MULTIVARIATE DISTRIBUTION CORRECTION OF CLIMATE MODEL OUTPUTS: A GENERALISATION OF QUANTILE MAPPING APPROACHES

Léonard Deckens, Sylvie Parey, Mathilde Grandjacques, D Dacunha-Castelle

MULTIVARIATE DISTRIBUTION CORRECTION OF CLIMATE MODEL OUTPUTS: A GENERALISATION OF QUANTILE MAPPING APPROACHES

Multivariate bias adjustment for climate change studies

Research article

L Dekens[1,2], S Parey[1], M Grandjacques[3], D Dacunha-Castelle[4]

1 : EDF Recherche et Développement site de Chatou, MFEE, France

2 : Ecole Normale Supérieure de Lyon et université Claude Bernard Lyon 1, France

3 : Lianes (Laboratoire d'Intelligence Artificielle pour les Nouvelles Energies), Institut LIST, CEA, Université Paris-Saclay, F-91120, Palaiseau, France

4 : Université Paris Sud, Orsay, France

Corresponding author : S Parey, EDF Recherche et Developpement 6 quai Watier 78401 Chatou France, sylvie.parey@edf.fr

Abstract: Climate change impact studies necessitate the estimation of climate variables evolution in the future. These are given by climate model simulations made under different greenhouse gas and aerosol emission scenarios agreed at the international level. However climate model outputs have biases, especially at the local scale, and need to be corrected against observations. Common bias-correction methods are distribution based and form the well-known quantile mapping approaches. This paper presents a generalization of such techniques to the consideration of multivariate distributions. This approach uses the basic lemma of Lévy-Rosenblatt which allows the transport of a distribution on another one, in every dimension. It needs convenient non parametric estimations of conditional repartitions. The approach is first tested in a controlled framework, by use of statistical simulations, then in the real setting of climate simulation, in the bivariate case. An important issue of these types of distribution corrections is the different kinds of hypotheses of stationarity over a long enough period: stationarity of the link between model and observations whatever the period or stationarity of the change between present and future for model and observations. This choice differentiates approaches like Quantile Mapping and CDFt for example in the univariate framework, and makes them more efficient, in the univariate as well as in the multivariate context, when the data to be corrected best verify the assumed hypothesis.

Keywords: climate change, statistics, multivariate distributions, bias adjustment

1 Introduction

Since climate change is now attested (IPCC, 2013), and mitigation still underway, adaptation has to be anticipated in parallel to mitigation. The first step in adaptation is an estimation of the possible consequences of climate change at the scale of human societies and their activities. These estimations are commonly done through impact studies, based for example on specific models run with climatic variables. Observed variables are used to represent current conditions,

while climate model outputs are used to project future conditions. Climate models are numerical tools based on physical representations of the dynamics of the components, atmosphere, ocean, ice or land surface, and of their interactions, through physical or biochemical processes. Although such tools are more and more sophisticated, including more and more detailed processes, their outputs may still differ significantly from the local observations commonly used by impact models. Therefore, bias correction and downscaling techniques have become an active area of research in the last decade or so. The approach here, like other quantile mapping approaches, can be used for bias adjustment or bias adjustment and downscaling depending on the spatial scale of the reference dataset. When used with local observations it aims at predicting, in statistical terms, local climate variables using more global data provided by a climate model working at a larger scale.

Statistical bias correction methods are widely used to correct the distribution of the climate model variables so that they match that of some local observations. Such techniques are commonly recommended for impact studies (Teutschbein and Seibert, 2012; Gudmundsson et al., 2012; Chen et al., 2013). The most used techniques are the so called quantile mapping approaches (Panofsky and Brier, 1958; Haddad and Rosenfeld, 1997; Wood et al. 2004; Déqué 2007; Piani et al. 2010), and their variants like CDFt (Michelangeli et al., 2009). A limitation of such techniques is however that the correction is applied independently to the different variables when more than one climatic variable is needed for an impact study, with the risk of degrading the consistency between them. Recent approaches have been proposed to tackle this caveat, and correct two variables, essentially temperature and rainfall, in a consistent way (Zhang and Georgakakos, 2012; Piani and Hearter, 2012; Li et al., 2014). Vrac and Friederichs, 2014 go further and propose an approach, based on the empirical copula (by reordering univariate bias-corrected variables), potentially able to tackle both the inter-variable and spatial consistencies. One issue with the approach is however that it can only reproduce the historical

temporal sequencing, which is an important limitation for climate impact studies. Cannon, 2016 suggests a methodology based on the correction of the marginal distributions again by quantile mapping with then an iterative scheme to push either the Pearson correlation dependence structure or the Spearman rank correlation dependence structure towards observed values.

The approach proposed and tested in this paper is a generalization of the quantile mapping techniques to the correction of multivariate distributions. The chosen setting is the typical problem faced with impact studies: over an historical period, time series of different climate variables are available from both climatic databases and climate model simulations, while for a future period, necessarily, only the climate model time series are available. Then, as climate model outputs have biases compared to the observations, the aim of the correction is to estimate for the future period, specific characteristics of time series at the desired location closer to that of the observations. Then, what is expected is not the precise sequencing in time of the variables, which is not an expected result of climate models, but rather characteristics as their distribution, which are quite invariant for time periods with adequate length, not too long to be able to neglect climate trends but not too short to be able to estimate characteristics like distributions. The methodology will be described in the fully multivariate context, considering p dimensions, but in practice, a dimension larger than 2 means much longer time series for the distribution estimations. This methodology uses as basic trick the transportation of a distribution on $\mathbb{R}^p$ onto another one fixed in advance. This is done by repeated applications of the lemma of Levy – Rosenblatt (Grandjacques 2015, Grandjacques et al. 2015). This approach allows to clarify which kind of stationarities are required and also gives a natural way of making clear which period lengths are concerned by all these approaches.

The theory underlying the methodology is presented in section 2, and section 3 explains the estimation choices made. Then, section 4 presents an application in a controlled framework, by use of bivariate Gaussian distributions, in order to evaluate when our bivariate bias-correction

is best justified. Bivariate Gaussian distributions are easy to handle, although climate variables are generally not normally distributed. The aim here is only to better understand the involved transformations. Finally, section 5 is devoted to an application to real climate variables, before coming to the conclusion and discussion in section 6.

2 Theoretical framework

2.1 The problem to be solved

As previously exposed, the aim here is to simultaneously bias correct different climatic variables according to available observations. We have then 2 time series with values in $\mathbb{R}^p$:

- a time series $(Y_t)_{t\in P0}$ given over the period $P_0=(1, \dots, n_0)$ only, and corresponding to the observations
- a time series (X_t) over a much larger period and given by a numerical climate model simulation for example.

The aim is then to obtain for a future period P, later than P_0, a projection of some characteristics of $(Y_t)_{t\in P}$, based on assumptions made about the link between X and Y. In order to make it easier in the following, we will suppose that P and P_0 are of same length and P corresponds to $P_k = \{kn_0+1, \dots, (k+1)n_0\}$.

The proposed methodology is based on the following assumptions.

Hypothesis H_1: *for each $k\in\mathbb{N}$, there is a process $X_{k,t}$, restriction of X_t when $t\in P_k$, which is stationary and weakly mixing. Similarly, it is supposed that there is a process $(Y_t)_{t\in N}$ stationary and weakly mixing, for which $(Y_t)_{t\in P0}$ is a restriction to the period P_0.*

H1 allows to deal with the intrinsic non stationarity of the data by selecting long enough periods over which it can reasonably be considered as stationary. If n_0 is large enough, this assumption allows valid estimations for some characteristics such as the distributions of $(Y_t)_{t\in P0}$ or $(X_{k,t})_{t\in Pk}$ for k in $\mathbb{N}$, since both processes have good ergodic properties. "Large enough" here means

sufficiently large to apply the law of large number for the needed estimations but not too large to avoid too strong trend effects. The distributions of $(Y_t)_{t \in P0}$ and $(X_{k,t})_{t \in Pk}$ are respectively noted F_0 and G_k.

Hypothesis H_2*: distributions* F_0 *and* G_k *are continuous with densities* f_0 *and* g_k *respectively, with* f_0 *and* g_k, $k \in \mathbb{N}$ *strictly positive on the interior of their support.*

H_2 avoids having to take conventions to compute inverse functions for the Cumulative Distributions Functions (CDFs), which we will have to consider in the methodology (in fact there is of course no inverse repartition in dimension>1 but a set of *p* one dimensional ones which are defined later using the Levy-Rosenblatt lemma). Such conventions would lead to very complex formulations. If one of the components of Y_t is continuous with the exception of a mass on one point, as is the case for rainfall, what follows can easily be extended to consider such a behavior.

2.2 Distribution transfer on $\mathbb{R}^p$

Definition: *If F and G are distributions defined on* $\mathbb{R}^p$ *and verifying* H_2, *and* $\mathbf{U}^p$ *the uniform distribution on* $\mathbb{R}^p$, *product of p uniform distributions on* $\mathbb{R}$, *then transferring distribution G to F will be defined as an application* $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ *such that T(G) = F, T(G) being the distribution image of G through T.*

In what follows *T* is first defined as an application $T : \mathbb{R}^p \rightarrow \mathbb{R}^p$ and then used as operator on distributions, defined for any borelian set I by $T(G)(I)=G(T_H^{-1} (I))$, H being the distribution of I.

The notations used for the conditional distributions and their inverses are the following:

If Z is a vector with distribution H on $\mathbb{R}^p$, $Z = (Z^1, Z^2, \dots, Z^p)$, F_1 is the distribution of Z^1 and for each k in $\{2, \dots, p\}$, $F_{k/1,\dots,k-1}(z_1, \dots, z_k)$ is the conditional distribution of Z^k for fixed $Z^1, .., Z^{k-1}$. Thus,

$$F_{k/_{1,..,k-1}}(z_1, \dots, z_{k-1}, z_k) = P(Z^k \leq z_k | Z^1 = z_1, \dots, Z^{k-1} = z_{k-1})$$

Then the inverse function of each strictly increasing function h: $\mathbb{R} \to \mathbb{R}$ is noted h^{-1}.

The following lemma proven by Paul Lévy and better known as the Rosenblatt lemma (Grandjaques, 2015, Grandjacques et al., 2015) defines a transfer function for each distribution H (verifying H_2 in our case) on $\mathbb{R}^p$. H is used here for genericity and stands for any distribution as F or G mentioned earlier.

Lemma 1: if U is defined as

$$\begin{cases} U^1 = H_1(Z^1) \\ U^k = H_{k/_{1,\dots,k-1}}(Z^1, \dots, Z^k) \\ U^p = H_{p/_{1,\dots,p-1}}(Z^1, \dots, Z^p) \end{cases}$$

then the distribution of $U = (U^k)_{k=1,..,P}$ is $\boldsymbol{U}^p$, uniform with independent marginal distributions.

Thus if T_H is the above described transformation, $T_H(H) = \boldsymbol{U}^p$ and its inverse T_H^{-1}: $\mathbb{R}^p \to \mathbb{R}^p$, such that $T_H^{-1}(\boldsymbol{U}^p) = H$ is defined as:

$$\begin{cases} Z^1 = H_1^{-1}(U^1) \\ Z^k = H_{k/_{1\dots k-1}}^{-1}(U^1, \dots, U^k) \\ Z^p = H_{p/_{1\dots p-1}}^{-1}(U^1, \dots, U^p) \end{cases}$$

Remark: T_H is obviously not the only transformation allowing a distribution transfer, different versions can be proposed depending for example on the order according to which each

158 component of Z is considered. T_H is sequential with respect to the space dimension; this is a
159 useful property in the applications. Note that in the Gaussian case it is not at all the classical
160 transformation obtained by diagonalization of the covariance matrix, which is not sequential.
161 The analogy is more with a Gramm-Schmidt orthogonalization (Greub, 1975).

162 Lemma 2: transfer of G onto F

163 *If* $T_{G,F} = T_F^{-1}(T_G)$ *then* $T_{G,F}$ *transfers the distribution* G *onto the distribution* F*, because*
164 $T_G(G)=\mathcal{U}$ *and* $T_F^{-1}(\mathcal{U})=F$.

165 2.3 Application to the projection of the distribution of Y over the future period
166 P_k

167 The projection of the distribution of Y_t over the future period P_k relies on another assumption
168 H_3 complementing H_2 and which relates the dynamics of X and Y.

169 At this stage, different hypotheses concerning time invariance can be made:

170 Hypothesis $H_{3\text{-}1}$*: let* $T_{Gk,Fk}$ $k\geq 1$ *be the transformation transferring the distribution* G_k *of*
171 $(X_{k,t})$ $t \in P_k$*, restriction of* X_t *over* P_k*, to that of* $(Y_{k,t})$*, restriction of* Y *over* P_k*, then* $T_{Gk,Fk}$ *does not*
172 *depend on k.*

173 $$T_{Gk,Fk} = T_{G,F} \text{ for each } k \text{ in } \mathbb{N}$$

174 *Thus:*

175 $$T_F^{-1}(T_G) = T_{F_k}^{-1}(T_{G_k}) \qquad \text{and} \qquad {T_{F_k}}^{-1} = {T_F}^{-1}(T_G\ ({T_{G_k}}^{-1}))$$

176 *and the projection over future period* P_k *can be obtained through:*

177 $$\hat{Y}_{k,t} = T_{F_k}^{-1}(T_{G_k}(X_{k,t})) = T_F^{-1}(T_G(X_{k,t})) \text{ with } k \in \mathbb{N},\ t \in P_k$$

178 This hypothesis means that the link between both series is invariant in time or that it does not
179 depend on trends.

180 For univariate bias-correction, this is the hypothesis leading implicitly to the same estimation
181 as with the techniques linked to Empirical Quantile Matching (Déqué et al., 2007).

182 Hypothesis $H_{3\text{-}2}$: *The transformation between two periods is the same for the*
183 *observation and for the climate model*

184 It is therefore quite different from our hypothesis $H_{3\text{-}1}$, which states that the transformation
185 between X and Y is invariant in time.

186 $$T_{F,Fk} = T_{G,Gk} \text{ for every } k \text{ in } \mathbb{N}$$

187 *Thus:*

188 $$T_F^{-1}(T_{F_k}) = T_G^{-1}(T_{G_k}) \qquad \text{and} \qquad \hat{Y}_{k,t} = T_G^{-1}(T_{G_k}(Y_{0,t}))$$

189 However, if $Y_{k,t}$ in the future is directly computed from the observations recorded over the
190 observation period, then it will keep the observed interannual variability, since the distribution
191 only is corrected. Thus to avoid this unrealistic behavior, because there is no reason that
192 interannual variability in the future will mimic that of the recent past period, one can rather
193 compute:

194 $$T_{F_k} = T_F(T_G^{-1}(T_{G_k}))$$

195 *and the desired bias corrected time series over future period P_k can be obtained through:*

196 $$\hat{Y}_{k,t} = (T_{F_k})^{-1}(T_{G_k}(X_{k,t})) = (T_{G_k})^{-1}(T_G(T_F^{-1}(T_{G_k}(X_{k,t})))) \quad \text{with } k \in \mathbb{N},$$

197 $t \in P_k$

For univariate bias correction, this estimation is the same as the one obtained by the CDFt correction for example (Cumulative Distribution Function transform, Michelangeli et al., 2009).

We use both approaches in the generalization to multivariate bias correction.

H_1 means that the distributions are stationary over long enough periods P_k. H_3 hypotheses mean either that the link between the deformations of the distributions of the modeled time series X_t and of the observed ones Y_t over these periods of stationarity does not vary over time ($H_{3\text{-}2}$), or that the link between the distributions of X_t and Y_t does not depend on time t ($H_{3\text{-}1}$). But, as H_3 only concerns instantaneous distributions, it does not imply the transfer of the entire dynamic from one time series to the other. This could be done partially for example if what has been previously described is not applied to F and G but to the distributions of (Y_{t-1},Y_t) and (X_{t-1},X_t), which could be possible according to H_2. Nevertheless, intuitively, it can be seen that some trajectory properties will bring additional consistency between F_0 and F_k. For example, if F has 2 components, Y_t^1 and Y_t^2, such consistency is due to the consideration of the whole temporal dependency, for example that of Y_t^1 and Y_{t-1}^2.

This justifies the idea proposed by different authors (Vrac and Friederich, 2014) of reordering the time series $Y_{k,t}$, $k \in \mathbb{N}$, $t \in P_k$ with regard to Y_t, $t \in P_0$. Remaining in dimension 2, if

$R_t = (r_t^1, r_t^2)$, $t \in P_0$ is the rank vector computed for each component independently,

then the transformations to period P_k are the simple permutations σ^1, σ^2 such that

$\sigma^1(r_t^1, t \in P_k) = (r_t^1, t \in P_0)$ and $\sigma^2(r_t^2, t \in P_k) = (r_t^2, t \in P_0)$.

However, this implies that the temporal sequencing of the variables over period P_0 is imposed to the future period P_k, which means in particular that the interannual variability in the period P_k remains that of period P_0. This is a very strong assumption, probably too strong because it is

physically very unlikely that this will be the case for climate time series. This is not the case for two historical periods, and it can be anticipated that climate change will impact not only the mean, but also the variability, and even the whole dynamics, both daily and interannually.

With our proposed approach, we have:

$$Y_t = T_F^{-1}(U_t) \text{ for } t \in P_0$$

$$Y_{k,t} = T_{Fk}^{-1}(U_{k,t}) \text{ for } t \in P_k, \text{ with:}$$

$$U_{k,t} = T_{F_k}(Y_t) \text{ and } V_{k,t} = T_{Gk}(X_t)$$

The components of U_t and V_t are independent, which allows reordering and applying the reordering transformation independently for each component. Such a reordering could be added, but there is still a strong risk of over fitting. This will not be considered here but rather left for a forthcoming paper.

2.4 Validation

Once the estimations have been made using functional estimations described in section 2, validation is undertaken in the following way.

The historical period P_0 is divided into two sub-periods Q_0 and Q_1. Then, Q_0 is used to calibrate F and G and Q_1 is used for validation. If $\hat{Y}_t$, $t \in Q_1$ is the corrected time series, the aim is to validate the bias-correction using both hypotheses $H_{3\text{-}1}$ and $H_{3\text{-}2}$, then either

$$\hat{F}_1^{-1} = T_{F_0}^{-1}(T_{G_0}(G_1^{-1}))$$

or

$$\hat{F}_1 = T_{F_0}(T_{G_0}^{-1}(G_1))$$

F_0 and G_0 being the distribution functions over Q_0 and G_1 over Q_1.

Validation consists then in comparing $\hat{F}_1$ to F_1 which can be estimated here. If d is a distance between distributions in $\mathbb{R}^p$, the level of correction will be defined as:

$$r = \frac{d(F_1, \hat{F}_1)}{d(F_1, G_1)}$$

Then the choice of d may be quite arbitrary. If $H(z) = P(Z^1 \leq z^1, \ldots, Z^p \leq z^p)$ is the distribution associated to the distribution H of Z on $\mathbb{R}^p$, the distance which naturally generalizes the distance on $\mathbb{R}$ for continuous distributions is given by $d(H, K) = sup_{z \in \mathbb{R}^p} |H(z) - K(z)|$. Under the assumption that H and K have densities φ and ψ, a distance L^p like

$$d_p(H, K) = \left(\int_{\mathbb{R}^p} |\varphi(z) - \psi(z)|^p \right)^{1/p}$$ can be used.

These distances have been chosen here but others could have been used like the "divergences" proposed by Rust et al, 2010.

3 Distribution estimations

The distributions F, G and G_k are at first unknown and still to be estimated so that the transformations T_F, T_G and T_{Gk} can be obtained explicitly. This implies the estimation of one dimensional conditional distributions, but with a conditioning on 2 to p-1 dimensions. According to hypothesis H_1 and H_2, all considered distributions have a probability density function on $\mathbb{R}^p$, which allows the use of non-parametric smoothing methods like kernel density estimation techniques. Two main approaches exist to estimate conditional distributions:

- direct use of a kernel estimator adjusted through an indicator function (Hall et al., 1999)
- estimation of a conditional density and integration afterwards (Hyndmann et al., 1996)

We will use this last approach, which allows the computation of different validation criteria. The numerical examples in parts 3 and 4 are given for 2 variables ($p=2$). Even if the theory is general and valid regardless of the number of dimensions, in practice the number of values necessary for a reliable estimation of the densities increases in a polynomial way with dimension p.

Let us consider a time series $X_t = (X_t^1, X_t^2)$ which verifies hypotheses H_1 and H_2. The estimation of the density of X_t^1 is classical and has good asymptotical properties (for a sufficiently large number of data) thanks to hypothesis H_1. This density will be noted $g_1(x^1)$ and its estimator $\hat{g}_1(x^1)$.

Then, $g_{2/1}(x^1, x^2) = \frac{g(x^1,x^2)}{g_1(x^1)}$. The classical kernel density estimator is given by:

$$\hat{g}_{2/1}(x^1, x^2) = \frac{\frac{1}{nh^{(1)}h^{(2)}} \sum_{i=1}^{n} K^{(1)}\frac{|x^1 - X_i^1|}{h^{(1)}} K^{(2)}\frac{|x^2 - X_i^2|}{h^2}}{\frac{1}{nh^1} \sum_{i=1}^{n} K^{(1)}\frac{|x^1 - X_i^1|}{h^{(1)}}}$$

where the kernels $K^{(1)}$ and $K^{(2)}$ are positive functions $\mathbb{R} \rightarrow \mathbb{R}$, with integral 1 and whose square can be integrated. The smoothing parameters $h^{(1)}$ and $h^{(2)}$ are chosen according to the data used for the estimation.

Different types of kernels are generally used:

- Gaussian: $K(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2})$,
- Epanechnikov: $K(x) = \frac{3}{4}(1 - x^2)1_{\{|x| \leq 1\}}$,
- Student: $\frac{1}{\sqrt{\pi}} \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$

We have chosen the Gaussian kernel.

To estimate parameters $h^{(1)}$ and $h^{(2)}$ we use the R package "hdrcde" developed by R. Hyndmann and based on the methods described in Hyndmann et al., 1996, Bashtannyck and Hyndmann, 2001, Hall et al., 1999, Fan et al., 1996, De Gooijer and Gannoun, 2000, Liebscher, 1996 and Hyndmann and Yao, 2002, valid under hypotheses H_1 and H_2. The choice of $h^{(1)}$, $h^{(2)}$ is then made either by using asymptotical results based on H_2 to which a condition C^2 (stating that all partial derivatives of g have to be twice continuously differentiable) has to be added, or by cross-validation using blocks in order to deal with the weak dependence. The criterion to be minimized in order to choose $h^{(1)}$ and $h^{(2)}$ is generally the L^2 norm between the densities and their estimates. We have also used methods based on sub-sampling and regression (Hall et al., 1999) or kernels with polynomial weights (Bashtannyck and Hyndmann, 2001), and in practice, the approach used in the R package "hdrcde" (which mixes regression and sub-sampling for kernel estimation Hyndmann et al., 1996, Bashtannyck and Hyndmann, 2001, Hyndmann and Yao, 2002) is efficient, that is easy to handle and giving good results in a reasonable computing time.

4 When is it justified?

The previously described methodology aims at bias-correcting the joint distribution of 2 or more variables at the same time. This brings more complexity and needs more data than the usual univariate distribution correction. Thus, in order to judge if there are best suited conditions for the use of such a technique, it has been decided to first test the approach with statistically simulated data.

4.1 Design of the simulation study

Since shuffling is not considered here, the corrections are only based on the distribution of the data. So, it has been decided to test the approach with data generated by chosen parametric distributions.

306 The aim is thus to simulate 4 bivariate distributions:

307 - 1 corresponding to the observations over the current period: F_0
308 - 1 corresponding to the model simulation over the same current period: G_0
309 - 1 corresponding to the model simulation over a future period: G_k
310 - 1 corresponding to the observations over a future period, which is not known a
311 priori and only used to validate the approach: F_k

312 It has been decided to simulate bivariate Gaussian distributions, which is a simple design.

313 4.2 Design of the tests

314 The parameters for the bivariate Gaussian distribution of the current period observations F_0
315 have been chosen based on summer temperature distributions for 2 distant points in Europe,
316 chosen arbitrarily as Hamburg and Orly. The means, variances and co-variances are estimated
317 from the EOBS daily mean temperature time series over the period 1979-2014 for the month of
318 July:

319 $m_1 = 17.5 \quad m_2 = 20.0 \quad v_{11} = 10 \quad v_{22} = 9 \quad v_{12} = v_{21} = 6$

320 with m_1 and m_2 respectively the means for variables 1 and 2, v_{11} and v_{22} their variances and v_{12}
321 their co-variance. This leads to a linear correlation $\rho = 0.63$ between both variables.

322 The values corresponding to the observations over current period are thus simulated by a
323 bivariate Gaussian distribution with the above mentioned parameters and 2000 values are
324 produced.

325 Then, some hypotheses have to be made to simulate the data for the model simulations (current
326 and future periods) and for the observations over future period used for verification.

327 To do so, model errors have first been postulated, additive for the means and multiplicative for
328 the variances and co-variances, noted em_1 and em_2 for the mean errors and es_{11}, es_{22} and es_{12}

for the variance, co-variance errors. The data corresponding to the model simulations for each variable over current period (2000 values) are thus produced by a bivariate Gaussian distribution G_0 with parameters:

$$m_1+em_1 \qquad m_2+em_2 \qquad v_{11} \times es_{11} \qquad v_{22} \times es_{22} \qquad v_{12} \times es_{12} \ (=v_{21} \times es_{21})$$

Then, climate shifts due to climate change are postulated in the same way, that is as additive for the means and multiplicative for the variances and co-variances, and noted dm_1, dm_2, ds_{11}, ds_{22} and ds_{12} respectively. The 2000 values corresponding to the model simulation over future period are produced by a bivariate Gaussian distribution G_k with parameters:

$$m_1+em_1+dm_1 \ ; \ m_2+em_2+dm_2 \ ; \ v_{11} \times es_{11} \times ds_{11} \ ; \ v_{22} \times es_{22} \times ds_{22} \ ; \ v_{12} \times es_{12} \times ds_{12}$$

and the data corresponding to the observations over future period (2000 values) are produced by a bivariate Gaussian distribution F_k with parameters:

$$m_1+dm_1 \qquad m_2+dm_2 \qquad v_{11} \times ds_{11} \qquad v_{22} \times ds_{22} \qquad v_{12} \times ds_{12} \ (=v_{21} \times ds_{21})$$

The validation criterion is the ratio r defined in section 1.4 with both L_1 and L_∞ norms, which is the distance between bias-corrected and observed distributions divided by the distance between simulated and observed distributions, both estimated over the future period.

The aim of the applied bias corrections, either univariate or bivariate, is to estimate the bivariate distribution of the observations for a future period ($\hat{F}_k$) from that of the observations and model simulation over current period (F_0 and G_0) and of model simulation over future period (G_k). Then, the previously defined ratio r measures the performance of the correction in making the corrected distribution closer to that of the observations over future period F_k than was the distribution of the model simulation over future period G_k if $r<1$.

Six bias corrections are applied and compared:

- each variable is independently corrected under hypothesis $H_{3\text{-}1}$ (stationarity of the transformation between model and observations)
- each variable is independently corrected under hypothesis $H_{3\text{-}2}$ (stationarity of the transformation between present and future periods)
- bivariate correction under hypothesis $H_{3\text{-}1}$ with variable 1 corrected first
- bivariate correction under hypothesis $H_{3\text{-}1}$ with variable 2 corrected first
- bivariate correction under hypothesis $H_{3\text{-}2}$ with variable 1 corrected first
- bivariate correction under hypothesis $H_{3\text{-}2}$ with variable 2 corrected first

denoted respectively UH, UV, B12H, B21H, B12V, B21V.

UH and UV are similar to empirical quantile mapping and CDFt respectively. Here since we only deal with distributions, the bivariate distribution correction under hypothesis $H_{3\text{-}2}$ is computed directly from Y_0 through $T_G^{-1}(T_{G_k})$. The univariate bias corrections are computed in the same way as the correction of the first variable in the bivariate corrections, and not taken from the R packages for Quantile Mapping or CDFt, in order to remain consistent in the comparisons.

Two cases have been considered in order to better discriminate hypotheses $H_{3\text{-}1}$ and $H_{3\text{-}2}$ and the consequences of using a correction technique which may not be the best adapted. As a matter of fact, when dealing with climate model simulations, it is difficult to test which stationarity is best verified (because we do not have the observations in the future), and Quantile Mapping or CDFt are generally indifferently used. The chosen cases correspond to:

- one case with model errors larger than climate change
- one case with climate change larger than model errors

In order to be able to test the order of variable corrections in the bivariate corrections, the errors and climate change shifts are not equal for each variable. The test cases are made with:

- em_1=4; em_2=5; es_{11}=es_{22}=es_{12}= 0.5 and dm_1=2; dm_2=1.5; ds_{11}=ds_{22}=1.2 ; ds_{12}= 1
- em_1=1; em_2=1.5; es_{11}=0.8 ;es_{22}=0.75 ;es_{12}= 0.9 and dm_1=4; dm_2=5; ds_{11}=ds_{22}=1.5; ds_{12}= 1

The choices have been made in considering observed model errors and climate change changes: climate models tend to underestimate the variances and in summer, both the means and the variances increase. Then this behavior has been exaggerated to produce large errors and large climate shifts.

4.3 Results

As previously described, we first simulate 2000 values with Gaussian bivariate distributions using fixed means and variances/co-variances, from which we then non-parametrically estimate bivariate distributions to compute distances and compare the different bias corrections. The non-parametrical estimation is based on the R function “kde2d” of package MASS, and with 2000 values only, such an estimation is uncertain. It has thus first been verified that the distances between the bivariate distributions with the chosen errors and climate change shifts are significantly larger than the distances between 2 sets of 2000 points taken from the same distribution. This is preferred to the consideration of a much larger number of values, because firstly, this considerably increases computing time for the corrections and secondly, in climate change studies, when corrections have to be made it is generally done on a monthly basis, we do not dispose of much more values. Thus 4 sets of 2000 points are produced by use of a bivariate Gaussian distribution: one mimicking 2 variables as observed under current climate conditions, another for current climate as simulated by a climate model and the 2 other sets mimicking observed and modeled variables under future climate conditions. Then the previously defined 6 bias corrections are applied to the set corresponding to the modeled variables under future climate conditions so that its distribution gets closer to that of the

variables corresponding to the observed ones under future climate conditions (as we are here in an academic situation where all 4 distributions are simulated). This is done 30 times and the distributions of the *r* ratios are examined.

4.3.1 Errors larger than climate change

In order to test for the best suited stationarity hypothesis, the transformations based on hypotheses $H_{3\text{-}1}$ and $H_{3\text{-}2}$ are applied to the same sample. $H_{3\text{-}1}$ states that $T_F^{-1}(T_G) = T_{F_k}^{-1}(T_{G_k})$ thus both $T_F^{-1}(T_G)$ and $T_{F_k}^{-1}(T_{G_k})$ are applied to the sample obtained with G_0. Then, the distance between the obtained distributions is computed (as L_1 or L_∞ norm). Similarly, $H_{3\text{-}2}$ postulates that $T_F^{-1}(T_{F_k}) = T_G^{-1}(T_{G_k})$ thus both $T_F^{-1}(T_{F_k})$ and $T_G^{-1}(T_{G_k})$ are applied to the sample obtained with F_0 and the distance between the obtained distributions is computed. Each transformation is computed with variable 1 first and with variable 2 first, and 30 generations are considered. Then, the distance corresponding to $H_{3\text{-}1}$ is compared to that corresponding to $H_{3\text{-}2}$ to infer the best verified hypothesis. In this case, the average L_1 distances corresponding to each transformation for the 30 tests are as follows:

distance according to $H_{3\text{-}1}$ with variable 1 first: $6.2\ 10^{-4}$

distance according to $H_{3\text{-}1}$ with variable 2 first: $6.3\ 10^{-4}$

distance according to $H_{3\text{-}2}$ with variable 1 first: $7.5\ 10^{-4}$

distance according to $H_{3\text{-}2}$ with variable 2 first: $8.4\ 10^{-4}$

The distances after $H_{3\text{-}1}$ are lower, so hypothesis $H_{3\text{-}1}$ seems best verified. Each of the six corrections are then applied to estimate F_k and figure 1 presents the boxplots of the *r* ratio for each correction using either L_1 (top panel) or L_∞ (bottom panel) distances. It shows that, as expected, both univariate and bivariate corrections based on hypothesis $H_{3\text{-}1}$ (UH, B12H and B21H) perform better than those based on hypothesis $H_{3\text{-}2}$ (UV, B12V and B21V), although

these last corrections bring some improvement too. In this case, the correction is very efficient (mean ratios around 0.2) and the bivariate correction does not bring substantial improvement.

4.3.2 Climate change larger than model errors

In this case, the mean distances between the different transformations are as follows:

distance according to $H_{3\text{-}1}$ with variable 1 first: $6.5\ 10^{-4}$

distance according to $H_{3\text{-}1}$ with variable 2 first: $8.2\ 10^{-4}$

distance according to $H_{3\text{-}2}$ with variable 1 first: $2.9\ 10^{-4}$

distance according to $H_{3\text{-}2}$ with variable 2 first: $3.6\ 10^{-4}$

Hypothesis $H_{3\text{-}2}$ is here best verified. Again, the obtained results after applying each of the six distribution corrections (figure 2) confirm that the corrections based on hypothesis $H_{3\text{-}2}$ perform better in this case. The corrections are however much less efficient than in the previous setting, with best ratios around 0.5 while they were around 0.2 in the previous case, but this may be due to the chosen parameters. Actually, the choices made for the model errors and climate change shifts are not symmetrical. Here, choosing the least verified hypothesis may lead to ratios higher than 1 (no corrections). Here, bivariate correction, at least when starting with variable 1, brings some improvement compared to independent univariate corrections.

In both cases, the improvements due to the most suited correction appear more clearly with distance $L\infty$. Distance L_1 is an average over the whole distribution (or at least its estimation on some grid) while distance $L\infty$ corresponds to a single value: the maximal one. As here, the introduced errors imply distribution shifts, the corrections reduce the shift and thus have a greater impact on $L\infty$ than on L_1.

4.3.3 The role of correlation strength

Another question arising in this context is the role of the correlation strength between the variables in the importance and performance of multivariate bias correction. In order to investigate this point, the previous tests have been performed again with lower covariances between both variables:

$v_{12}=v_{21}=3$ which leads to a linear correlation coefficient ρ around 0.3

$v_{12}=v_{21}=1$ which leads to a linear correlation coefficient ρ around 0.1

respectively called medium and low correlation, while the previous test is called high correlation. The results are presented for L_1 and L_∞ norms in figure 3, in keeping only the best corrections in each case (H for hypothesis $H_{3\text{-}1}$ in the case of larger model errors than climate shifts and V for hypothesis $H_{3\text{-}2}$ in the reverse case). They show that the correlation strength does not have any significant impact on the performance of the corrections, even though, in the case of climate shifts larger than model errors, bivariate correction performs better than univariate correction in the strong correlation case but equally well otherwise. Correlation has been compared too (figure 4). While univariate correction does not really change the model correlation, bivariate correction does, and generally in the right way. However, bivariate correction seems to underestimate the correlation coefficient when covariance (and correlation) is high, especially when hypothesis $H_{3\text{-}2}$ is concerned. Moreover, bivariate correction shows more variable results than univariate correction among the 30 bivariate Gaussian distribution generations. This is most probably due to the fact that bivariate correction needs more distribution estimations than univariate correction, which increases statistical errors.

Now, when both model errors and climate shifts are moderate:

$em_1=1$; $em_2=1.5$; $es_{11}=0.8$; $es_{22}=0.75$; $es_{12}= 0.9$ and $dm_1=2$; $dm_2=1.5$; $ds_{11}=ds_{22}=1.2$; $ds_{12}= 1$

then both corrections perform as well, as can be seen in figure 5, because both hypotheses are quite equally verified. However, the dispersion for the bivariate correction is higher again,

which may be linked to the uncertainty in the bivariate distribution estimation when the correlation is high. As a matter of fact, with these parameters, the correlation between both variables is 0.63 for F_0 and goes to 0.73 for G_0. This was the same for the case with higher climate shifts than model errors, which showed a similar behavior (figure 2).

5 Test with climate data

The goal of the developed approach is the bias-correction of climate data, thus the next step consists in testing it with observed and model variables. As all here is defined for continuous distributions, and in order to test the approach with different distributions for each variable, it has been decided to work with temperature and wind speed. Long time series of observed daily mean temperature and wind speed have been obtained for the city of Hamburg in Germany from the ECA&D project web site (http://eca.knmi.nl/dailydata/) (Klein Tank et al., 2002). In order to maximize the chance that the time series are homogeneous, they will be considered from 1950 to 2015 (66 years). Then, the climate model simulation of the IPSL-CM5-MR model (Dufresne et al., 2013) has been arbitrarily chosen in the CMIP5 database as a test model, and the time series of the nearest grid box to Hamburg has been extracted from the historical run (1950-2005) and the RCP 8.5 projection run (2006-2100), to compute time series of daily mean temperature and wind speed over the same period P_0=1950-2015. This period is then divided into two 33-year sub-periods, Q_1=1950-1982, chosen as calibration period, and Q_2=1983-2015 chosen as validation period.

As the approach is valid for stationary time series over the defined periods, the correction is applied on a monthly basis, in order to get rid of the annual cycle. It is always difficult to consider that climatic variables are stationary over a defined period, because of both climate change and interannual variability. The World Meteorological Organization recommends to consider at least 30 years to define the climate of an area, thus considering 33-year periods seems a reasonable choice.

First, the distances after transformations of the same sample according to each $H_{3\text{-}1}$ and $H_{3\text{-}2}$ hypothesis are computed for each month to infer which one is best verified. Table 1 summarizes the results for each month, together with an indication of the correlation between wind and temperature as observed over period Q_1 according to the linear Pearson correlation coefficient or the rank Spearman one. Except for July, for which $H_{3\text{-}1}$ seems clearly best verified, for all other months, either $H_{3\text{-}1}$ show a small preference or it is difficult to discriminate both hypothesis. Therefore, the multivariate correction will be applied according to $H_{3\text{-}1}$, which corresponds to estimations similar as those made by Empirical Quantile Mapping, since between 2 recent past periods the climate shift is not too high. Here, the functions of the R package *qmap* have been used for the univariate corrections in order to compare our proposed approach to standard ones used in climate studies. Daily temperature and wind speed are generally moderately correlated (positively in winter, negatively in summer) except in September.

5.1 Temperature first, then wind

The first test is made by correcting temperature first, and then wind according to temperature as described in sections 1 and 2. The ratios of the distance between the bivariate distribution of the corrected variables and the observed ones divided by the distance between the modeled variables and the observed ones over the validation period are compared for both bivariate and independent univariate bias-corrections for each month (table 2). Here only the L_1 norm is considered as both used norms lead to the same conclusions in the previous section. The first outcome is that, generally bias-correction improves the distance to the observations (ratio <1), and bivariate correction gives slightly better results than univariate corrections for 7 months. The worst correction is obtained for the month of May, while the best occurs in July, both for univariate and bivariate corrections. Figure 6 allows the comparison of the bivariate distributions in May, for observations and model over the calibration (Q_1) and validation (Q_2)

periods and for Q_1 and Q_2 for the model and the observations. It shows that the model errors are rather low, and that there is very little change between both periods, for the model as well as for the observations. It is thus not surprising that the correction is modest in this case, as shown in figure 7. On the contrary, in July (figure 8) the model errors are quite large, and climate change is modest. The situation is more similar to that of our academic case with large model errors and moderate climate shift which previously lead to the best corrections based on hypothesis $H_{3\text{-}1}$ (section 4.3.1). Figure 9 illustrates the distributions before and after corrections and shows that bivariate correction brings some improvement. As far as correlation is concerned, univariate bias correction does not have any impact on the model correlation, whereas bivariate correction does, and generally improves the correlation. This can be seen in table 2 for example for May, when the model anti-correlation is stronger than observed and this is better after bivariate correction or for July when model anti-correlation is weaker than observed and increased by the bivariate bias correction. This is shown for the Spearman rank correlation coefficient but the results are similar with the Pearson correlation coefficient. Thus bivariate corrections clearly improve the situation if the correlations for observation and model are different enough to allow the correction being larger than the statistical errors due to the dimension.

5.2 Wind first then temperature

The same corrections have then been tested again but by correcting wind speed first, and then temperature according to wind speed. The ratios of distances after and before correction are summarized in table 3**,** together with the Spearman rank correlation coefficients. The results are similar even though the months for which univariate correction gives slightly better results are not always the same. July shows again the best performance for the corrections while May remains the worst corrected, with a small advantage to bivariate correction though.

6 Conclusion and perspectives

In this paper, a new approach for bias-correcting multivariate distributions of climate model simulations according to observations has been proposed and tested in controlled conditions, by use of statistical simulations, and in real conditions, with a climate model simulation and observations for wind speed and temperature. This approach is based on the Lévy-Rosenblatt lemma and generalizes the univariate distribution corrections like empirical quantile mapping or CDFt.

The development of the technique first showed that depending on the hypothesis made (invariance of the transformation between model and observations in time or invariance of the transformation between two periods for observed and model variables), the correction is similar to empirical quantile mapping or to CDFt.

Then, the tests with bivariate Gaussian distributions allowed to compare the performances of the corrections in controlling the model errors and climate shifts, as well as the strength of the variables correlation. The parameters are based on observed and model July temperature for two distant points, and willingly exaggerated in order to better see the differences in the approaches. It shows that the verification of the chosen stationary hypothesis: $H_{3\text{-}1}$ (stationarity of the link between model and observations) or $H_{3\text{-}2}$ (stationarity of the link between present end future conditions) has a stronger importance for the correction performance than univariate or bivariate correction, whatever the correlation strength between both variables. Furthermore, the order of variable corrections does not seem to have important consequences in this framework.

Then a last test is made in a more real setting, with daily mean temperature and wind speed time series observed over period 1950-2015 and simulated by the IPSL-CM5-MR model in Hamburg. The corrections are applied on a monthly basis, in order to meet as closely as possible the hypothesis of stationarity, in a cross-validation setting, 1950-1982 being the calibration period and 1983-2015 the validation period. Daily temperature and wind speed are moderately

correlated in each month, with a positive correlation in winter and a negative one in summer, with lower correlation in fall and spring (and almost no correlation in September). As the periods are close, the conditions of a more reliable application of corrections based on hypothesis $H_{3\text{-}1}$ are generally met and this technique is applied. The results show that bivariate correction generally leads to a slightly better correction.

This study opens some questions before the application of such a correction can be generalized. First it can be extended to the case of at least one variable with non-continuous distribution. Most applications have to deal with temperature and rainfall rather than temperature and wind speed. Two ways can be tested to do so:

- Correction of the number of rainy days (based on the differences between model and observations if $H_{3\text{-}1}$ is considered or on the change between present and future if $H_{3\text{-}2}$ is considered), then of the amount of rainfall on rainy days and finally of temperature according to rainfall (in managing again rainy and non-rainy days)
- Transformation of the rainfall distribution so that it becomes more continuous as proposed in Vrac et al 2016

Then although the methodology is generic and theoretically works regardless of the number of variables, extension to more than two variables will need more data for the estimations to be reliable. Furthermore, the need to estimate more distributions increases the statistical errors and the improvement is more obvious if the discrepancies to be corrected are large.

Lastly, the very important question of stationarity remains. Applying the correction on a monthly basis is the easiest solution. However, because of interannual variability, it is necessary to calibrate the correction over a long enough period (at least 30 years). Then in the climate change context, the conditions of application of empirical quantile mapping like approaches (based on hypothesis $H_{3\text{-}1}$), that the distributions are invariant in time, cannot hold, and then,

CDFt like approaches (based on hypothesis $H_{3\text{-}2}$) are more adapted, but may bring a lower improvement. Besides, using Quantile Mapping like corrections in such cases may worsen the situation. It seems then important to think at techniques able to stationarize the distributions and let them be closer in time and between observation and model. This can be done by removing seasonalities and trends, at least in the mean and the variance. However, part of the bias is embedded in the estimation of such quantities for the model time series and they have to be corrected too. In a univariate context, this can be made in an additive or multiplicative way. But in a multivariate context, this implies to think at a way of consistently correcting these parts of the signal as well. Future work will consist in clarifying these questions of non stationarity by working with a parameterization of the two kinds of dynamic deformations we have formalized in hypotheses $H_{3\text{-}1}$ and $H_{3\text{-}2}$. Such parameterization should be more complex than a simple shift but still simple enough to be applied routinely.

Acknowledgements

The authors would like to acknowledge funding for the European Climatic Energy Mixes (ECEM) project by the Copernicus Climate Change Service, a programme being implemented by the European Centre for Medium-Range Weather Forecasts 5 (ECMWF) on behalf of the European Commission. The specific grant number is 2015/C3S_441_Lot2_UEA.

7 References

Bashtannyk D. M., Hyndmann R. J., 2001: Bandwidth selection for kernel conditional density estimation, Computational Statistics & Data Analysis

Cannon A. J., 2016: Multivariate Bias Correction of Climate Model Output: Matching Marginal Distributions and Intervariable Dependence Structure. Journal of Climate, 29, 7045-7064, DOI: 10.1175/JCLI-D-15-0679.1

Chen J., Brissette F. P., Chaumont D., and Braun M., 2013: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resources Research, 49 (7), 4187–4205, doi:10.1002/wrcr.20331

De Gooijer J.G. and Gannoun A., 2000: Nonparametric conditional predictive regions for time series, Computational Statistics and Data Analysis 33, 259–275.

Déqué M., 2007: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16– 26

Dufresne J-L, Foujols M-A, Denvil S, Caubel A, Marti O, Aumont O, Balkanski Y, Bekki S, Bellenger H, Benshila R, Bony S, Bopp L, Braconnot P, Brockmann P, Cadule P, Cheruy F, Codron F, Cozic A, Cugnet D, de Noblet N, Duvel J-P, Ethé C, Fairhead L, Fichefet T, Flavoni S, Friedlingstein P, Grandpeix J-Y, Guez L, Guilyardi E, Hauglustaine D, Hourdin F, Idelkadi A, Ghattas J, Joussaume S, Kageyama M, Krinner G, Labetoulle S, Lahellec A, Lefebvre M-P, Lefevre F, Levy C, Li ZX, Lloyd J, Lott F, Madec G, Mancip M, Marchand M, Masson S, Meurdesoif Y, Mignot J, Musat I, Parouty S, Polcher J, Rio C, Schulz M, Swingedouw D, Szopa S, Talandier C, Terray P, Viovy N, Vuichard N, 2013: Climate change projections using the IPSLCM5 earth system model: from CMIP3 to CMIP5. Climate Dynamics 40(9):2123–2165

Fan J., Yao Q and Tong H., 1996: Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems. Biometrika, 83 (1). pp. 189-206.

Grandjacques, M., 2015 : Analyse de sensibilité pour des modèles stochastiques à entrées dépendantes : application en énergétique du bâtiment. Thèse Université de Grenoble

Grandjacques M., Delinchant B., Adrot O. Pick and Freeze, 2015: estimation of sensitivity index for static and dynamic models with dependent inputs - hal.archives-ouvertes.fr

639 Greub W. H., 1975: Linear Algebra, 4th edition, Springer Verlag

640 Gudmundsson L., Bremnes J., Haugen J., and Engen-Skaugen T., 2012: Technical note:
641 Downscaling RCM precipitation to the station scale using statistical transformations–a
642 comparison of methods. Hydrology & Earth System Sciences, 16 (9), 3383–3390,
643 doi:10.5194/hess-16-3383-2012

644 Haddad, Z. and D. Rosenfeld, 1997: Optimality of empirical z-r relations. Q. J. R. Meteorol.
645 Soc., 123, 1283-1293

646 Hall P, Wolff R, Yao Q, 1999: Methods for estimating a conditional distribution function,
647 Journal of the American Statistical Association, 94 (445). pp. 154-163

648 Hyndmann R, Bashtannyk D M, Grunwald G K, 1996: Estimating and visualizing conditional
649 densities, Journal of Computational and Graphical Statistics, 1996

650 Hyndmann R, Yao Q, 2002: Nonparametric estimation and symmetry test for conditional
651 density functions, Nonparametric statistics

652 IPCC, "Climate change 2013; The physical basis - summary for policymakers," Fifth
653 Assessment Report of the Intergovernemental Panel on Climate Change, 2013

654 Klein Tank, A. M. G., et al., 2002: Daily dataset of 20th-century surface air temperature and
655 precipitation series for the European Climate Assessment, International Journal of Climatology,
656 22, 1,441–1,453, Data and metadata available at http://eca.knmi.nl.

657 Liebscher E., 1996: Strong convergence of sums of c-mixing random variables with
658 applications to density estimation. Stochastic Processes and their Applications 65

659 Li C., Sinha E., Horton D.E., Diffenbaugh N.S., and Michalak A.M., 2014: Joint bias correction
660 of temperature and precipitation in climate model simulations. Journal of Geophysical
661 Research: Atmospheres, 119 (23), 13–153, doi:10.1002/2014JD022514

662 Michelangeli P.-A., Vrac M., and Loukos H., 2009: Probabilistic downscaling approaches:
663 Application to wind cumulative distribution functions, Geophys. Res. Lett., 36, L11708,
664 doi:10.1029/2009GL038401

665 Panofsky H. and Brier G., 1958: Some applications of statistics to meteorology. Tech. rep.,
666 University Park, Penn. State Univ., 224 pp

667 Piani, C., Haerter J., and Coppola E., 2010: Statistical bias correction for daily precipitation in
668 regional climate models over Europe. Theoretical and Applied Climatology, 99, 187-192,
669 doi:10.1007/s00704-009-0134-9

670 Piani, C. and Haerter J.O., 2012: Two dimensional bias correction of temperature and
671 precipitation copulas in climate models. Geophys. Res. Lett., doi:10.1029/2012GL053839

672 Rust H., Vrac M., Lengaigne M., Sultan B.: Quantifying Differences in Circulation Patterns
673 Based on Probabilistic Models. J. Climate, 23:6573-6589, 2010

674 Teutschbein, C., and Seibert J., 2012: Bias correction of regional climate model simulations for
675 hydrological climate-change impact studies: Review and evaluation of different methods.
676 Journal of Hydrology, 456, 12–29, doi:10.1175/JAMC-D-11-0149.1

677 Vrac, M. and Friederichs P., 2015: Multivariate-intervariable, spatial, and temporal-bias
678 correction. Journal of Climate, 28 (1), 218–237, doi:10.1175/JCLI-D-14-00059.1

679 Wood A., Leung L., Sridhar V. and Lettenmaier D., 2004: Hydrologic implications of
680 dynamical and statistical approaches to downscaling climate model outputs. Clim. Change, 62
681 (189-216).

682 Vrac M., Noël T., Vautard R., 2016 : Bias correction of precipitation through Singularity
683 Stochastic Removal: Because occurrences matter. Journal of Geophysical Research
684 Atmosphere, doi: 10.1002/2015JD024511

685 Zhang F., and Georgakakos A.P., 2012: Joint variable spatial downscaling. Climatic Change,
686 111 (3-4), 945–972, doi:10.1007/s10584-011-0167-9

687

688

689 TABLES

690

Month	$H_{3\text{-}1}$ T first	$H_{3\text{-}1}$ w first	$H_{3\text{-}2}$ T first	$H_{3\text{-}2}$ w first	correlation			
					observations		IPSL simulation	
	L_1 (10^{-4})	L_1 (10^{-4})	L_1 (10^{-4})	L_1 (10^{-4})	Pearson	Spearman	Pearson	Spearman
January	7.0	6.6	9.5	9.0	0.47	0.50	0.50	0.54
February	7.0	6.3	10.5	8.1	0.36	0.35	0.40	0.43
March	14.7	13.2	14.0	13.8	0.11	0.10	0.17	0.18
April	18.0	17.1	17.1	17.3	-0.15	-0.15	-0.07	-0.06
May	13.7	12.6	19.0	17.6	-0.17	-0.17	-0.32	-0.33
June	21.5	20.8	21.6	21.3	-0.28	-0.28	-0.27	-0.26
July	13.2	12.2	26.2	28.5	-0.37	-0.39	-0.19	-0.15
August	15.0	16.2	17.3	20.0	-0.21	-0.22	-0.32	-0.30
September	14.5	14.8	19.3	22.9	-0.01	0.00	-0.14	-0.13
October	13.3	9.6	10.5	9.3	0.17	0.16	0.10	0.11
November	8.2	11.6	12.2	13.3	0.33	0.33	0.39	0.40
December	9.4	8.7	11.7	11.2	0.46	0.48	0.47	0.50

691 Table 1: L_1 distance between the bivariate distributions obtained with transformations based
692 on hypothesis $H_{3\text{-}1}$ with temperature (T) or wind (w) first (first two columns) and between the
693 bivariate distributions obtained with transformations based on hypothesis $H_{3\text{-}2}$ with
694 temperature (T) or wind (w) first (columns 3 and 4), correlation between temperature and
695 wind (Pearson and Spearman coefficients for the observations (columns 6 and 7) and for the
696 IPSL model simulation (columns8 and 9) for each month

697

Month	correlation		Bivariate correction		Univariate correction	
	observations	IPSL model	Distance ratio	correlation	Distance ratio	correlation
January	0.55	0.55	**0.522**	0.62	0.563	0.55
February	0.44	0.43	0.328	0.39	**0.301**	0.43
March	0.11	0.18	**0.758**	0.17	0.771	0.18
April	-0.08	-0.05	0.713	-0.14	0.712	-0.05
May	-0.11	-0.30	**0.992**	-0.21	*1.098*	-0.30
June	-0.26	-0.29	0.496	-0.33	**0.483**	-0.29
July	-0.31	-0.16	**0.179**	-0.40	0.244	-0.16
August	-0.20	-0.34	**0.600**	-0.36	0.606	-0.34
September	0.05	-0.12	0.855	-0.03	**0.845**	-0.12
October	0.15	0.12	**0.934**	0.17	0.952	0.12
November	0.32	0.36	**0.788**	0.36	0.849	0.36
December	0.49	0.47	0.671	0.45	**0.650**	**0.47**

Table 2: correlation (Spearman rank correlation coefficient) over the validation period and ratio of the distance (between the bivariate distributions of the corrected and observed variables divided by that of the modeled and observed variables) and correlation over the validation period after bivariate bias correction with temperature corrected first and independent univariate bias-correction. Bold indicates the lowest ratios obtained when correction is efficient

Month	correlation		Bivariate correction		Univariate correction	
	observations	IPSL model	Distance ratio	correlation	Distance ratio	correlation
January	0.55	0.55	**0.532**	0.48	0.563	0.55
February	0.44	0.43	0.316	0.33	**0.301**	0.43
March	0.11	0.18	**0.729**	0.12	0.771	0.18
April	-0.08	-0.05	**0.695**	-0.11	0.712	-0.05
May	-0.11	-0.30	**0.968**	-0.19	*1.098*	-0.30
June	-0.26	-0.29	0.523	-0.35	**0.483**	-0.29
July	-0.31	-0.16	**0.196**	-0.39	0.244	-0.16
August	-0.20	-0.34	**0.571**	-0.36	0.606	-0.34
September	0.05	-0.12	0.853	-0.09	**0.845**	-0.12
October	0.15	0.12	0.970	0.19	**0.952**	0.12
November	0.32	0.36	**0.819**	0.30	0.849	0.36
December	0.49	0.47	**0.601**	0.58	0.650	0.47

Table 3: same as table 2 but for wind corrected first, then temperature according to wind

FIGURES

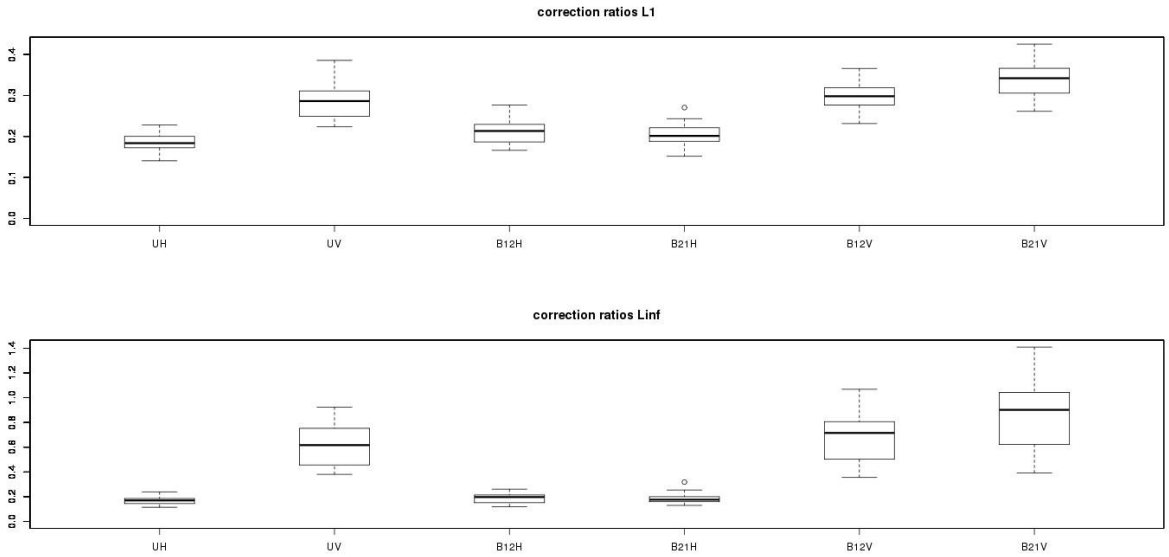


Figure 1: boxplots of the correction ratios obtained with 30 simulations with bivariate Gaussian distributions in the case of larger model errors than climate shifts, based on distance L_1 (top panel) and L_∞ (bottom panel). U refers to univariate correction, B to bivariate starting by variable 1 (12) or 2 (21) and H is for hypothesis $H_{3\text{-}1}$ and V for hypothesis $H_{3\text{-}2}$

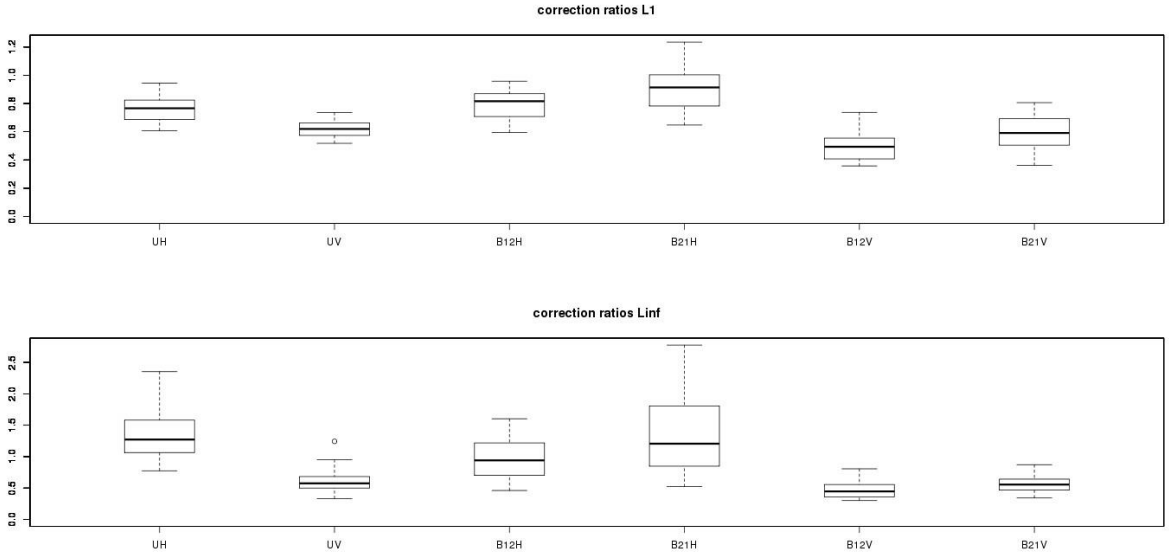


Figure 2: boxplots of the correction ratios obtained with 30 simulations with bivariate Gaussian distributions in the case of larger climate shifts than model errors, based on distance L_1 (top panel) and L_∞ (bottom panel). U refers to univariate correction, B to bivariate starting by variable 1 (12) or 2 (21) and H is for hypothesis $H_{3\text{-}1}$ and V for hypothesis $H_{3\text{-}2}$

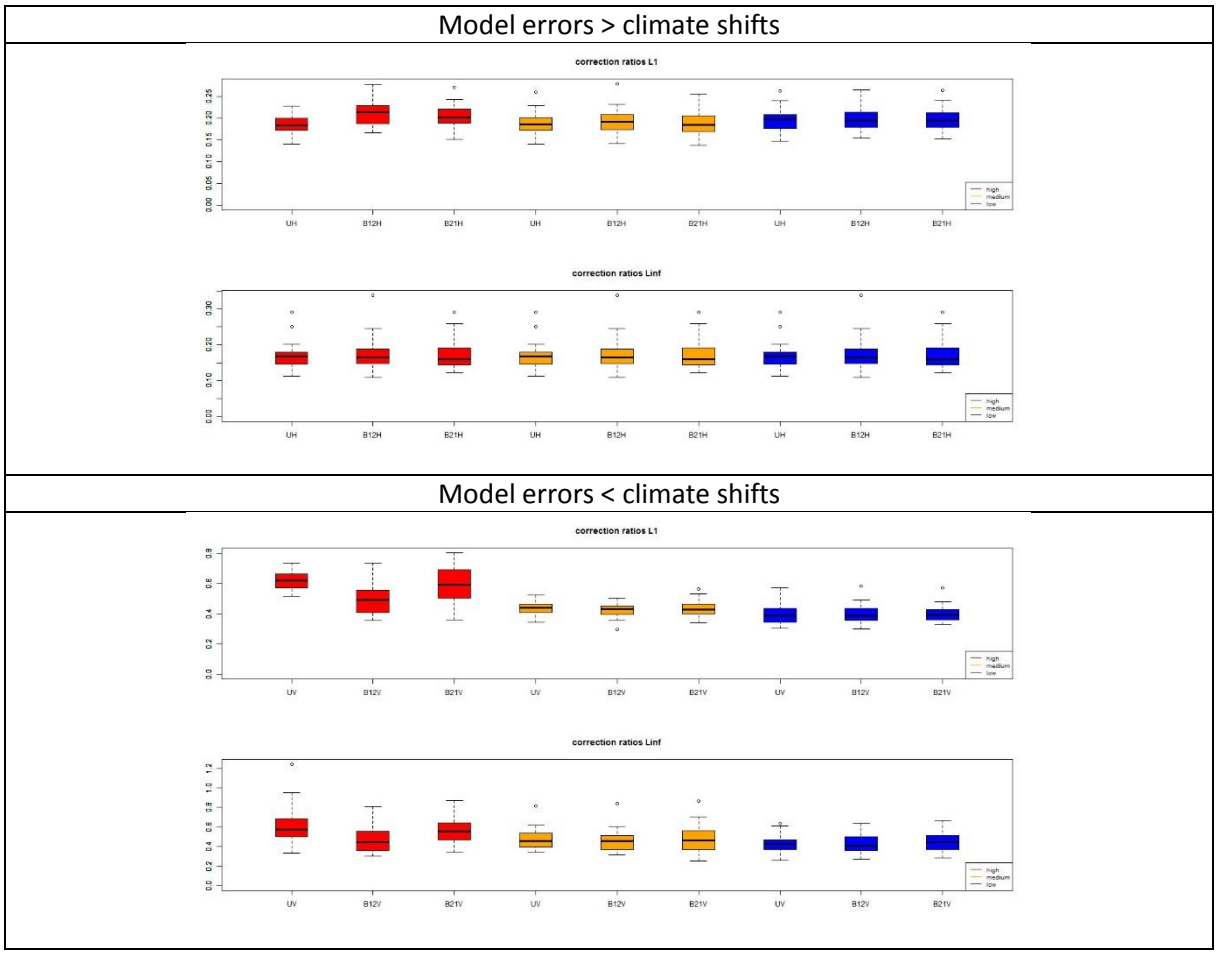
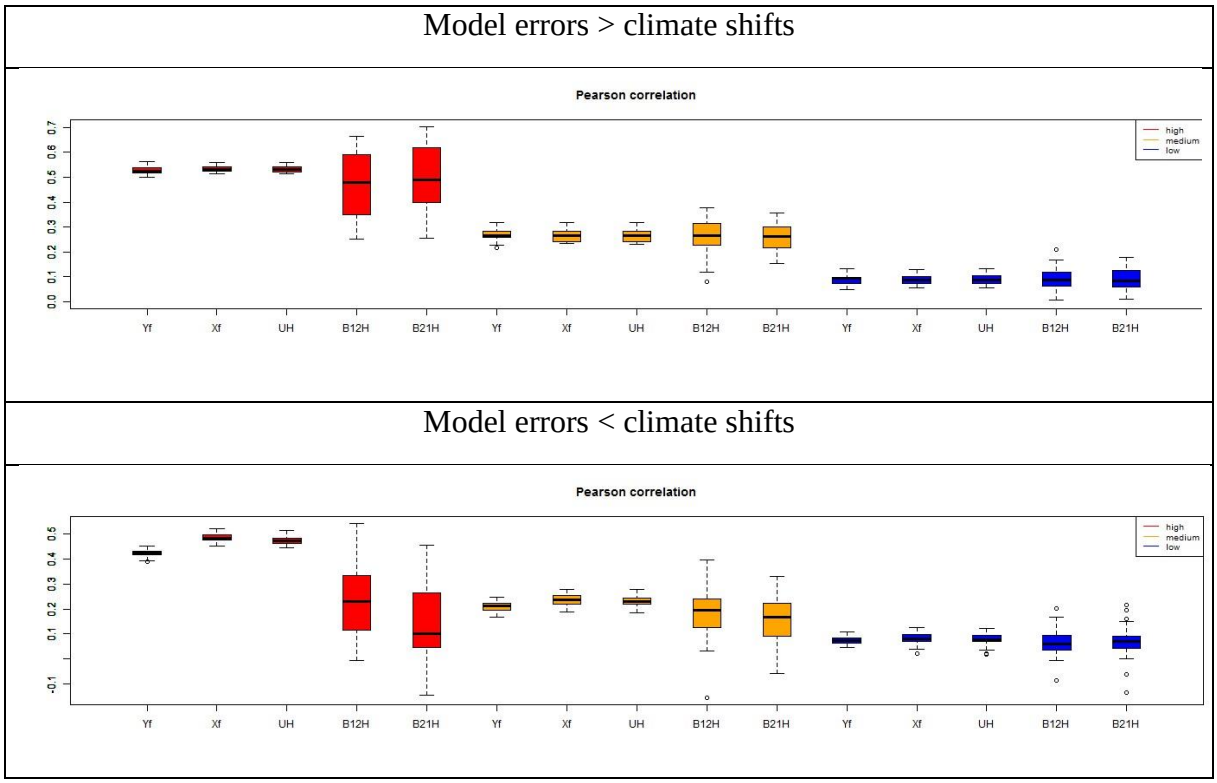


Figure 3: boxplots of the correction ratios obtained with 30 simulations with bivariate Gaussian distributions in both cases (larger model errors than climate shifts, top panel and larger climate shifts than model errors, bottom panel) based on distances *L1* and $L\infty$ and for different correlation strengths between the variables: high (red), medium (orange) and low (blue). U refers to univariate correction, B to bivariate starting by variable 1 (12) or 2 (21) and H is for hypothesis $H_{3\text{-}1}$ and V for hypothesis $H_{3\text{-}2}$



733 Figure 4: boxplots of the correlation coefficients obtained with 30 simulations with bivariate
734 Gaussian distributions in both cases (larger model errors than climate shifts, top panel and
735 larger climate shifts than model errors, bottom panel) and for different covariance strengths
736 between the variables: high (red), medium (orange) and low (blue). Yf refers to the
737 observations over the validation period, Xf to the simulation over the same period, U refers to
738 univariate correction, B to bivariate starting by variable 1 (12) or 2 (21) and H is for
739 hypothesis $H_{3\text{-}1}$ and V for hypothesis $H_{3\text{-}2}$

740

741

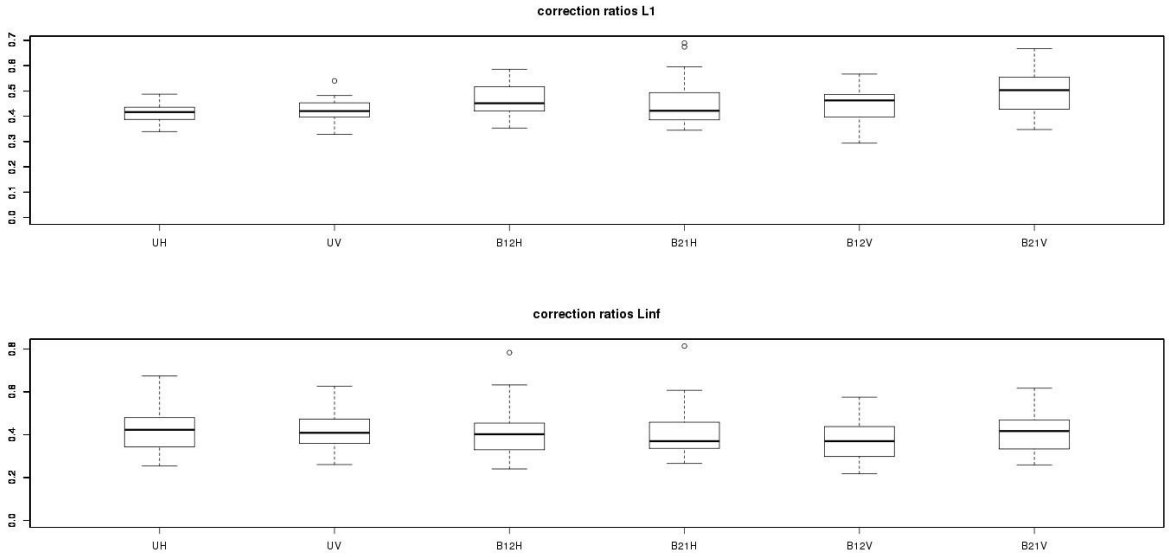


Figure 5: boxplots of the correction ratios obtained with 30 simulations with bivariate Gaussian distributions in the case of moderate climate shifts and model errors, based on distance L_1 (top panel) and L_∞ (bottom panel). U refers to univariate correction, B to bivariate starting by variable 1 (12) or 2 (21) and H is for hypothesis $H_{3\text{-}1}$ and V for hypothesis $H_{3\text{-}2}$

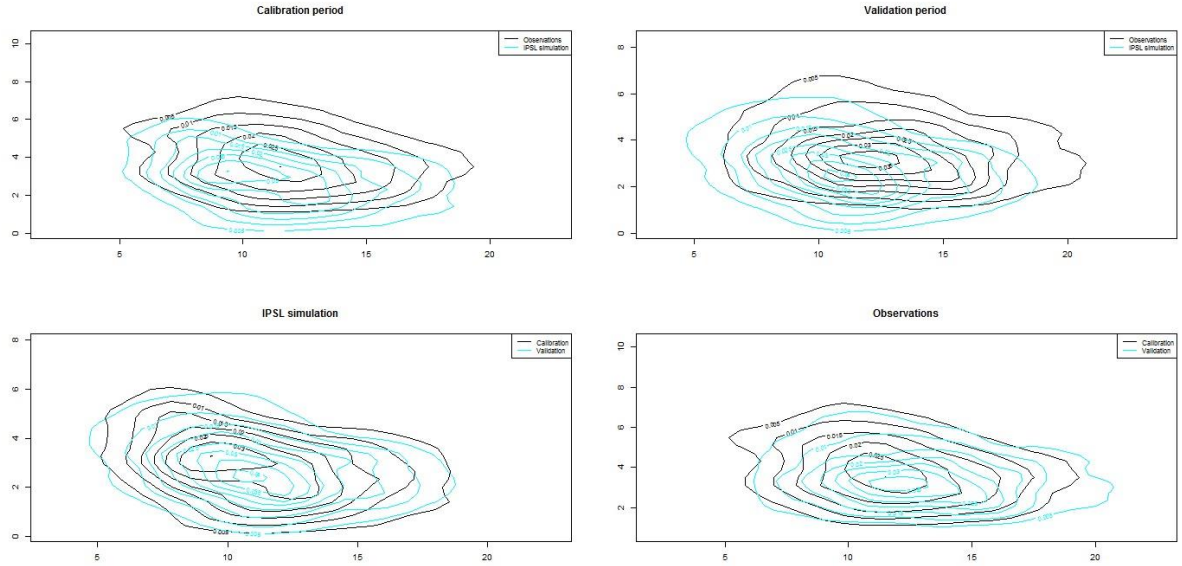


Figure 6: Top line: comparison between observed (black) and modeled (cyan) bivariate distributions of temperature and wind speed in Hamburg for each period (calibration period: left panel, validation period: right panel); bottom line: comparison between calibration (black) and validation (cyan) bivariate distributions of temperature and wind speed in Hamburg for the model and the observations (IPSL model: left panel; observations: right panel) for the month of May

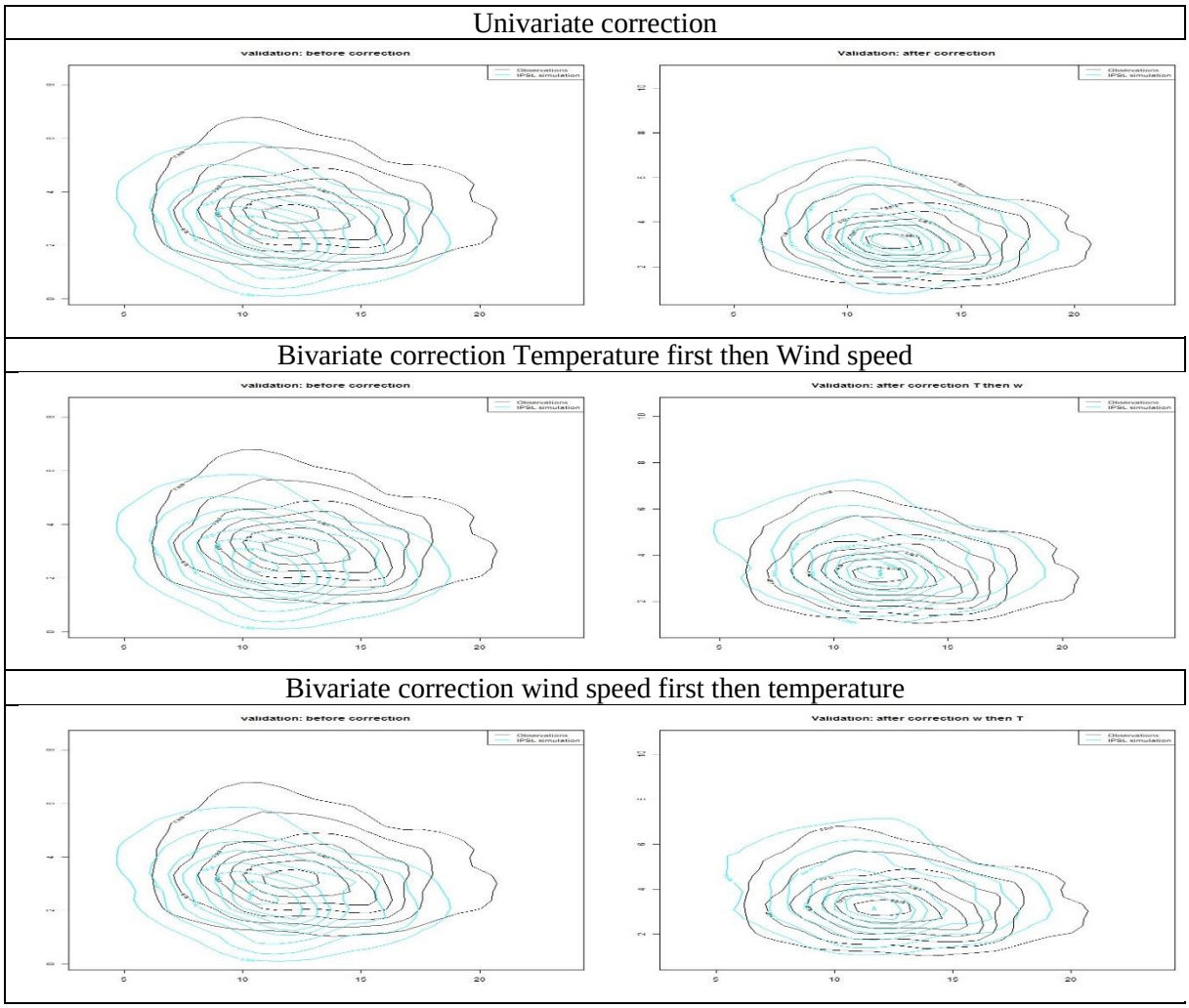
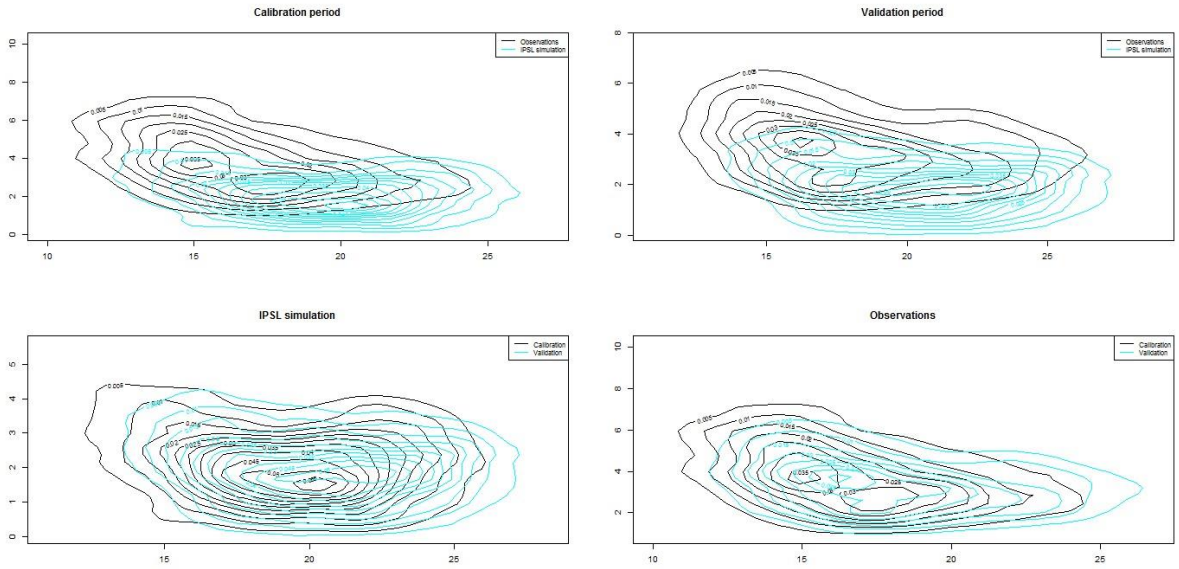


Figure 7: comparison between observed (black) and modeled (cyan) bivariate distributions of temperature and wind speed in Hamburg for the validation period before correction (left panels) and after univariate correction (top right), bivariate correction with temperature first (medium right) and wind speed first (bottom right) for the month of May

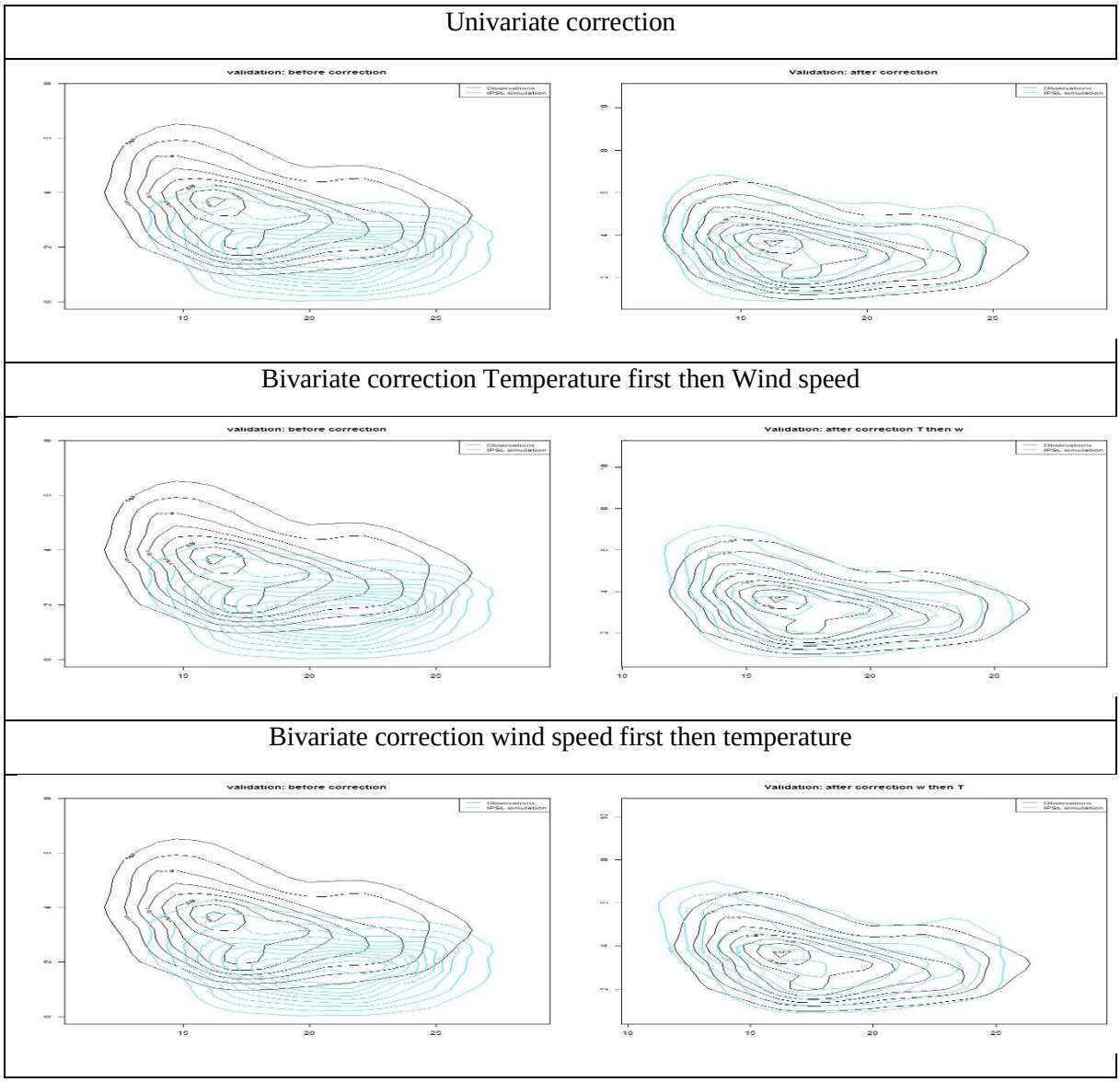


764

765 Figure 8: same as figure 6 but for the month of July

766

767



768 Figure 9: same as figure 7 but for the month of July