* [2013] requires the simultaneous choice of pair-copula families as well as parameter estimation. At the ground level, the pairwise dependence measures may be estimated empirically from the data. From the second level on, the variables are conditional distributions. To be able

to compute them using the formula of *Joe* [1996], one needs estimates of the pair-copulas in the preceding level. The steps (ii) and (iii) of the inference procedure (see section 3) must therefore be part of the selection procedure.

A fast, stable, and convenient way of estimating the parameters is to use the stepwise semiparametric (SSP) estimator, which is designed for PCCs. The SSP estimator proceeds as follows. The univariate margins F_i are estimated in a separate step. The estimation of the copula parameters is based on the empirical distribution functions

$$F_{in}(y) = 1/(n+1)\sum_{k=1}^{n} I(x_{ki} \le y).$$

The log likelihood of an R-vine for n independent d-dimensional samples is given by

$$\sum_{k=1}^{n}\sum_{i=1}^{d} \log f_i(x_{ik}) + \sum_{k=1}^{n}\sum_{i=1}^{d-1}\sum_{e\in E_i} \log c_{j(e),k(e)|D(e)}\left(F\left(x_{j(e),k}|\boldsymbol{x}_{D(e),k}\right), F\left(x_{k(e),k}|\boldsymbol{x}_{D(e),k}\right)\right),$$

where E_i is the set of edges $e = j(e), k(e)|D(e)$ in level i, $j(e)$ and $k(e)$ being the conditioned set and $D(e)$ the conditioning set, and $\boldsymbol{x}_{D(e)}$ denotes the subvector of $\boldsymbol{x} = (x_1, \ldots, x_d)$ determined by the indices in $D(e)$. The SSP procedure consists in estimating the ground-level parameters first, by maximizing the corresponding part of the log likelihood function. Subsequently, the second-level parameters are estimated in a similar manner, inserting the ground-level estimates into the relevant terms of the log likelihood function. It continues in this stepwise manner until it reaches the top of the structure. For instance, for the three-dimensional example, the univariate margins F_1, F_2, and F_3 are estimated separately. Next, c_{13} and c_{23} are estimated, by inserting the empirical margins F_{in}, $i = 1, 2, 3$, into the terms of the log likelihood function, corresponding to the ground level. The conditional distributions $F_{1|3}$ and $F_{2|3}$ are then estimated using the formula of *Joe* [1996] on the estimated copulas $\hat{c}_{13}$ and $\hat{c}_{23}$. Finally, $c_{12|3}$ is estimated based on the second-level terms of the log likelihood function. For more details, see *Hobæk Haff* [2013, 2012].

Appendix B: Temporal Independence

An initial study revealed that the daily series of wet days are weakly temporally dependent and that this dependence is well modeled by an AR(1) process for each series. In agreement with *Serinaldi* [2009], the resulting AR coefficients were, however, small and moreover very similar across grid boxes and data sets. We therefore choose to ignore the temporal dependence. Without correcting for autocorrelation, the resulting spatial dependence models are easier to interpret, since they are related directly to the amounts of precipitation and not to the residuals from a time series model.

In a general model including temporal dependence, the measures we are interested in, such as Spearman rank correlation $\rho_S(X_{i,t}, X_{j,t})$ and the coefficients of upper tail dependence $\lambda_U(X_{i,t}, X_{j,t})$ between two neighboring grid boxes i and j, would in general be time-varying. Our assumption of temporal independence, however, enforces these measures to be constant in time. Thus, in order to justify the choice of temporal independence, we have to assess whether these measures are indeed reasonably constant in time. To this end, we have estimated the time-varying Spearman's ρ and upper tail dependence coefficients of three different pairs of grid boxes during each of the two seasons for the observations and the simulations by METNO-HIRHAM and DMI-HIRHAM5. They are computed in a moving window of 500 days that is shifted ahead 50 days at a time. Moreover, we have estimated the measures parametrically, based on the bivariate copula that fits each pair best. Variations in time turned out to be small. In those cases where there appears to be a difference between the constant measures, the ordering between the time-varying estimates remains constant (see Figures S29 and S30 in the supporting information for details). In other words, if our approach favors one RCM above another based on the constant measures, the same RCM would be preferred based on the corresponding time-varying measures. We conclude that the above statistical model with serial independence is sensible and adequate for our purpose.

Acknowledgments
This work is sponsored by Statistics for Innovation, (sfi)^2. We would like to thank the referees and Associate Editor for their help in improving this paper with their good comments and suggestions. The observed precipitations come from the Climatology Division of the Norwegian Meteorological Institute. These data may be obtained from Ole Einar Tveito (ole.einar.tveito@met.no). The remaining data supporting this paper may be obtained from Ingrid Hobæk Haff (ingrid@nr.no).

References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009), Pair-copula constructions of multiple dependence, *Insur. Math. Econ.*, *44*(2), 182–198.
Ailliot, P., C. Thompson, and P. Thomson (2009), Space-time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions, *Appl. Stat.*, *58*, 405–426.
Bedford, T., and R. Cooke (2001), Probabilistic density decomposition for conditionally dependent random variables modeled by vines, *Ann. Math. Artif. Intell.*, *32*, 245–268.
Bedford, T., and R. Cooke (2002), Vines—A new graphical model for dependent random variables, *Ann. Stat.*, *30*(4), 1031–1068.
Berg, D. (2009), Copula goodness-of-fit testing: An overview and power comparison, *Eur. J. Finan.*, *15*, 675–701.
Böhm, U., M. Kücken, W. Ahrens, A. Block, D. Hauffe, K. Keuler, B. Rockel, and A. Will (2006), CLM-the climate version of LM: Brief description and long-term applications, *COSMO Newsl.*, *6*, 225–235.
Bronstert, A., V. Kolokotronis, D. Schwandt, and H. Straub (2007), Comparison and evaluation of regional climate scenarios for hydrological impact analysis: General scheme and application example, *Int. J. Climatol.*, *27*(12), 1579–1594.
Casati, B., G. Ross, and D. Stephenson (2004), A new intensity-scale approach for the verification of spatial precipitation forecasts, *Meteorol. Appl.*, *11*, 141–154.
Chandler, R. E., and H. S. Wheater (2002), Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland, *Water Resour. Res.*, *38*(10), 1192, doi:10.1029/2001WR000906.
Christensen, J., O. Christensen, P. Lopez, E. van Meijgaard, and M. Botzet (1996), The HIRHAM4 regional atmospheric climate model, *Tech. Rep. Sci. Rep. 96-4*, Dan. Meteorol. Inst., Copenhagen, Den.
Christensen, J. H., and O. B. Christensen (2007), A summary of the PRUDENCE model projections of changes in European climate by the end of this century, *Clim. Change*, *81*, 7–30.
Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values, Springer Ser. in Stat.*, Springer, London.
Collins, M., B. Booth, G. Harris, J. Murphy, D. Sexton, and M. Webb (2006), To quantifying uncertainty in transient climate change, *Clim. Dyn.*, *27*, 127–147.
Davison, A. C., S. Padoan, and M. ribatet (2012), Statistical modeling of spatial extremes, *Stat. Sci.*, *27*(2), 161–186.
Deser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn (2011), Regional climate models add value to global model data: A review and selected examples, *Bull. Am. Meteorol. Soc.*, *92*, 1181–1192.
Dißmann, J., E. Brechmann, C. Czado, and K. Kurowicka (2013), Selecting and estimating regular vine copulae and application to financial returns, *Comput. Stat. Data Anal.*, *59*, 52–69.
Diggle, P., and P. Ribeiro (2007), *Model-Based Geostatistics, Springer Ser. in Stat.*, Springer, New York.
Eden, J., M. Widmann, D. Grawe, and S. Rast (2012), Skill, correction, and downscaling of GCM-simulated precipitation, *J. Clim.*, *25*, 3970–3984.
Eden, J., M. Widmann, D. Maraun, and M. Vrac (2014), Comparison of GCM- and RCM-simulated precipitation following stochastic postprocessing, *J. Geophys. Res. Atmos.*, *119*, 11,040–11,053, doi:10.1002/2014JD021732.
Fowler, H. J., S. Blenkinsop, and C. Tebaldi (2007), Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling, *Int. J. Climatol.*, *27*, 1547–1578.
Frahm, G., M. Junker, and R. Schmidt (2005), Estimating the tail-dependence coefficient: Properties and pitfalls, *Insur. Math. Econ.*, *37*, 80–100.
Frei, C., J. H. Christensen, M. Deque, D. Jacob, R. G. Jones, and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, *108*(D3), 4124, doi:10.1029/2002JD002287.
Frei, C., R. Schöll, S. Fukutome, J. Schmidli, and P. L. Vidale (2006), Future change of precipitation extremes in Europe: An intercomparison of scenarios from regional climate models, *J. Geophys. Res.*, *111*, D06105, doi:10.1029/2005JD005965.
Genest, C., and A.-C. Favre (2007), Everything you always wanted to know about copula modeling but were afraid to ask, *J. Hydrol. Eng.*, *12*(4), 347–368.
Genest, C., B. Rémillard, and D. Beaudoin (2009), Goodness-of-fit tests for copulas: A review and power study, *Insur. Math. Econ.*, *44*, 199–213.
Gilleland, E., D. Ahijevych, B. Brown, B. Casati, and E. Ebert (2009), Intercomparison of spatial forecast verification methods, *Weather Forecasting*, *24*, 1416–1430.
Gilleland, E., D. Ahijevych, B. Brown, and E. Ebert (2010), Verifying forecasts spatially, *Bull. Am. Meteorol. Soc.*, *91*, 1365–1373.
Giorgi, F., C. Jones, and G. Asrar (2009), Addressing climate information needs at the regional level: The CORDEX framework, *WMO Bull.*, *58*(3), 175–183.
Goodess, C., et al. (2010), An intercomparison of statistical downscaling methods for Europe and European regions—Assessing their performance with respect to extreme weather events and the implications for climate change applications, *Tech. Rep.*, Clim. Res. Unit, Sch. of Environ. Sci., Univ. of East Anglia.
Grønneberg, S. (2011), The copula information criterion and its implications for the maximum pseudo-likelihood estimator, in *Dependence Modeling: Vine Copula Handbook*, edited by D. Kurowicka and H. Joe, chap. 6, pp. 113–138, World Sci., Singapore.
Grotch, S. L., and M. C. MacCracken (1991), The use of general circulation models to predict regional climate change, *J. Clim.*, *4*, 286–303.
Haugen, J., and H. Haakenstad (2006), Validation of HIRHAM version 2 with 50 km and 25 km resolution, *RegClim. Gen. Tech. Rep. 9*, Norw. Meteorol. Inst., Oslo, Norway.
Haylock, M. R., G. C. Gawley, C. Harpham, R. L. Wilby, and C. M. Goodess (2006), Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios, *Int. J. Climatol.*, *26*(10), 1397–1415.
Hobæk Haff, I. (2012), Comparison of estimators for pair-copula constructions, *J. Multivariate Anal.*, *110*, 91–105.
Hobæk Haff, I. (2013), Parameter estimation for pair-copula constructions, *Bernoulli*, *19*, 462–491.
Jacob, D. (2001), A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin, *Meteorol. Atmos. Phys.*, *77*(1–4), 61–63.
Jacob, D., et al. (2007), An inter-comparison of regional climate models for Europe: Model performance in present-day climate, *Clim. Change*, *81*, 31–52.
Joe, H. (1996), Families of m-variate distributions with given margins and $m(m-1)/2$ dependence parameters, in *Distributions with Fixed Marginals and Related Topics*, edited by L. Rüschendorf, B. Schweizer, and M. D. Taylor, pp. 120–141, Inst. Math. Stat., Hayward, Calif.
Joe, H. (1997), *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.
Jansson, A., O. E. Tveito, P. Pirinen, and M. Scharling (2007), NORDGRID—A preliminary investigation on the potential for creation of a joint Nordic gridded climate dataset, *Met. Rep. 03/2007 Climate*, p. 48, Norw. Meteorol. Inst., Oslo, Norway.

Kjellström, E., L. Bärring, S. Gollvik, U. Hansson, C. Jones, and P. Samuelsson (2005), A 140-year simulation of European climate with the new version of the Rossby Centre Regional Atmospheric Climate model (RCA3), *Tech. Rep.*, SMHI, Norrköping, Sweden.
Kleiber, W., R. W. Katz, and B. Rajagopalan (2012), Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes, *Water Resour. Res.*, *48*, W01523, doi:10.1029/2011WR011105.
Kotlarski, S., et al. (2014), Regional climate modelling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble, *Geosci. Model. Dev. Discuss.*, *7*, 217–293.
Kurowicka, D. (2011), Optimal truncation of vines, in *Dependence Modeling: Vine Copula Handbook*, edited by D. Kurowicka and H. Joe, pp. 233–248, World Sci., Singapore.
Maraun, D. (2013), Bias correction, quantile mapping and downscaling: Revisiting the inflation issue, *J. Clim.*, *26*, 2137–2143.
Maraun, D., et al. (2010), Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, *48*, RG3003, doi:10.1029/2009RG000314.
Maraun, D., T. J. Osborn, and H. Rust (2012), The influence of synptic airflow on UK daily precipitation extremes. Part II: Regional climate model and E-OBS data validation, *Clim. Dyn.*, *39*, 287–301.
Maraun, D., M. Widmann, J. M. Gutiérrez, S. Kotlarski, R. E. Chandler, E. Hertig, J. Wibig, R. Huth, and R. A. I. Wilcke (2015), VALUE: A framework to validate downscaling approaches for climate change studies, *Earth's Future*, *3*, 1–14, doi:10.1002/2014EF000259.
Mearns, L., W. Gutowski, R. Jones, L.-Y. Leung, S. McGinnis, A. Nunes, and Y. Qian (2009), A regional climate change assessment program for North America, *Eos Trans. AGU*, *90*(36), 311–312.
Nelsen, R. (1999), *An Introduction to Copulas, Lect. Notes in Stat.*, vol. 139, Springer, New York.
Nikulin, G., E. Kjellström, U. Hansson, G. Strandberg, and A. Ullerstig (2011), Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations, *Tellus A*, *63*, 41–55.
Prein, A., A. Gobiet, M. Suklitsch, H. Truhetz, N. Awan, K. Keuler, and G. Georgievski (2013), Added value of convection permitting seasonal simulations, *Clim. Dyn.*, *41*, 2655–2677.
Rasmussen, S., J. Christensen, M. Drews, D. Gochis, and J. Refsgaard (2012), Spatial scale characteristics of precipitation simulated by regional climate models and the implications for hydrological modelling, *J. Hydrometeorol.*, *13*, 1817–1835.
Roberts, N., and H. Lean (2007), Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events, *Mon. Weather Rev.*, *136*, 78–97.
Rummukainen, M. (2010), State-of-the-art with regional climate models, *WIREs Clim. Change*, *1*, 82–96, doi:10.1002/wcc.8.
Sansó, B., and L. Guenni (2000), A nonstationary multisite model for rainfall, *J. Am. Stat. Assoc.*, *95*, 1089–1100.
Schindler, A., D. Maraun, and J. Luterbacher (2007), Validation of the present day annual cycle in heavy precipitation over the British Islands simulated by 14 RCMs, *J. Geophys. Res.*, *117*, D18107, doi:10.1029/2012JD017828.
Schölzel, C., and P. Friederichs (2008), Multivariate non-normally distributed random variables in climate research—Introduction to the copula approach, *Nonlinear Processes Geophys.*, *15*, 761–772.
Serinaldi, F. (2009), A multisite daily rainfall generator driven by bivariate copula-based mixed distributions, *J. Geophys. Res.*, *114*, D10103, doi:10.1029/2008JD01125.
Sklar, A. (1959), Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Stat. Univ. Paris*, *8*, 229–231.
Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, *106*, 7183–7192.
Uppala, S. M., et al. (2005), The ERA-40 Re-analysis, *Q. J. R. Meteorolog. Soc.*, *131*, 2961–3012.
van der Linden, P., and J. F. B. Mitchell (2009), ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project, *Tech. Rep.*, Met Off. Hadley Cent., Exeter, U. K.
van Meijgaard, E., L. H. van Ulft, W. J. van de Berg, F. C. Bosveld, B. J. J. M. van den Hurk, G. Lenderink, and A. P. Siebesma (2008), The KNMI regional atmospheric climate model RACMO version 2.1, *Tech. Rep. 302*, R. Dutch Meteorol. Inst., KNMI, De Bilt, Netherlands.
Wernli, H., M. Paulat, M. Hagen, and C. Frei (2008), SAL—A novel quality measure for the verification of quantitative precipitation forecasts, *Mon. Weather Rev.*, *136*, 4470–4487.
Widmann, M., C. S. Bretherton, and E. P. Salathe (2003), Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor, *J. Clim.*, *16*(5), 799–816.
Wilks, D. (1998), Multisite generalization of a daily precipitation generation model, *J. Hydrol.*, *210*, 178–191.
Wong, G., D. Maraun, M. Vrac, M. Widmann, J. Eden, and T. Kent (2014), Stochastic model output statistics for bias correcting and downscaling precipitation including extremes, *J. Clim.*, *27*, 6940–6959.
Yang, C., R. E. Chandler, and V. S. Isham (2005), Spatial-temporal rainfall simulation using generalized linear models, *Water Resour. Res.*, *41*, W11415, doi:10.1029/2004WR003739.