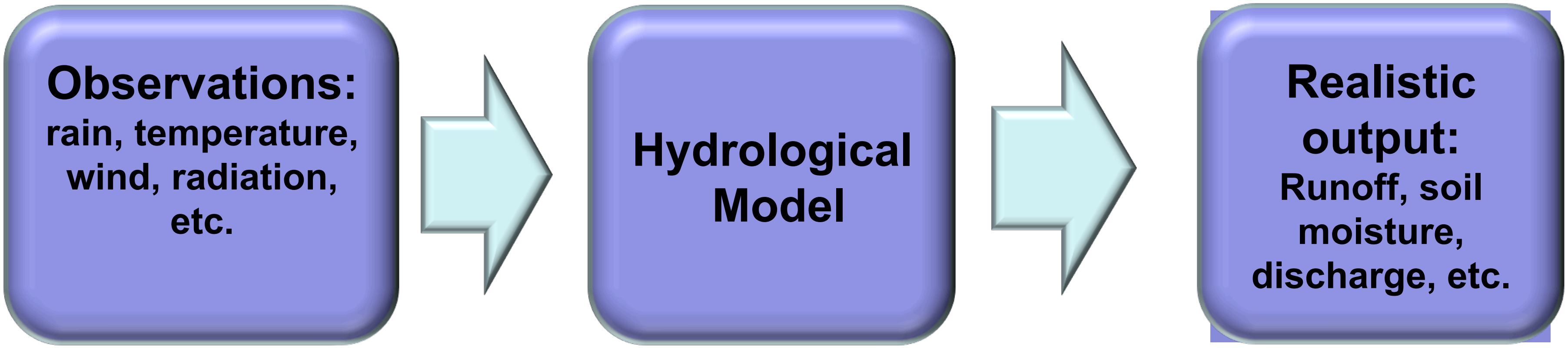
Statistical Bias Correction of hydrological forcing fields from GCMs: *basic concepts*.

C. Piani

Why do we bias correct GCM output before using it to force hydrological models?



Force a hydrological model, that performs well when forced with observations, with unprocessed GCM output and you don't get an acceptable result…

Why?

•Gridded precipitation from CGMs is not the same physical variable as the observed:

- •Temporal and spatial averaging.
- •Under-catch corrections
- •Sampling error
- •Other?

•The GCM daily temperature cycle is physically closer to the observed but there are still differences:

- •Temporal and spatial averaging
- •Ground effects
- •Evapotranspiration terms are very sensitive to temperature

Before statistical bas correction

- The difference in simulated present day and future climate was calculated.

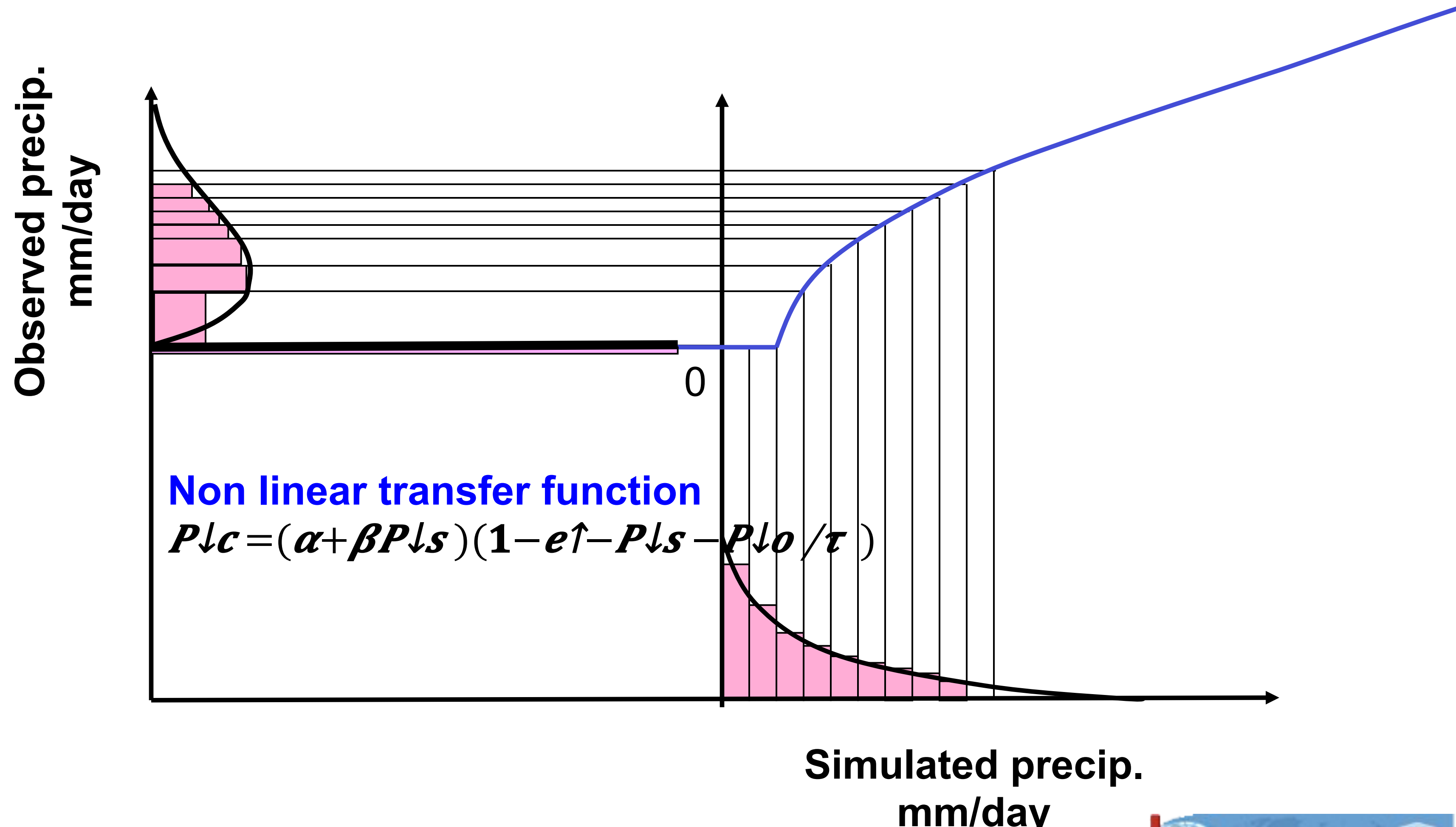
- Then it was simply added to present day climate.

- This is tantamount to applying an ADDITIVE BIAS CORRECTION and improves only the mean values of the climate variables.

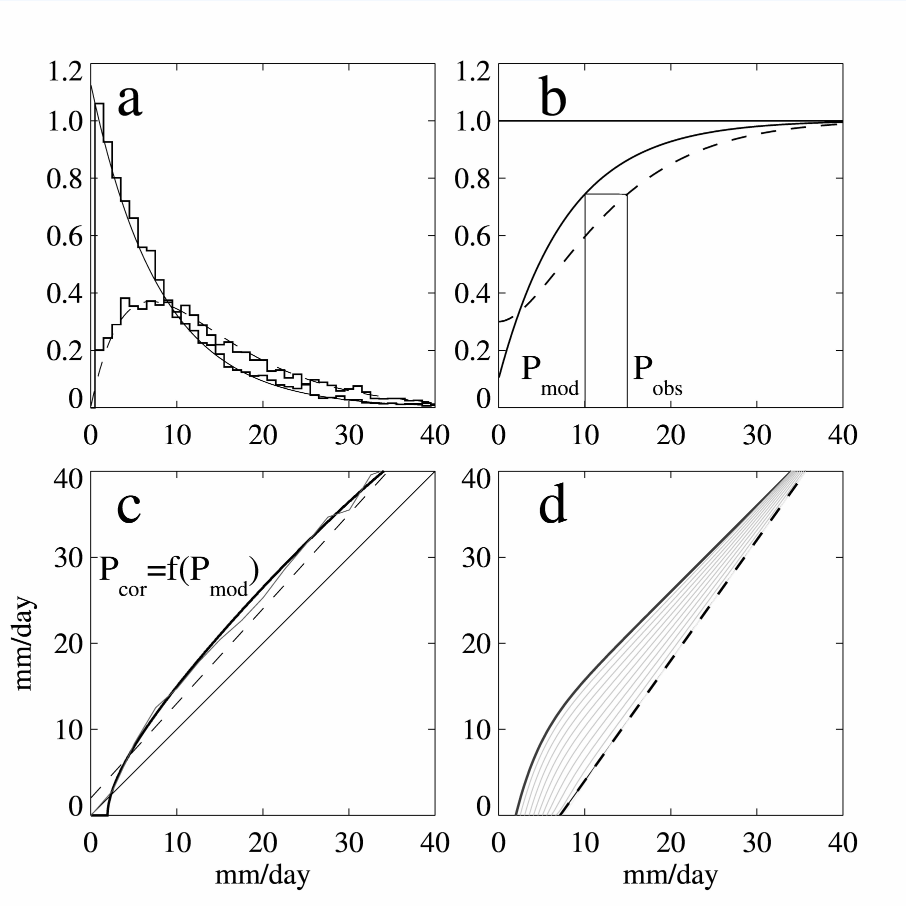
- **The statistical 'histogram matching' bias correction method developed here potentially corrects all moments of the statistical distribution.**

- **It uses all available information from both simulations and observations.**

How histogram matching works



Deriving the transfer function



Yes it does work:

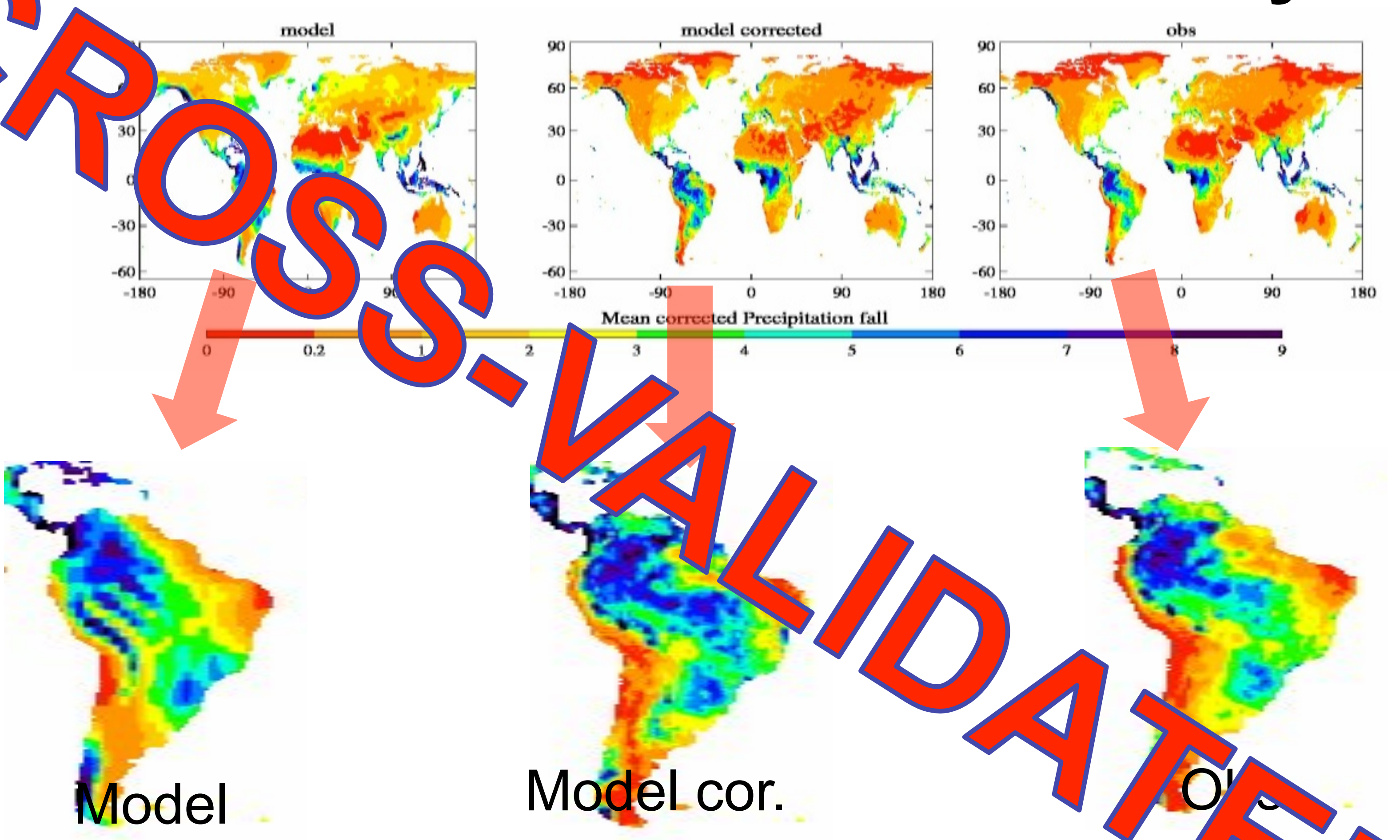
a) Idealized histograms of simulated (solid line) and observed (dashed line) daily precipitation.

b) Cumulative distributions.

c) Transform function. Is determined by few (< 3) parameters.

d) Transitional daily transform functions

Does the method work?… well yes.



CROSS-VALIDATE!!!

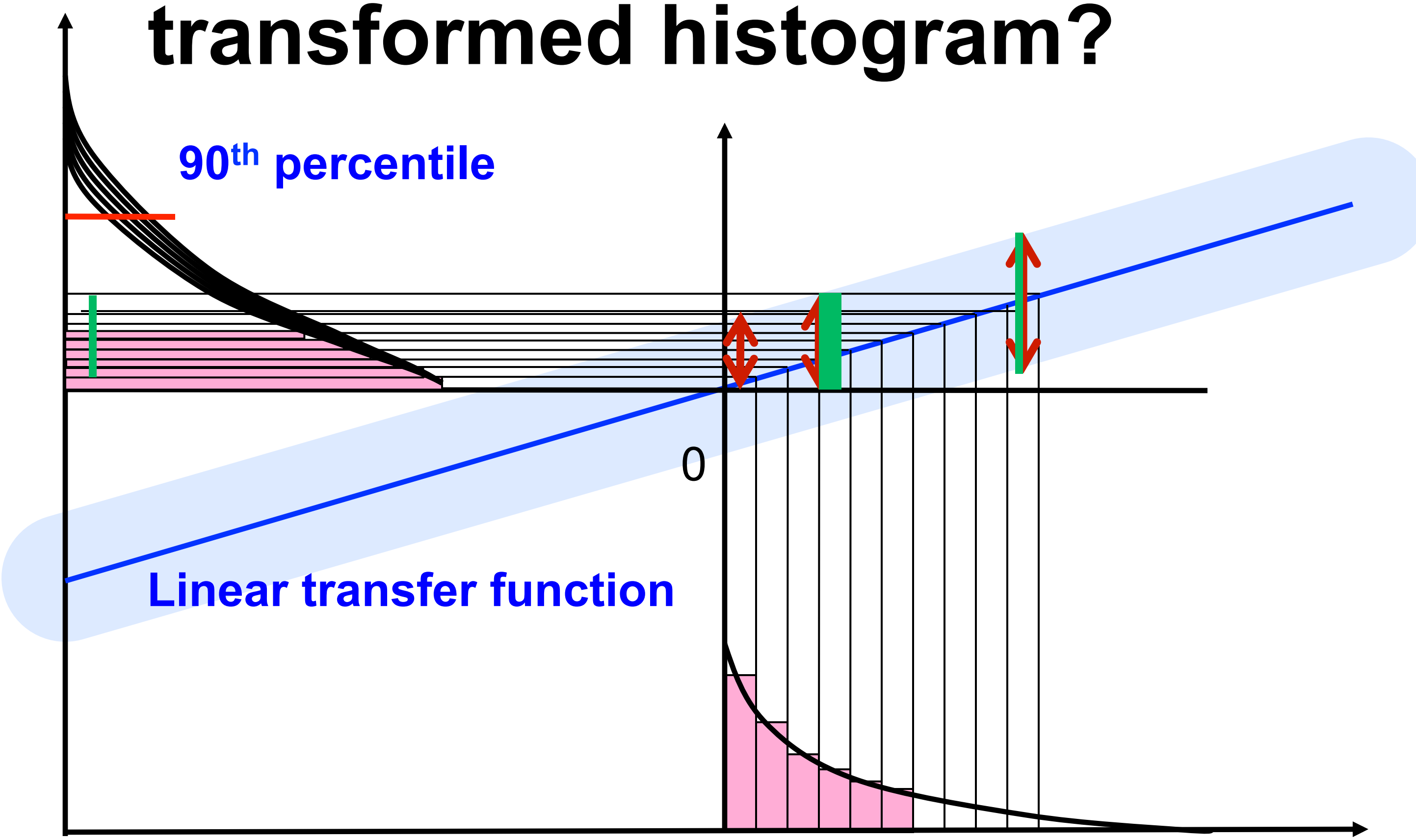
1990-2000 January precipitation over South America corrected using 1960-1970 transfer function.

Uncertainty in the bias correction (TF)

•Fits to the transform function are associated with uncertainty from different sources:

- Standard error associated with fit (*negligible*).
- Choice of fitting function (*can be made negligible, trade-off with robustness*).
- Decadal variability of fit parameters (*This is the big one… and this is why you cross-validate!!!!!!*).

How does uncertainty in the TF affect the transformed histogram?



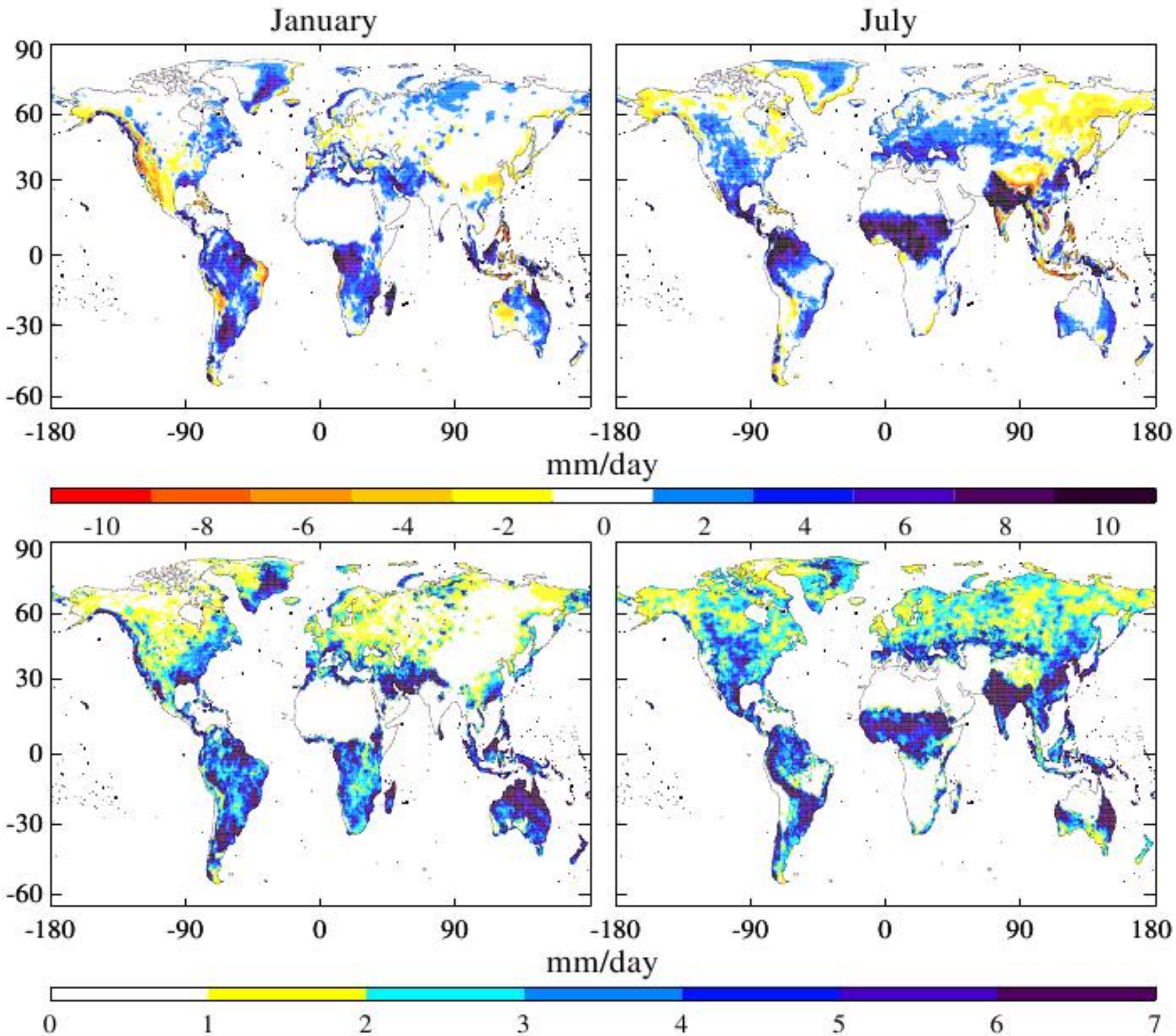
•How can we produce a horizontal mapping of the bias-correction-induced uncertainty? (*ex.: precipitation*):

•Plot the average additive correction for the 90th percentile of the local precipitation intensity distribution in mm/day.

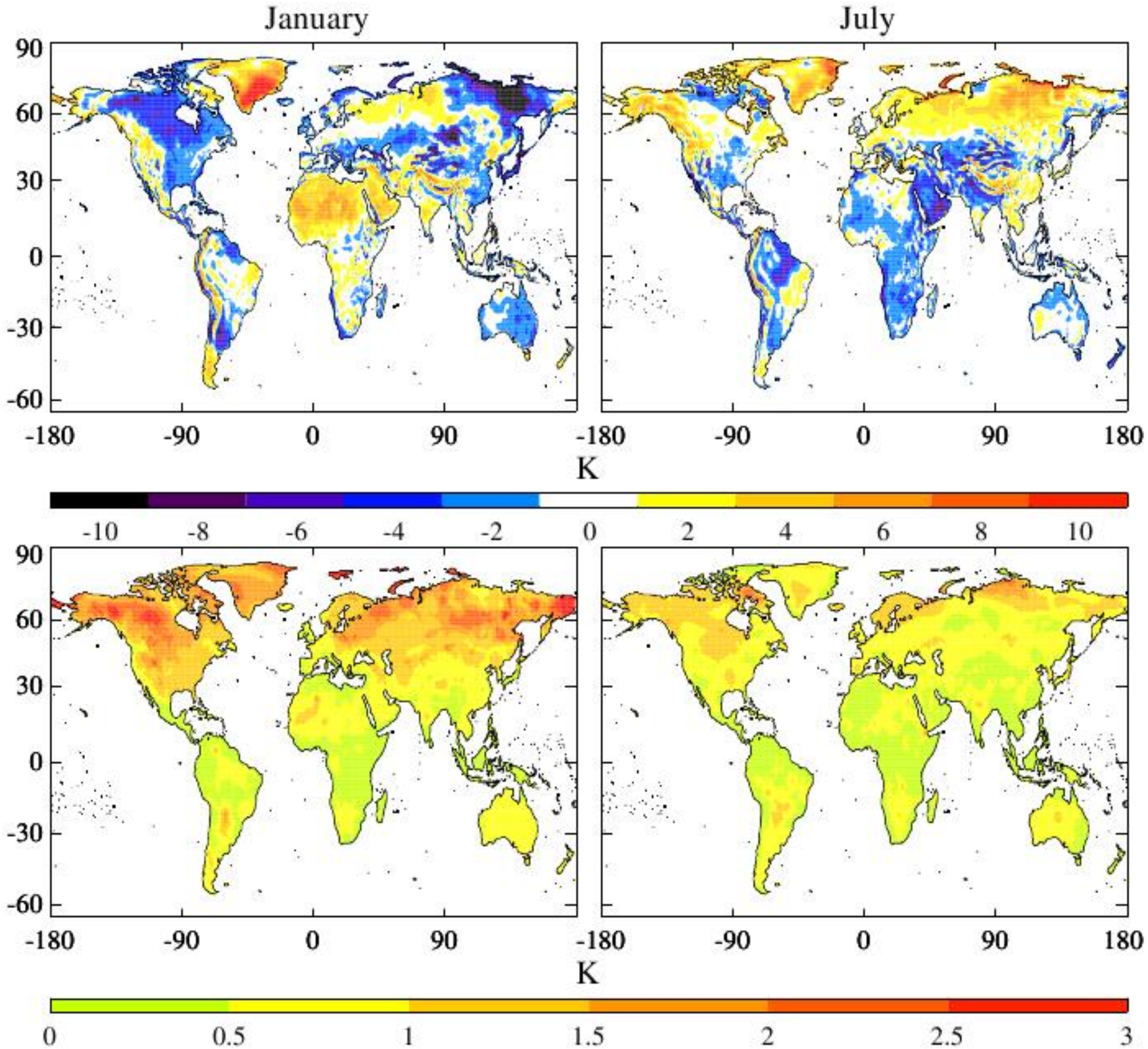
• The average is computed over the 12 separate *TFs* obtained using the 3 members of the ECHAM5 ensemble alternatively with the 4 decadal periods form 1960 to 1999.

• Plot the standard deviation across the 12 *TFs* for the same intensity percentile.

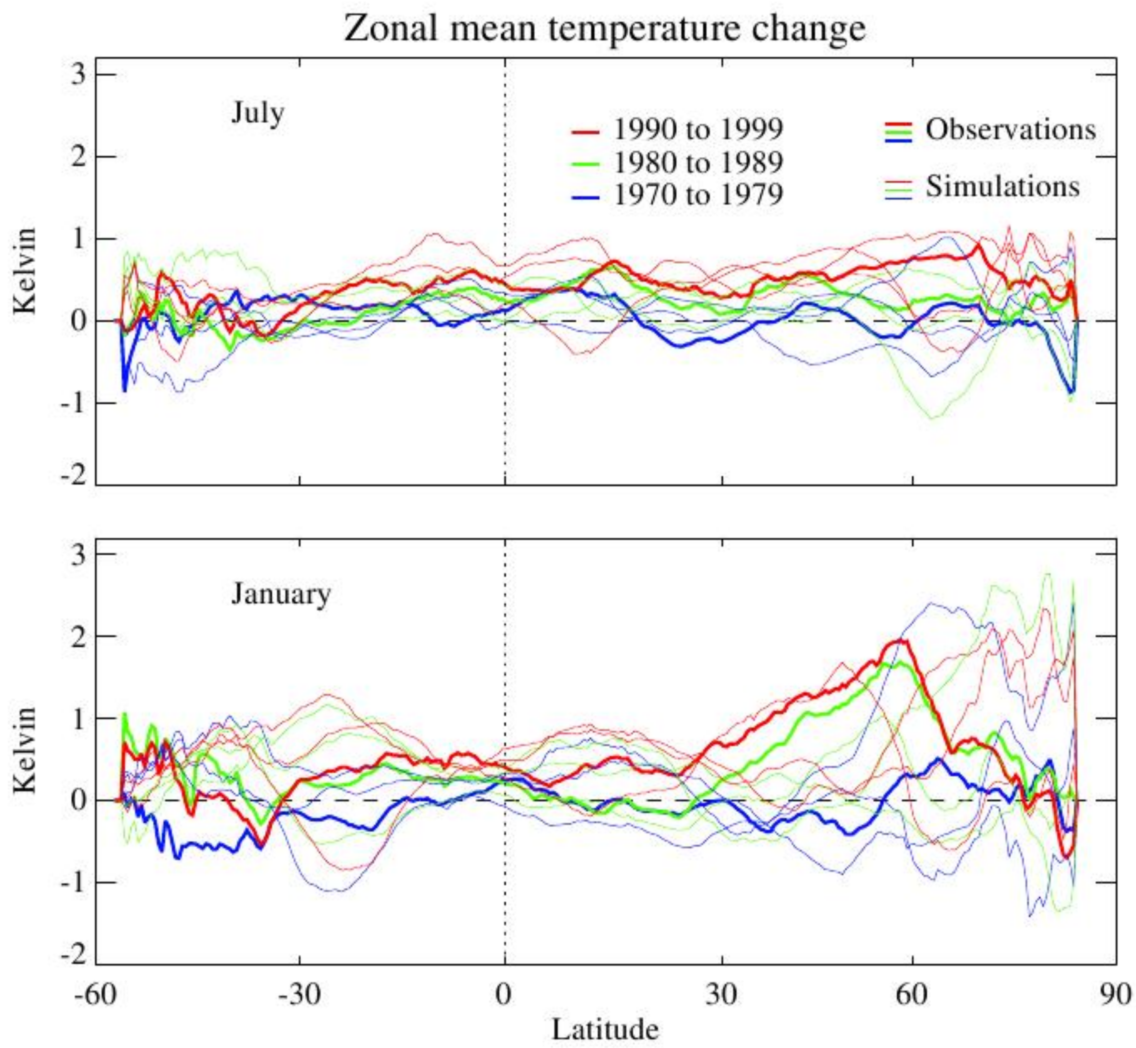
Uncertainty in the bias correction for daily precipitation.



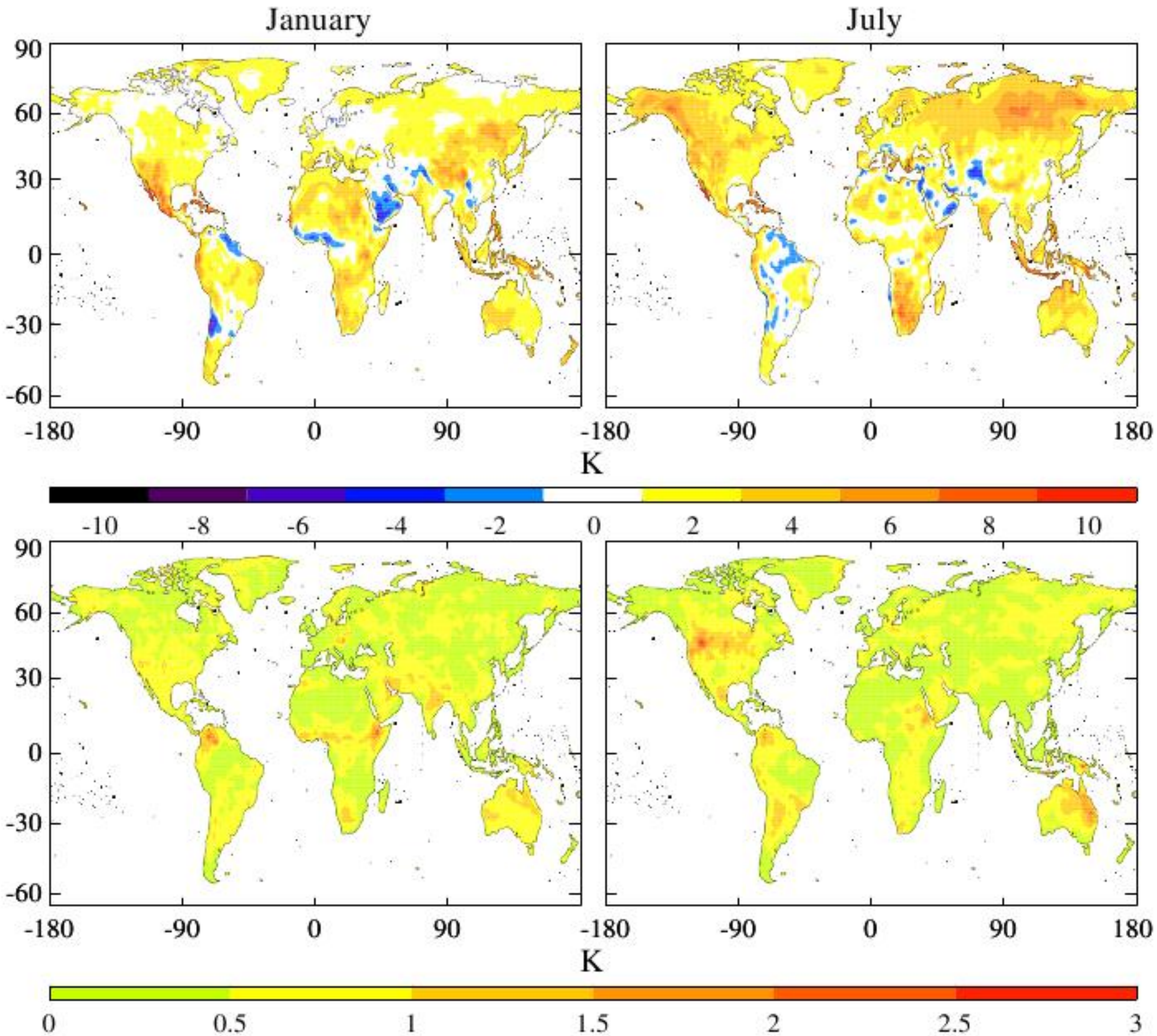
Uncertainty in the bias correction for daily temperarure.



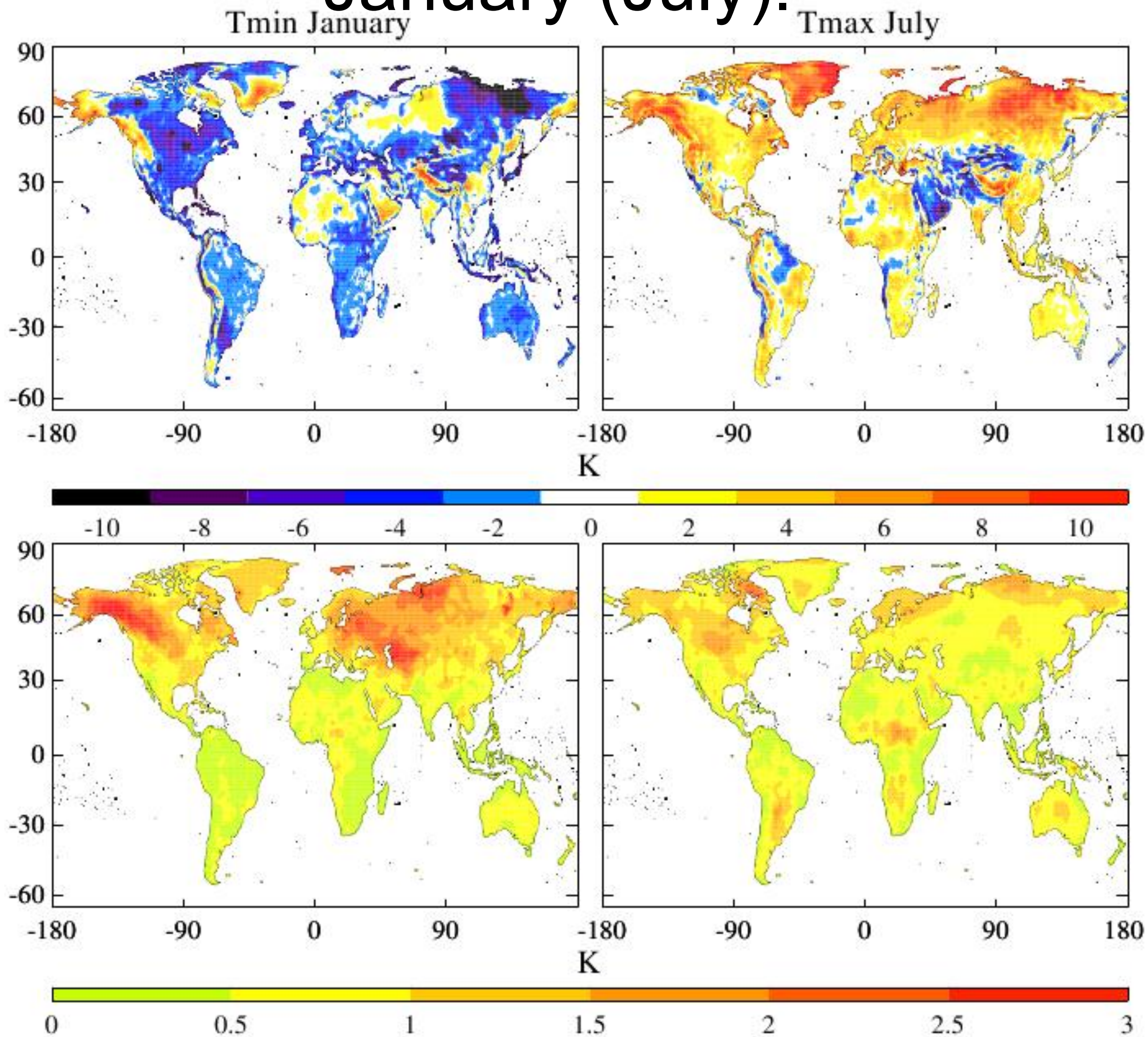
Decadal variability of zonal mean temperature.



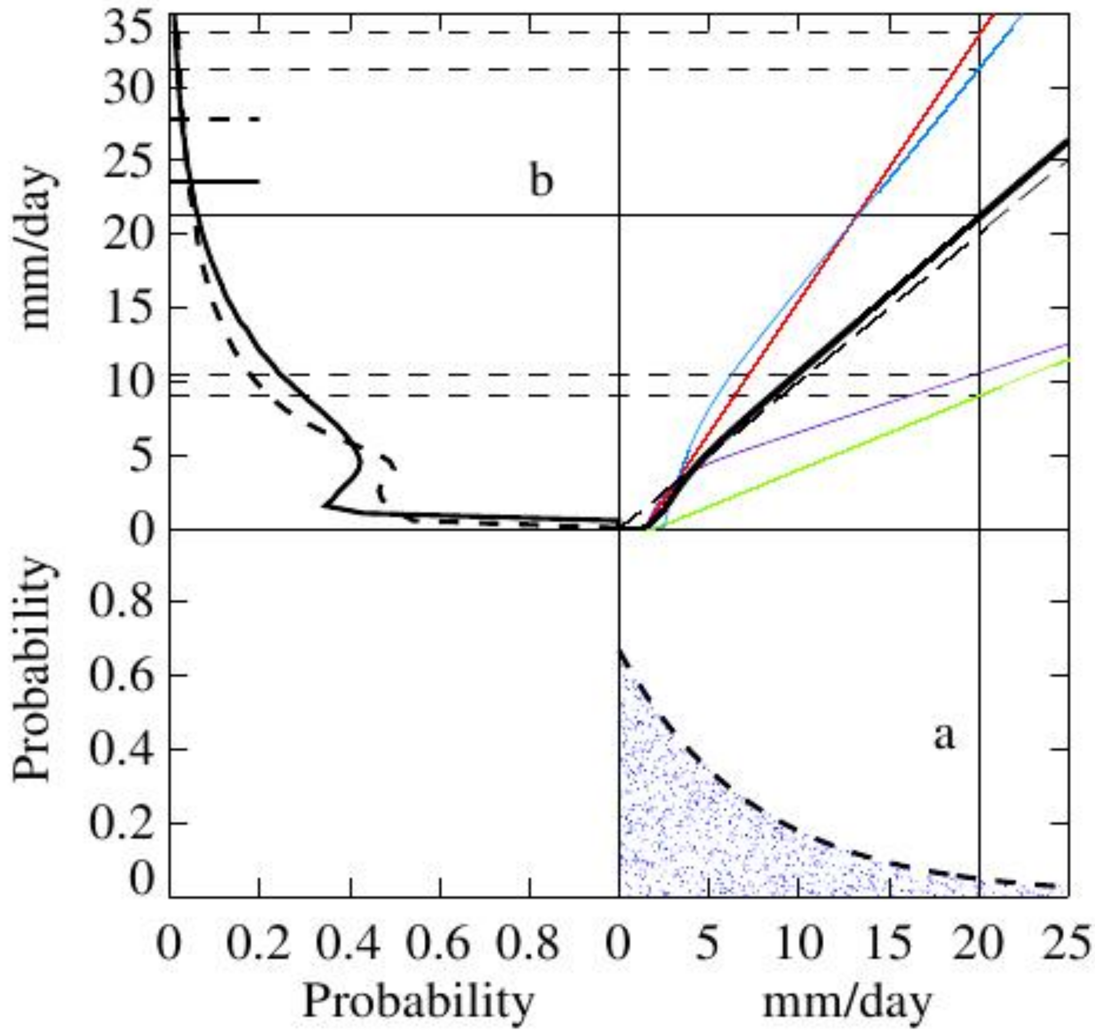
Uncertainty in the bias correction for daily temperature cycle.



Uncertainty in the bias correction for mean daily minimum (maximum) temperatures for the month of January (July).



Accounting for uncertainty in the bias correction.



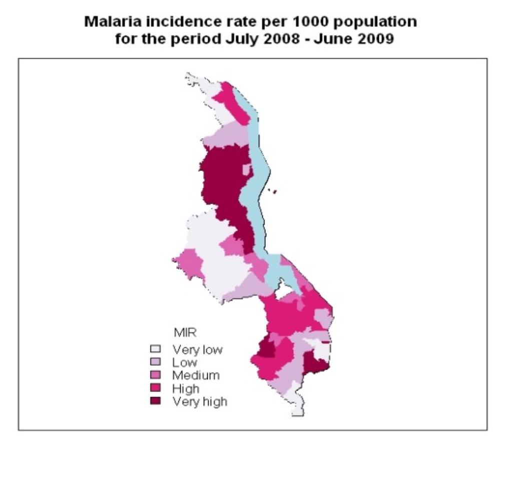
conclusions

•Statistical BC allows all the information in both observations and model to be taken in consideration.

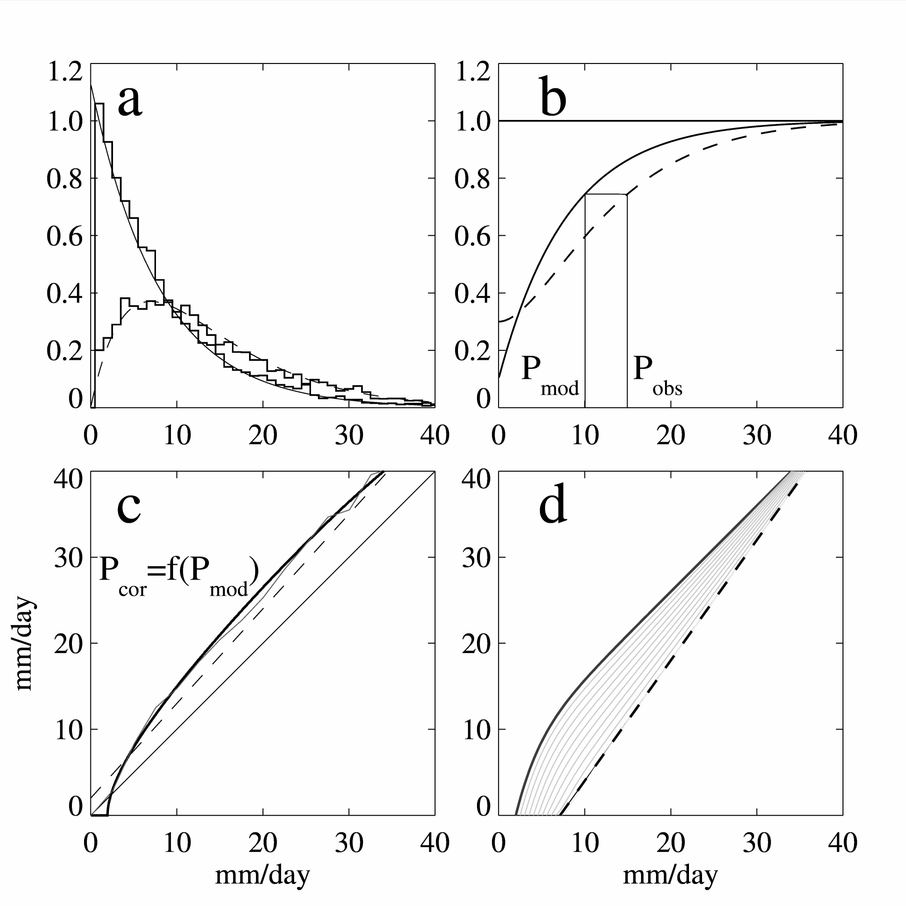
•Cross-validation is essential.

•Using different observational periods one can give qualitative descriptions of the uncertainty associated with the BC.

Statistical Bias Correction: operational challenges and common blunders.



Histogram matching methodology



a) Idealized histograms of simulated (solid line) and observed (dashed line) daily precipitaiton.

b) Cumulative distributions.

c) Transform function. Is determined by few (< 3) parameters.

d) Transitional daily transform functions

What are some of the remaining problems?

•Statistical bias corrections are couched in uncertainty. The difference between Bias and Error depends on the length of the simulation.

•**Suggested solution**: Analysis of transfer function spread.
(Piani et al., 2010b, Chen et al. 2012)

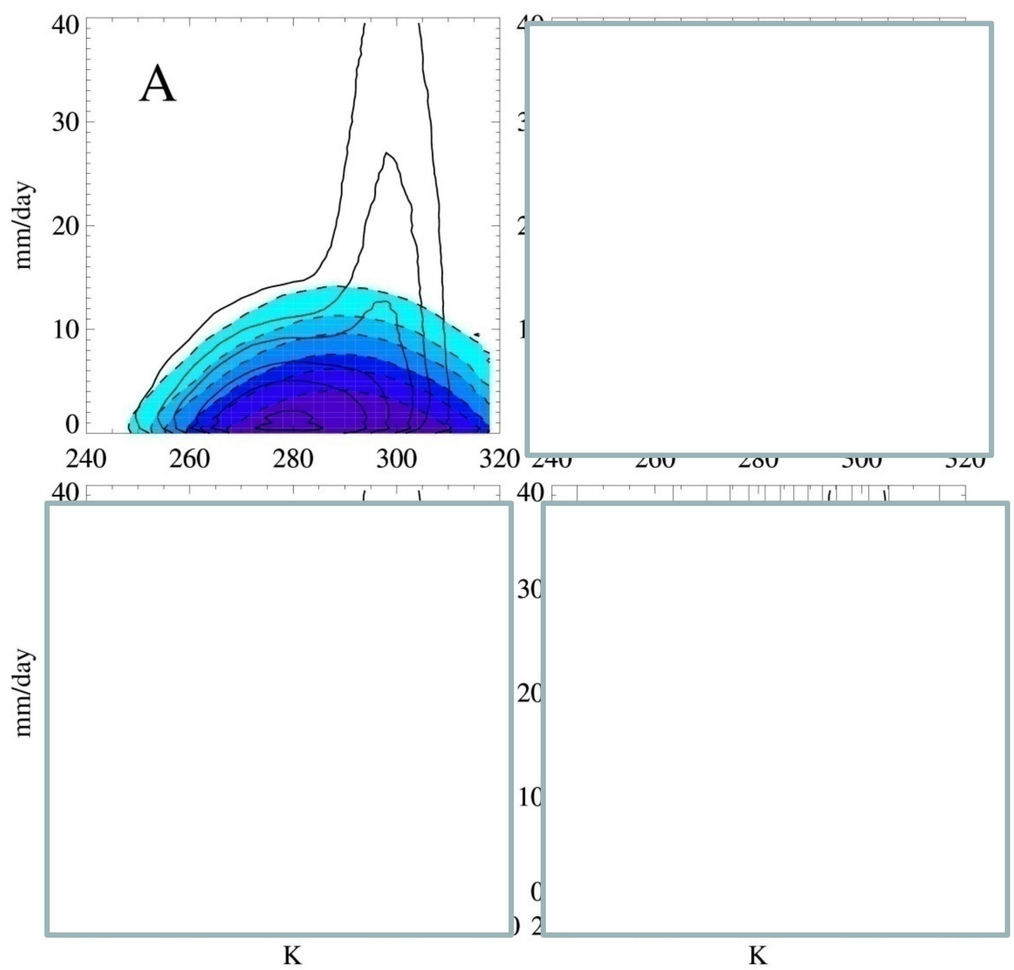
•So far we have corrected temperature and precipitation separately. No improvements are made in the representation of the dynamical relations between the two variables.

•**Suggested solution**: undertake full 2D statistical bias correction.
(Piani and Haerter, 2012)

•Corrections are not independent on time scale: if you correct the daily variance you do not correct the monthly variance.

•**Suggested solution**: the cascade statistical bias correction.
(*Haerter and Piani, 2011*)

2D statistical bias correction of temperature and precipitation (2D histogram matching).



A) Idealized 2D histograms of simulated (colored contours) and observed (solid contours) daily precipitation and temperature.

B) Like **A**, but the simulations have been independently corrected with a linear Transfer Function.

C) Like **B**, but the simulations have been independently corrected with a perfect Transfer Function.

D) Like **B**, but the simulations have been corrected with a 2D linear Transfer Function.

Application 2D statistical bias correction

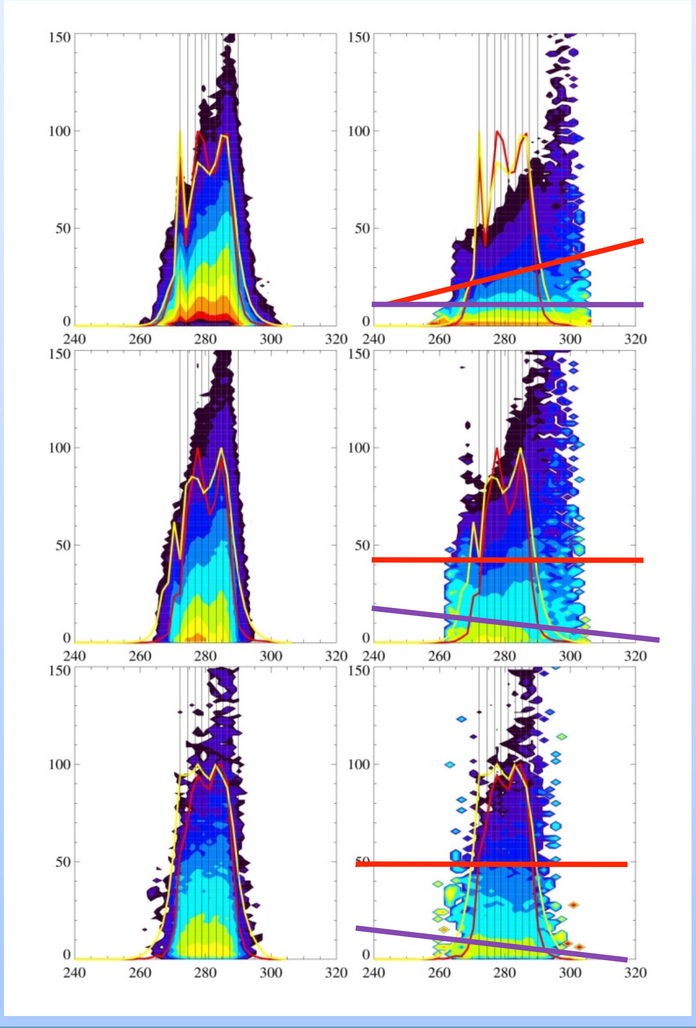
Simulations: Max Planck Institute for Meteorology regional climate model (REMO)

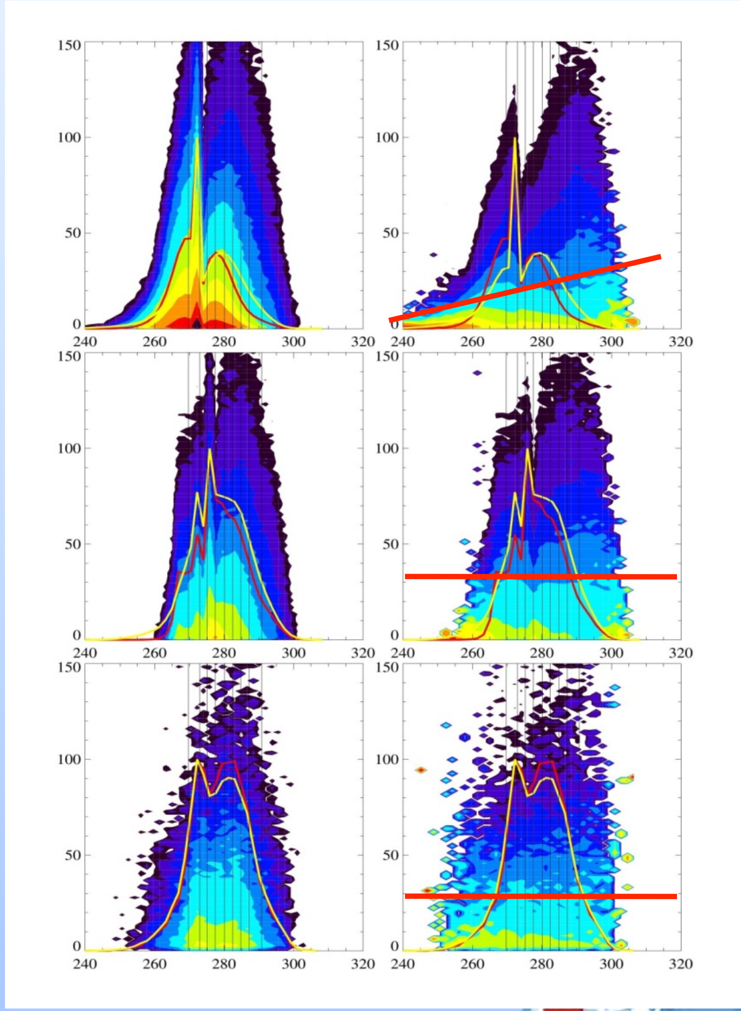
Observations: station data (Kempten, Schleswig).

Application 2D statistical bias correction

Simulations: Max Planck Institute for Meteorology regional climate model (REMO)

Observations: station data (Kempten, Schleswig).





Copulas

- **Definition of 2D Copula**

C: $[0,1]^2 \to [0,1]^2$ is a *2*D **copula** if *C* is a joint cumulative distribution function of a 2D random vector on the unit square with uniform marginals.

- **Definition of 2D Copula probability density function**

- *Cpdf*: $[0,1]^2 \to [0,1]$ is a *2D* **copula probability density function** if *Cpdf* is a joint PDF of a *2*D random vector on the unit square with uniform marginals.

- **Marginal** is the resulting PDF for one of the variables after integrating over the other….

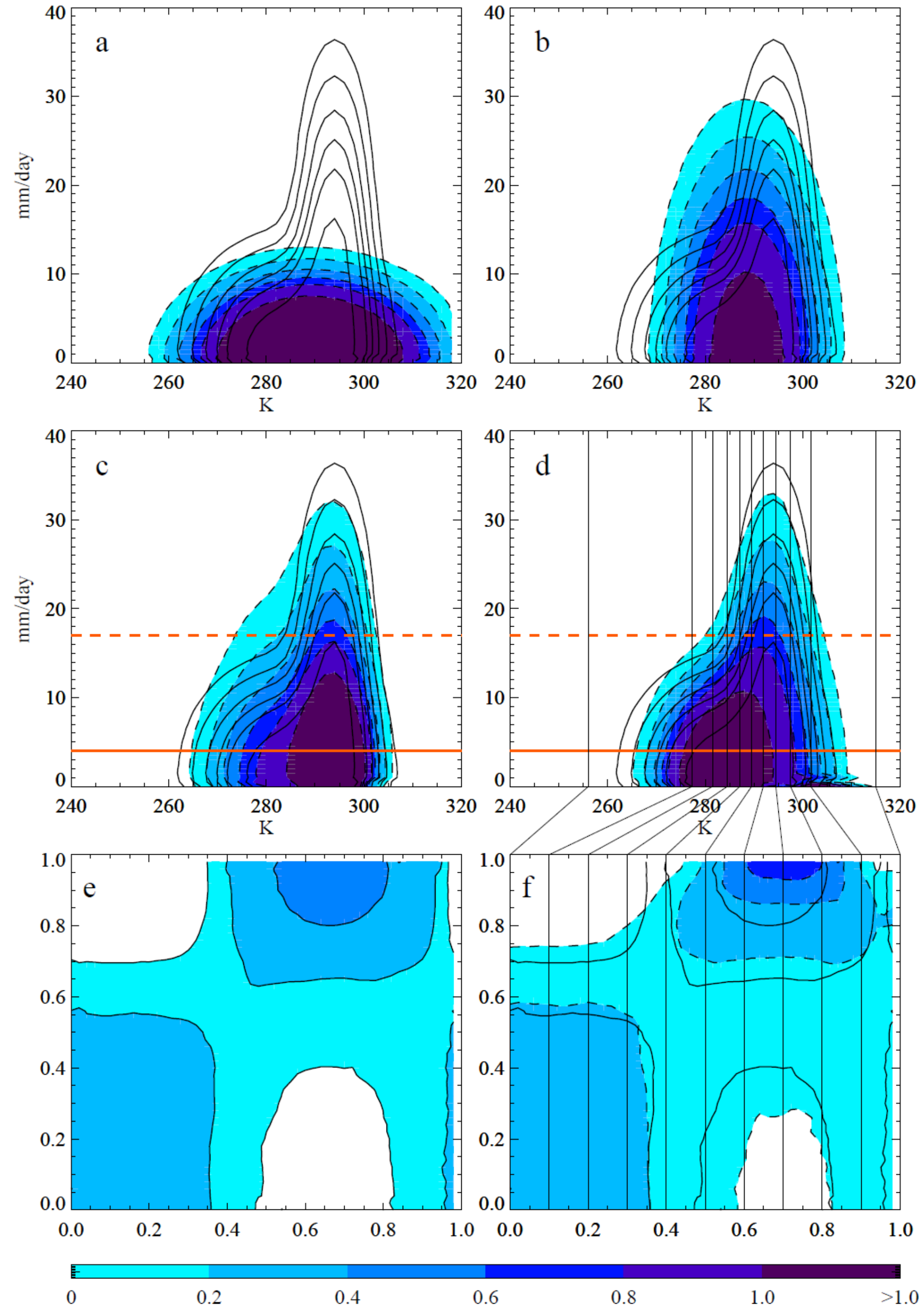
Why Copulas?

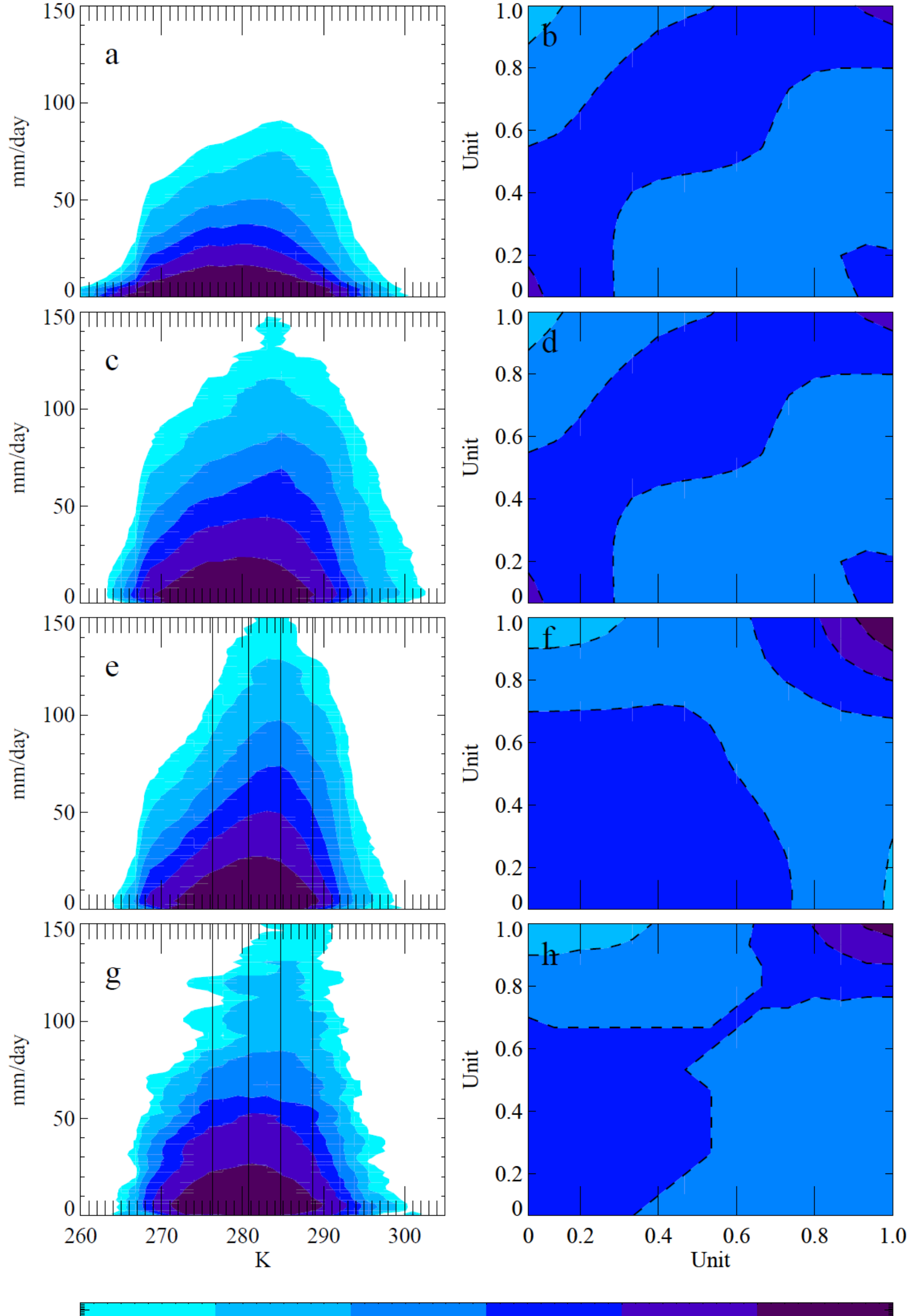
- Copulas are a comprehensive graphic representation of the statistical link between two variables unobscured by the particular shape of the distributions of the individual variables (marginals).

- The calculation of a CPDFs is difficult to explain to a human but extremely easy to explain to a machine.

- To a machine: plot the 2D PDF of the rankings.

- To a human: ….For example, given a data set of 100 measurements of daily precipitation and temperature, simply substitute every pair (P,T) with their rankings, i.e. (0.94,0.52) if it was the 94TH driest day and the 52nd coldest. Now plot the 2D PDF of the rankings.


a
b
c
d
e
f
mm/day
mm/day
K
K
K
K
0
0.2
0.4
0.6
0.8
1.0
>1.0



a
b
c
d
e
f
g
h
mm/day
Unit
150
100
50
0
1.0
0.8
0.6
0.4
0.2
260
270
280
290
300
K
Unit

Conclusions 2D correction

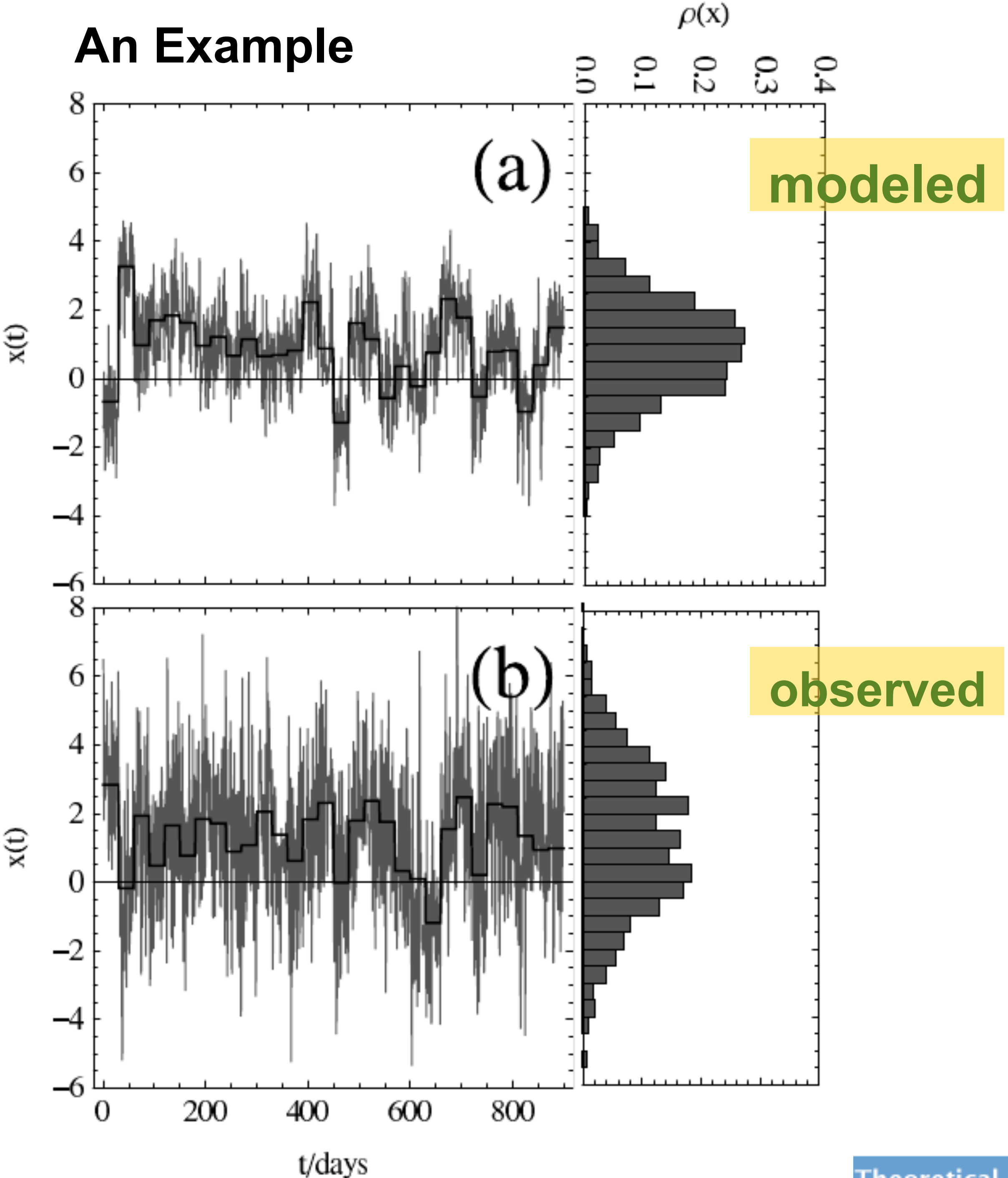
•2D Bias equalizations effectively reproduces the structure of the observed Copula.

•2D Bias correction has very high observational requirements which limit its applicability to gridded output.

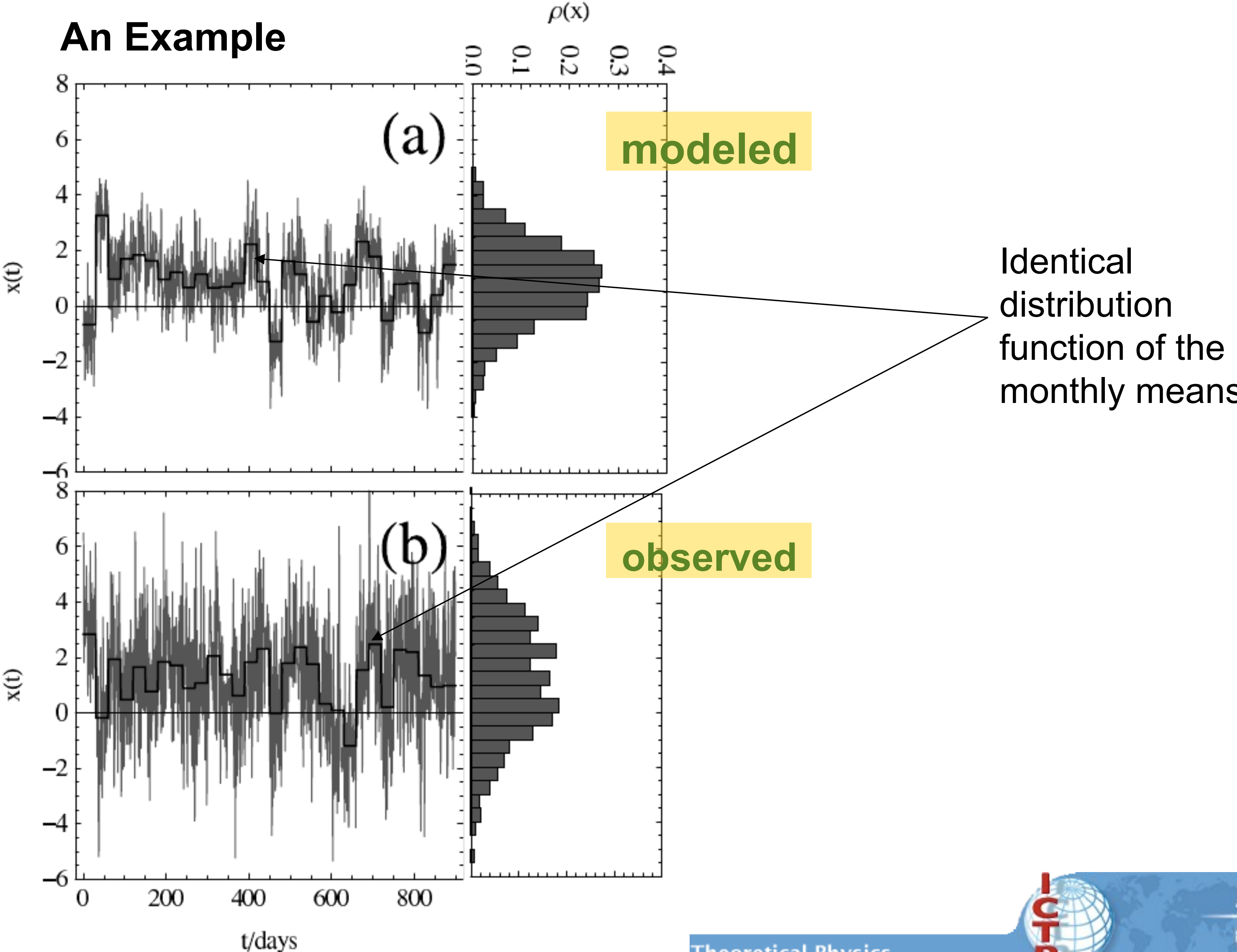
Bias Corrections are dependent on time scale.

i.e. if you correct the daily variance you do not correct the monthly variance

An Example

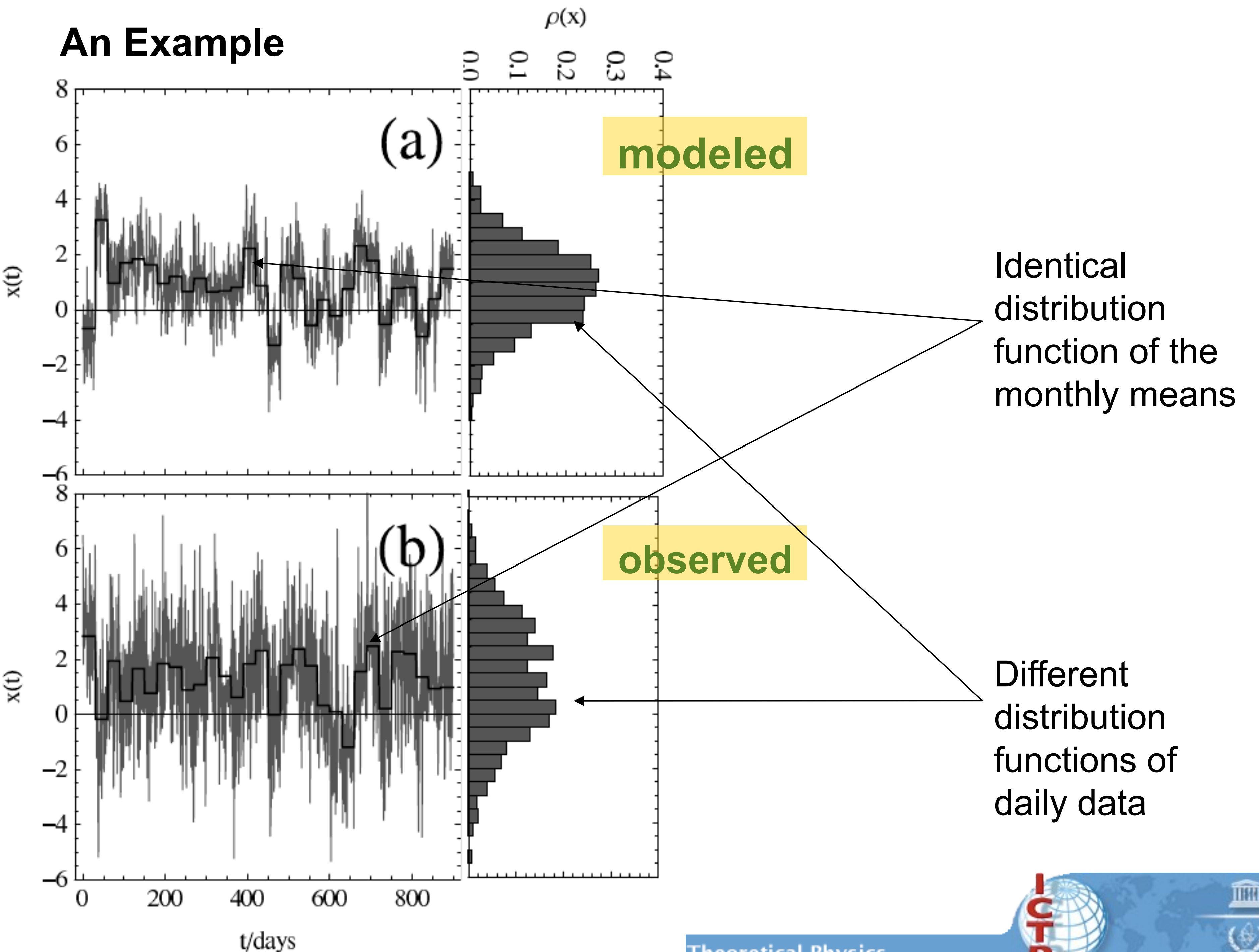

ρ(x)
0.0
0.1
0.2
0.3
0.4
(a)
modeled
(b)
observed
x(t)
t/days
8
6
4
2
0
−2
−4
−6
0
200
400
600
800

An Example



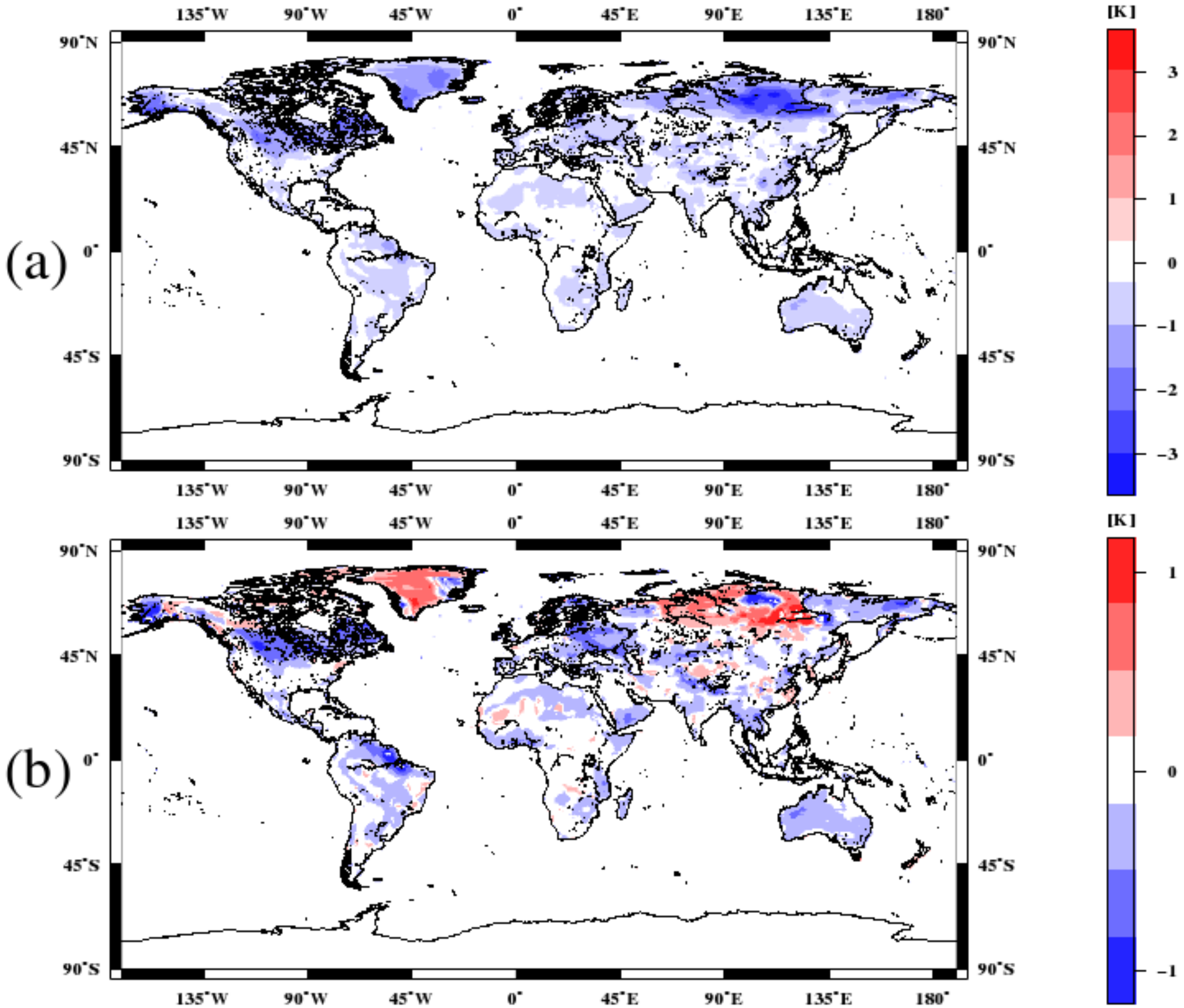
ρ(x)
0.0
0.1
0.2
0.3
0.4
(a)
modeled
x(t)
(b)
observed
t/days
Identical distribution function of the monthly means

An Example



ρ(x)
0.0
0.1
0.2
0.3
0.4
8
6
4
2
0
−2
−4
−6
(a)
x(t)
modeled
8
6
4
2
0
−2
−4
−6
(b)
x(t)
observed
0
200
400
600
800
t/days
Identical distribution function of the monthly means
Different distribution functions of daily data

Improvement of Variance through standard bias correction

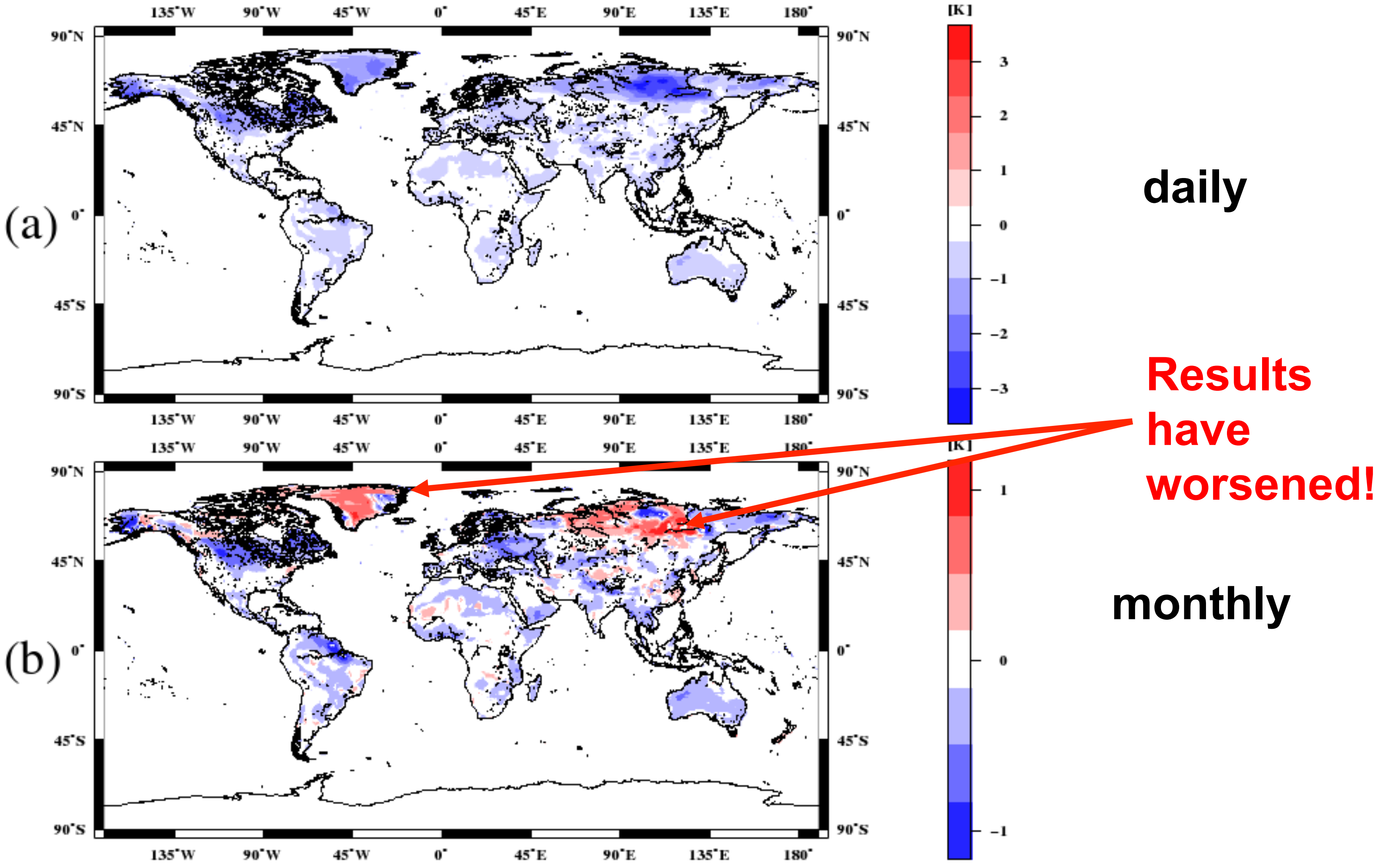


daily

monthly

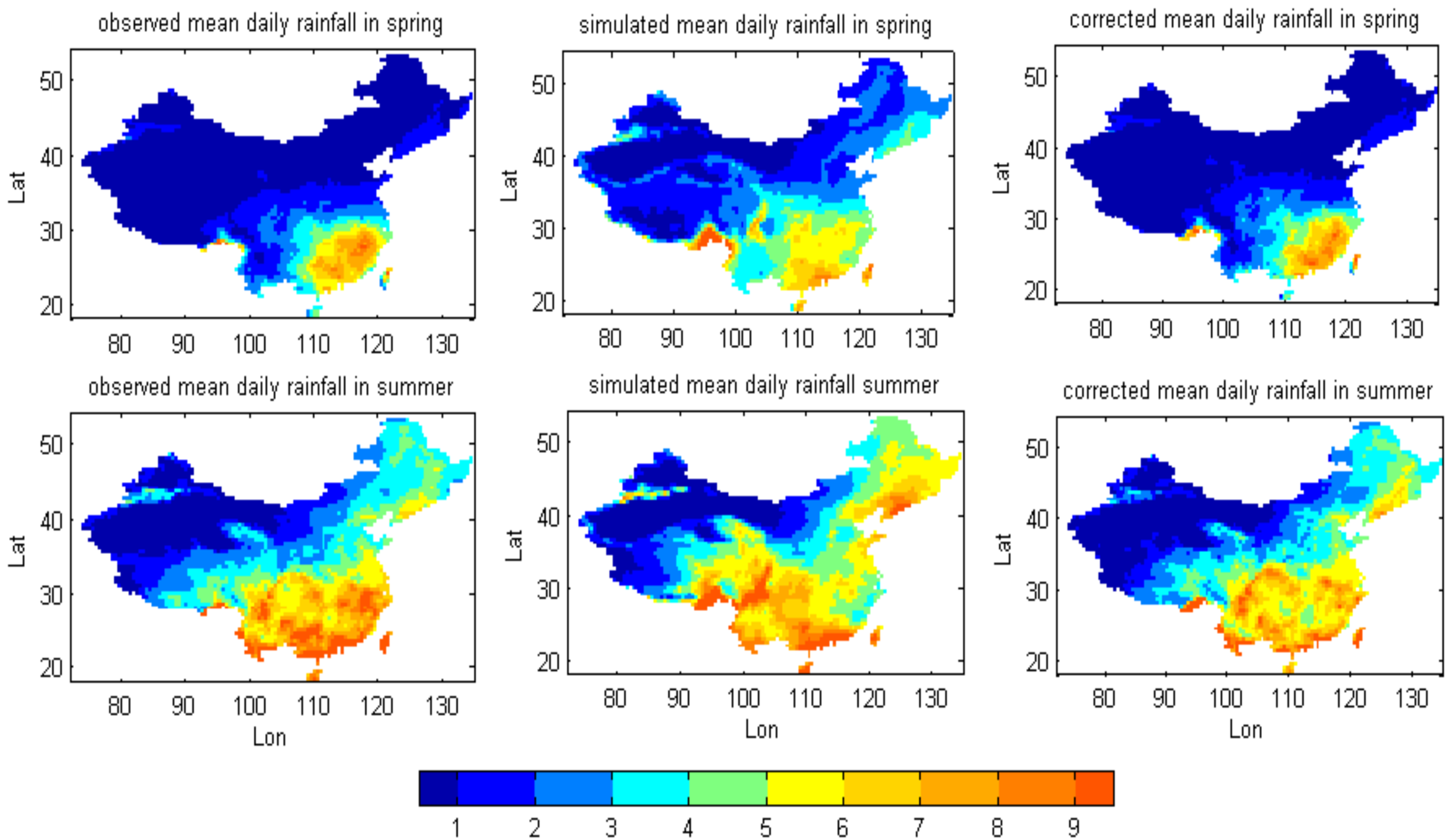
$$\Delta SD(T) = |SD(T_{mod,cor}) - SD(T_{obs})| - |SD(T_{mod,org}) - SD(T_{obs})|$$

Improvement of Variance through standard bias correction



$$\Delta SD(T) = |SD(T_{mod,cor}) - SD(T_{obs})| - |SD(T_{mod,org}) - SD(T_{obs})|$$

Crop yield model input over China.



Cascade bias correction method (Haerter et al.)

1. produce relative fluctuations

$$T'_{i,j} \equiv T_{i,j} - \bar{T}_j$$

Temperature at day i of month j

Mean of month j

Cascade bias correction method (Haerter et al.)

1. produce relative fluctuations

$$T'_{i,j} \equiv T_{i,j} - \bar{T}_j$$

Temperature at day i of month j

Mean of month j

Transfer function for daily fluctuations

$$T'^{cor}_{l,k} = f_{daily}(T'_{l,k})$$

2. produce bias correction to daily relative fluctuations

Cascade bias correction method (Haerter et al.)

1. produce relative fluctuations

$$T'_{i,j} \equiv T_{i,j} - \bar{T}_j$$

Temperature at day i of month j

Mean of month j

2. produce bias correction to daily relative fluctuations

Transfer function for daily fluctuations

$$T'^{cor}_{l,k} = f_{daily}(T'_{l,k})$$

3. produce bias correction to monthly mean fluctuations

$$\bar{T}^{cor}_{l} = f_{monthly}(\bar{T}_l)$$

Cascade bias correction method (Haerter et al.)

1. produce relative fluctuations

$$T'_{i,j} \equiv T_{i,j} - \bar{T}_j$$

Temperature at day i of month j

Mean of month j

2. produce bias correction to daily relative fluctuations

Transfer function for daily fluctuations

$$T'^{cor}_{l,k} = f_{daily}(T'_{l,k})$$

3. produce bias correction to monthly mean fluctuations

$$\bar{T}^{cor}_{l} = f_{monthly}(\bar{T}_l)$$

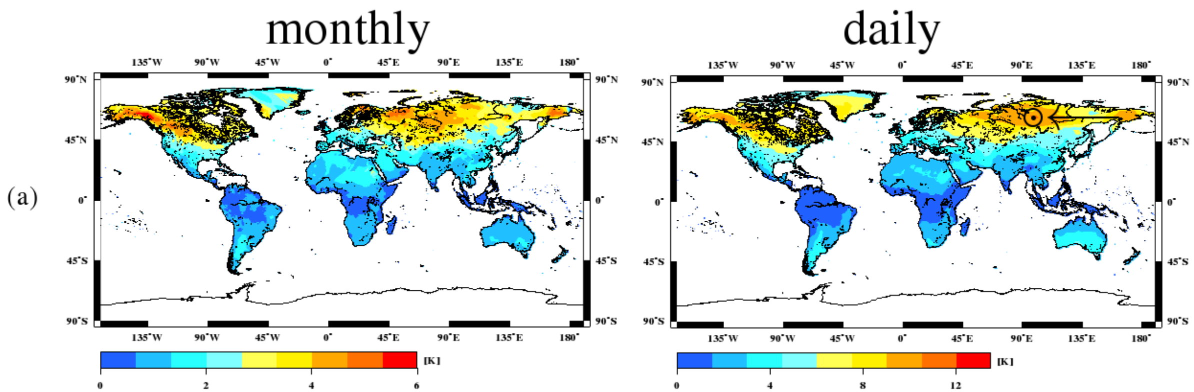
4. re-assemble the bias corrected time series

$$T^{cor}_{l,k} = \bar{T}^{cor}_{l} + T'^{cor}_{l,k}$$

Corrected monthly

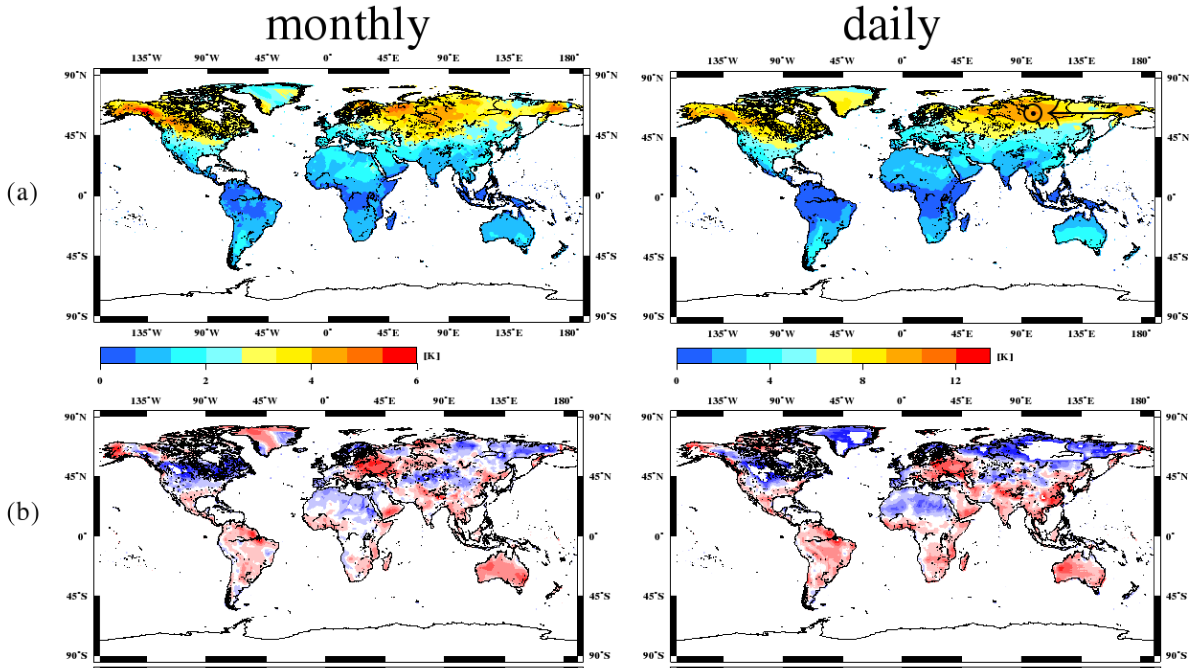
Corrected daily

Standard deviation of WFD, difference to model, standard corrected model, cascade corrected model



WFD

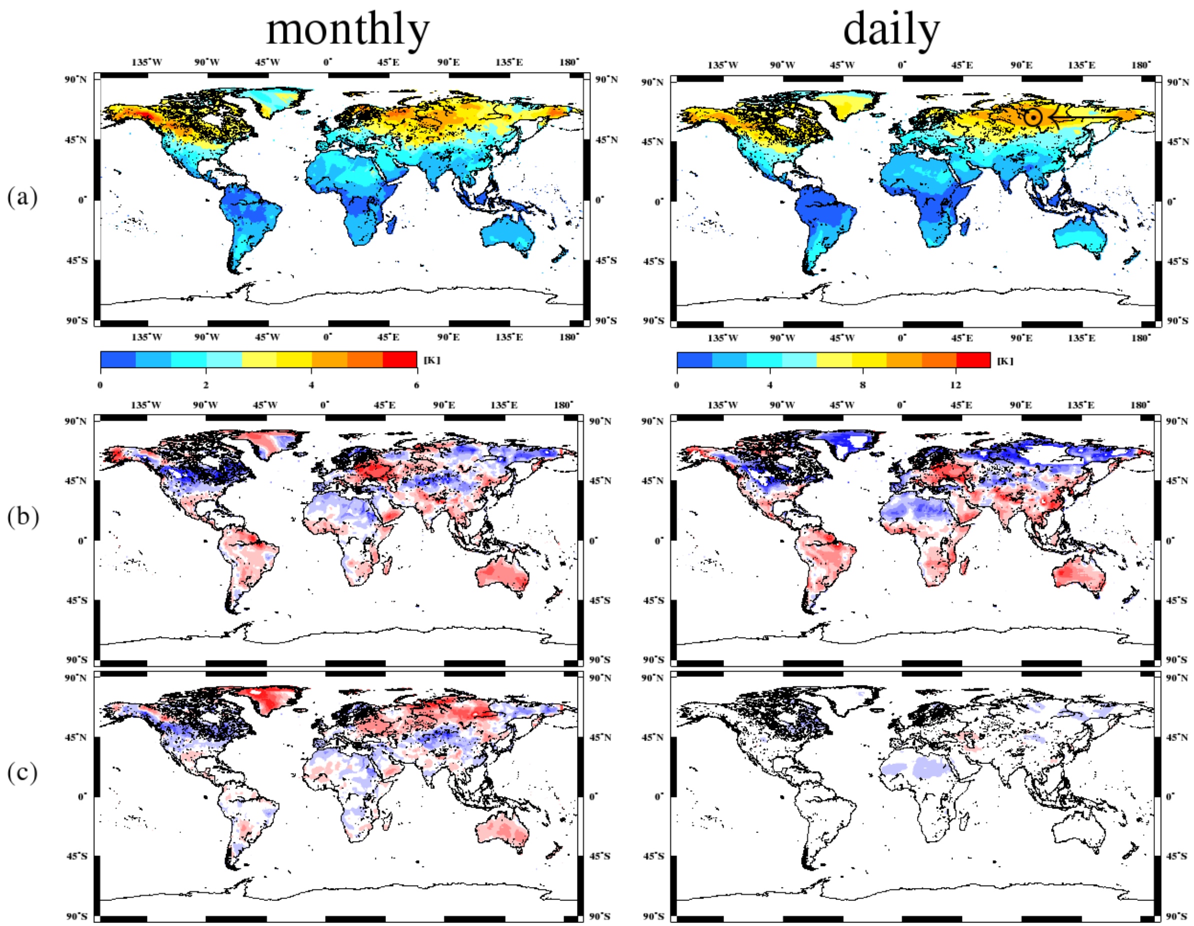
Standard deviation of WFD, difference to model, standard corrected model, cascade corrected model



WFD

uncorrected model

Standard deviation of WFD, difference to model, standard corrected model, cascade corrected model

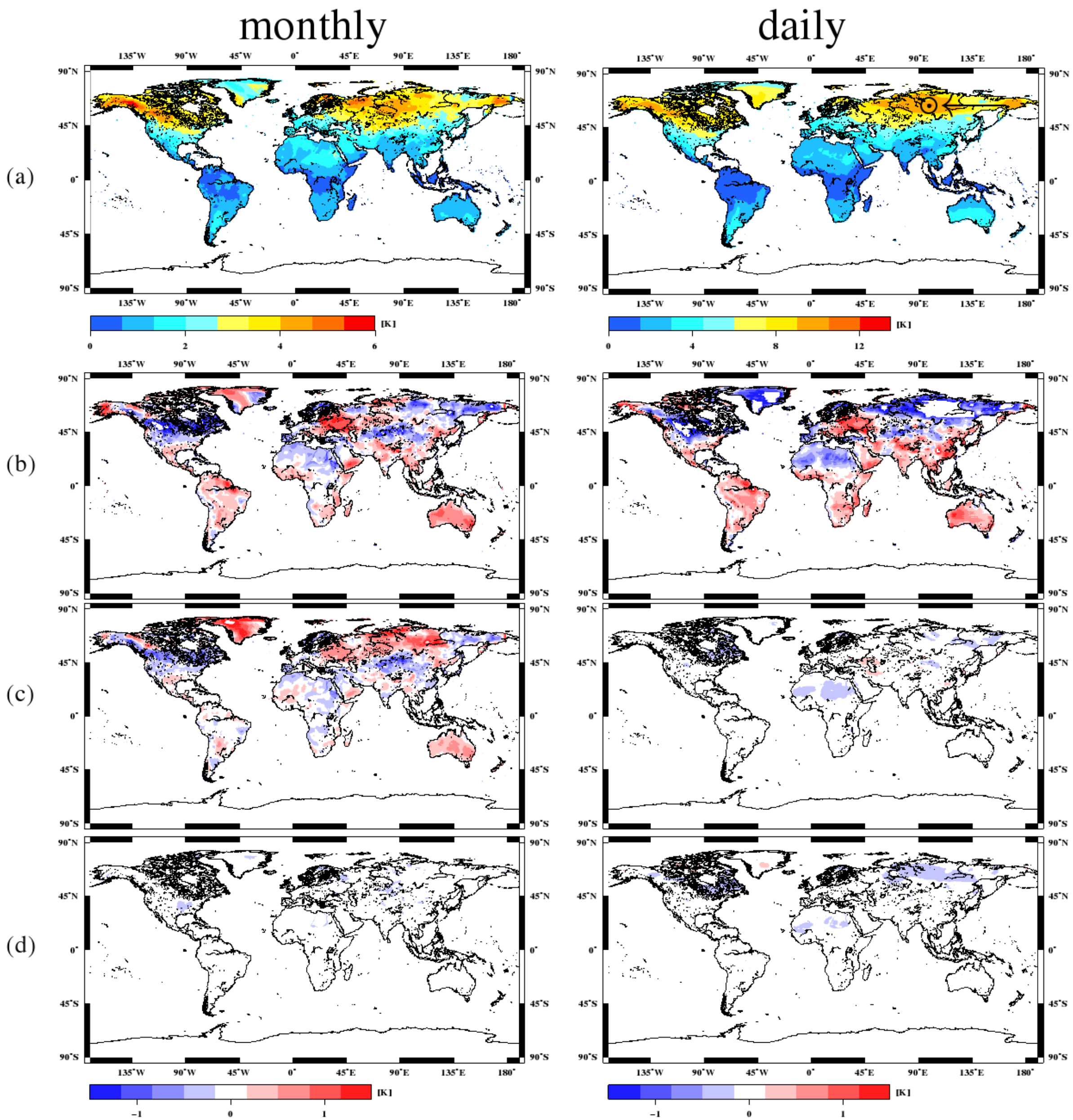


WFD

uncorrected model

Standard corrected model

Standard deviation of WFD, difference to model, standard corrected model, cascade corrected model



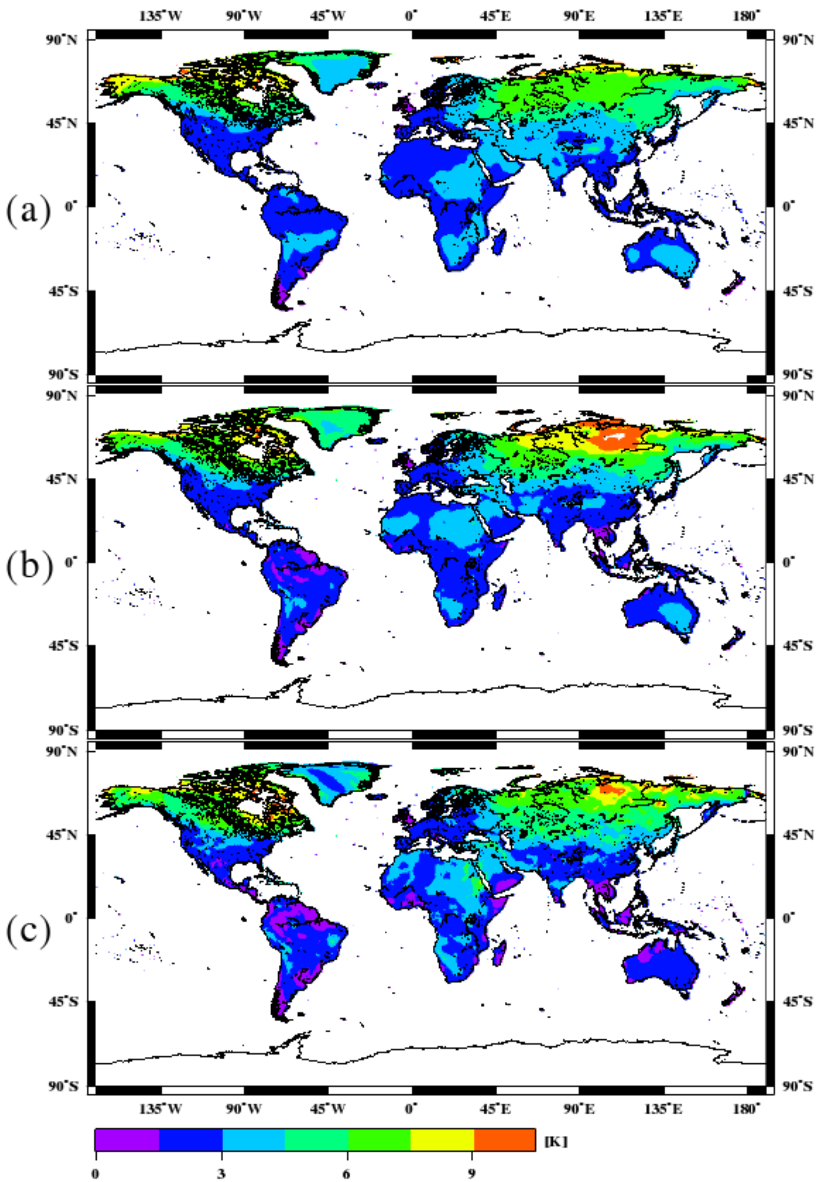
WFD

uncorrected model

Standard corrected model

Cascade corrected model

The big question: How do the different methods impact on the climate change signal?

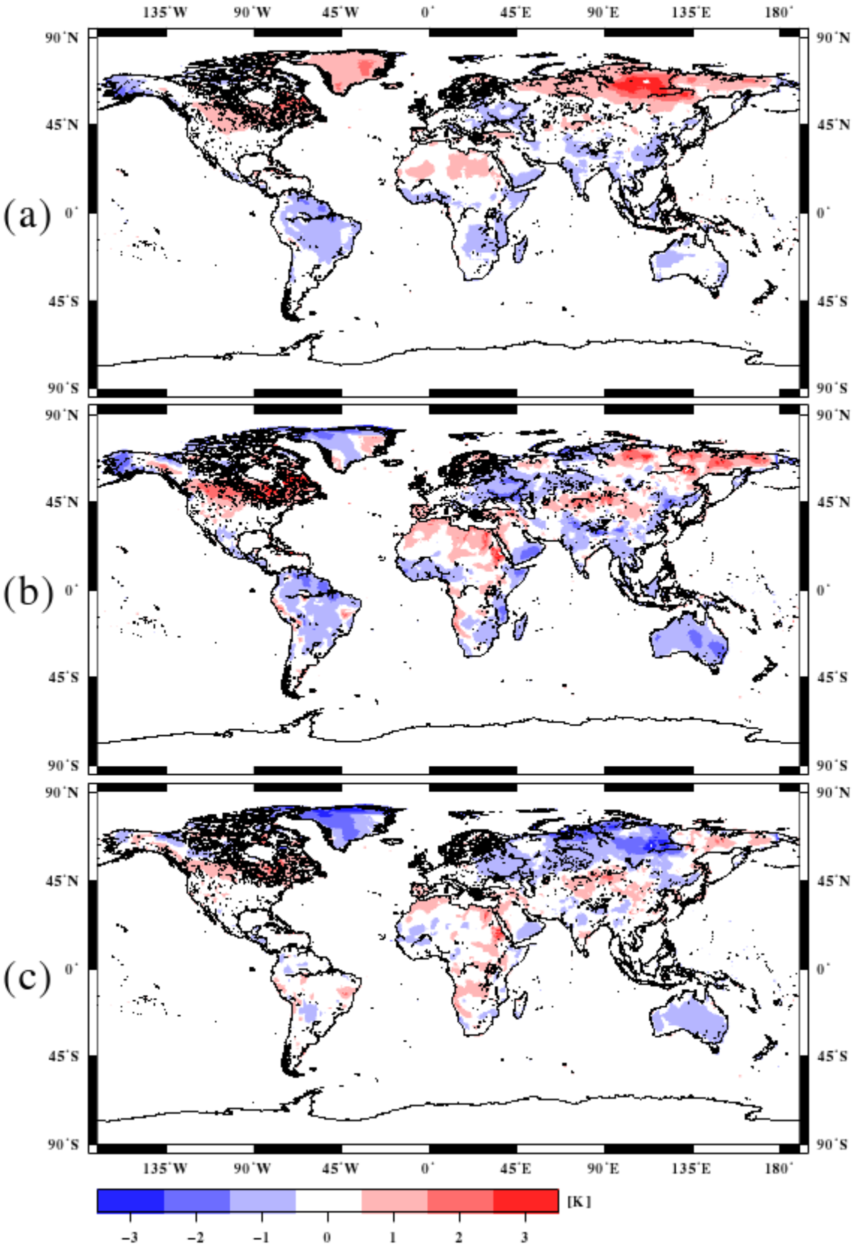


No bias correction

Standard bias correction

Cascade bias correction

The big question: How do the different methods impact on the climate change signal?



Change with standard BC

Change with cascade BC

Cascade-standard

Conclusions

- Statistical Bias Corrections perform transformations to entire PDF, consequently **mixing timescales**
- **Cascade bias correction** which **keeps timescales separate**
- Future climate change signal is impacted upon by bias correction

THANK YOU