ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR-
UND MEERESFORSCHUNG

Introduction of the BiasAdjustCXX bias correction command-line tool (... and "python-cmethods")

Alfred Wegener Institute for Polar and Marine Research

Benjamin Thomas Schwertfeger

benjamin.schwertfeger@awi.de
development@b-schwertfeger.de
https://b-schwertfeger.de
https://github.com/btschwertfeger

January 19, 2023

Content

1. Introduction

Motivation

Within my bachelor thesis, different correction methods were implemented, applied and evaluated to gain a better understanding of the impact on the data, as previous analyses were based on data whose creation process was not clearly traceable.

Achivements

- implementation, application, and analysis of different bias correction methods and their effects on modeled climate data
- development of two independand tools for adjusting climate data in C++ and Pyhton

1.2 What are bias corrections?

Statistical methods/techniques to adjust for deviations in distribution characteristics between modeled and observed time series climate data. (Can be applied to both past and future time periods)

When are these procedures applied?

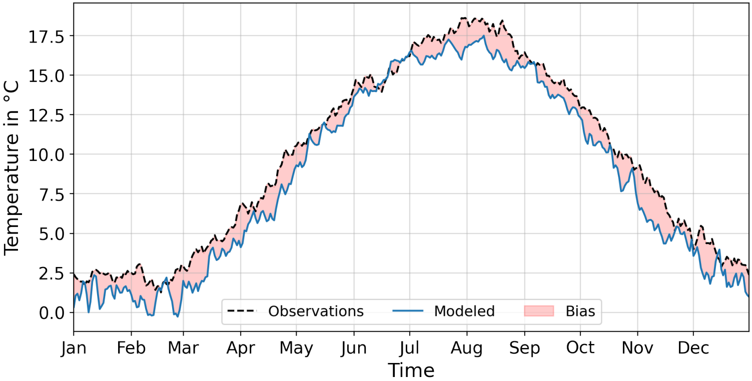
post-processing modeled climate data for climate studies, threshold analyses, model optimizations, risk assessments, ...

What can be approximated using bias corrections?

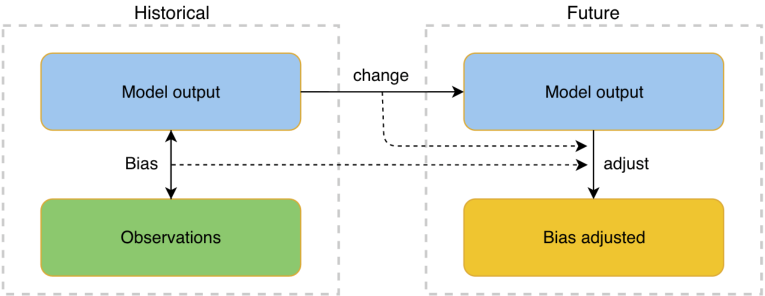
mean, variance, density, and also correlation

1.3 How does bias look like?

Figure: mean air temperatures at 2 m height per day per year in Bremerhaven between 1981-2010



Schematic bias correction overview



Assumption: There is an existing relationship between the generated data of modeled climate scenarios and the real, actual climatic processes and conditions, and that topographic influences in observed, reanalyzed as well as modeled time series are also related to each other.

2. Methods

- scaling-based methods
- distribution-based methods
- ... machine learning

Requirements

- 1x modeled time series & 1x observed/reanalyzed time series of the control period (e.g. 1951-1980)
- 1x modeled time series of the period to adjust (e.g. 1981-2010)

Methods and references

Scaling-based methods

- Delta Method +/* (Beyer et al., 2020)
- Linear Scaling +/* (Teutschbein and Seibert, 2012)
- Variance Scaling + (Teutschbein and Seibert, 2012)

Additional variation: custom long-term 31-day moving window variant

Distribution-based methods

- Quantile Mapping +/* (Tong et al., 2021)
- Quantile Delta Mapping +/* (Cannon et al., 2015)

Additive variant is always shown, for the multiplicative variant, replace + and - with * and /

2.1 Delta Method

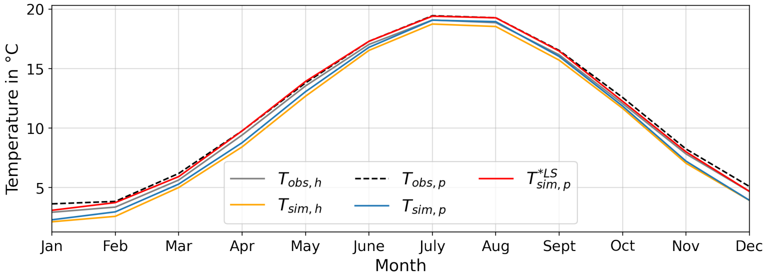
Adjustment of long-term (a) monthly or (b) long-term 31-day moving window means using the difference between the modeled time series:

$$
\begin{aligned}
T^{*DM}_{sim,p}(i) &= T_{obs,h}(i) + \Big(\mu_m\big(T_{sim,p}(i)\big) - \mu_m\big(T_{sim,h}(i)\big)\Big) && (1)\\
&= 2.93°\mathrm{C} + (2.44°\mathrm{C} - 3.0°\mathrm{C}) && (2)\\
&= 2.93°\mathrm{C} + (-0.56°\mathrm{C}) && (3)\\
&= 2.37°\mathrm{C} && (4)
\end{aligned}
$$

2.2 Linear Scaling

Adjustment of long-term (a) monthly or (b) long-term 31-day moving window means using the difference between the time series of the control period:

$$T^{*LS}_{sim,p}(i) = T_{sim,p}(i) + \Big(\mu_m\big(T_{obs,h}(i)\big) - \mu_m\big(T_{sim,h}(i)\big)\Big) \tag{5}$$



2.3 Variance Scaling

1. Shift mean to zero:

$$T_{sim,h}^{VS(1)}(i) = T_{sim,h}^{*LS}(i) - \mu_m\left(T_{sim,h}^{*LS}(i)\right) \tag{6}$$

$$T_{sim,p}^{VS(1)}(i) = T_{sim,p}^{*LS}(i) - \mu_m\left(T_{sim,p}^{*LS}(i)\right) \tag{7}$$

2. Adjust/scale the standard deviation (so variance is scaled too):

$$T_{sim,p}^{VS(2)}(i) = T_{sim,p}^{VS(1)}(i) \cdot \left[\frac{\sigma_m\left(T_{obs,h}(i)\right)}{\sigma_m\left(T_{sim,h}^{VS(1)}(i)\right)}\right] \tag{8}$$

2.3 Variance Scaling

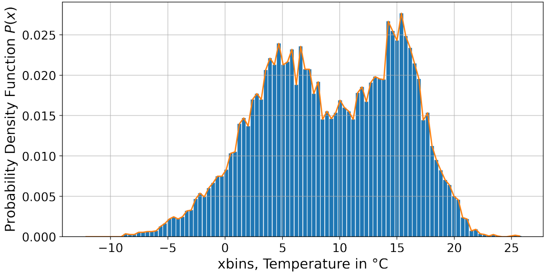
3. Get the mean back:

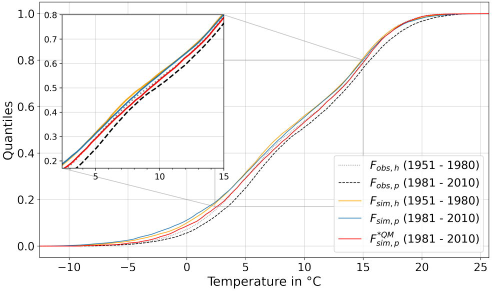
$$T^{*VS}_{sim,p}(i) = T^{VS(2)}_{sim,p}(i) + \mu_m\left(T^{*LS}_{sim,p}(i)\right) \tag{9}$$

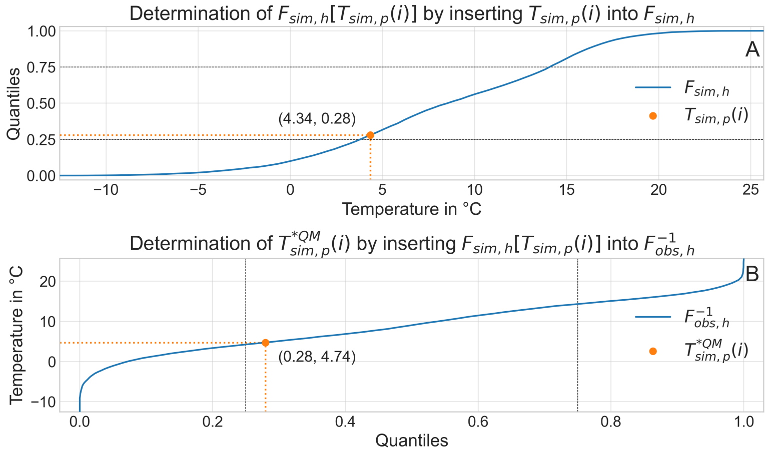
Note: No multiplicative variant for VS

2.4 Quantile Mapping

$$T^{*QM}_{sim,p}(i) = F^{-1}_{obs,h}\left\{F_{sim,h}\left[T_{sim,p}(i)\right]\right\} \quad (10)$$






Determination of $F_{sim,h}[T_{sim,p}(i)]$ by inserting $T_{sim,p}(i)$ into $F_{sim,h}$
A
Quantiles
Temperature in °C
(4.34, 0.28)
$F_{sim,h}$
$T_{sim,p}(i)$
Determination of $T^{*QM}_{sim,p}(i)$ by inserting $F_{sim,h}[T_{sim,p}(i)]$ into $F^{-1}_{obs,h}$
B
Temperature in °C
Quantiles
(0.28, 4.74)
$F^{-1}_{obs,h}$
$T^{*QM}_{sim,p}(i)$

2.5 Quantile Delta Mapping

1. Adjustment only on the basis of the observations of the control period

$$\varepsilon(i) = F_{sim,p}\left[T_{sim,p}(i)\right], \quad \varepsilon(i) \in \{0,1\} \tag{11}$$

$$T_{sim,p}^{QDM(1)}(i) = F_{obs,h}^{-1}\left[\varepsilon(i)\right] \tag{12}$$

2. Compute the delta:

$$\Delta(i) = F_{sim,p}^{-1}\left[\varepsilon(i)\right] - F_{sim,h}^{-1}\left[\varepsilon(i)\right] \tag{13}$$

$$= T_{sim,p}(i) - F_{sim,h}^{-1}\left[\varepsilon(i)\right] \tag{14}$$

2.5 Quantile Delta Mapping

3. Add the delta:

$$T_{sim,p}^{*QDM}(i) = T_{sim,p}^{QDM(1)}(i) + \Delta(i) \tag{15}$$

3. Results

Data

- CMIP6.CMIP.MPI-M.MPI-ESM1-2-HR.historical
- NOAA/CIRES/DOE 20th Century Reanalysis V3
- data sets contain daily mean 2 m air temperatures

Periods

- reference/control period: 1951-1980
- scenario period: 1981-2010

Region und resolution

- Region: Europe from -45°W to 65°E, from 23°N to 72°N
- Resolution: 1.0° x 1.0°

3.1 Mean Bias Error

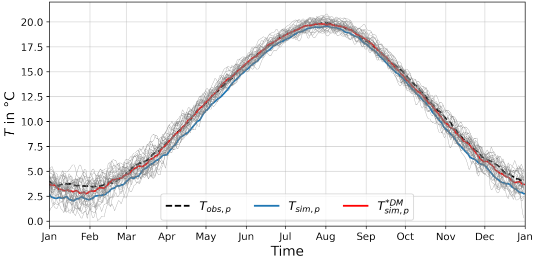


Figure 1: Mean 2 m air temperatures per day of year in observed, modeled and delta method-adjusted data in Europe (1981-2010)

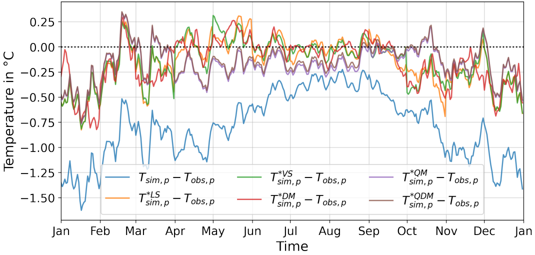
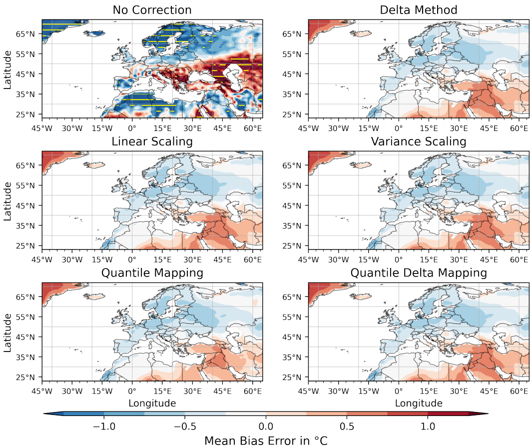
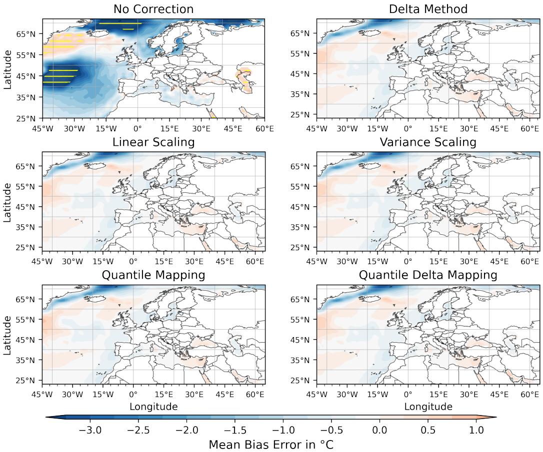


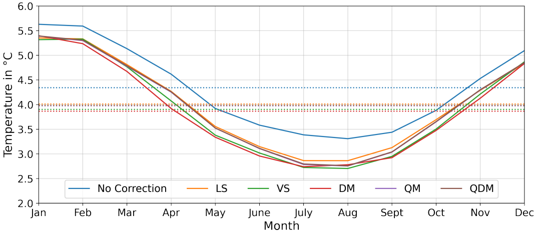
Figure 2: Mean anomaly between modeled, adjusted and observed 2 m air temperatures in Europe (1981-2010)

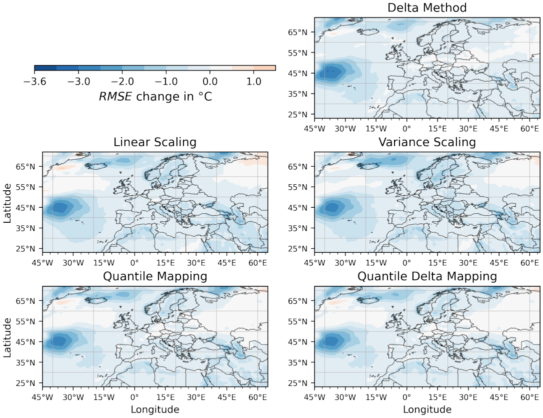

No Correction
Delta Method
Linear Scaling
Variance Scaling
Quantile Mapping
Quantile Delta Mapping
Latitude
Longitude
65°N
55°N
45°N
35°N
25°N
45°W 30°W 15°W 0° 15°E 30°E 45°E 60°E
−1.0 −0.5 0.0 0.5 1.0
Mean Bias Error in °C


No Correction
Delta Method
Linear Scaling
Variance Scaling
Quantile Mapping
Quantile Delta Mapping
Latitude
Longitude
65°N
55°N
45°N
35°N
25°N
45°W 30°W 15°W 0° 15°E 30°E 45°E 60°E
−3.0 −2.5 −2.0 −1.5 −1.0 −0.5 0.0 0.5 1.0
Mean Bias Error in °C

3.2 Root Mean Square Error

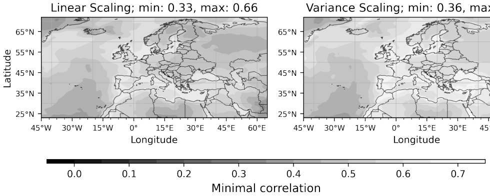
Figure 3: Mean *RMSE* per month

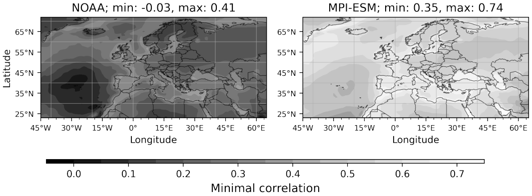


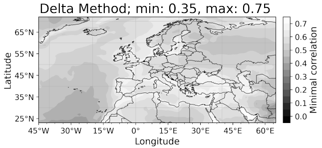


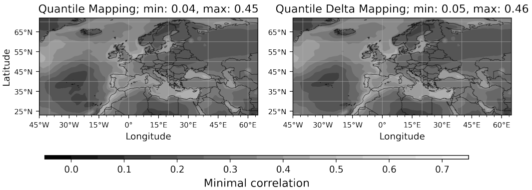
3.3 Minimum correlation matrices

Minimum correlation matrices on the basis of monthly means between 1981-2010









4. Implementation and modules

Repositories and installation

- BiasAdjustCXX command-line tool (https://github.com/btschwertfeger/BiasAdjustCXX)

```
git clone https://github.com/btschwertfeger/BiasAdjustCXX.git
cd BiasAdjustCXX && mkdir build && cd build
cmake .. && cmake --build .
```

- python-cmethods module (https://github.com/btschwertfeger/python-cmethods)

```
python3 -m pip install python-cmethods
```

python-cmethods

```
from cmethods.CMethods import CMethods; import xarray as xr
cm = CMethods()
obsh = xr.open_dataset('observations_historical.nc')['tas']
simh = xr.open_dataset('modeled_historical.nc')['tas']
simp = xr.open_dataset('scenario.nc')['tas']

qdm_res = cm.quantile_delta_mapping(
    obs = obsh, simh = simh, simp = simp,
    kind = '+', n_quantiles = 100
); print(qdm_res)

> array([-1.12788787, -1.53094775, -2.3066464 , ..., 3.70744536,
         4.89798346, 4.84614564])
```

BiasAdjustCXX

```
BiasAdjustCXX                 \
   --ref observations.nc \ # referenece data (control period)
   --contr control.nc    \ # modeled data    (control period)
   --scen scenario.nc    \ # data to adjust  (scenario period)
   -o tas_dm_adjusted.nc \ # output file     (path)
   -v tas                \ # variable of interest
   -m delta_method       \ # adjustment method
   -k "+"                  # additive variant
```

Now you ...

- login to ollie (ssh username@ollie0.awi.de)
- execute the following commands:

```
cp -r /home/ollie/bschwert/public/HandsOnBiasAdjust ~/
cd ~/HandsOnBiasAdjust
module load python3
jupyter notebook --no-browser --ip=0.0.0.0
```

- And open the notebook using the displayed URL

Help: https://spaces.awi.de/display/HELP/Jupyter+Notebook

References I

Beyer, R., Krapp, M., & Manica, A. (2020). An empirical evaluation of bias correction methods for palaeoclimate simulations. *Climate of the Past*, *16*(4), 1493–1508. https://doi.org/10.5194/cp-16-1493-2020

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. https://doi.org/10.1175/JCLI-D-14-00754.1

Teutschbein, C., & 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. https://doi.org/10.1016/j.jhydrol.2012.05.052

Tong, Y., Gao, X., Han, Z., Xu, Y., Xu, Y., & Giorgi, F. (2021). Bias correction of temperature and precipitation over china for rcm simulations using the qm and qdm methods. *Climate Dynamics*, *57*(5), 1425–1443. https://doi.org/10.1007/s00382-020-05447-4