Evaluating the structure of bias-corrected climate variables

Measuring joint extremity with vine copulas

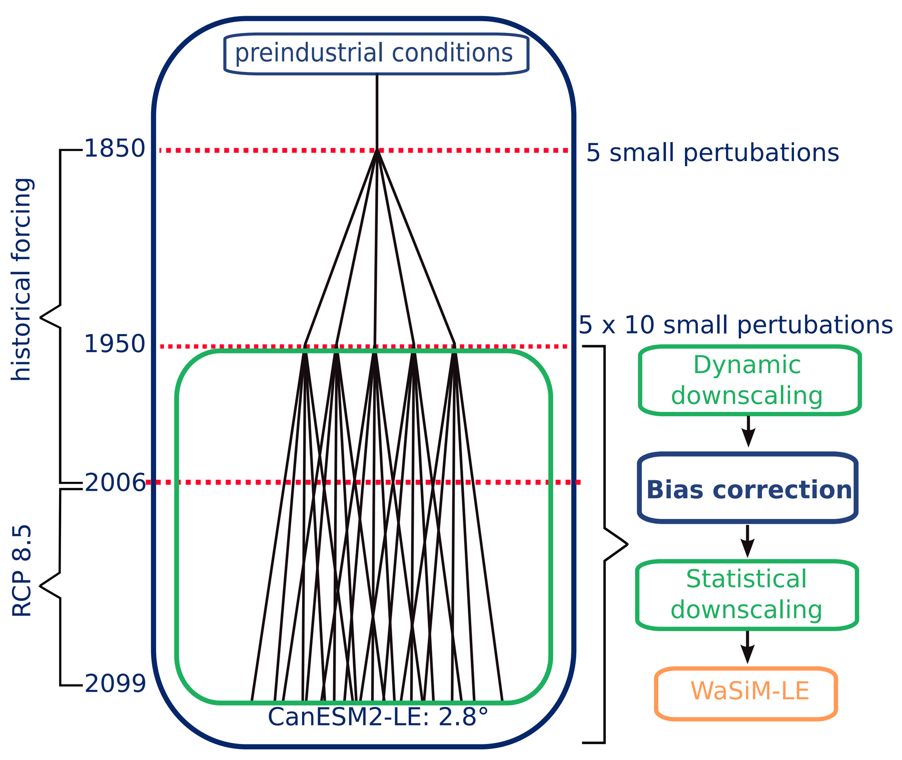
Henri Funk[1,2], Helmut Küchenhoff[1], Ralf Ludwig[2], Thomas Nagler[3]

[1] *Statistical Consulting Unit StaBLab, LMU Munich, Germany*
[2] *Department of Geography, LMU Munich, Germany*
[3] *Computational Statistics & Data Science, LMU Munich, Germany*

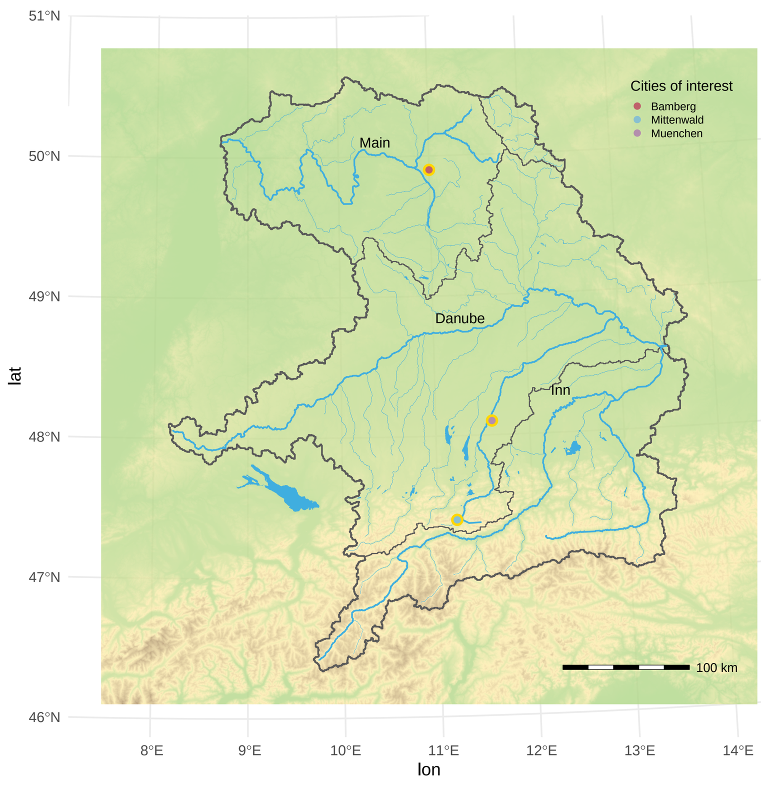
h.funk@lmu.de

From climate data to hydrological data

Hydrological modelling for Bavaria



(a) CRCM5-LE and Workflow



(b) Wasim domain of hydrological Bavaria

Figure 1: *In Figure* **(a)***, we illustrate the SMILE CRCM5-LE based on Leduc et al. (2019). Within the ClimEx workflow, each of its 50 transient simulations, spanning from 1950 to 2099, undergoes downscaling and correction for hydrological modelling using WaSim. Our primary focus encompasses Bavaria and its associated river sources, as depicted in Figure* **(b)***. We exemplify the process during summer in Mittenwald (alpine climate).*

Climate variables

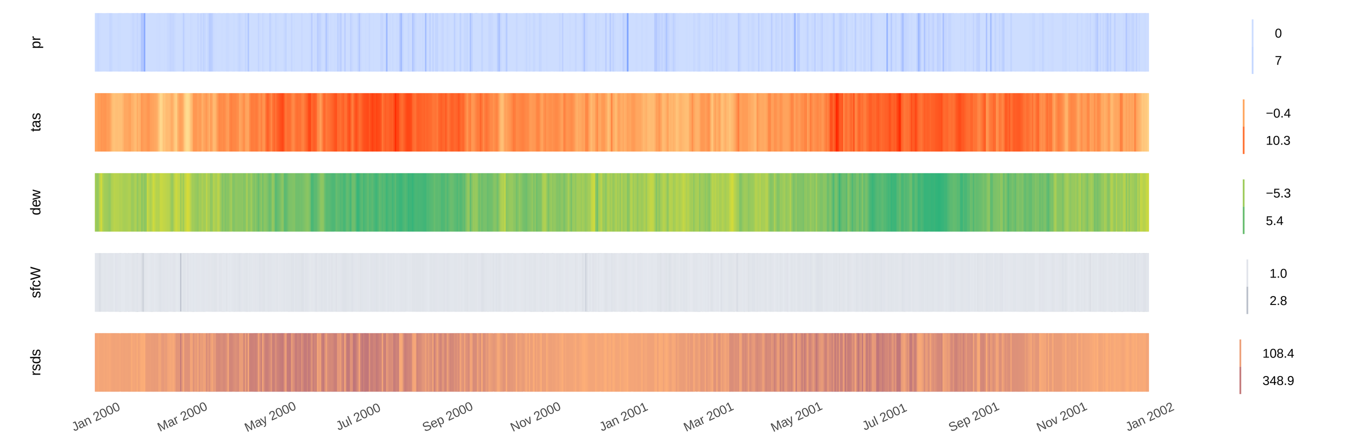


Figure 2: *Concerned climate variables in a daily resolution over two years for Mittenwald.* **pr**- *Precipitation (mm day-1);* **tas**- *Temperature (deg. C);* **dew**- *Dewpoint temperature (deg. C);* **sfcW**- *Wind speed (m/s);* **rsds**- *Downwelling shortwave radiation (W)*

Data types

Name	Abbrev.	Explanation
SDCLIREF	oc	Measured and interpolated reference data
CRCM5 calibration	mc	Simulation data from CRCM5 member *kba*
CRCM5 univariate BC	$\hat{mc}_{UBC}$	Univariate corrected mc (Cannon 2015)
CRCM5 multivariate BC	$\hat{mc}_{MBCn}$	Multivariate corrected mc (Cannon 2018)
CRCM5 multivariate BC II	$\hat{mc}_{VBC}$	Multivariate corrected mc (**Authors'** correction)

Table 1: The calibration period spans from 1981 to 2010. All corrections are implemented on seasonal subsets. VBC is a work in progress vine copula bias correction of the authors.

Research Question: How accurate do the corrected data $\hat{mc}$ represent the measured, true distribution $F^{(oc)}$?

Approaching evaluation measures for bias correction

Bias correction and Evaluation

$$\hat{mc} = F^{(oc)-1}(F^{(mc)}(mc))$$

If the bias correction was successful, we assume

$$\hat{mc} \sim F^{oc}$$

Analysis by vine copula

To analyse, the joint distribution we need to understand its constitution. According to Sklar's theorem, the joint distribution F can be expressed w.t.h. of the marginal distribution functions F_i and the (vine) copula function C.

$$F(x_{pr}, \dots, x_{rsds}) = C(F_{pr}(x_{pr}), \dots, F_{rsds}(x_{rsds}))$$

$\Rightarrow$ The efficacy of the correction can be asserted by the congruence between $C^{(oc)}$ and $C^{(\hat{mc})}$.

Bivariate evaluation

The advantage of vine copulas is their deduction in a product of marginal bivarite copulas which can be analysed individually.

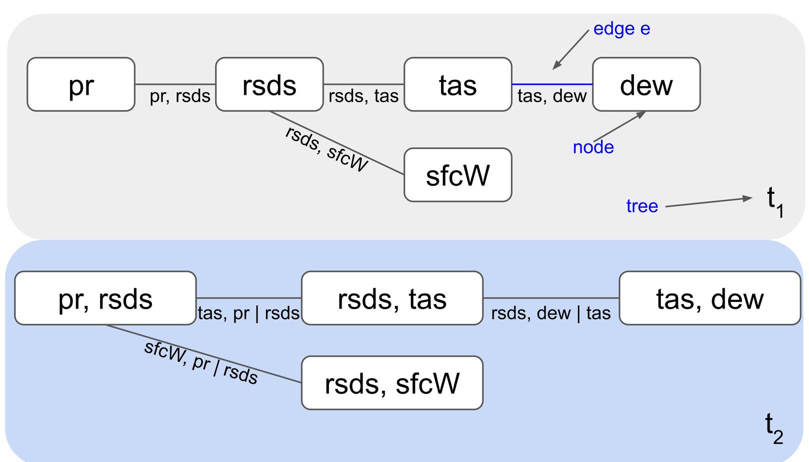
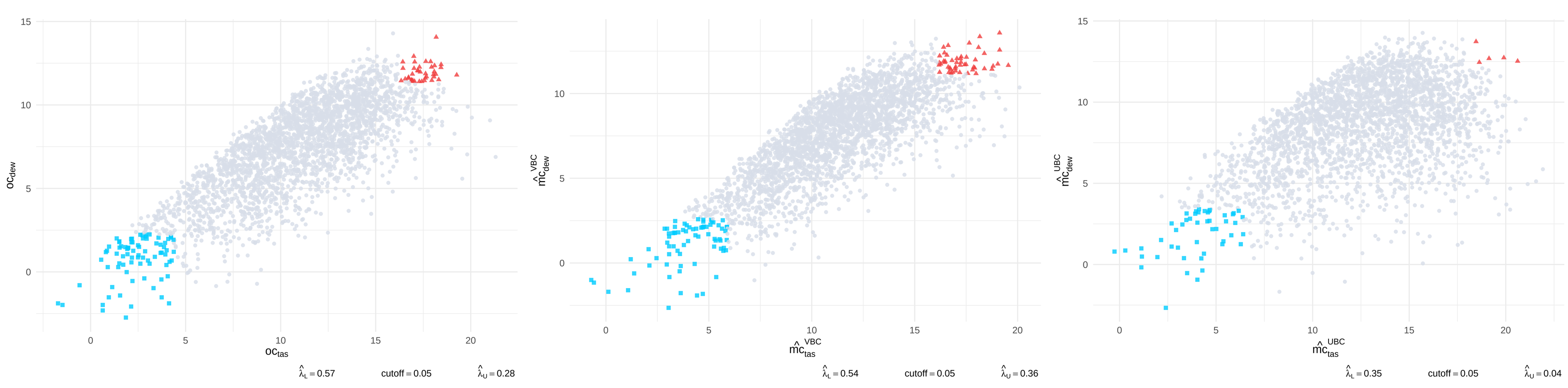


Figure 3: *Initial two trees from the oc vine graph as identified by Dissman (2013). Subsequent bivariate analyses utilize edge $e = (dew, tas)$.*

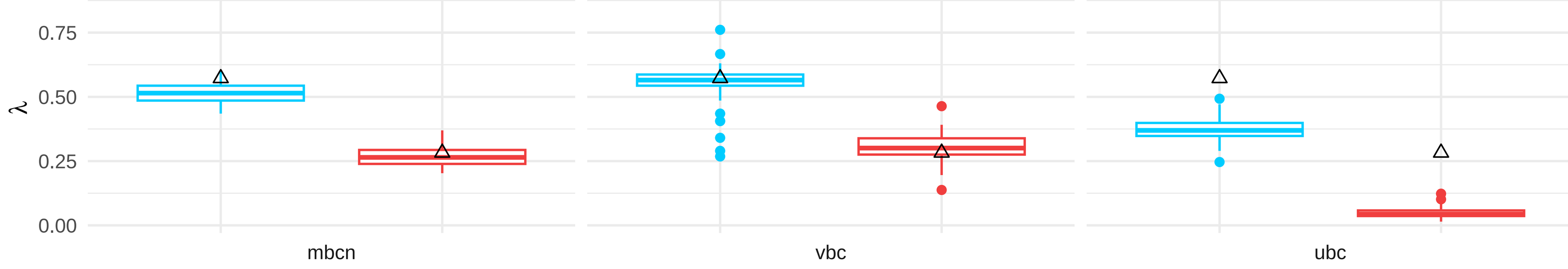
Analysis of joint extremes- tail dependence

Tail dependency measures the probability λ that one variable (**dew**) is extreme, given that another variable (**tas**) is also extreme. There is lower (λ_L) and upper (λ_U) tail dependency:

$$\lambda_U = \lim_{q\to 1} P\left(u_{e_{dew}} > q \mid u_{e_{tas}} > q\right); \lambda_L = \lim_{q\to 0} P\left(u_{e_{dew}} < q \mid u_{e_{tas}} < q\right)$$



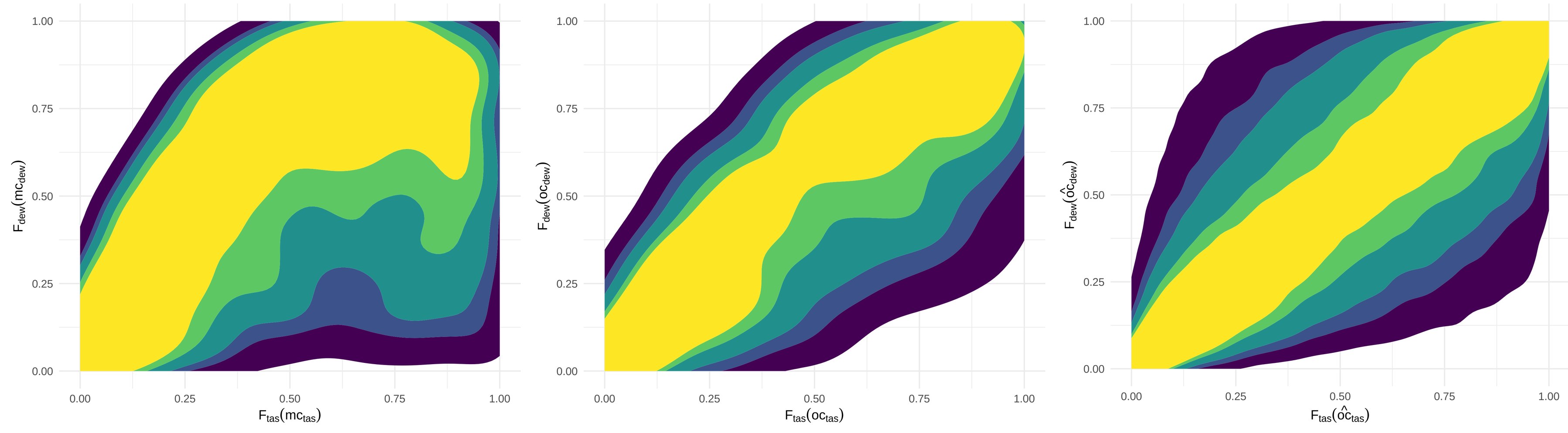
(a) empirical tails oc (b) empirical tails $\hat{mc}_{VBC}$ (c) empirical tails $\hat{mc}_{UBC}$



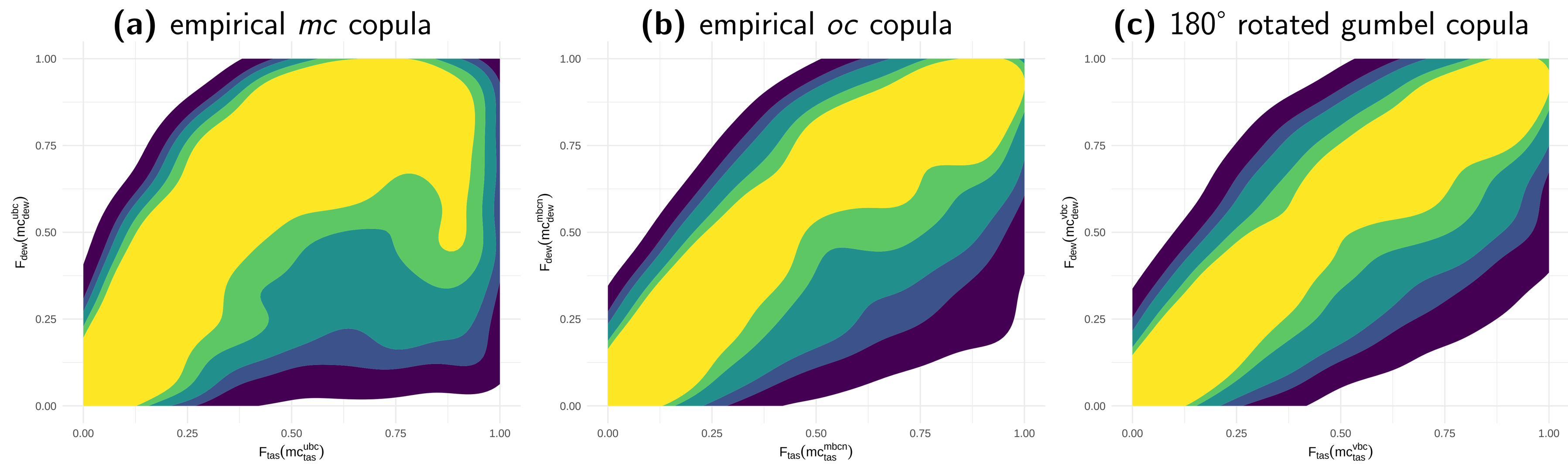
(d) empirical tail probabilities λ over all 50 CRCM5 members

Figure 4: *Depicts tail dependencies between* **tas** *and* **dew***: upper tail dependencies λ_U in red and lower tail dependencies λ_L in blue. Panels* **(a)** *to* **(c)** *illustrate scatterplots of the marginal joint distribution for observed data oc, multivariate correction $\hat{mc}_{VBC}$, and univariate correction $\hat{mc}_{UBC}$. The multivariate correction in* **(b)** *effectively reproduces the joint distribution of oc in* **(a)** *shape, whereas the univariate approach in* **(c)** *inadequately captures the distribution, particularly underrepresenting the tails. Panel* **(d)** *presents the boxplots of tail dependencies across all members. Consistent with* **(c)***, UBC (right) undervalues the tails. While MBCn (lower left) improves representation, it remains an underestimate. VBC (mid), despite its higher variance, more accurately captures the tails in median terms.*

Descriptive evaluation of the relationship between tas and dew



(a) empirical mc copula (b) empirical oc copula (c) 180° rotated gumbel copula

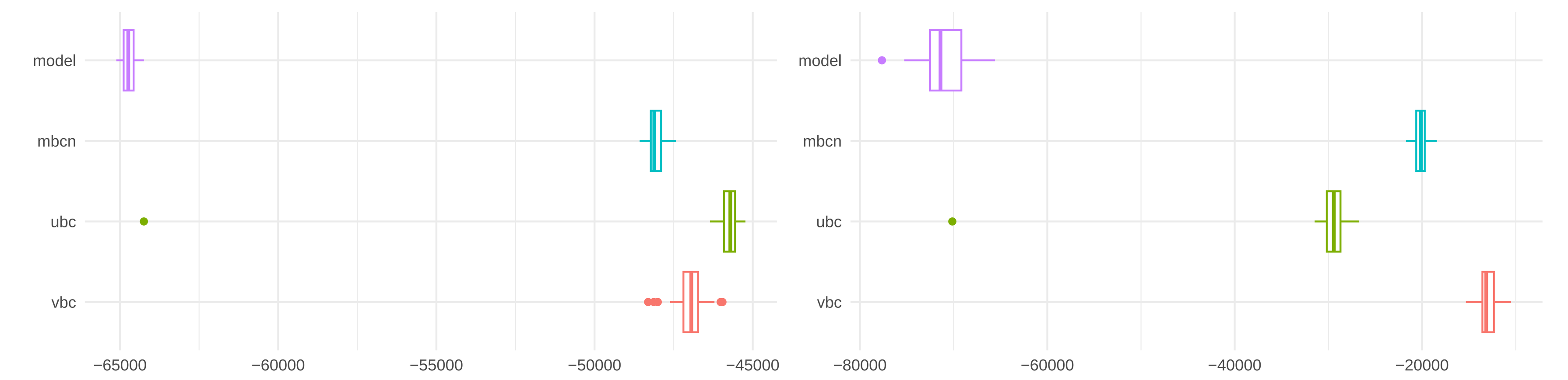


(d) empirical $\hat{mc}_{UBC}$ copula (e) empirical $\hat{mc}_{MBCn}$ copula (f) empirical $\hat{mc}_{VBC}$ copula

Figure 5: *Depicts the summer relationship between* **tas** *and* **dew** *in Mittenwald using the empirical copulas of mc* **(a)***, oc* **(b)***, $\hat{mc}_{UBC}$* **(d)***, $\hat{mc}_{MBCn}$* **(e)***, $\hat{mc}_{VBC}$* **(f)** *and the parametric copula $\hat{oc}$ by their high-density regions (HDR $\in \{0.6, 0.75, 0.9, 0.95, 0.99\}$). Compared to oc in* **(b)** *and* **(c)***, the empirical copula for mc* **(a)** *spans a larger area within the hypercube $[0,1]^2$. This suggests a diminished correlation, particularly evident in regions of lower* **tas** *and elevated* **dew***. Notably, the univariate correction UBC in panel* **(d)** *fails to rectify this relationship. Only the multivariate methods in* **(e)** *and* **(f)** *succeed in addressing this discrepancy, revealing more realistic bivariate rank correlations.*

Modelling - Infer the global likelihood of the model

1. Model the (parametric) copula of oc, $\{(C_e^{(oc)}, \hat{\theta}_e^{(oc)}) \mid \forall e \in E\}$ (see Fig. 3)
2. Evaluate the log-likelihood of $\hat{mc}$ given oc to check if $\hat{mc} \sim F^{(oc)}$
   - 2.1 for univariate margins: $\mathcal{L}(x^{(\hat{mc})} \mid \hat{f}^{(oc)}) := \sum_{i\in U} ln(\hat{f}_i^{(oc)}(x_i^{(\hat{mc})}))$
   - 2.2 and bivariate margins separately: $\mathcal{L}(u^{(\hat{mc})} \mid (C^{(oc)}, \hat{\theta}^{(oc)})) := \sum_{e\in E} ln(c_e^{(oc)}(u_e^{(\hat{mc})} \mid \hat{\theta}_e^{(oc)}))$



(a) univariate log-likelihood $\mathcal{L}(x^{(\hat{mc})} \mid \hat{f}^{(oc)})$ (b) bivariate log-likelihood $\mathcal{L}(u^{(\hat{mc})} \mid (C^{(oc)}, \hat{\theta}^{(oc)}))$

Figure 6: *In Panel* **(a)***, the likelihood of all univariate margins is depicted. All correction methodologies enhance the original model data's fit, with the univariate correction showing the most improvement. MBCn appears to exert more influence on the original rank structure compared to VBC. Panel* **(b)** *delineates the likelihood across all bivariate margins. Given the modeled copulas, VBC optimally refines the model data fit, while UBC offers the minimal enhancement.*

References

Cannon, A. J., Sobie, S. R., & Murdock, T. Q. (2015). Bias correction of GCM precipitation by quantile mapping: how well do methods preserve changes in quantiles and extremes?. Journal of Climate, 28(17), 6938-6959.
Cannon, A. J. (2018). Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables. Climate dynamics, 50, 31-49.
Czado, C. (2019). Analyzing Dependent Data with Vine Copulas. In Lecture Notes in Statistics. Springer International Publishing. https://doi.org/10.1007/978-3-030-13785-4.
Dissmann, J., Brechmann, E. C., Czado, C., & Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis, 59, 52-69.
Leduc, M., Mailhot, A., Frigon, A., Martel, J. L., Ludwig, ... & Scinocca, J. (2019). The ClimEx project: A 50-member ensemble of climate change projections at 12-km resolution over Europe and northeastern North America with the Canadian Regional Climate Model (CRCM5). Journal of Applied Meteorology and Climatology, 58(4), 663-693.
Wood, R. R., Willkofer, F., Schmid, F. J., Trentini, F., Komischke, H., & Ludwig, R. (2017, April). SDCLIREF-A sub-daily gridded reference dataset. In EGU General Assembly Conference Abstracts (p. 15739).