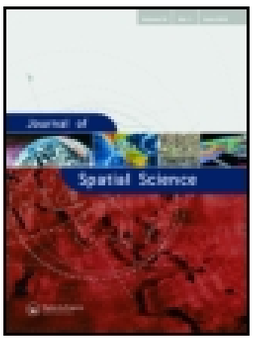

Journal of Spatial Science

ISSN: 1449-8596 (Print) 1836-5655 (Online) Journal homepage: https://www.tandfonline.com/loi/tjss20

Multivariate copula quantile mapping for bias correction of reanalysis air temperature data

Fakhereh Alidoost, Alfred Stein, Zhongbo Su & Ali Sharifi

To cite this article: Fakhereh Alidoost, Alfred Stein, Zhongbo Su & Ali Sharifi (2019): Multivariate copula quantile mapping for bias correction of reanalysis air temperature data, Journal of Spatial Science, DOI: 10.1080/14498596.2019.1601138

To link to this article: https://doi.org/10.1080/14498596.2019.1601138

JOURNAL OF SPATIAL SCIENCE
https://doi.org/10.1080/14498596.2019.1601138

Taylor & Francis
Taylor & Francis Group

Multivariate copula quantile mapping for bias correction of reanalysis air temperature data

Fakhereh Alidoost, Alfred Stein, Zhongbo Su and Ali Sharifi

ITC, University of Twente, Enschede, the Netherlands

ABSTRACT

Reanalysis data retrieved from the European Centre for Medium-range Weather Forecasts (ECMWF) are commonly used for hydrological studies. Their use requires bias correction, defined as the difference between reanalysis values and measurements. We propose three multivariate copula quantile mappings (MCQMs) to predict bias-corrected values at unvisited locations. We apply the methods to the Qazvin Plain, Iran, for daily air temperature retrieved from weather stations and the ECMWF archive. Results showed that MCQMs reduced bias by 46% as compared with classical quantile mapping. The study concludes that MCQMs are well able to represent the spatial and temporal variation of air temperature.

KEYWORDS

Bias correction; copula; conditional; mean temperature; data scarce

1. Introduction

Hydrological studies often require air temperature as a key variable to support water management in an irrigation network. A particular example concerns daily air temperature data that are needed in hydrological models for assessing near-real-time crop and irrigation water requirements and for improving crop water productivity in agricultural areas. At local scales (Sarma 2005), sparsely and irregularly distributed data from weather stations are a challenge to be used in hydrological models at unvisited locations in irrigation networks. To address this problem, additional spatially distributed data may be included as proxies. A typical example is gridded reanalysis weather data from the European Centre for Medium-Range Weather Forecasts (ECMWF). These data are available from the ERA-Interim data assimilation system, one of the most commonly used reanalysis archives over the last decades (Berrisford et al. 2009, Dee et al. 2011, Persson 2013). The coarse resolution of the models, the mutual dependence of weather parameters and the variability of these parameters in space and time are major sources of uncertainties when using these data (Dee *et al*. 2011, Durai and Bhradwaj 2014). One typical type of uncertainty is the presence of systematic differences between the gridded reanalysis data and the true temperature values at the ground surface. Due to the relatively coarse spatial resolution of the reanalysis data, a systematic bias may occur locally.

CONTACT Fakhereh Alidoost f.alidoost@utwente.nl
Supplemental data for this article can be accessed here.

In our paper, we aim to address bias in gridded reanalysis data. We consider weather station measurements as benchmarks, representing the true values of the near-surface air temperature. Hence, bias is defined as the difference between the reanalysis values and the measurements from weather stations (Hannah and Valdes 2001, Persson 2013). Reanalysis data are centred at grid cells, i.e. we consider an unvisited location at the centre of a grid cell characterised by a reanalysis value. Positions of weather stations, however, probably do not closely coincide with the grid centres and hence a bias correction method has to be applied.

Various bias correction methods have been proposed in the literature: quantile mapping (Ines and Hansen 2006), linear-scaling factor methods (Lenderink *et al*. 2007) and nonlinear methods (Lafon *et al*. 2013). The Gamma and empirical distributions have been used for bias correction of precipitation data and the Gaussian distribution for bias correction of air temperature data (Teutschbein and Seibert 2012, Lafon *et al*. 2013, Kum *et al*. 2014).

Recently, copula-based methods have been developed for deriving bias-corrected weather data (Vogl *et al*. 2012, Mao *et al*. 2015). A copula links univariate distributions with a multivariate distribution based upon Sklar's theorem (Sklar 1973, Nelsen 2006). So far, copula-based bias correction methods have mainly been applied to precipitation time series retrieved from regional climate models under the assumption of temporal stationarity. Laux *et al*. (2011) employed bivariate copulas to describe dependences between daily precipitation time series retrieved from a regional climate model and measurements at three locations where data are available. They fitted a bivariate copula to daily time series at one location, ignoring the temporal variation of the copula parameters as well as any spatial dependency. In addition, the estimation of distributions required to remove autocorrelation and heteroscedasticity, which may exist in any climate time series (Laux *et al*. 2011). Mao *et al*. (2015) investigated bias correction methods of daily precipitation data and showed that copula-based bias correction methods perform better than quantile mapping. To the best of our knowledge, copula-based bias correction for temperature data has not been developed yet.

The aim of our study is to obtain bias-corrected daily air temperature values at unvisited locations in a data-scarce area. These values have to be applied in a hydrological model to manage water resources in an irrigation network. We developed three multivariate copula quantile mappings for this purpose and used those to estimate the joint multivariate distributions of air temperature. Near-surface air temperature has a strong and predictable dependence on terrain according to elevation, land surface temperature and land cover (Hengl *et al*. 2012, Kilibarda *et al*. 2014, Parmentier *et al*. 2015). Our methods include elevation as a quantitative covariate. Previous studies focused on bias correction using vertical lapse rates, serving as a simple model of the atmosphere within a vertical column at equilibrium. This is useful especially in mountainous areas (Gao *et al*. 2012, 2017), but in agricultural areas the relation between elevation and near-surface air temperature does not follow the lapse rate. We investigated three new methods including two types of dependences: first, we used the dependence between air temperature and elevation at a single location, second we used the dependence between air temperatures at a single location and its nearest

neighbour, and third we integrated these dependences. The three methods are compared with classical quantile mapping.

The structure of this paper is as follows. Copulas and bias correction methods are presented in Section 2. The study area and data are introduced in Section 3. The results are discussed in Section 4, followed by the conclusion in Section 5.

2. Method

2.1. Multivariate copula quantile mappings

Multivariate copula quantile mapping (MCQM) is a d-dimensional quantile mapping method that relies on two conditional copula distributions (Gräler 2014, Verhoest *et al.* 2015). From two random variables X and Y over the same spatial domain, n samples $\{x_1, \ldots, x_n\}$ are obtained from weather station measurements and m samples $\{y_1, \ldots, y_m\}$ from reanalysis weather data. Bias b_i at location i is:

$$b_i = x_i - y_i. \tag{1}$$

The joint distribution function $H(X, Y)$ is written in terms of a copula as $C(U, V)$, where U and V are uniformly distributed random variables (Nelsen 2006). The empirical marginal probability u_i using the rank-order transformation equals:

$$u_i = \frac{\text{rank}(x_i)}{n+1}, i = 1, \quad \ldots, n. \tag{2}$$

A monotone cubic spline is fitted to the pairs (x_i, u_i) to obtain a continuous approximation of the marginal distribution F_X as $u_i = F_X(x_i)$ (Fritsch and Carlson 1980). The marginal distribution F_Y is estimated in a similar way. Use of an empirical distribution avoids estimating theoretical marginal distributions that might otherwise affect the estimation of copula parameters. Further note that the marginal distribution is assumed to be stationary (see Appendix).

The purpose of quantile mapping is to predict u_i at an unvisited location i. The inverse transformation of the marginal distribution F_X^{-1} provides the bias-corrected value $\hat{x}_i$:

$$\hat{x}_i = F_X^{-1}(\hat{u}_i) \tag{3}$$

where the notation $\widehat{\ }$ denotes that $\hat{x}$ and $\hat{u}$ are predicted values. To obtain $\hat{u}_i$, we develop three MCQMs including d-dimensional joint distributions where $2 \leq d \leq 3$.

MCQM-I: let Z be a covariate for X and Y, e.g. elevation. Then two conditional distributions $C(U|W = w_i)$ and $C(V|W = w_i)$ are obtained based upon bivariate joint distributions $C(U, W)$ and $C(V, W)$ describing non-spatial dependences, where the distributions can belong to different families and $w_i = F_Z(z_i)$. The marginal probability $\hat{u}_i$ is obtained using the inverse transformation of $C(U|W = w_i)$ as:

$$\hat{u}_i = C^{-1}(C(v_i|W = w_i)|W = w_i). \tag{4}$$

Distributions can be extended to higher dimensions if more than one covariate is available.

MCQM-II: we consider two bivariate joint distributions $C(U, U_{-i})$ and $C(V, V_{-i})$ that describe spatial dependences between air temperatures at location i and its nearest neighbour $-i$, and two conditional distributions $C(U|U_{-i} = u_{-i})$ and $C(V|V_{-i} = v_{-i})$ are based upon the joint distributions, where $u_{-i} = F_X(x_{-i})$ and $v_{-i} = F_Y(y_{-i})$. The marginal probability $\hat{u}_i$ is then obtained as:

$$\hat{u}_i = C^{-1}(C(v_i|V_{-i} = v_{-i})|U_{-i} = u_{-i}). \tag{5}$$

Distributions can be extended to higher dimensions using more than one neighbour where the number of observations is sufficient to obtain a correlogram that describes dependences in space (Oden 1984).

MCQM-III: the third method combines MCQM-I and MCQM-II. We consider two conditional distributions $C(U|U_{-i} = u_{-i},\ W = w_i)$ and $C(V|V_{-i} = v_{-i}, W = w_i)$ based upon trivariate joint distributions $C(U, U_{-i},\ W)$ and $C(V, V_{-i},\ W)$ describing non-spatial and spatial dependences. The marginal probability $\hat{u}_i$ is then obtained as:

$$\hat{u}_i = C^{-1}(C(v_i|V_{-i} = v_{-i}, W = w_i)|U_{-i} = u_{-i},\ W = w_i). \tag{6}$$

As for MCQM-II, distributions can be extended to higher dimensions. For MCQMs, it is assumed that the conditional probability of X conditioned on its covariate $F_X(X|.)$ is equal to the conditional probability of Y conditioned on that covariate $F_Y(Y|.)$.

2.2. Copula estimation in MCQMs

A bivariate copula describes the dependences between two variables. We used five copula families among several families available in the literature because other families lead to computational limitations (Gräler 2014): the Gaussian, Student's t, Clayton, Gumbel and Frank families (Nelsen 2003, Demarta and McNeil 2005). These families describe different tail dependences and have one parameter that specifies the correlation between variables (Table 1) (Joe 1993, Manner 2007). We apply maximum likelihood to estimate the parameter for each family using starting values obtained by Kendall's τ, being a measure for association between variables (Nelsen 2006). The best-fitting family

Table 1. Five families of copulas to describe dependences in MCQMs.

Index	Name	$C_\theta(u,v)$	Property index
1	Gaussian	$\emptyset_R(\emptyset^{-1}(u), \emptyset^{-1}(v)); R = \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix}$	1, 2, 6
2	Student's t	$t_{R,\vartheta}\left(t_\vartheta^{-1}(u), t_\vartheta^{-1}(v)\right);\ R = \begin{bmatrix} 1 & \theta \\ \theta & 1 \end{bmatrix};\ \vartheta = degree\ of\ freedom$	1, 2, 6
3	Clayton	$\left[max\left\{\left(u^\theta + v^\theta - 1\right), 0\right\}\right]^{\frac{-1}{\theta}}$	1, 2,4,5,6
4	Gumbel	$\exp\left(-\left[(-lnu)^\theta + (-lnv)^\theta\right]^{\frac{1}{\theta}}\right)$	1,2,3,6
5	Frank	$\frac{-1}{\theta}\ln\left(1 + \frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{e^{-\theta}-1}\right)$	1,2,6
1	Property	Permutation symmetry	
2		Symmetry about medians	
3		Extreme value copula	
4		Lower tail dependence	
5		Upper tail dependence	
6		Extendibility to multivariate copula	

is selected according to Akaike's Information Criteria (AIC) (Akaike 1974). The selection of families depends upon both the number of observations and the probabilistic nature of the dependence between variables.

In MCQM-III, a multivariate copula describes dependences between three or more variables. Following Aas *et al*. (2009), we estimate the conditional distribution $C(U|U_{-i} = u_{-i},\ W = w_i)$. To do so, the copula density $c(U, U_{-i},\ W)$ is first decomposed into bivariate copulas based upon a canonical vine or C-vine structure: $c(U, W),\ c(U, U_{-i})$ and $c(C(W|U), C(U_{-i}|U))$. Then, each bivariate copula is estimated in a similar way to the two-dimensional case. The copula density is the product of all bivariate copula densities in the structure. The conditional distribution $C(V|.)$ is estimated in a similar way.

We use the Cramér–von Mises statistic $S_n^{(B)}$ to obtain goodness-of-fit for Gaussian, Clayton, Gumbel and Frank families (Genest *et al*. 2009). The $S_n^{(B)}$ has practical limitations for the Student's *t* family. Hence, we use the White statistic for the Student's *t* copula (Huang and Prokhorov 2014). The *p* values are obtained for the tests using 100 bootstraps.

2.3. Quantile mapping

A comprehensive study carried out by Teutschbein and Seibert (2012) showed that quantile mapping (QM) performs best among the classical bias correction methods. QM is implemented as:

$$\hat{x}_i = F_X^{-1}(v_i). \tag{7}$$

QM assumes that there is a perfect dependence between variables, i.e. $\hat{u}_i = v_i$. It is sensitive to the number of quantile divisions when using an empirical marginal distribution. There are several names in the literature for this method, such as probability mapping, CDF matching and quantile-quantile mapping.

2.4. Comparison and evaluation of the bias correction methods

We compare MCQMs with quantile mapping using leave-*K*-out cross-validation (Lafon *et al*. 2013). To this end, the observations in *K* successive years on day *j* at station *i* are removed from the dataset and the bias-corrected values are predicted using the remainder of the observations. The mean absolute error $MAE_{i,j}$ equals:

$$MAE_{i,j} = \frac{1}{K}\sum_{k=1}^{K}\left|x_{i,j,k} - \hat{x}_{i,j,k}\right|. \tag{8}$$

We determine total mean absolute error *MAE* and spatial and temporal error scores, i.e. *SES* and *TES*, for *t* days and *n* stations as:

$$MAE = \frac{1}{t}\sum_{j=1}^{t}\left(\frac{1}{n}\sum_{i=1}^{n} MAE_{i,j}\right) \tag{9}$$

$$SES = \sum_{i=1}^{n}\left(rank\left(\frac{1}{t}\sum_{j=1}^{t} MAE_{i,j}\right)\right), \tag{10}$$

$$TES = \sum_{j=1}^{t} \left(rank\left(\frac{1}{n} \sum_{i=1}^{n} MAE_{i,j} \right) \right). \tag{11}$$

The lowest score indicates the best method (Durai and Bhradwaj 2014). In addition, we define correlations r_i and r_j, which indicate temporal and spatial dependences between measurements and bias-corrected values, respectively, as:

$$r_i = corr\left(\left\{ \hat{x}_{i,1}, \hat{x}_{i,j}, \ldots, \hat{x}_{i,t} \right\}, \left\{ x_{i,1}, x_{i,j}, \ldots, x_{i,t} \right\} \right), \tag{12}$$

$$r_j = corr\left(\left\{ \hat{x}_{1,j}, \hat{x}_{i,j} \ldots, \hat{x}_{n,j} \right\}, \left\{ x_{1,j}, x_{i,j}, \ldots, x_{n,j} \right\} \right). \tag{13}$$

Spatial and temporal correlation scores, i.e. *SCS* and *TCS*, are then obtained as:

$$SCS = \sum_{i=1}^{n} (rank(r_i)), \tag{14}$$

$$TCS = \sum_{j=1}^{t} (rank(r_j)). \tag{15}$$

The highest score indicates the best method.

All implementations were done with the statistical packages R using the libraries gstat (Pebesma 2004), copula (Kojadinovic and Yan 2010), VineCopula (Brechmann and Schepsmeier 2013) and the basic packages.

3. Case study

Our methods are applied to the Qazvin irrigation network located in the Qazvin Plain, Iran (Figure 1). Iran is a water-scarce country with limited rainfall in many places and irrigation is widely used to support the agricultural activities in the country. Qazvin Plain is one of the oldest and most advanced plains in Iran. The climate is arid with an average annual precipitation of about 327 mm, average temperature 14°C and mean annual reference evapotranspiration varying from 1300 to 1400 mm. The network has been part of a pilot study for a project aiming at irrigation management (Sharifi 2013). In the area, there are agricultural fields, dominated by wheat, barley, maize, orchards, urban areas and natural vegetation. The study area extends between 35.44° and 36.68° latitude (N) and 49.09° and 50.92° longitude (E) to include as many weather stations as possible, i.e. 24 stations (Figure 1, and Supplementary material, Table 1). The European Centre for Medium-Range Weather Forecasts (ECMWF) provides reanalysis weather data using the ERA-Interim data assimilation system (Berrisford et al. 2009, Dee et al. 2011, Persson 2013). ECMWF products are available at a wide range of spatial resolutions, e.g. regular and rotated lat/lon grids and reduced Gaussian grid. For the dissemination, air temperature is bi-linearly interpolated to a 0.125° lat/lon at three hourly intervals and 10×15 grid cells cover the study area (Figure 1). There are some advantages and disadvantages in every interpolation method (Persson 2013). Only if the interpolation point is in the centre of a grid cell does an interpolated value represent the mean value over the grid-box area. The disseminated data at resolutions coarser than 0.125° are interpolated values in 0.5°, 1.0° or 1.5°. This may have an undesired effect (Persson 2013). Further applications of the new copula-based methods in other case studies including air temperature at different resolutions should provide more insight on bias in the future. Daily mean air temperature is determined by averaging the minimum and maximum temperatures at each station in June from 2004 to 2014 (Figure 2) considering the

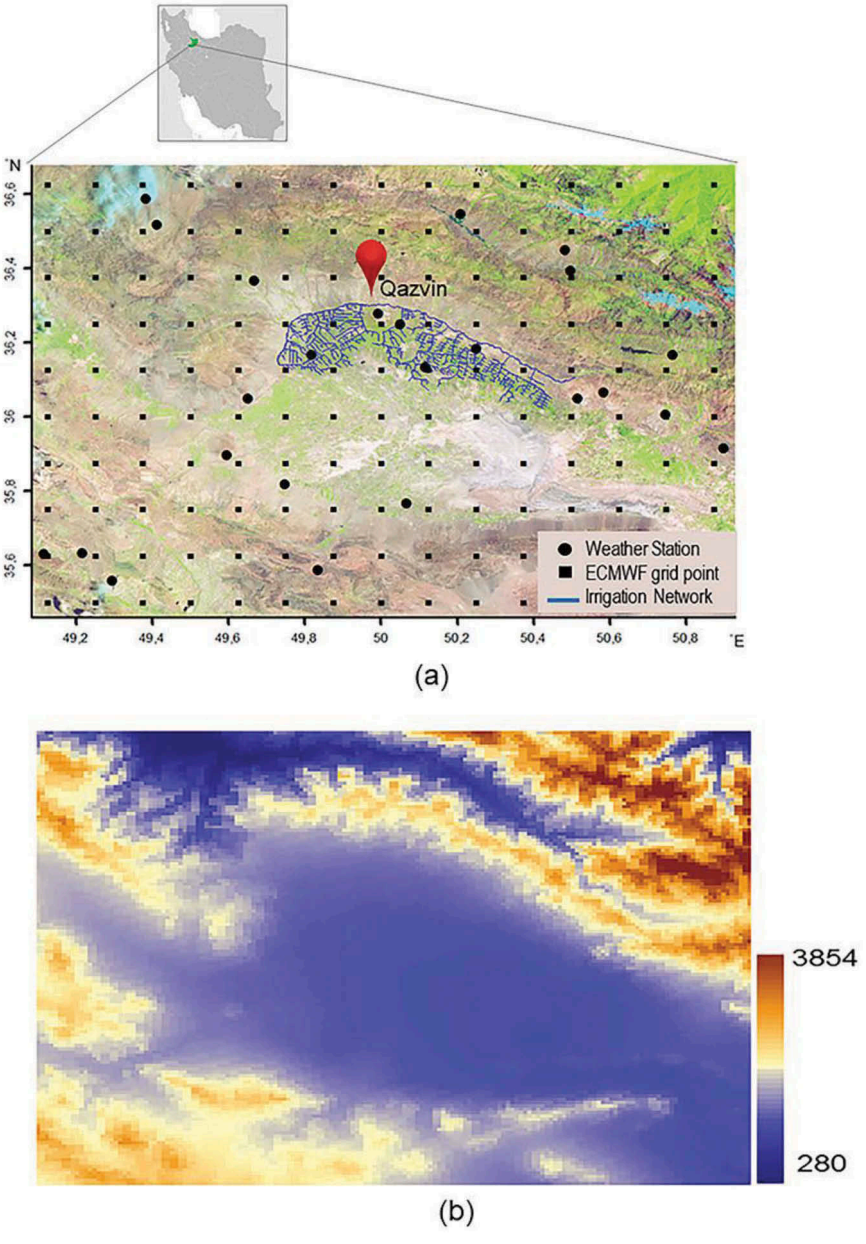

Figure 1. Study Area located in Qazvin Plain, Iran. (a) The area covers an irrigation network, 24 weather stations and a sample subset of 10 × 15 grid cells of the ECMWF dataset at 0.125° lat/lon distances. The background image is obtained by Landsat 8 RGB bands. (b) Elevations (m) are covariates for air temperature in MCQM-I. It was retrieved from MODIS products at a 1 km spatial resolution.

importance of this month in the crop calendar of the irrigation network: it is the end of winter crops and the beginning of summer crops, especially maize.

The measurements at the stations are assigned to the reanalysis values at the nearest grid cells (Figure 3). For instance, the measurements at stations number 4 and 11 are

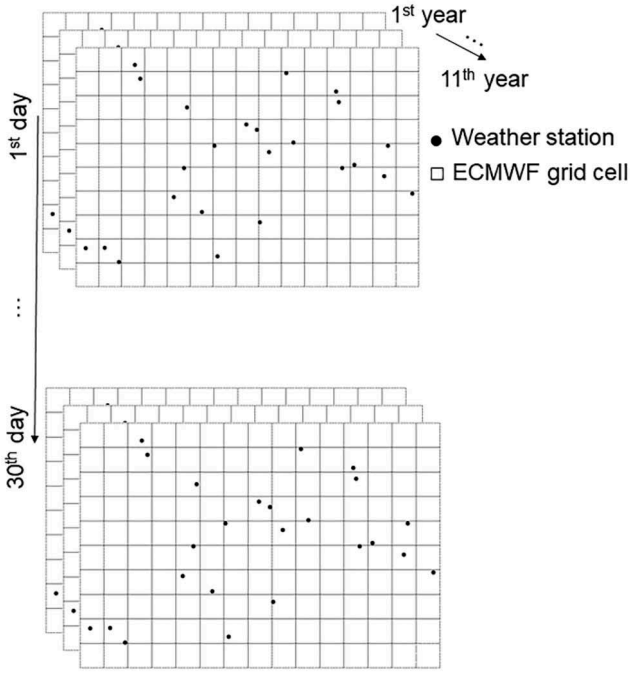

Figure 2. The data frame. The measurements from 24 weather stations and the reanalysis data from 150 ECMWF grid cells are available in June during 11 years from 2004 to 2014.

assigned to the reanalysis value at a grid cell. There are 150 grid cells × 11 years = 1650 reanalysis air temperatures and 24 stations × 11 years = 264 measurements on each day of June. Missing values in the measurements from weather stations may occur; their number differs between stations and days.

A comparison of the time series of the measurements and reanalysis values revealed systematic overestimation and underestimation (Figure 5 and Supplementary material, Figure 2). We noted that the time series at stations 13 and 21 have a lower correlation with the time series of reanalysis air temperature than the other stations (Figure 3(b)). The time series at those stations revealed that the quality of their measurements, in particular their accuracy, is low (Supplementary material, Figure 2). In addition, spatial correlations between the measurements and reanalysis air temperature are weak for most of the days in June 2014 (Figure 3(a)).

The NASA Land Processes Distributed Active Archive Center (LPDAAC) provides the Moderate Resolution Imaging Spectroradiometer (MODIS) products. The MOD03 product provides per-pixel digital elevation model values in a sequence of swath-based products at 5-minute increments. This resulted in 22,410 elevations at a 1 km spatial resolution (Figure 1).

This study focuses on obtaining the bias-corrected daily air temperature at unvisited locations on each day in June 2014. We applied the methods on each day of June 2014, separately. In other words, the methods are tested 30 times. The presented methods are generic and could be applied to a different time domain. In this area, the temperature is quite different in summer and winter. The performance of the methods for another

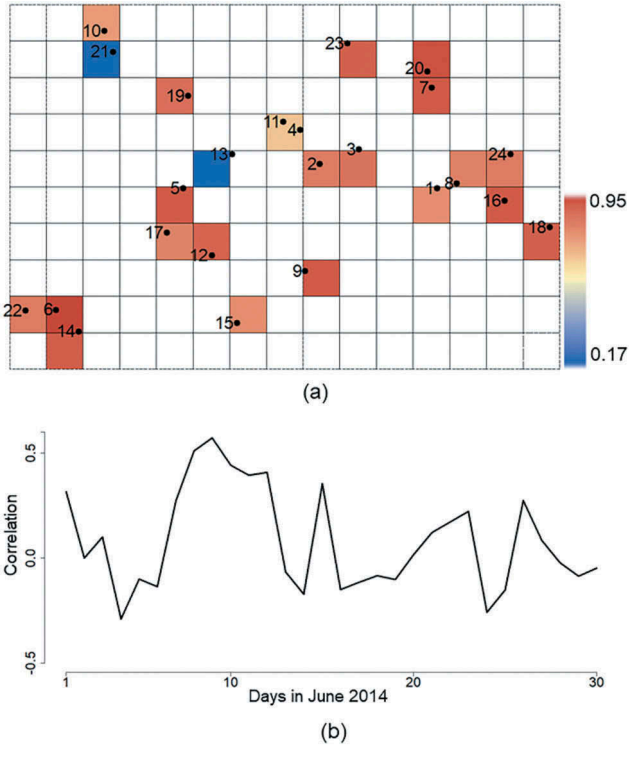

Figure 3. Correlations r_i and r_j that indicate temporal and spatial dependences between measurements and ECMWF ERA-interim reanalysis air temperature: (a) r_i at each weather station, (b) r_j on each day in June 2014.

month in winter needs to be further investigated. This, however, is outside the scope of the study.

The total mean absolute bias was equal to 3.6°C for all stations and all days. We did not consider predicting the bias-corrected air temperature at an unvisited location using the mean absolute bias since there is both spatial and temporal variation. We used the same elevations for 11 years assuming that elevation remains the same (Figure 1). Dependences are studied in a relatively small area and are thus unlikely to change spatially in a non-stationary way. The elevations are homogeneous, except for the north-eastern side of the study area (Figure 1).

4. Results and discussion

4.1. Marginal distributions and copulas

Marginal distributions and copulas are estimated for each day in June 2014, separately. The empirical marginal distribution on the first day is shown in Figure 4. The method to estimate empirical marginal distribution is not unique and a more generally applicable sensitivity analysis might help to show the effects of other methods on the results. For instance, we also used kernel density estimation and noticed that

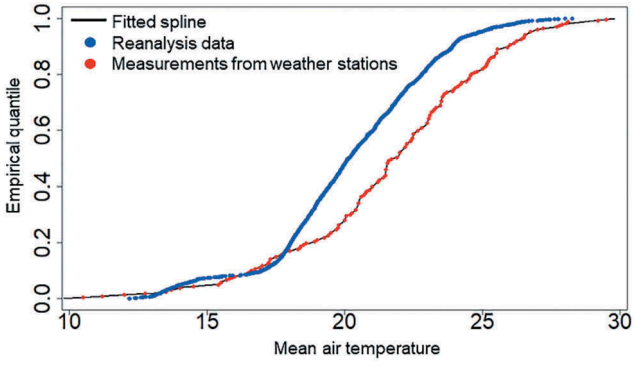

Figure 4. Empirical marginal probabilities on June 1st. A monotone cubic spline is fitted to obtain the marginal distribution function. Marginal distribution functions are estimated at each day of June, separately.

the final results of the bias correction methods changed only slightly (results not shown). To assess spatial stationarity, a trend surface was fitted to the measurements (Appendix). The β_1 parameter has p values in the range [0.02, 0.80] with mean value equal to 0.19, whereas the β_2 parameter has p values in the range [0.02, 0.99] with a mean value of 0.45. We were thus safe to assume spatial stationarity when estimating the marginal distributions.

The parameters of five copula families and the number of data for fitting purposes are listed in Table 2. We considered the elevation as the covariate in MCQM-I. We found that the best-fitting family was the Frank family for the joint distribution of the measurements and the elevation for all days and for the joint distribution of the reanalysis air temperature and the elevation for 18 days (Table 2). The p values of the Cramér–von Mises statistic $S_n^{(B)}$ were larger than 0.05 for all days, showing that the best-fitting family describes the dependences well (Table 2).

We considered spatial dependences in MCQM-II. The Student's t family dominates the dependences of the measurements for 14 days and the dependences of the reanalysis air temperature for 15 days (Table 2). The p values of The Cramér–von Mises statistic $S_n^{(B)}$ and the White statistic were larger than 0.1 except for the Gumbel family at five days, showing that the best-fitting family describes the spatial dependences well (Table 2). The p values were close to zero and the best-fitting family was the Gumbel family at days 1, 10, 17, 21 and 22. The low p values are related either to the limitation of the test or to the inflexibility of those five families. The p values were close to one for the Student's t.

For MCQM-III, the parameters of three bivariate copulas were estimated (Supplementary material, Table 2). Best-fitting families turned out to be non-Gaussian families for most of the days. The p values of the Cramér–von Mises statistic $S_n^{(B)}$ were larger than 0.2 for most of the days, showing that the best-fitting family describes the dependences well.

4.2. Bias-corrected values

In the following, we present the bias-corrected values at the first station for all days in June 2014 (Figure 5) and on 1 June 2014 for all grid cells in the study area (Figure 6).

Table 2. The *p* value and best-fitting families in MCQM-I and MCQM-II. The copula families are: N = Gaussian, T = Student's *t*, C = Clayton, G = Gumbel and F = Frank. Number of data denotes the number of marginal probabilities of each variable used for fitting purposes and equals the number of weather station measurements on each day in June during the years 2004 to 2014.

		MCQM-I				MCQM-II			
		$C(U,W)$		$C(V,W)$		$C(U,U_{-i})$		$C(V,V_{-i})$	
Day	Number of data	*p*-value	Best	*p*-value	Best	*p*-value	Best	*p*-value	Best
1	226	0.36	F	0.45	F	0.99	T	0.00	G
2	224	0.29	F	0.42	F	0.99	T	1.00	T
3	226	0.26	F	0.32	F	1.00	T	0.99	T
4	226	0.18	F	0.25	F	1.00	T	0.29	G
5	226	0.31	F	0.44	F	1.00	T	0.98	T
6	226	0.21	F	0.28	F	0.59	F	0.92	F
7	226	0.15	F	0.33	F	0.51	F	0.98	T
8	225	0.39	F	0.41	F	1.00	T	0.93	F
9	226	0.28	F	0.31	N	0.44	F	0.62	F
10	226	0.27	F	0.46	N	0.44	G	0.00	G
11	226	0.26	F	1.00	T	0.66	G	0.93	F
12	226	0.37	F	0.27	F	1.00	T	0.99	T
13	226	0.29	F	0.25	F	1.00	T	1.00	T
14	226	0.19	F	0.51	N	1.00	T	0.96	F
15	226	0.09	F	0.45	N	1.00	T	0.98	T
16	226	0.27	F	0.20	F	1.00	T	0.97	T
17	226	0.17	F	0.25	F	0.40	G	0.01	G
18	226	0.10	F	0.32	C	0.60	F	0.98	T
19	226	0.34	F	0.37	F	0.04	C	0.96	T
20	226	0.39	F	0.55	N	0.31	C	0.95	T
21	226	0.27	F	0.36	N	1.00	T	0.00	G
22	226	0.31	F	0.30	F	0.86	G	0.06	G
23	225	0.25	F	0.35	N	0.63	F	0.99	T
24	226	0.18	F	0.28	F	0.44	N	0.97	T
25	226	0.07	F	0.22	N	1.00	T	0.99	T
26	226	0.10	F	0.36	F	1.00	T	0.07	G
27	226	0.22	F	0.50	F	0.37	F	0.98	T
28	226	0.22	F	0.20	N	0.39	C	0.10	G
29	226	0.21	F	0.20	F	0.64	C	0.15	G
30	225	0.09	F	0.12	C	0.61	F	0.34	G

Detailed comparisons for all days and all grid cells are given in the supplementary material (Figures 2 and 3).

Time series of the bias-corrected values obtained by MCQM-I at the first station (Figure 5(a)) showed that MCQM-I successfully corrects for bias on most of the days, as well as the days with high extremes in comparison with time series obtained by QM (Figure 5(d)). Mean absolute bias was equal to 4.52°C at this station. Mean absolute error and mean absolute prediction error were equal to 1.46°C and 1.40°C for MCQM-I, whereas for QM they were equal to 2.84°C and 2.82°C, respectively. MCQM-I resulted in a heterogeneous map on 1 June 2014 (Figure 5(c)) in comparison with the map obtained by QM (Figure 5(f)). The spatial variation obtained by QM was similar to the spatial variation of the reanalysis air temperature, as shown in Figure 5(b), due to the assumption of a perfect dependence between variables in QM. The visual comparison of the spatial variation of the elevation (Figure 6) with the spatial variation of the map obtained by MCQM-I (Figure 6(c)) revealed that this method was able to describe the co-variability between the air temperature and the elevation. Mean absolute bias was equal to 2.83°C on this day. Mean absolute error and mean absolute prediction error were

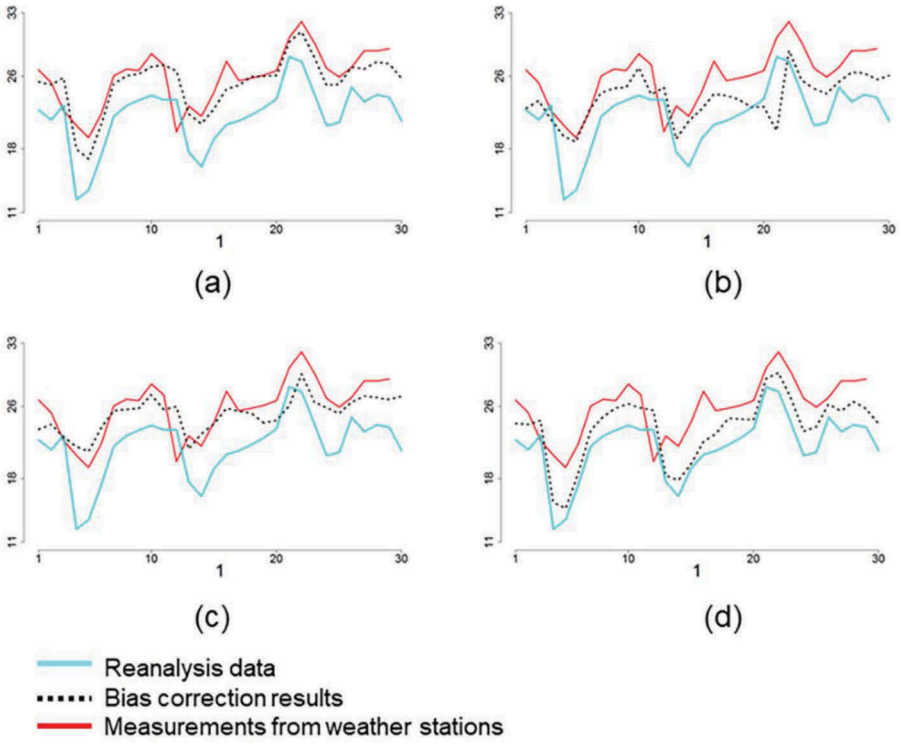

Figure 5. Time series of the daily mean air temperature obtained from: weather stations, ECMWF ERA-interim reanalysis data and bias correction methods at the first station in June 2014: (a) MCQM-I, (b) MCQM-II, (c) MCQM-III and (d) QM. The vertical axis is daily mean air temperature.

equal to 2.07°C and 1.55°C for MCQM-I, whereas for QM they were equal to 2.62°C and 1.93°C, respectively.

Time series of the bias-corrected values obtained by MCQM-II at the first station (Figure 6(b)) showed that MCQM-II successfully corrects for bias on most days except for days with extreme temperature, in comparison with time series obtained by MCQM-I and QM (Figure 6(a,d)). Mean absolute error and mean absolute prediction error were equal to 2.62°C and 2.67°C for MCQM-II at this station, whereas for QM they were equal to 2.84°C and 2.82°C, respectively. MCQM-II resulted in a more heterogeneous map on 1 June 2014 (Figure 6(d)) than the maps obtained by MCQM-I and QM (Figure 6(c,f)). Mean absolute error and mean absolute prediction error were equal to 2.66°C and 2.15°C for MCQM-II on this day, whereas for QM they were equal to 2.62°C and 1.93°C, respectively.

Time series of the bias-corrected values obtained by MCQM-III (Figure 5(c)) at the first station showed that MCQM-III performed better than MCQM-I (Figure 5(a)) in correcting for bias for most days except for the days with extremes. Mean absolute error and mean absolute prediction error were equal to 1.77°C and 1.68°C for MCQM-III at this station, whereas for QM they were equal to 2.84°C and 2.82°C, respectively. The Figure 6(e) showed that MCQM-III resulted in a heterogeneous map as compared with the maps obtained by other methods on 1 June 2014. Mean absolute error and mean absolute prediction error were equal to 2.36°C and 1.84°C for MCQM-III on this day, whereas for QM they were equal to 2.62°C and 1.93°C, respectively.

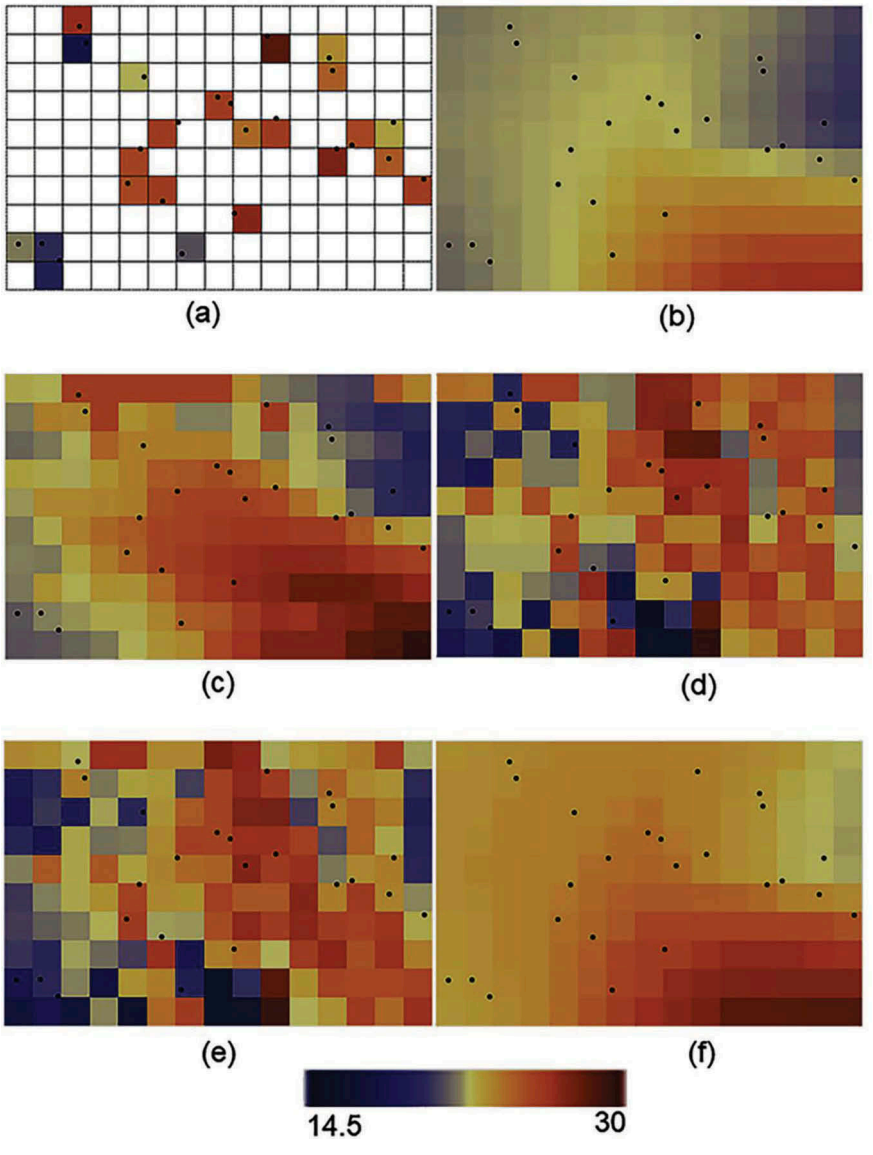

Figure 6. Daily mean air temperature obtained from: (a) weather stations, (b) ECMWF ERA-interim reanalysis data, and the bias correction methods on 1 June 2014, (c) MCQM-I, (d) MCQM-II, (e) MCQM-III and (f) QM. For experimentation in our study, a sample subset of 10 × 15 grid cells of the ECMWF dataset was selected at 0.125° lat/lon distances.

4.3. Evaluation and comparison

Leave-*K*-out cross-validation was carried out where *K* has values in the range 1 to 11, denoting the number of measurements from a weather station on one day for 11 years. MCQM-III was superior to MCQM-I, MCQM-II and QM as shown by *MAE* (Table 3). The average of absolute bias was equal to 3.6°C whereas *MAE* were slightly above 2°C. *SES* showed that MCQM-I resulted in more precise predictions in time, i.e. 30 days in June

(Table 3, second column), whereas *TES* indicated that MCQM-III resulted in more precise predictions in space (Table 3, third column). To extend the evaluation of the bias correction methods beyond the cross-validation, we can perform a random-split sampling validation in a well-monitored study area. This allows potentially more reliable uncertainty assessments. It is, however, beyond the scope of this paper. We treated the available measurements as benchmarks during the cross-validation. The horizontal distances, height differences and differences in land cover between the location of a station and the centre of a grid cell are associated with uncertainties.

MCQM-I resulted in stronger correlations in time as shown by *SCS* (Table 3, fourth column) and correlations r_i (Supplementary material, Figure 4(b)) whereas MCQM-III resulted in stronger correlations in space as shown by *TCS* (Table 3, last column) and correlations r_j (Supplementary material, Figure 4(a)). A comparison based upon *TCS* showed that the new methods perform better than QM in correcting reanalysis air temperature at unvisited locations in a data-scarce area. It further revealed that MCQM-II including only one nearest neighbour was unable to represent the spatial variation of daily air temperature. In order to do so, MCQM-II needs to be extended towards more nearest neighbours, allowing the use of a correlogram (see Section 2.1). A correlogram, however, faces the balancing issue between the number of spatial bins and the number of observations. The effect of the number of nearest neighbours on MCQM-II needs to be further investigated in a well-monitored area. Correlations r_j and r_i between the measurements and bias-corrected values obtained by QM were close to the correlations between the measurements and the reanalysis values (Supplementary material, Figure 4(a,b)). This was expected because of the assumption of a perfect dependence between variables in QM (Section 2.3).

The previous comparisons showed the performance of the methods based upon an individual criterion. To evaluate the performance based upon all criteria, we ranked the methods in each column of Table 3, where the lowest rank value denotes the best method (Table 4). Then, the overall score based upon the sum of the rank values showed that MCQM-I, MCQM-II and MCQM-III reduced bias by 58%, 16% and 63%, respectively as compared with QM (Table 4).

A practical advantage of MCQM-III is that it predicts the spatial variation of the bias-corrected air temperature maps in a data-scarce area (Supplementary material, Figure 3). The use of MCQM-III, however, is limited to the availability of the covariate at unvisited locations. We applied MCQMs to correct for bias in reanalysis air temperature, highlighting the potential of the methods for other weather data. Further comparison to other bias correction methods, e.g. triple collocation analysis (Stoffelen 1998), might help to assess the performance of MCQMs.

Table 3. Total mean absolute error (*MAE*), spatial error scores (*SES*), temporal error scores (*TES*), spatial correlation scores (*SCS*) and temporal correlation scores (*TCS*), obtained by the quantile mapping (QM), and the multivariate quantile mappings (MCQM-I, MCQM-II and MCQM-III). The underlined values denote the best method.

Method	*MAE*	*SES*	*TES*	*SCS*	*TCS*
MCQM-I	2.23	51	58	77	85
MCQM-II	2.40	63	88	46	65
MCQM-III	2.13	54	38	61	112
QM	2.68	72	116	56	38

Table 4. Overall score based upon Table 3. The methods are ranked based upon each criterion, i.e. each column in Table 3, where the lowest rank value denotes the best method. Then, an overall score based upon the sum of the rank values is obtained for each method. The underlined value denotes the best method.

Method	Rank value based on *MAE*	Rank value based on *SES*	Rank value based on *TES*	Rank value based on *SCS*	Rank value based on *TCS*	Overall score
MCQM-I	2	1	2	1	2	8
MCQM-II	3	3	3	4	3	16
MCQM-III	1	2	1	2	1	<u>7</u>
QM	4	4	4	3	4	19

5. Conclusions

This study addressed bias correction in ECMWF reanalysis air temperature using its covariates in a data-scarce area. We developed three multivariate copula quantile mappings to do so. We concluded the following:

- The new methods are beneficial for the local refinement of reanalysis weather data at grid cells without weather station measurements.
- The new methods are advantageous as they can treat co-variability, i.e. both weather data and covariates, and hence increase the precision of the mapping.

We see two ways to further extend the current study. First, we selected the number and type of covariates based upon our experience. A more general sensitivity analysis might help to show the effects of other covariates, e.g. land surface temperature and land cover. Second, it might be of interest to study the ability of the new methods to reproduce other statistical moments of the theoretical marginal distribution of air temperature. This could help to further model extremes in air temperature.

Acknowledgments

The authors acknowledge the European Centre for Medium-Range Weather Forecasts (ECMWF) for providing weather forecast data, the SAJ Consulting firm in Iran for providing weather station data, and Land Processes Distributed Active Archive Center (LPDAAC) of the US Geological Survey for MODIS elevation products.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Aas, K., *et al*., 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44 (2), 182–198.

Akaike, H., 1974. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19 (6), 716–723. doi:10.1109/TAC.1974.1100705

Berrisford, P., *et al*. 2009. The ERA-Interim archive (version 1.0), ERA Report Series: European Centre for Medium Range Weather Forecasts, Reading, UK.

Brechmann, E.C. and Schepsmeier, U., 2013. Modeling Dependence with C- and D-Vine Copulas: the R Package CDVine. *Journal of Statistical Software*, 52 (3), 1–27. doi:10.18637/jss.v052.i03

Dee, D.P., *et al*., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137 (656), 553–597. doi:10.1002/qj.828

Demarta, S. and McNeil, A.J., 2005. The *t* copula and related copulas. *International Statistical Review/Revue Internationale de Statistique*, 73 (1), 111–129.

Durai, V.R. and Bhradwaj, R., 2014. Evaluation of statistical bias correction methods for numerical weather prediction model forecasts of maximum and minimum temperatures. *Natural Hazards*, 73, 1229–1254. doi:10.1007/s11069-014-1136-1

Fritsch, F.N. and Carlson, R.E., 1980. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17, 238–246. doi:10.1137/0717021

Gao, L., Bernhardt, M., and Schulz, K., 2012. Elevation correction of ERA-Interim temperature data in complex terrain. *Hydrology and Earth System Sciences*, 16 (12), 4661–4673. doi:10.5194/hess-16-4661-2012

Gao, L., *et al*., 2017. Elevation correction of ERA-Interim temperature data in the Tibetan Plateau. *International Journal of Climatology*, 37, 3540–3552. doi:10.1002/joc.4935

Genest, C., Remillard, B., and Beaudoin, D., 2009. Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44 (2), 199–213.

Gräler, B., 2014. *Developing spatio-temporal copulas*. Doctoral. Westfälische Wilhelms-Universität Münster.

Hannah, E. and Valdes, P., 2001. Validation of ECMWF (re)analysis surface climate data, 1979-1998, for Greenland and implications for mass balance modelling of the Ice Sheet. *International Journal of Climatology*, 21 (2), 171–195. doi:10.1002/(ISSN)1097-0088

Hengl, T., *et al*., 2012. Spatio-temporal prediction of daily temperatures using time-series of MODIS LST images. *Theoretical and Applied Climatology*, 107, 265–277. doi:10.1007/s00704-011-0464-2

Huang, W. and Prokhorov, A., 2014. A Goodness-Of-Fit test for copulas. *Econometric Reviews*, 33 (7), 751–771. doi:10.1080/07474938.2012.690692

Ines, A.V.M. and Hansen, J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. *Agricultural and forest meteorology*, 138, 44–53. doi:10.1016/j.agrformet.2006.03.009

Joe, H., 1993. Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, 46, 262–282. doi:10.1006/jmva.1993.1061

Kilibarda, M., *et al*., 2014. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *Journal of Geophysical Research: Atmospheres*, 119, 2294–2313.

Kojadinovic, I. and Yan, J., 2010. Modeling Multivariate Distributions with Continuous Margins Using the copula R Package. *Journal of Statistical Software*, 34 (9), 1–20. doi:10.18637/jss.v034.i09

Kum, D., *et al*., 2014. Projecting Future Climate Change Scenarios Using Three Bias-Correction Methods. *Hindawi Publishing Corporation Advances in Meteorology*, 2014, 1–13. doi:10.1155/2014/704151

Lafon, T., *et al*., 2013. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. *International Journal of Climatology*, 33, 1367–1381. doi:10.1002/joc.3518

Laux, P., *et al*., 2011. Copula-based statistical refinement of precipitation in RCM simulations over complex terrain. *Hydrology and Earth System Sciences*, 15 (7), 2401–2419. doi:10.5194/hess-15-2401-2011

Lenderink, G., Buishand, A., and van Deursen, W., 2007. Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrology and Earth System Sciences*, 11 (3), 1145–1159. doi:10.5194/hess-11-1145-2007

Manner, H., 2007. *Estimation and Model Selection of Copulas with an Application to Exchange Rates*. Maastricht, MD: Universiteit Maastricht, Maastricht research school of Economics of TEchnology and ORganizations.

Mao, G., *et al*., 2015. Stochastic bias correction of dynamically downscaled precipitation fields for Germany through Copula-based integration of gridded observation data. *Hydrology and Earth System Sciences*, 19 (4), 1787–1806. doi:10.5194/hess-19-1787-2015

Nelsen, R., 2003. Properties and applications of copulas: A brief survey. *In*: J. Dhaene, N. Kolev, and P. Morettin eds. *Proceedings of the First Brazilian Conference on Statistical Modeling in Insurance and Finance*. Sao Paulo: University Press USP.

Nelsen, R.B., 2006. *An Introduction to Copulas*. Portland, USA: Springer.

Oden, N.L., 1984. Assessing the Significance of a Spatial Correlogram. *Geographical analysis*, 16, 1–16. doi:10.1111/j.1538-4632.1984.tb00796.x

Parmentier, B., *et al*., 2015. Using multi-timescale methods and satellite-derived land surface temperature for the interpolation of daily maximum air temperature in Oregon. *International Journal of Climatology*, 35, 3862–3878. doi:10.1002/joc.4251

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, 30 (7), 683–691. doi:10.1016/j.cageo.2004.03.012

Persson, A., 2013. *User guide to ECMWF forecast products*. Livelink 4320059. UK: European Centre for Medium Range Weather Forecasts, Reading.

Sarma, A.A.L.N., 2005. Scales of Climate. *In*: J.E. Oliver, ed. *Encyclopedia of World Climatology*. Dordrecht, Netherlands: Springer, 637–639.

Sharifi, M., 2013. *Development of planning and monitoring system supporting irrigation management in the Ghazvin irrigation network*. Tehran, Iran: SAJ Co.

Sklar, A., 1973. Random variables, joint distribution functions, and copulas. *Kybernetika*, 9 (6), 449–460.

Stoffelen, A., 1998. Toward the true near-surface wind speed: error modeling and calibration using triple collocation. *Journal of Geophysical Research*, 103 (C4), 7755–7766. doi:10.1029/97JC03180

Teutschbein, C. and Seibert, J., 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. *Journal of Hydrology*, 456–457, 12–29. doi:10.1016/j.jhydrol.2012.05.052

Verhoest, N.E.C., *et al*., 2015. Copula-Based Downscaling of Coarse-Scale Soil Moisture Observations With Implicit Bias Correction. *IEEE Transactions on geoscience and remote sensing*, 53 (6), 3507–3521. doi:10.1109/TGRS.2014.2378913

Vogl, S., *et al*., 2012. Copula-based assimilation of radar and gauge information to derive bias-corrected precipitation fields. *Hydrology and Earth System Sciences*, 16 (7), 2311–2328. doi:10.5194/hess-16-2311-2012

Appendix

To test for the assumption of second-order stationarity, we considered the null hypothesis H_0 as:

$$E[X_i] = \mu \tag{1}$$

where X_i is a random variable at location i and $E[]$ denotes the mathematical expectation. The alternative hypothesis H_1 is that there is a trend of degree one as:

$$E[X_i] = \beta_0 + \beta_1.x_i' + \beta_2.y_i', \tag{2}$$

where x_i' and y_i' are coordinates of location i and the β_j denote regression parameters. The parameters are estimated using a generalised linear model followed by their p values from a t test. We applied this trend to the measurements from 24 weather stations on each day of June 2014. The values of β_1 and β_2 were found to be not significantly different from zero, with their p values above 0.05 on most of the days (Supplementary material, Figure 1). On the six days (out of 30 days) when the p value was below 0.05, it was still above 0.01. Based on this evidence, and the limited effects of including non-stationarity, we felt confident in assuming stationarity.