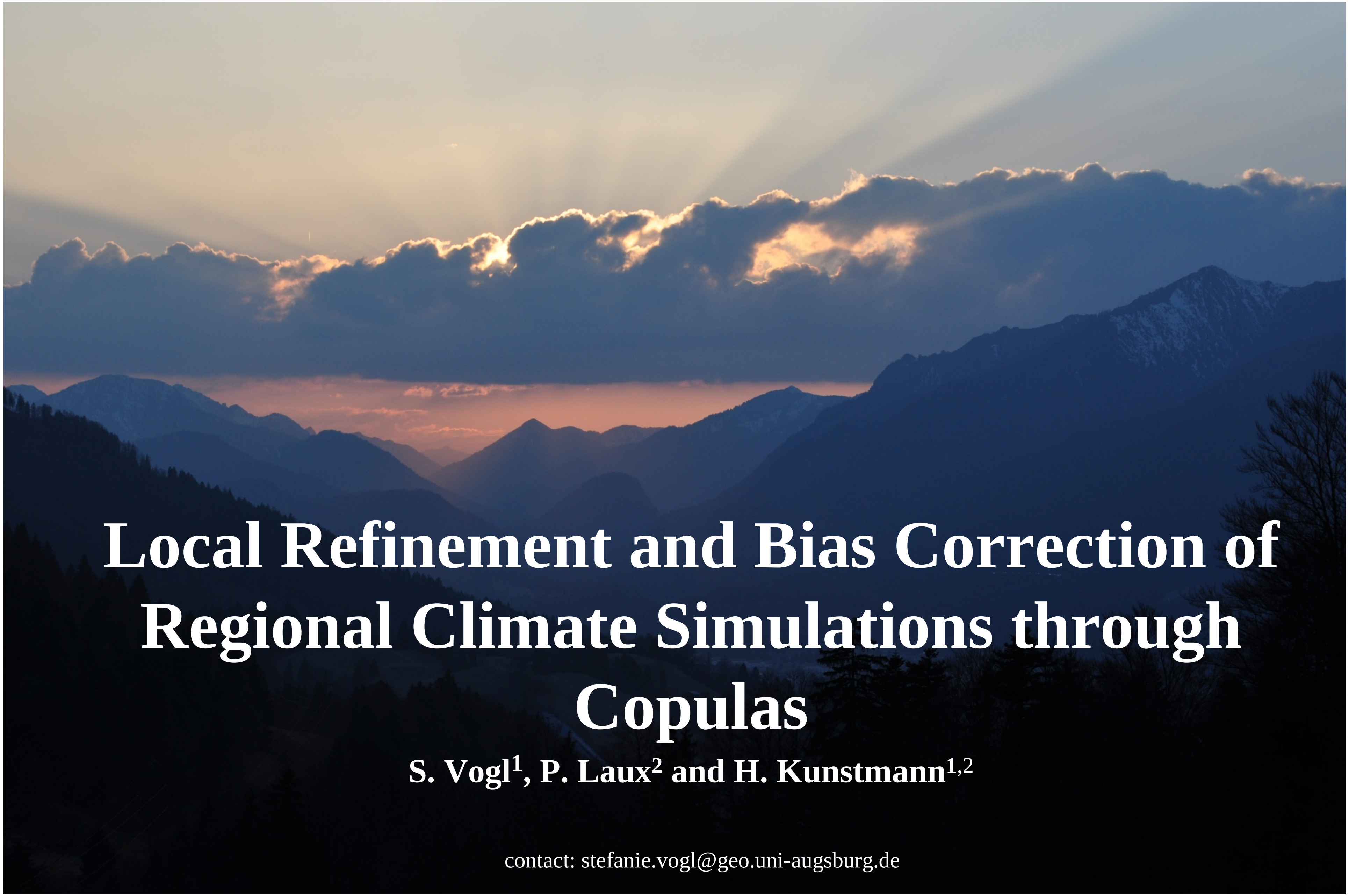


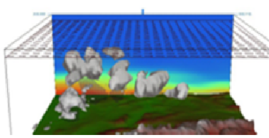
Local Refinement and Bias Correction of Regional Climate Simulations through Copulas

S. Vogl[1], P. Laux[2] and H. Kunstmann[1,2]

contact: stefanie.vogl@geo.uni-augsburg.de

[1] University of Augsburg, Institute for Geography, Regional Climate and Hydrology, Augsburg

[2] Karlsruhe Institute of Technology (KIT), Institute for Meteorology and Climate Research



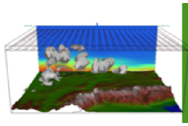
Why regional climate modelling?

In the light of warming climate there is a need for adaption to changing environmental conditions especially in climate-sensitive regions as the alpine space

Now decisions have to be taken for the future!

- flood prevention measures
- drinking water supply
- hydroelectric power production
- water availability for agricultural purposes
- tourism

Decision makers need regional clima impact studies with a high spatio-temporal resolution to provide regional precipitation and temperature fields.

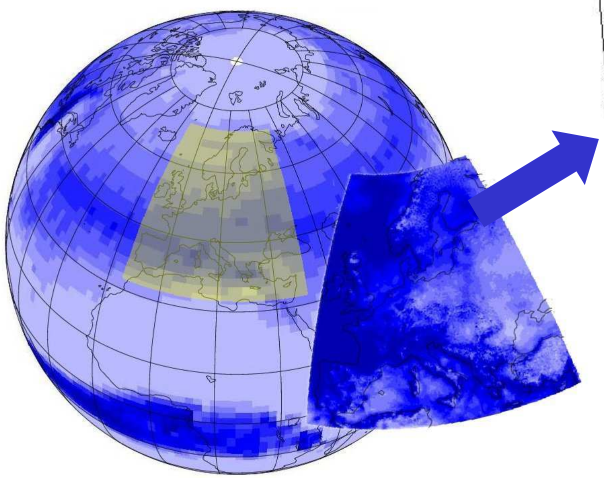


Downscaling - from Global Climate Model (GCM) to Regional Climate Model (RCM)

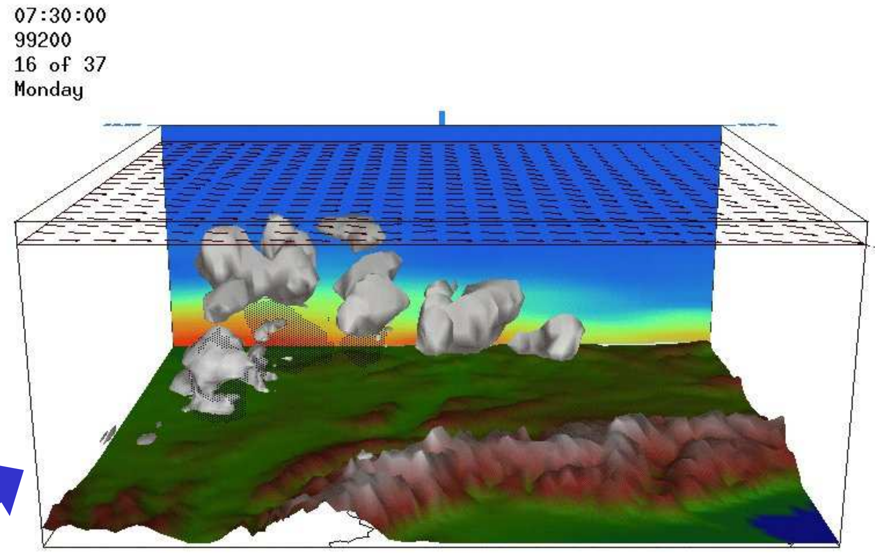
GCM

downscaling

RCM



dynamical downscaling



the boundary conditions for the regional model are taken from GCM results.

Three different approaches:

Dynamical downscaling

- Local variable of interest is predicted from values of the corresponding variable simulated at the **closest grid point of the GCM**
- **Empirical adjustment** compensates simulation errors and small-scale effects

- Output from the GCM is used to **drive a nested high resolution regional climate model** (RCM)
- The prediction is now based on the result at the nearby RCM grid point

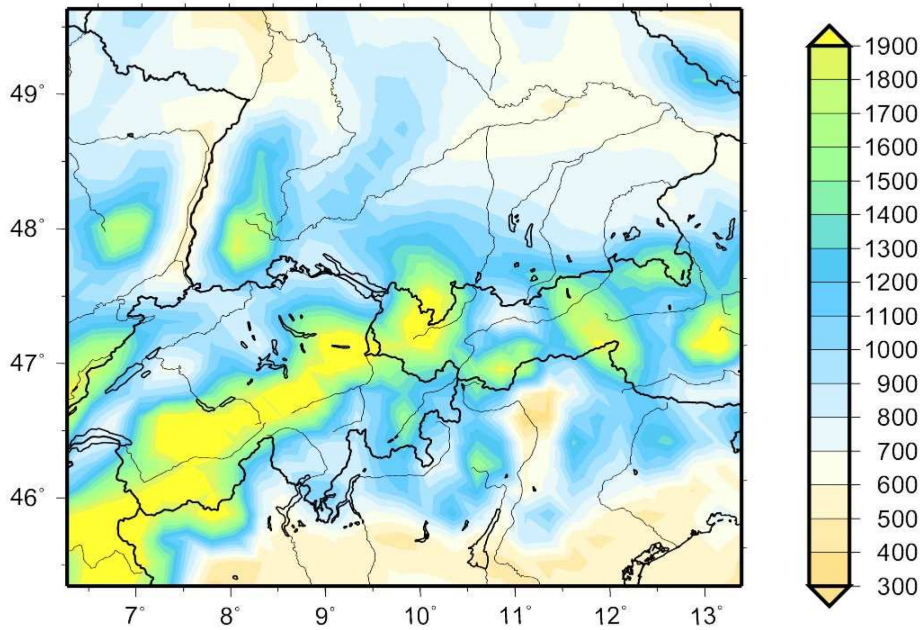
Statistical downscaling

- **Statistic relationships are developed** to link the local variable to predictor variables (stochastic weather generators)

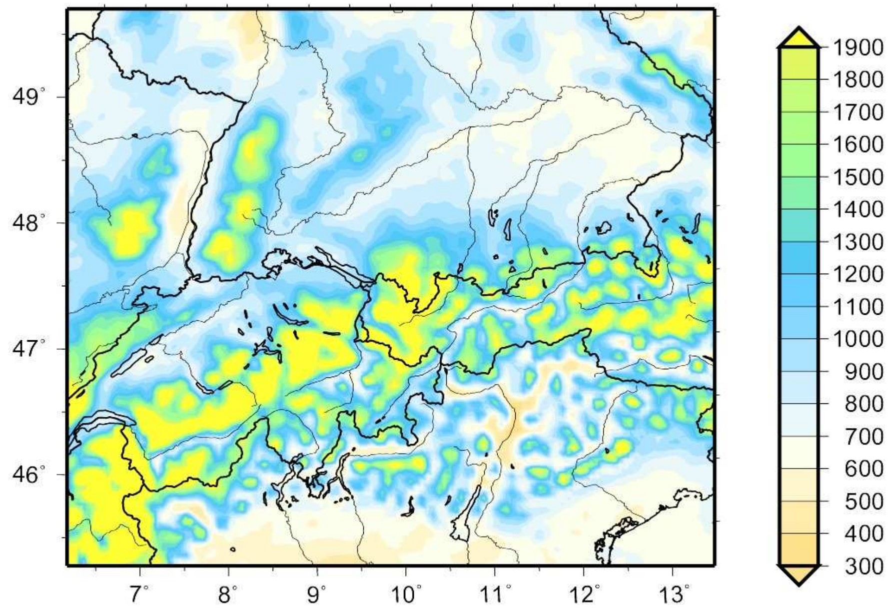
In practice the downscaling techniques are often mixtures of this approaches and show a mixture of their attributes

Advantages of regional climate modelling

resolution 19.2 km



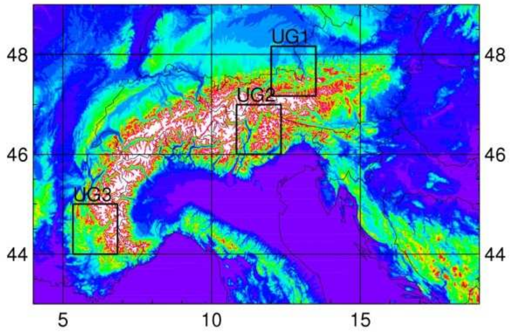
resolution 4.8 km

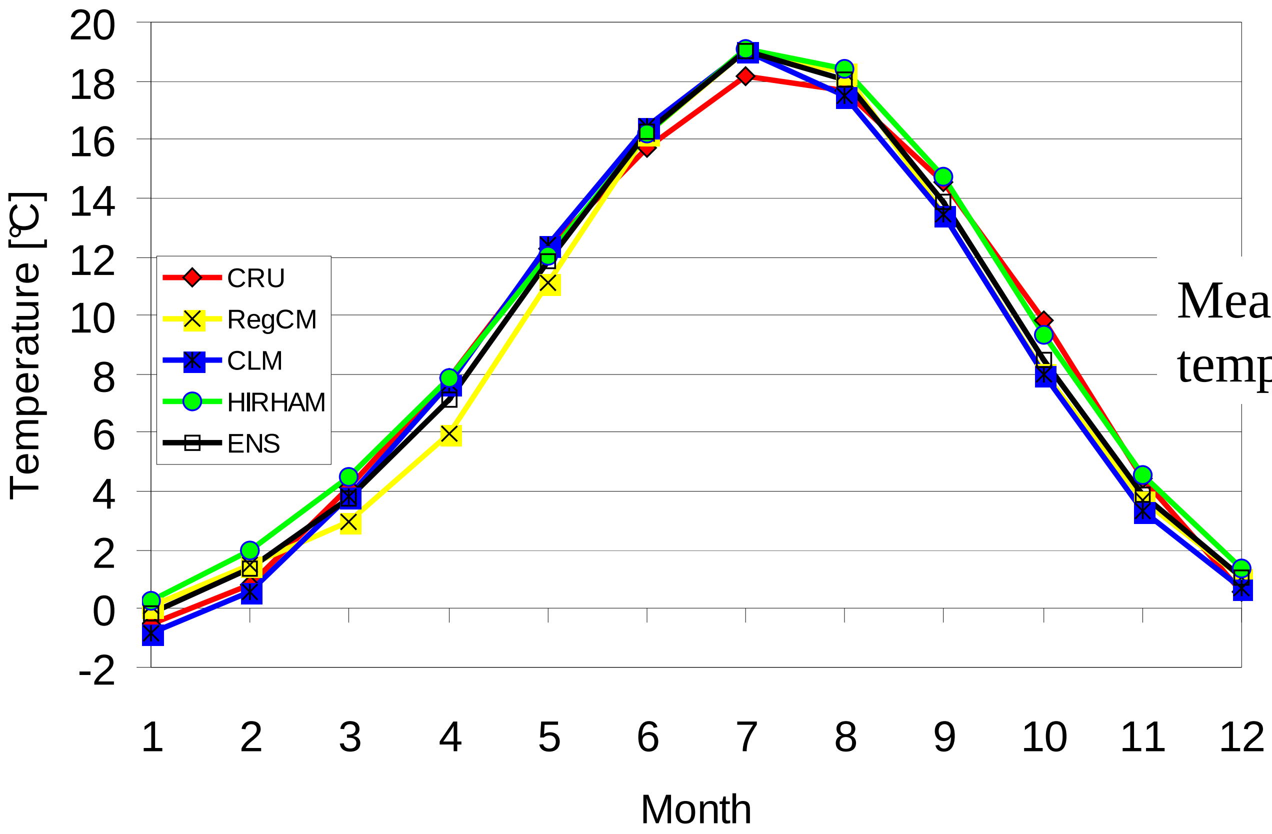


R.Knoche, IMK-IFU, KIT

But how good do RCMs reproduce the observed regional climate?

Example in the Alpine Space



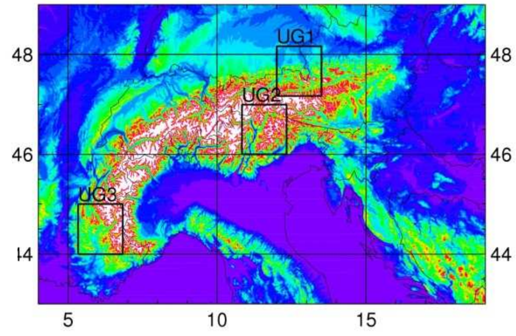


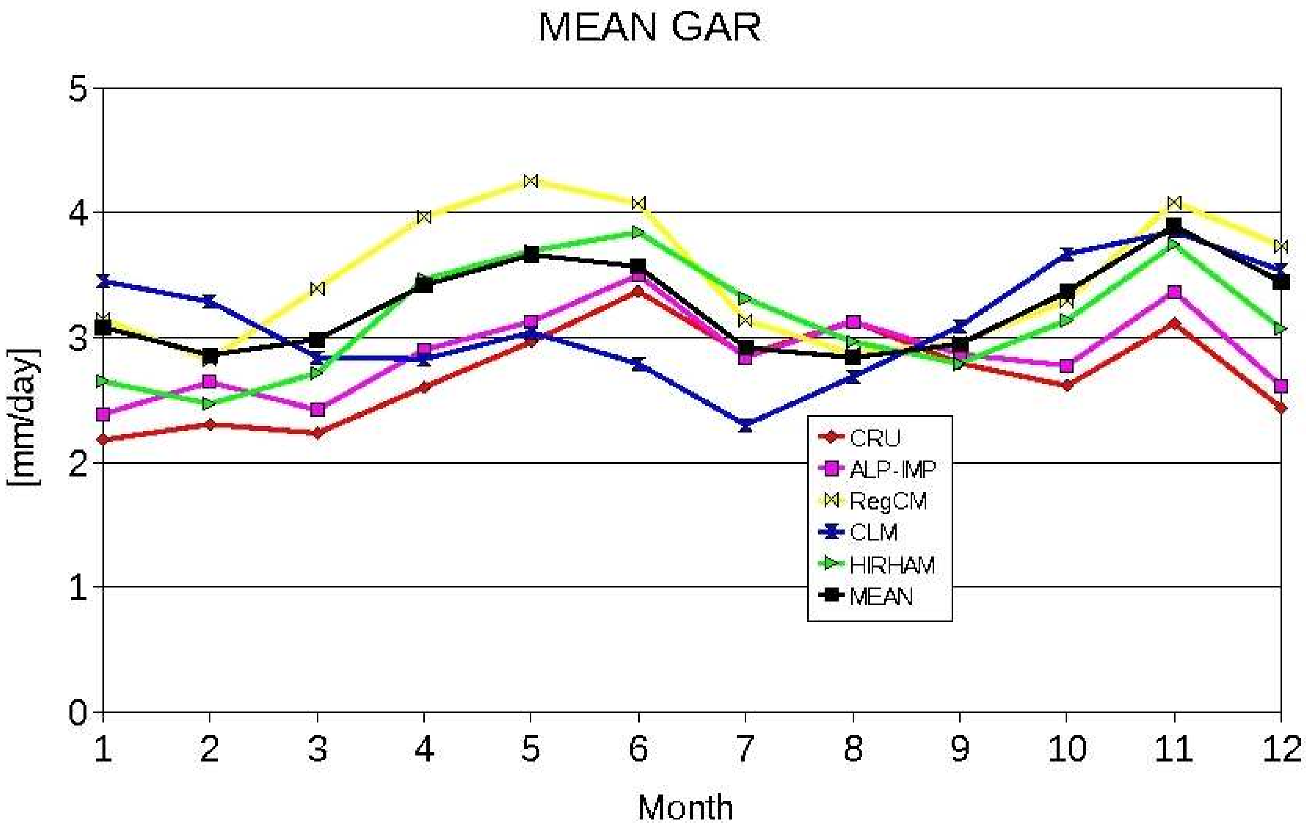
Mean monthly temperature bias +/-1°C

Smiatek, Kunstmann, Knoche, Marx, 2009 (JGR)

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

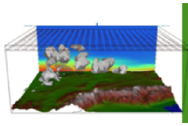
Example in the Alpine Space





Precipitation bias up to 1.5 mm/day :-|

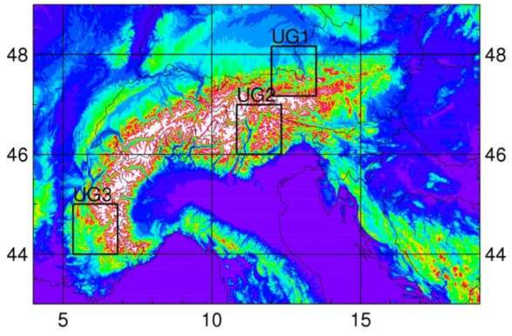
Smiatek, Kunstmann, Knoche, Marx, 2009 (JGR)

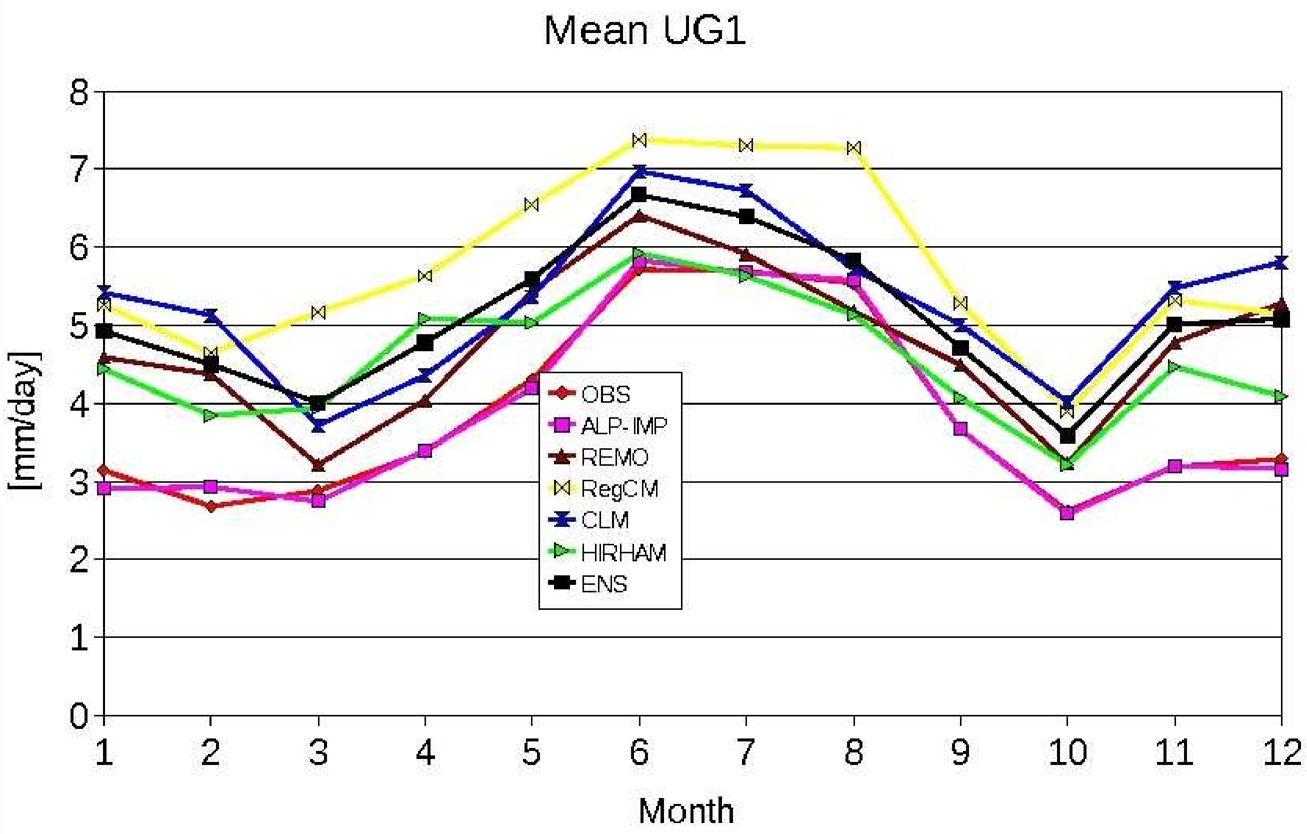


Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann

Universität Augsburg University

Example in the Alpine Space





Precipitation bias up to 2.5 mm/day :-(

Smiatek, Kunstmann, Knoche, Marx, 2009 (JGR)

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann

Universität Augsburg University

- The validation of the RCM results shows the **need for bias correction (BC)** before coupling the model results to the hydrological models
- Bias correction methods:
  - **Linear BC (corrects difference in the mean)**
  - **Non-linear BC (adjustment of mean and standard deviation)**
  - **Histogram equalization techniques**
  - **Bootstrap BC**
  - **Quantile regression**
  - **...**

- often correlation based
- lack of correcting specific types of systematic errors
- mostly correction for single variables (decoupling)

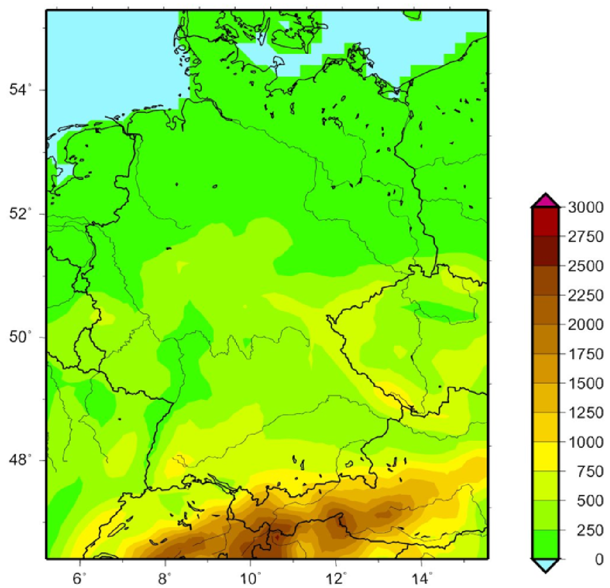
- **But still:** improvement of bias correction/downscaling will influence the final results of the hydrological model

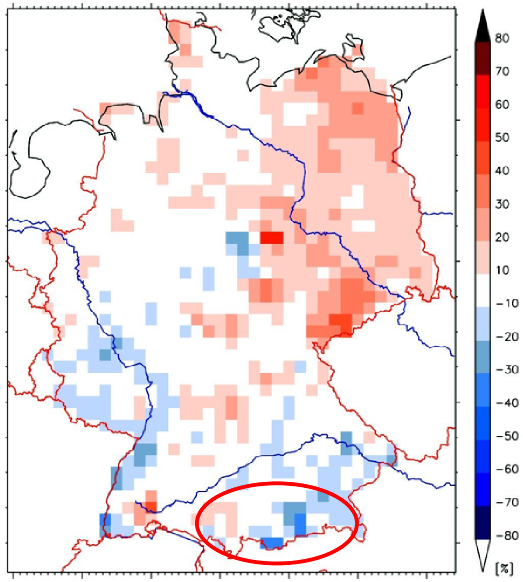
➡ Develop new methods

➡ Copula theory could show new ways

<u>Application:</u>

Bias-correction of regional climate modeling in the alpine space





Domain and topography of regional climate simulations with MM5, 19.2 km spatial resolution (left).

Bias of mean annual total precipitation for the MM5 with respect to the DWD reference data set [%] (right)

- Dry bias (wet bias) for eastern part of Germany (alpine space & Rhine valley)
- Underestimation in the alpine region possibly due to the complex terrain

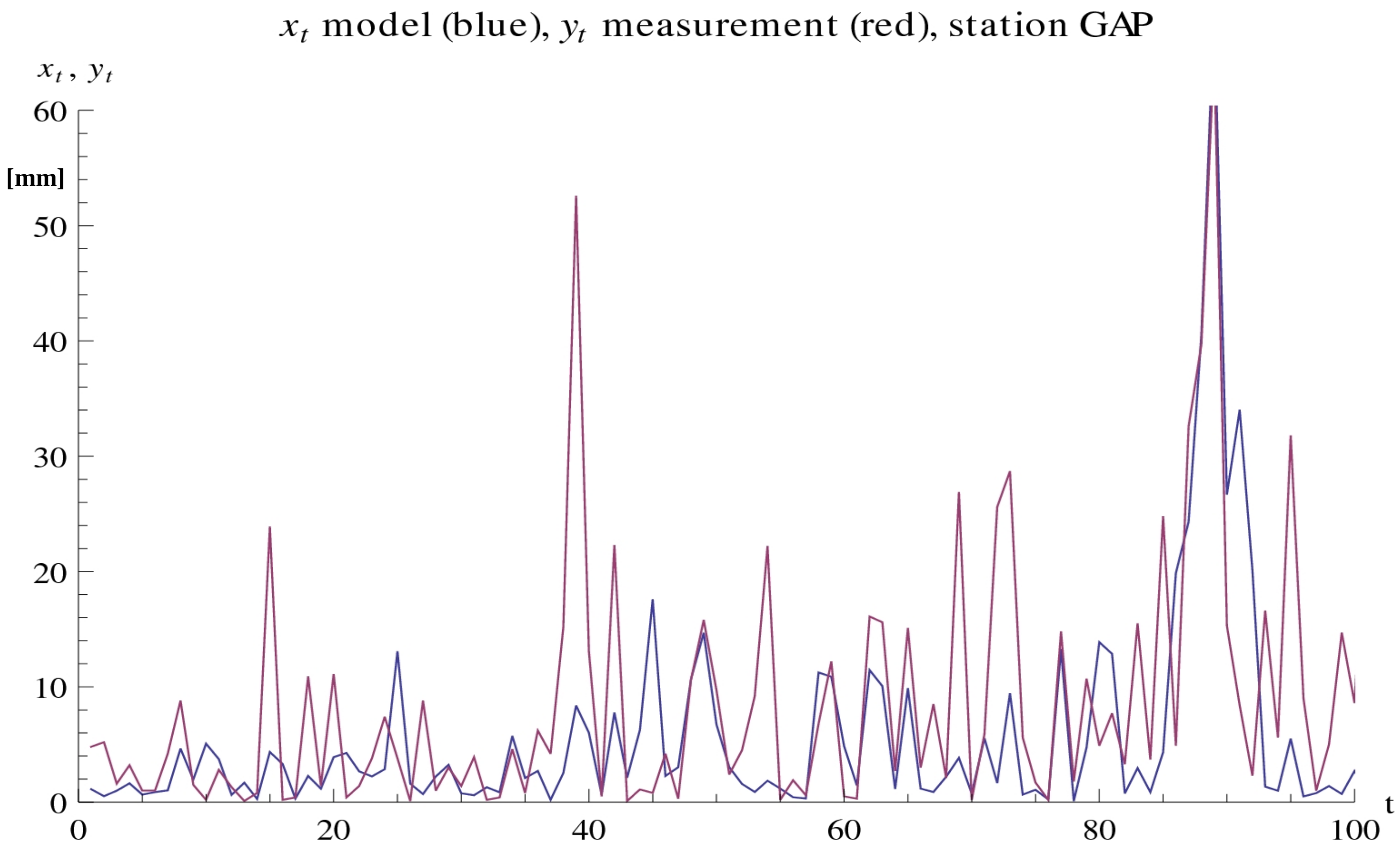
Theory:

random sample from (X,Y):

$$(X_1, Y_1), \ldots, (X_n, Y_n)$$

In practice we have two time series: e.g.

- modeled temperature and measured temperature
- modeled precip. and measured precip.



The joint dependence between these variables is fully characterized by their Copula C(x,y).

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

- every Copula is the representation of the dependence structure of the two (or more) variables

- by using a Copula it is possible to derive a bi- or multivariate PDF f(x,y) just by knowing the single marginal distributions $F_X(x)$ and $F_Y(y)$

$$\boxed{\underbrace{f(x,y)}_{\textbf{joint density}} = \underbrace{c(F_X(x),F_Y(y))}_{\textbf{pdf of Copula (copula density)}} \cdot \underbrace{f_X(x)\cdot f_Y(y)}_{\textbf{pdfs of the marginals}}}$$

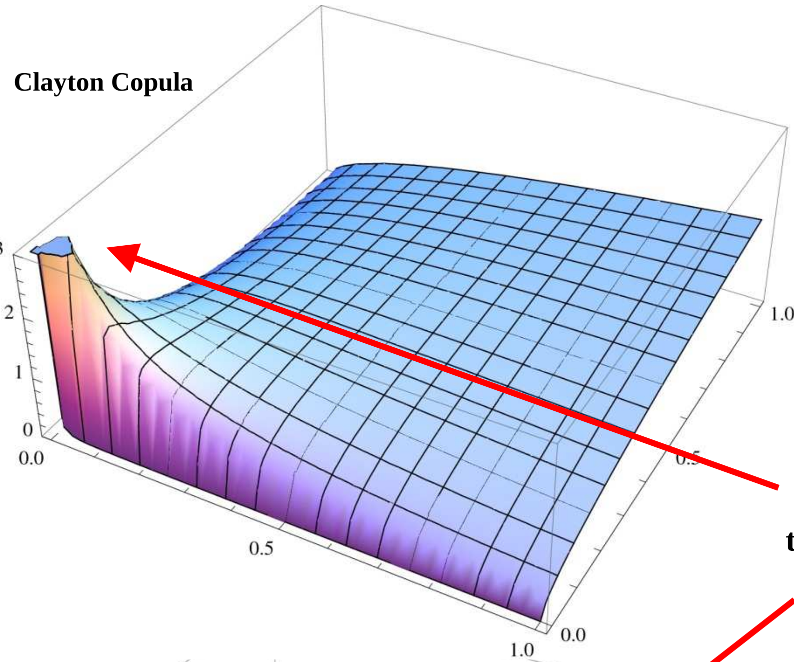
$\Downarrow$

$$c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}$$

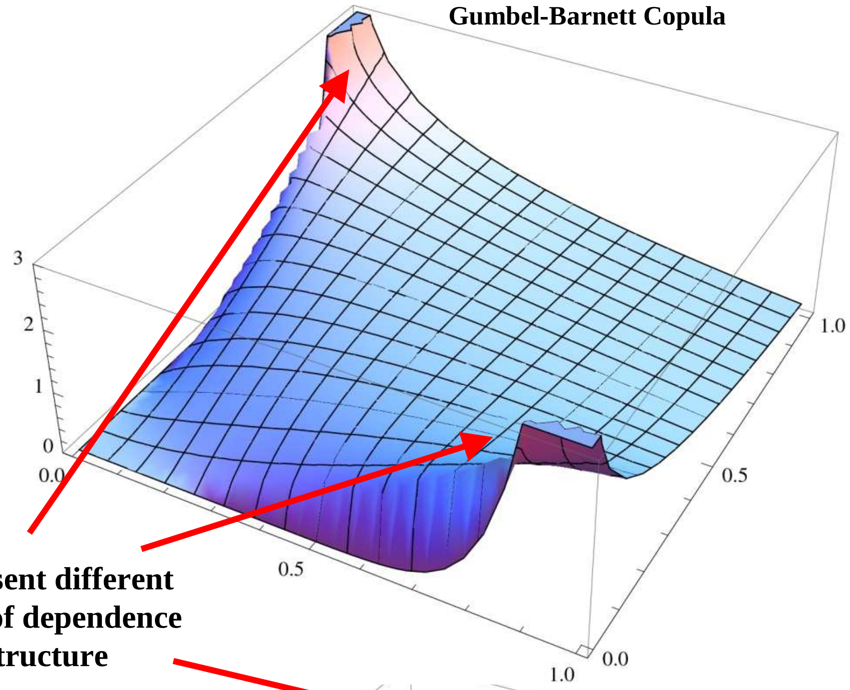
$\Rightarrow$ **the copula density c(u,v) is often called“dependence function”**

PDFs of different Copula families

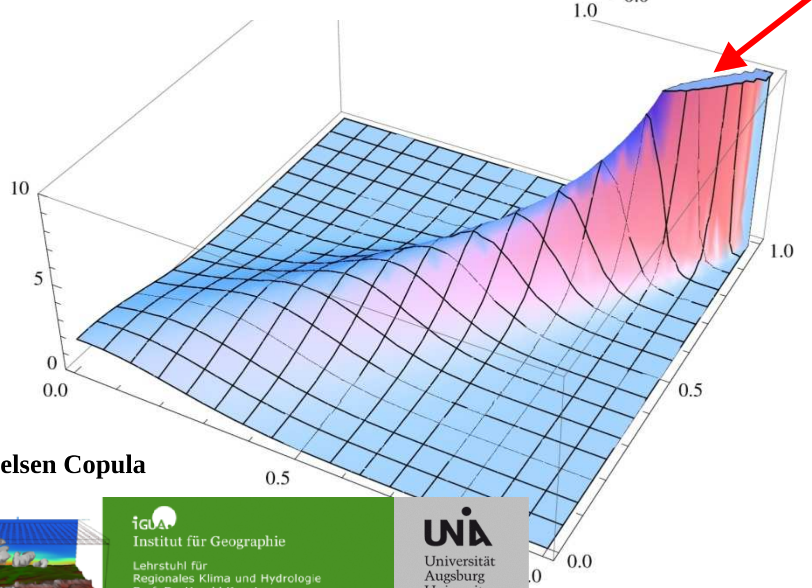
Clayton Copula



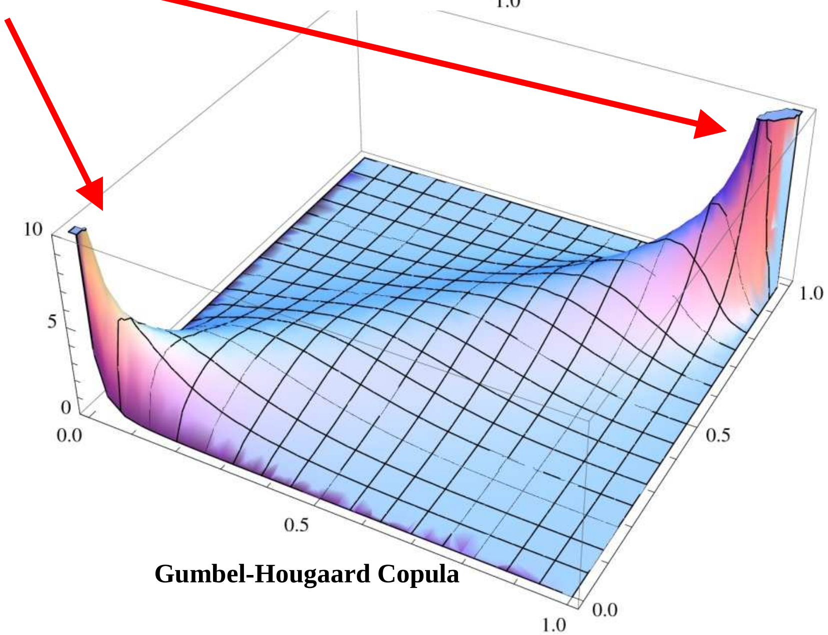
Gumbel-Barnett Copula



represent different types of dependence structure



Nelsen Copula



Gumbel-Hougaard Copula

Connection of the Copula parameter to rank based dependence estimators - Kendalls tau

(+) $$\boxed{\tau = 2\frac{n_c - n_d}{n(n-1)} = 4\iint_{I^2} C(u,v)\,dC(u,v) - 1}$$

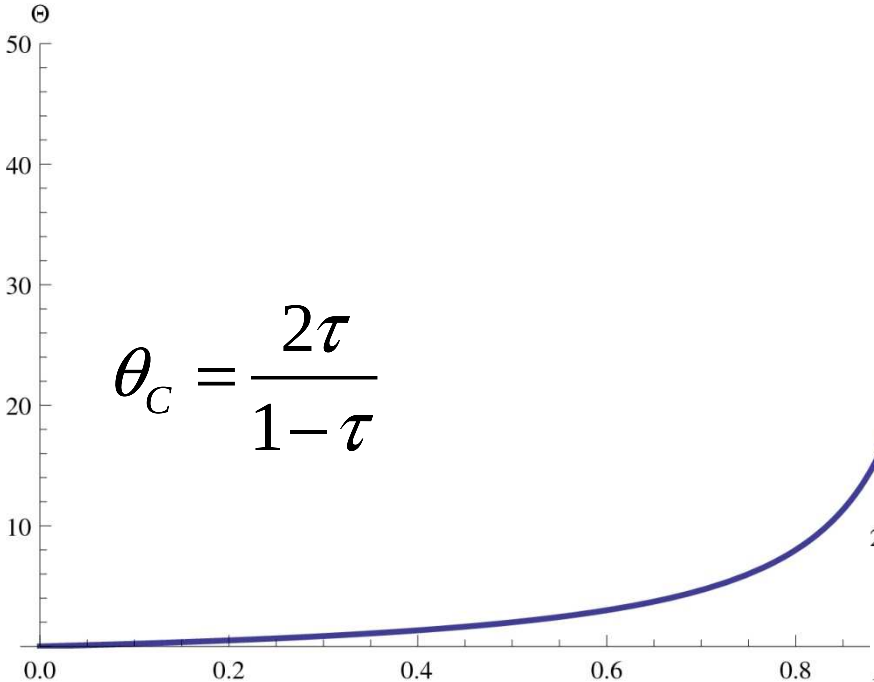
τ is a measure of dependence based on ranks

$$C_\theta(u,v) = \max\left(\left[u^{-\theta} + v^{-\theta} - 1\right]^{-\frac{1}{\theta}}, 0\right)$$ Clayton Copula

$$C_\theta(u,v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right)$$ Gumbel-Hougaard Copula

There is a relationship between Kendalls τ and the Copula parameter θ via (+)

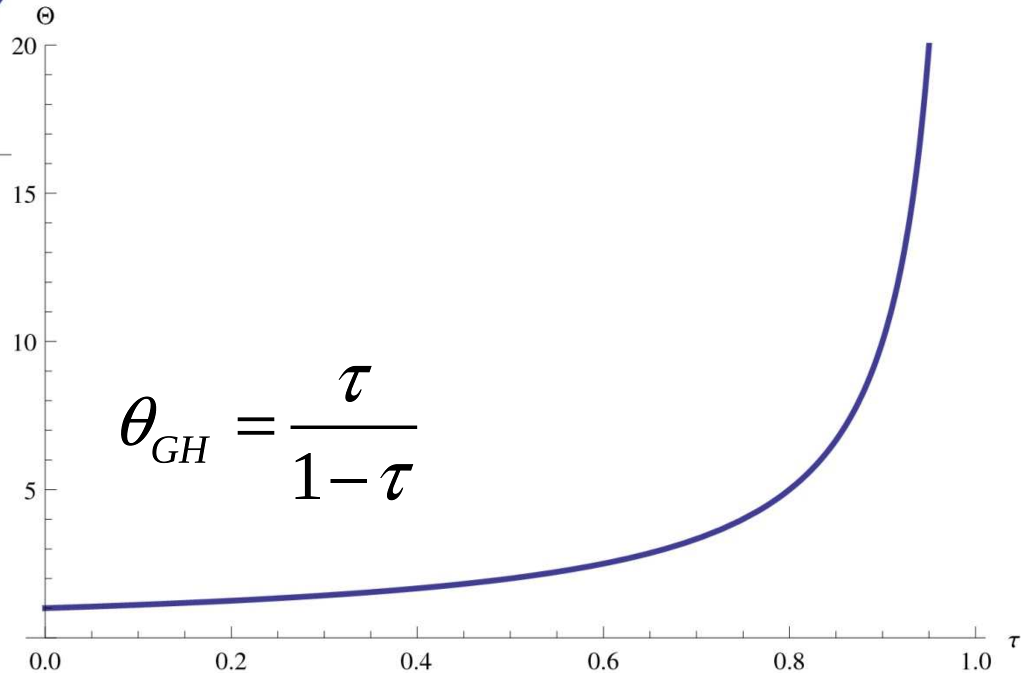
θ is a monotonically increasing function of τ



$$\theta_C = \frac{2\tau}{1-\tau}$$

Clayton Copula

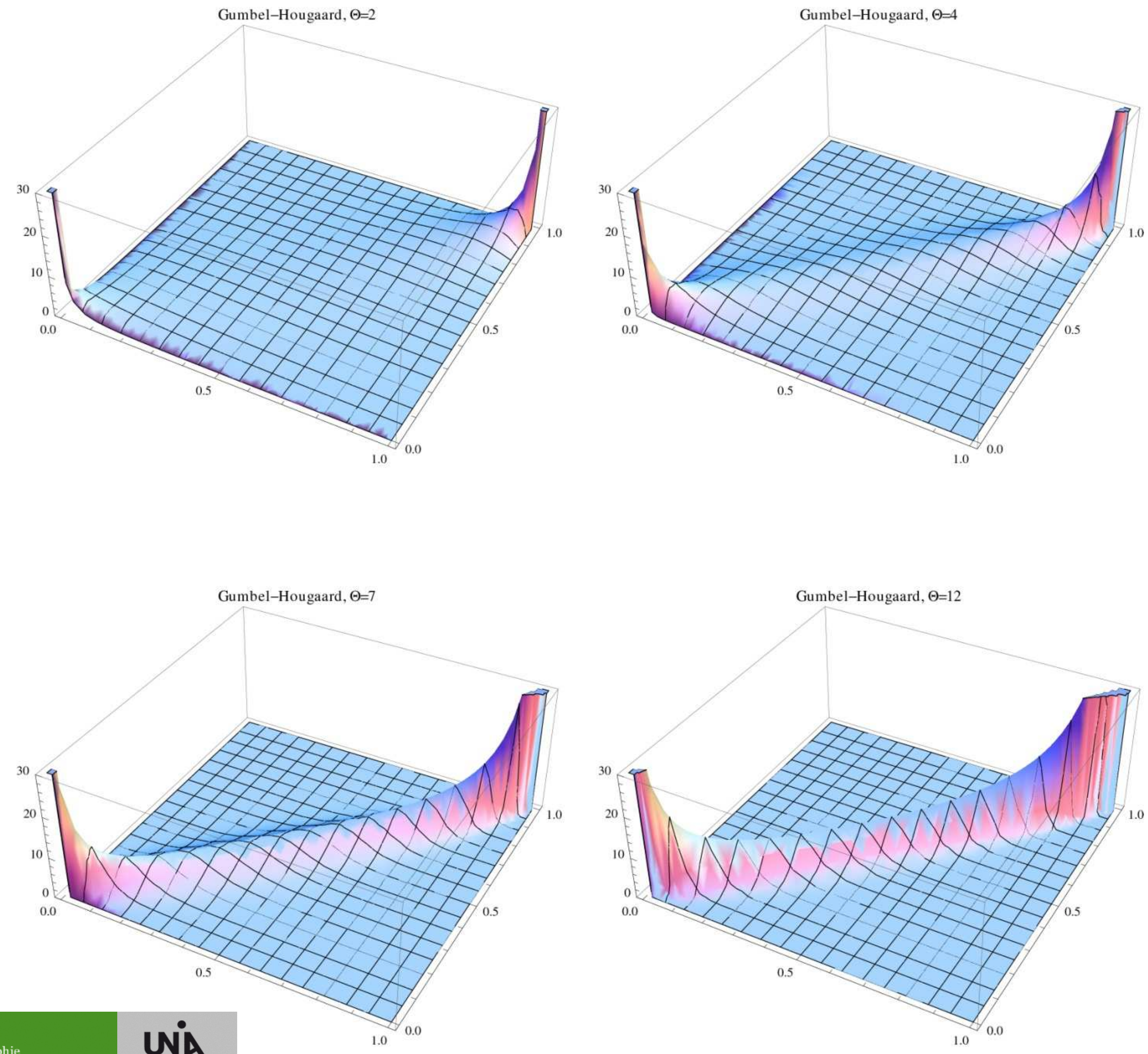
high Copula parameters indicate strong correlation



$$\theta_{GH} = \frac{\tau}{1-\tau}$$

Gumbel-Hougaard Copula

larger Copula parameter - higher dependence

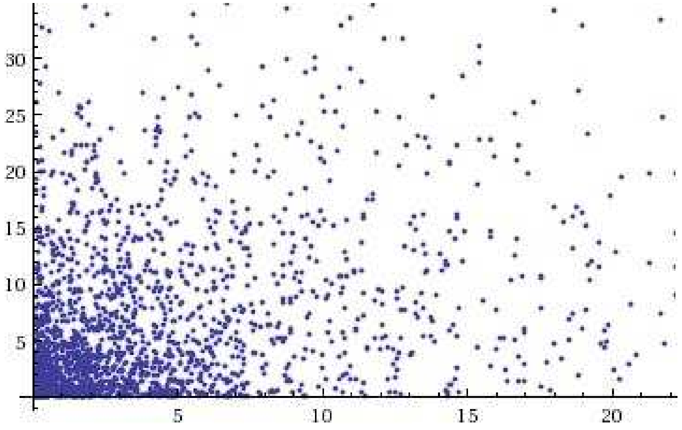


PDF of GH Copula for 4 different values of θ

Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i)

needs to be i.i.d. (independent and identically distributed)



scatter plot of the original data

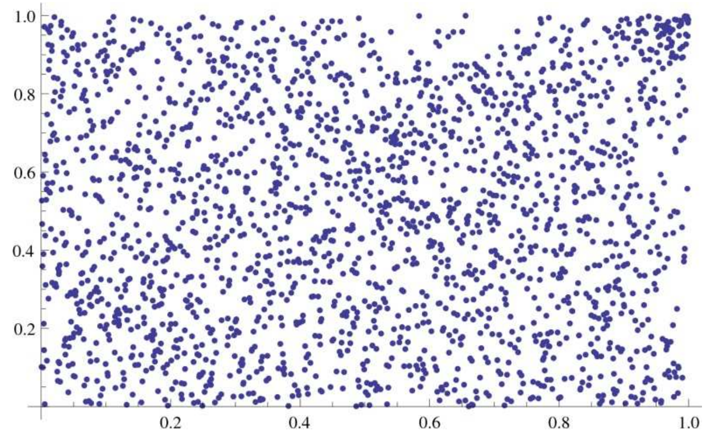
Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i) — needs to be i.i.d. (independent and identically distributed)

↓

transform data to an i.i.d. data set (X_i^*, Y_i^*) — ARMA-GARCH transformation is applied to get i.i.d. data

↓



scatter plot of the iid residuals

Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i)

needs to be i.i.d. (independent and identically distributed)

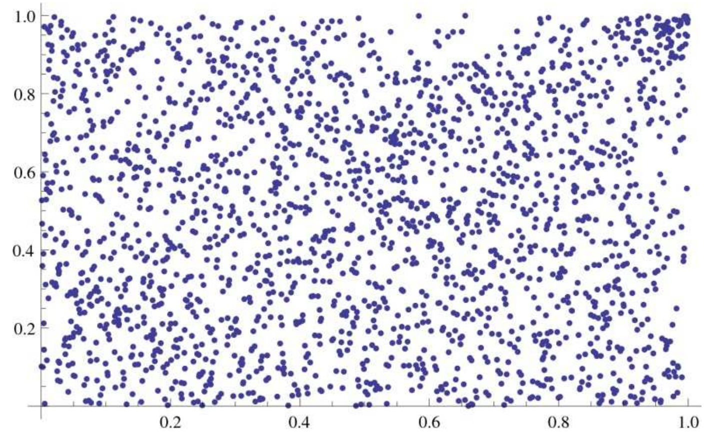
transform data to an i.i.d. data set (X_i^*, Y_i^*)

ARMA-GARCH transformation is applied to get i.i.d. data

ranks (R_i, S_i) of the i.i.d. data

ranks have the same dependence structure as original data

➡ same Copula



scatter plot of the ranks of the iid residuals

Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i)

needs to be i.i.d. (independent and identically distributed)

transform data to an i.i.d. data set (X_i^*, Y_i^*)

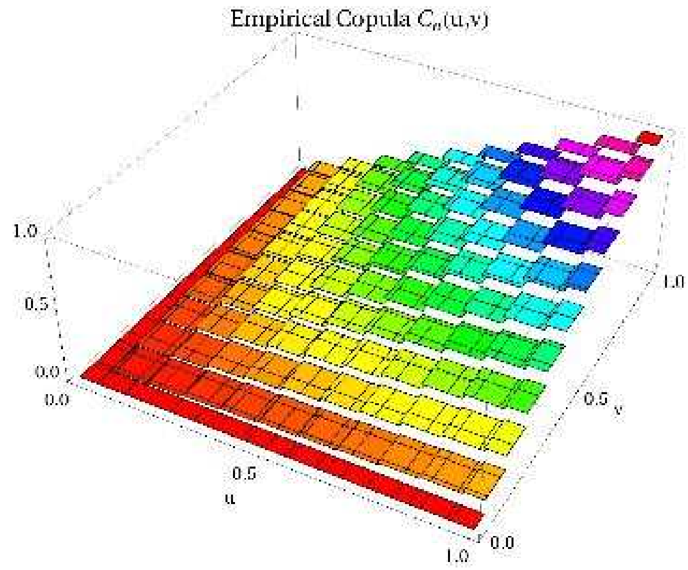
ARMA-GARCH transformation is applied to get i.i.d. data

ranks (R_i, S_i) of the i.i.d. data

ranks have the same dependence structure as original data

➡ same Copula

calculate the empirical copula $C_n(u,v)$



CDF of the empirical Copula

Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i) — needs to be i.i.d. (independent and identically distributed)

↓

transform data to an i.i.d. data set (X_i^*, Y_i^*) — ARMA-GARCH transformation is applied to get i.i.d. data

↓

ranks (R_i, S_i) of the i.i.d. data — ranks have the same dependence structure as original data

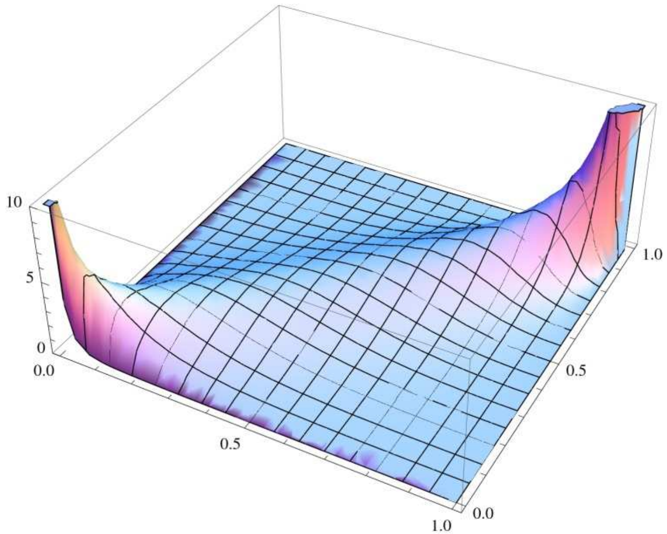
➡ same Copula

↓

calculate the empirical copula $C_n(u,v)$

↓

estimate the Copula parameter and decide which copula family is appropriate



PDF of the Gumbel-Hougaard Copula

Find a theoretical Copula to model the dependency between model output and measurement

data (X_i, Y_i)

needs to be i.i.d. (independent and identically distributed)

transform data to an i.i.d. data set (X_i^*, Y_i^*)

ARMA-GARCH transformation is applied to get i.i.d. data

ranks (R_i, S_i) of the i.i.d. data

ranks have the same dependence structure as original data

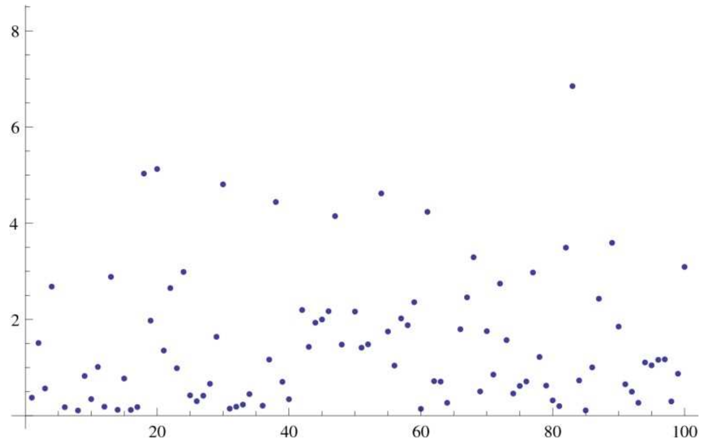
➡ same Copula

calculate the empirical copula $C_n(u,v)$

estimate the Copula parameter and decide which copula family is appropriate

Simulate data from Copula

Back transformation of simulated data to a time series

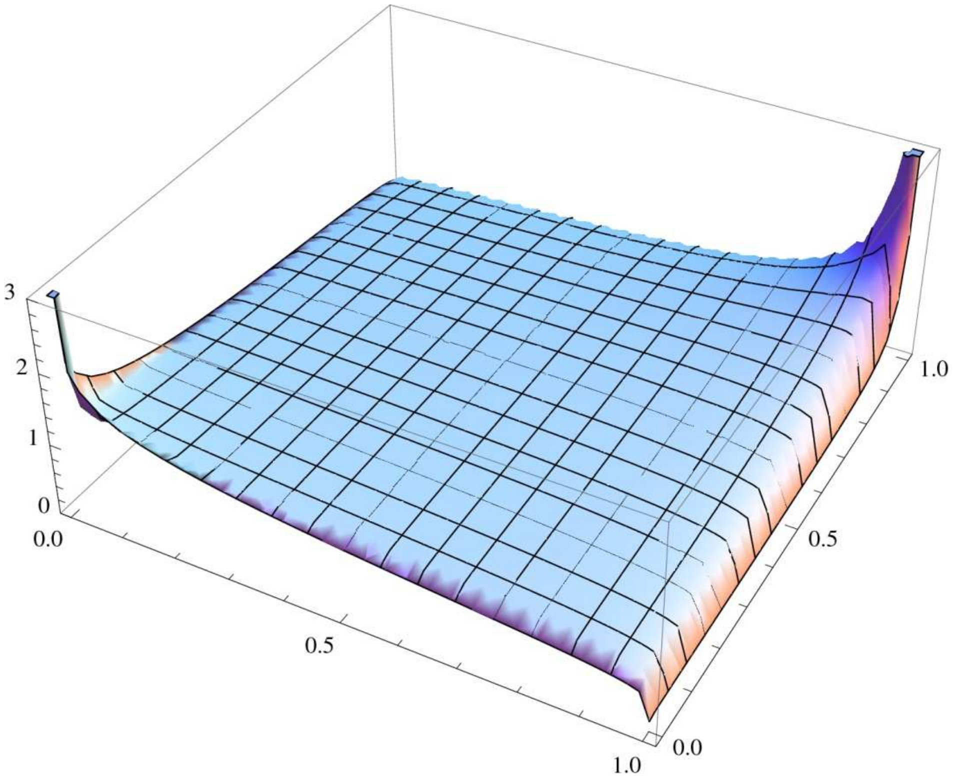


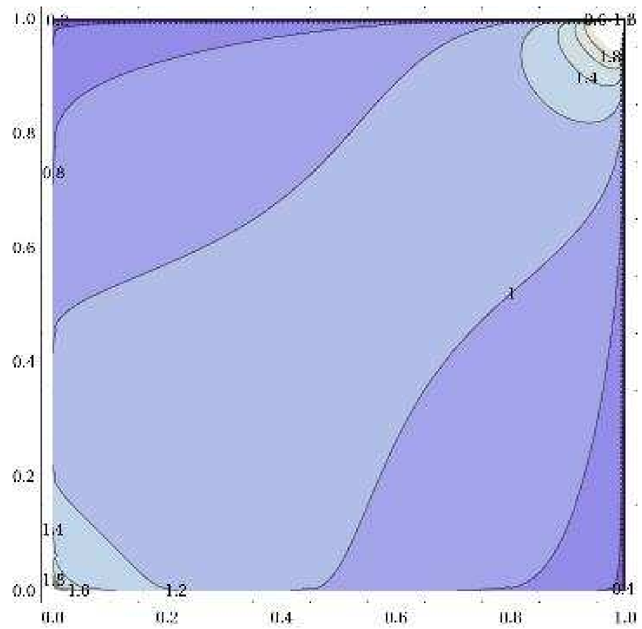
scatter plot of simulated data

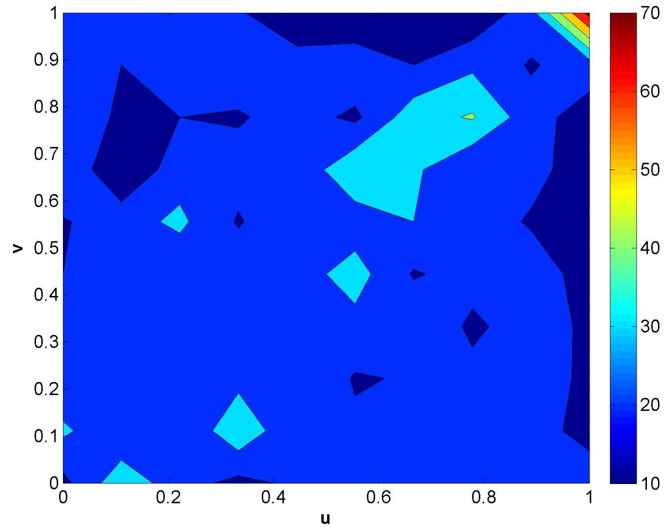
Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

After GOF-tests the Gumbel-Hougaard Copula was identified as appropriate

$\theta = 1.1$







PDF of the Gumbel-Hougaard Copula with $\theta = 1.1$

Modelling data from Copula

RCM

(X_i, Y_i)

observations

ARMA-GARCH:
iid residuals
$(\tilde{X}_i, \tilde{Y}_i)$

marginal
$F_X(x)$

Ranks:
$(\tilde{R}_i, \tilde{S}_i)$

marginal
$F_Y(y)$

Copula, θ

joint density

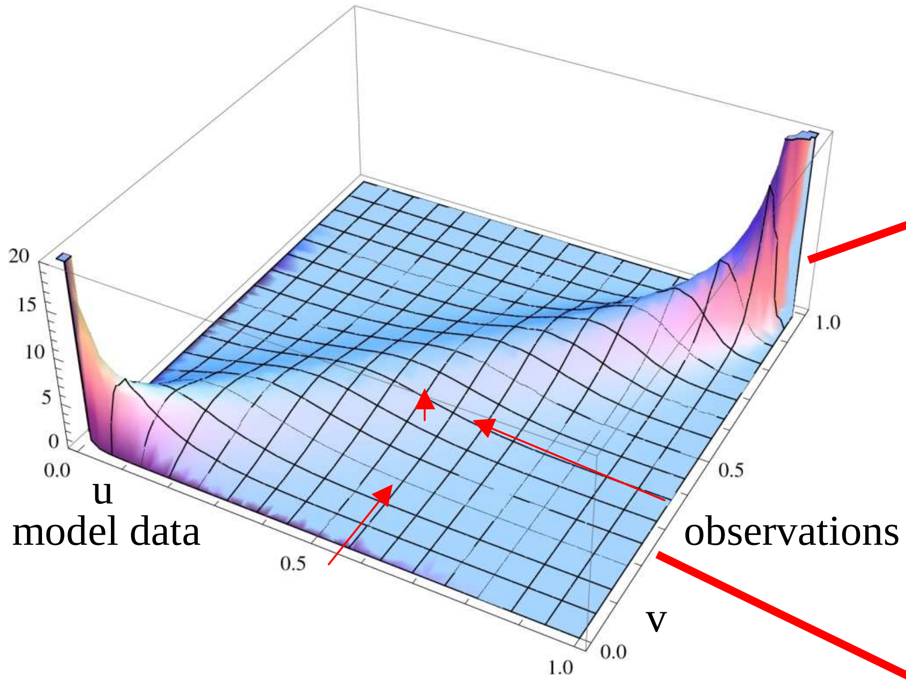
$$f_{XY}(x,y) = c_\theta(F_X(x), F_Y(y)) \cdot f_X(x) \cdot f_Y(y)$$

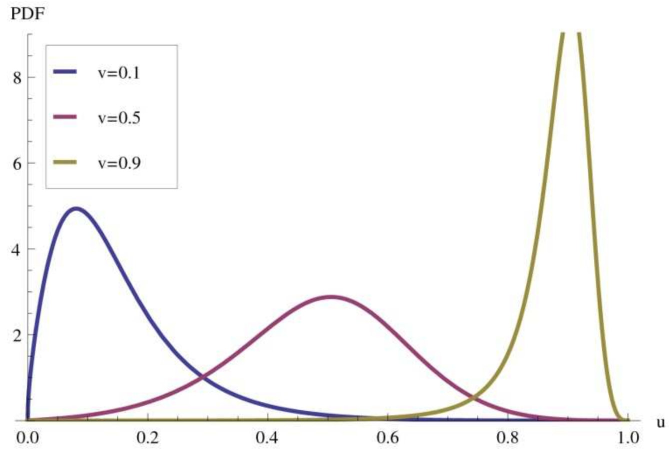
New features

Allows to apply the procedure to the whole range of data!

Expansion to the multivariate case possible

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University



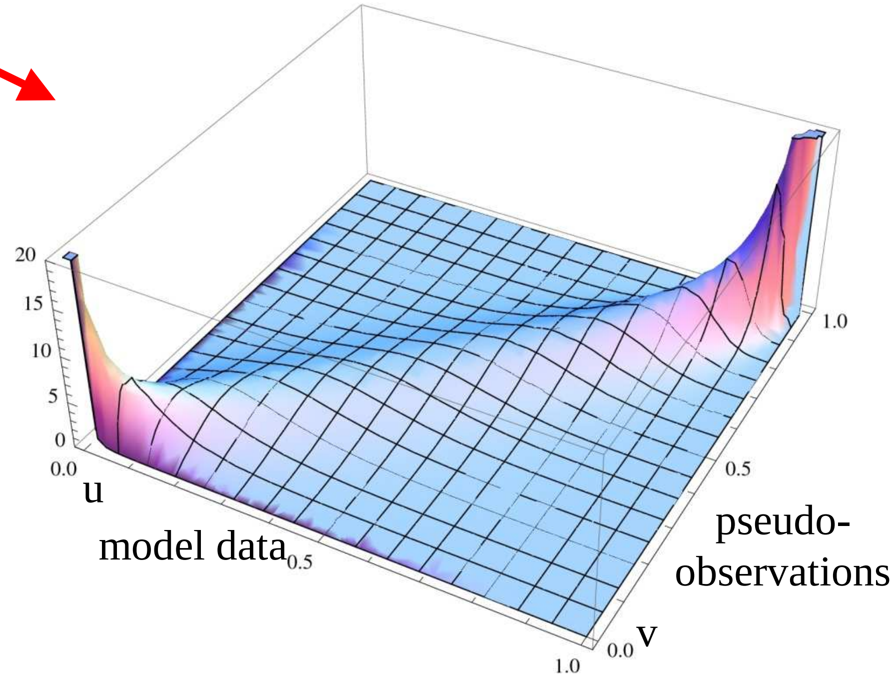


For a fixed v the PDF gives the probability that the corresponding u takes a certain value

Assumption:

•The joint PDF of model data and observation is the same as that of model data and pseudo-observation (BC)

•Pseuodo-obs. can be constructed from the PDF and therefore from the Copula!

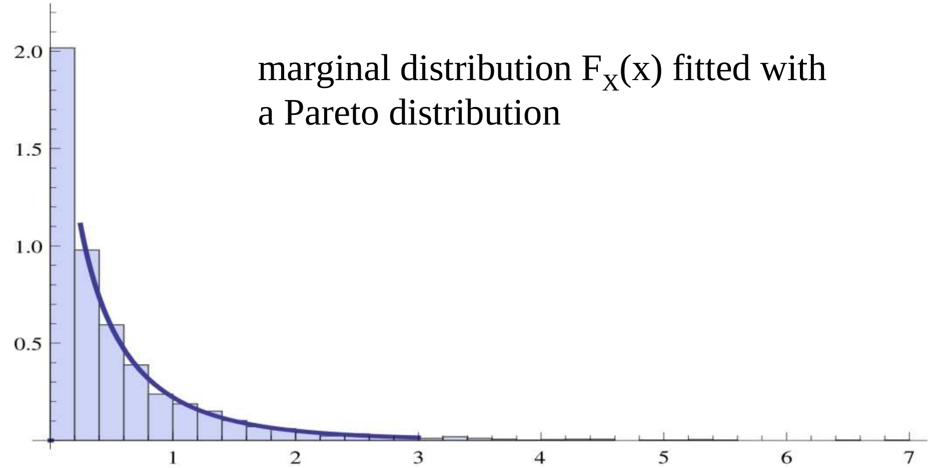


Algorithm for a conditional simulation of (X,Y)

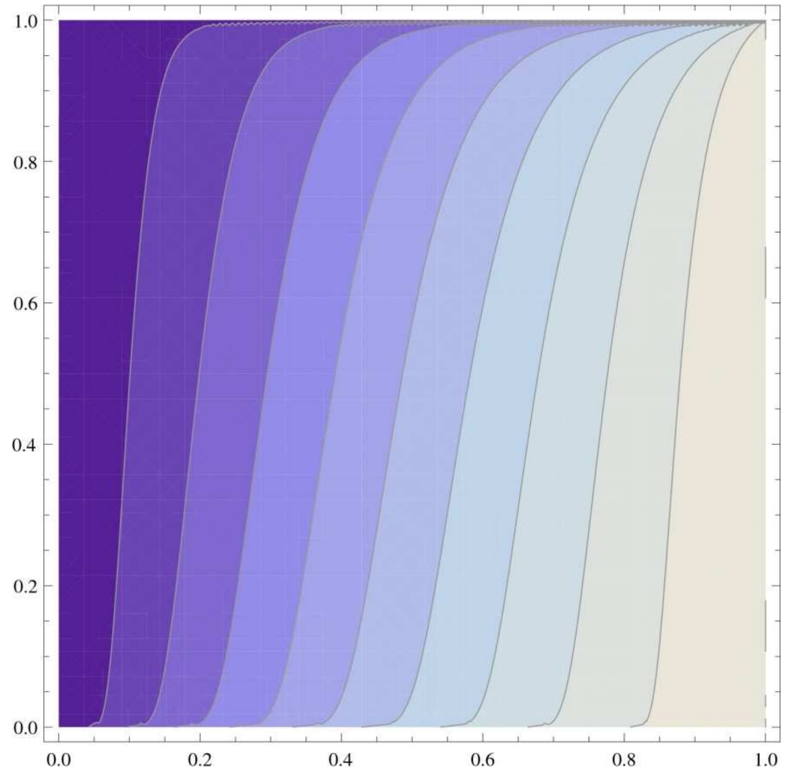
- calculate $u = F_X(x)$
- create random samples of v under the condition u using

$$\boxed{\frac{\partial}{\partial u} C(u,v) = P(V \le v, U = u)}$$

- use $F_Y^{-1}(v) = y$ to calculate a sample of model values y
- based on the conditional CDF there is a **range of possible values for the variable y**

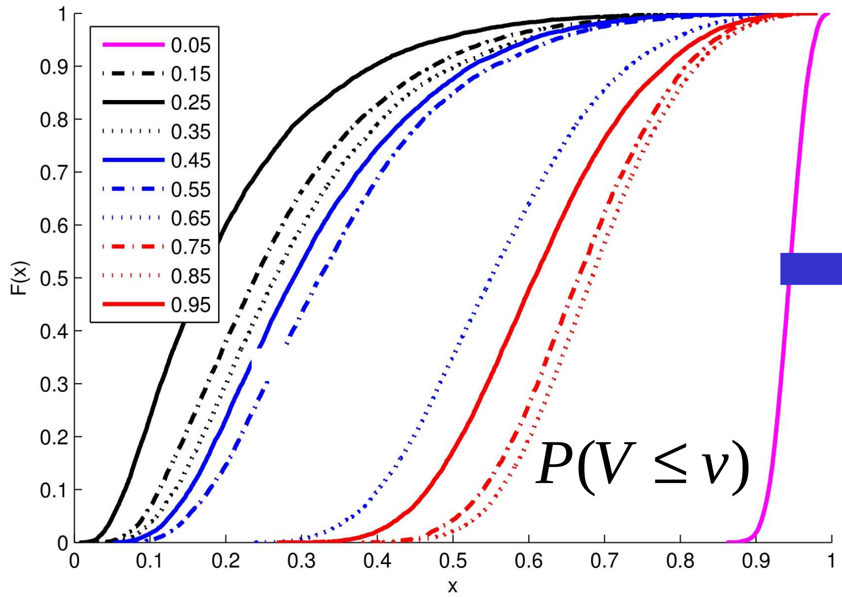


marginal distribution $F_X(x)$ fitted with a Pareto distribution

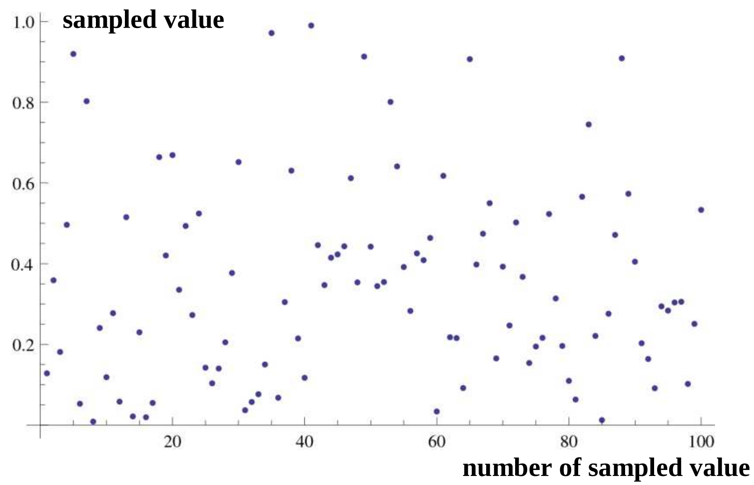


Conditional CDFs for the whole range of data

• now for each u in the time series a random sample is drawn from the conditional CDF

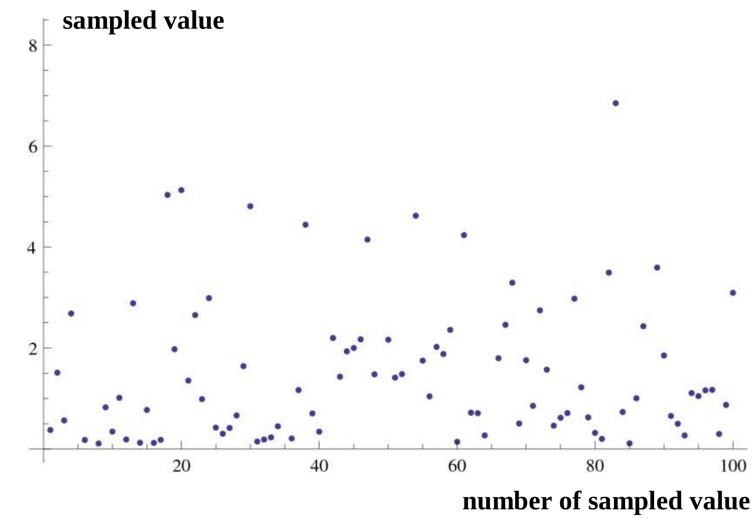


conditional CDFs for different values of u



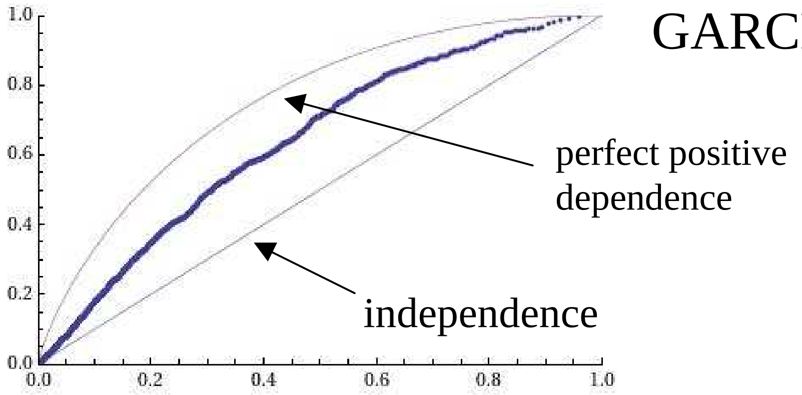
• the corresponding y-values are derived using the inverse marginal CDF

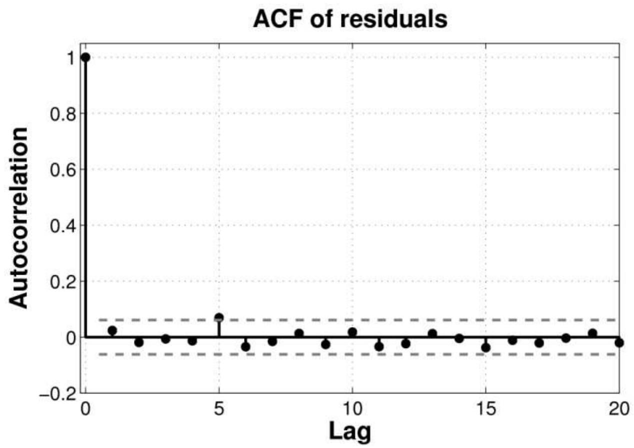
$$F_Y^{-1}(v) = y$$



Results

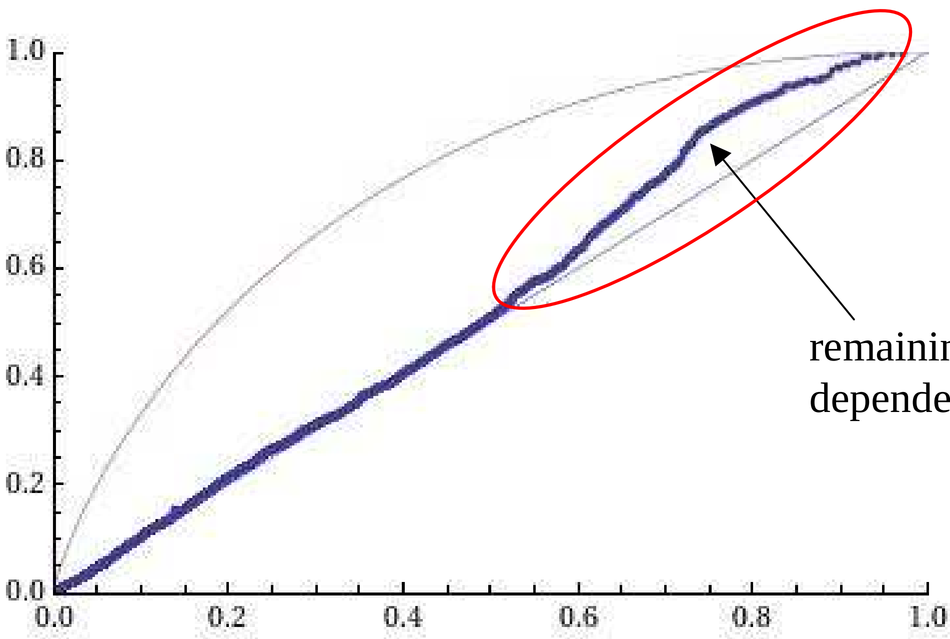
before ARMA-GARCH





after ARMA-GARCH

K-plots of time series before and after ARMA-GARCH transformation



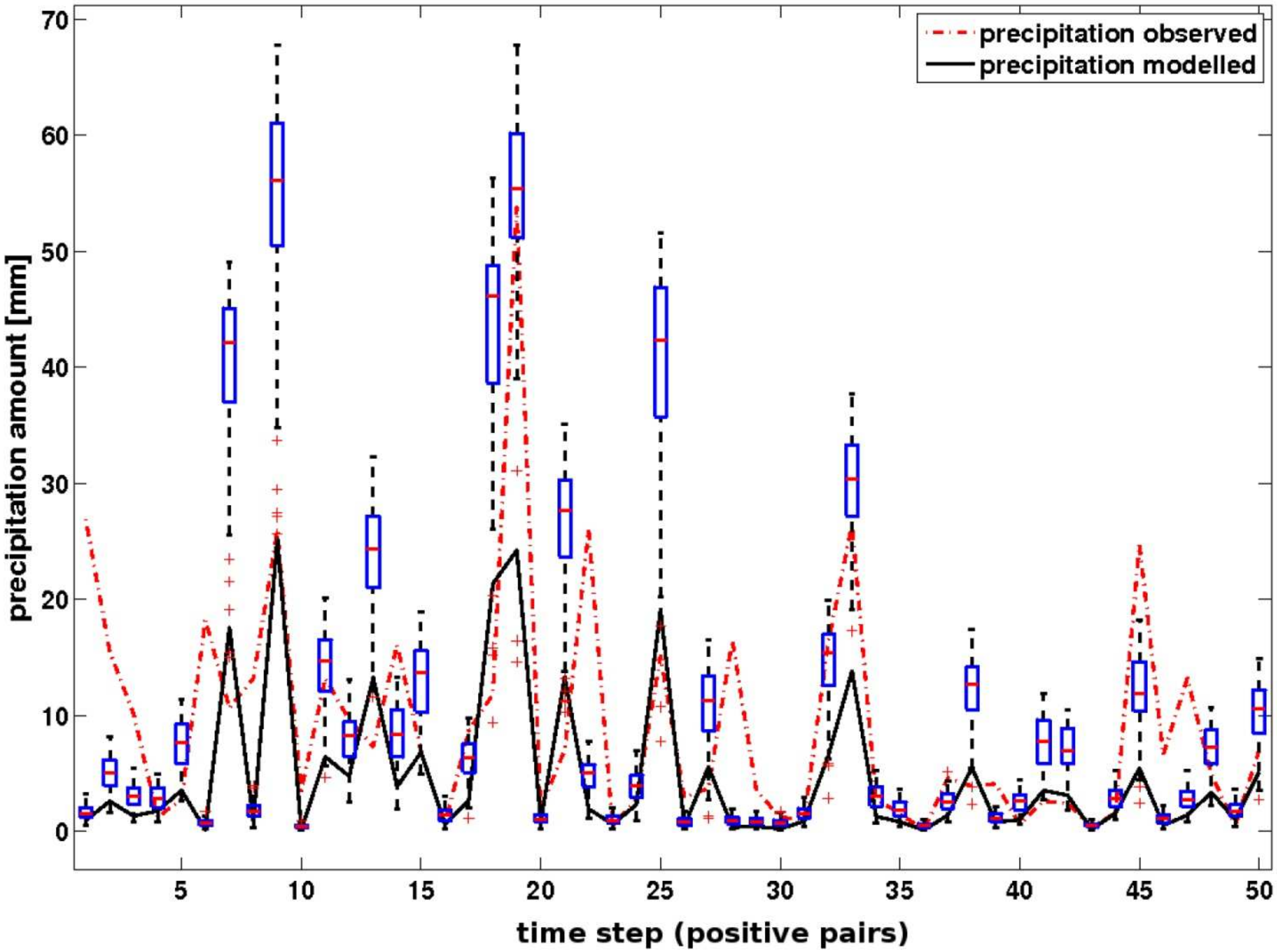
remaining positive dependence for high ranks

Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H. (2011): Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, HESS D 8, 3001–3045, 2011.

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann

Universität Augsburg University

Uncond. Pseudo-observations



Institut für Geographie
Lehrstuhl für
Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann

Universität
Augsburg
University

Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H. (2011): Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, HESS D 8, 3001–3045, 2011.

Weather classification - an improvement?

40 predefined types of weather classes
CP1 ....CP40

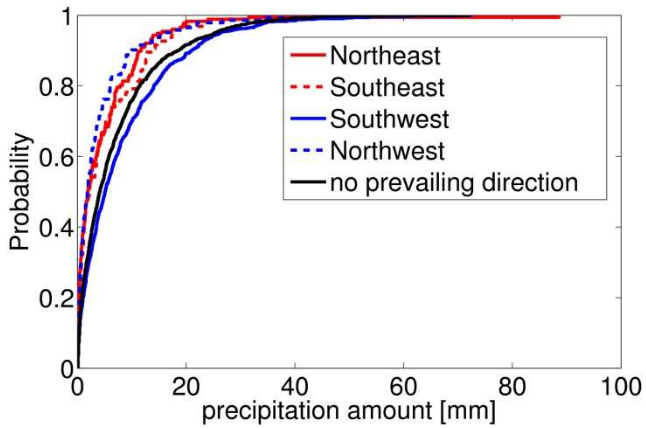
Direction of advection (700 hPa)		Cyclonality 950 hPa and 500 hPa)		Humidity of troposphere
northeasterly	southeasterly	AA	AC	dry
northwesterly	southwesterly	CA	CC	wet

Bissouli and Dittmann, (2001)

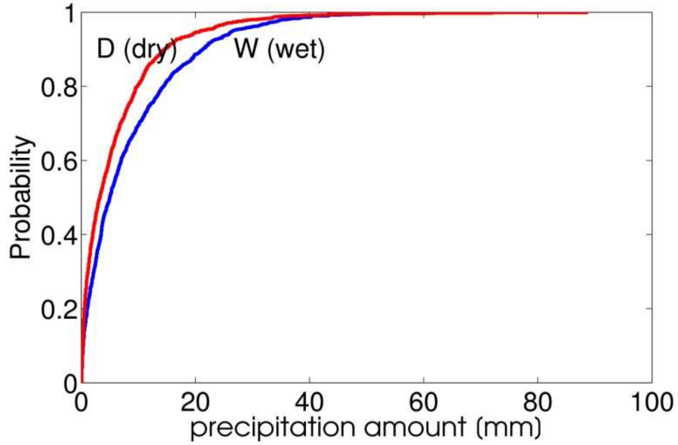
For each group of weather types a theoretical Copula model is estimated separately

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

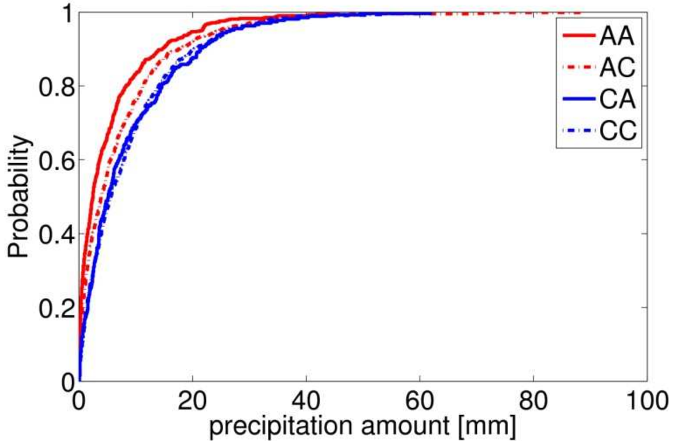
Dependence on large scale-weather situation



different advection types



humidity types



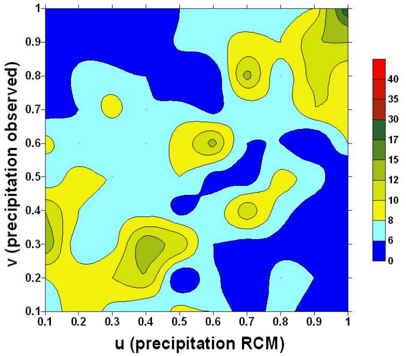
cyclonality

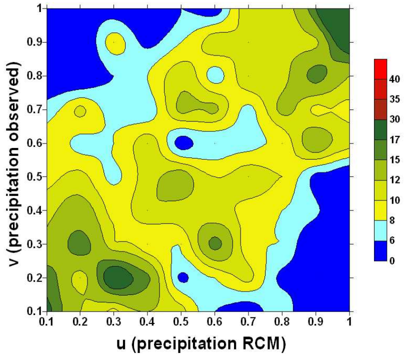
Marginal distributions for different weather conditions

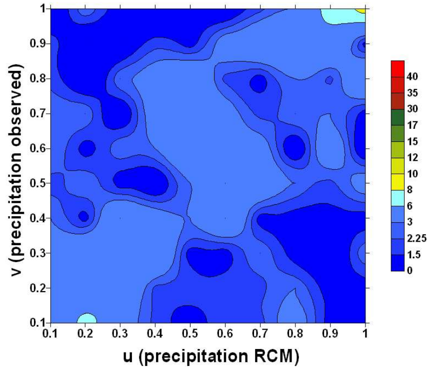
Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H. (2011): Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, HESS D 8, 3001–3045, 2011.

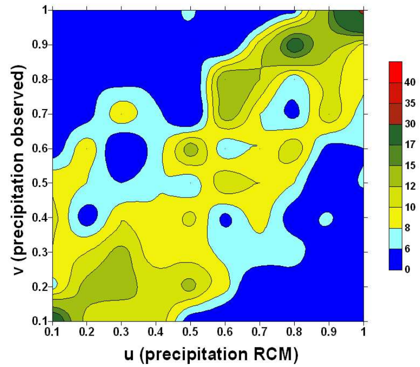
Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

Empirical Copula densities

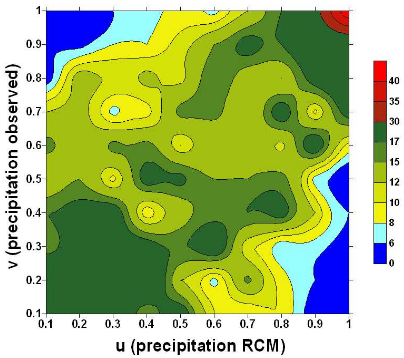


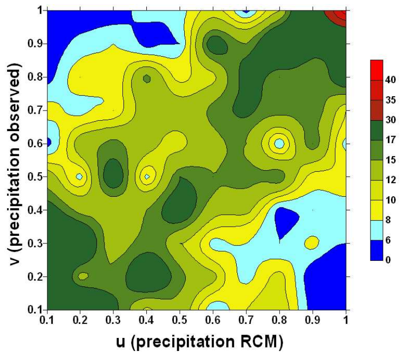


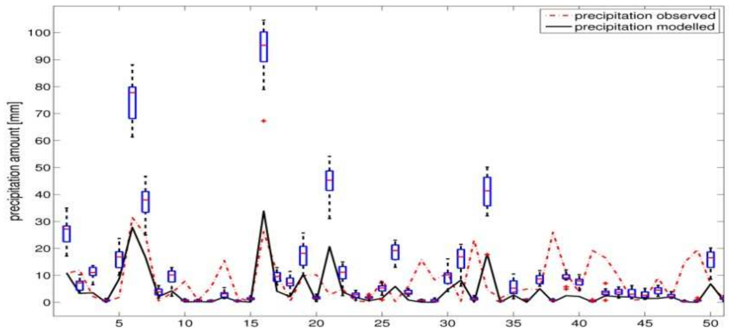




cyclonality





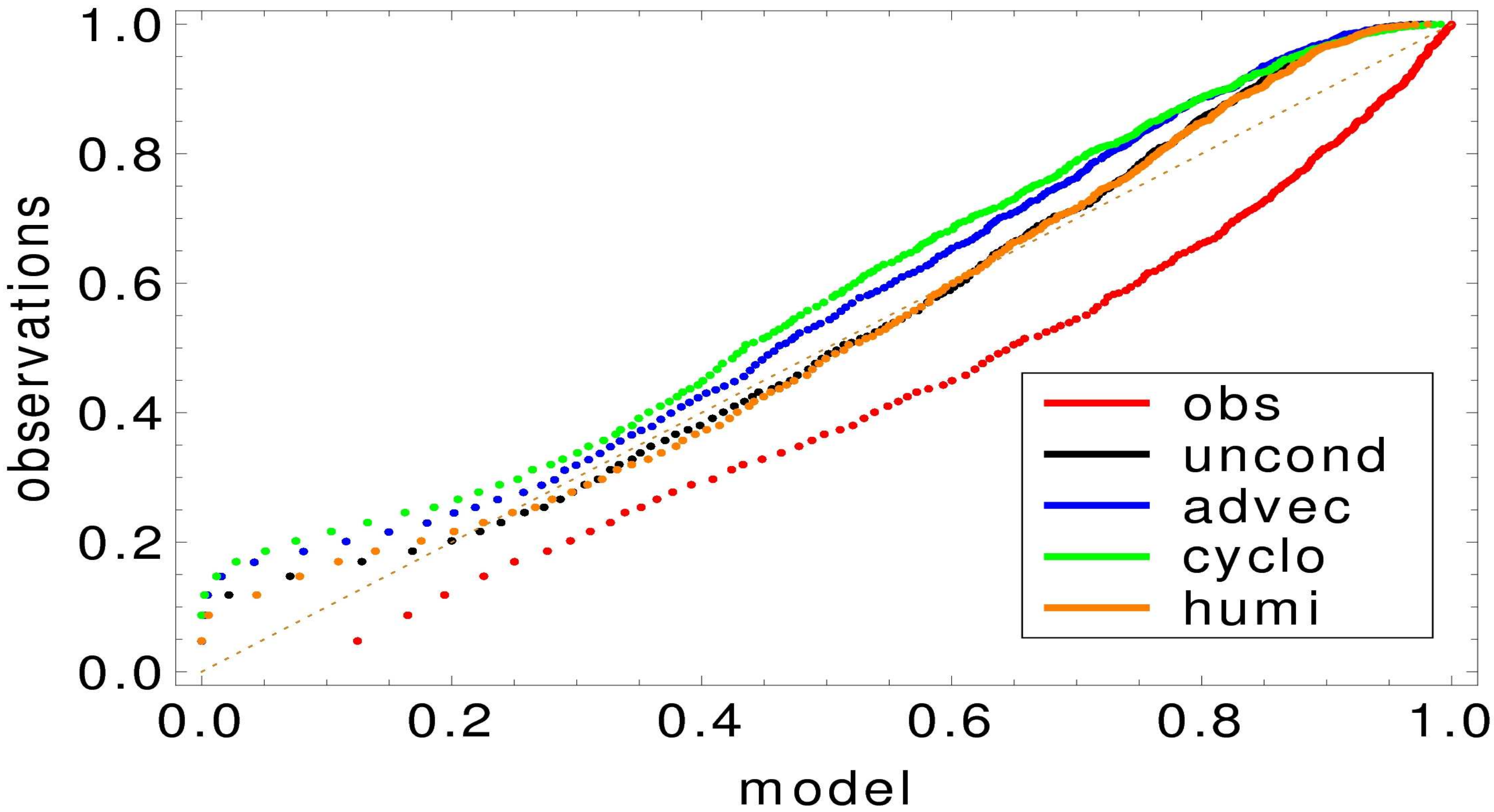


humidity

Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H. (2011): Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, HESS D 8, 3001–3045, 2011.

IGUA Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
UNIA Universität Augsburg University

Probability plots for different cond. simulations



	uncond	advec	cyclo	humi
obs	0.36	0.43	0.45	0.37

Pearson Correlation coefficients

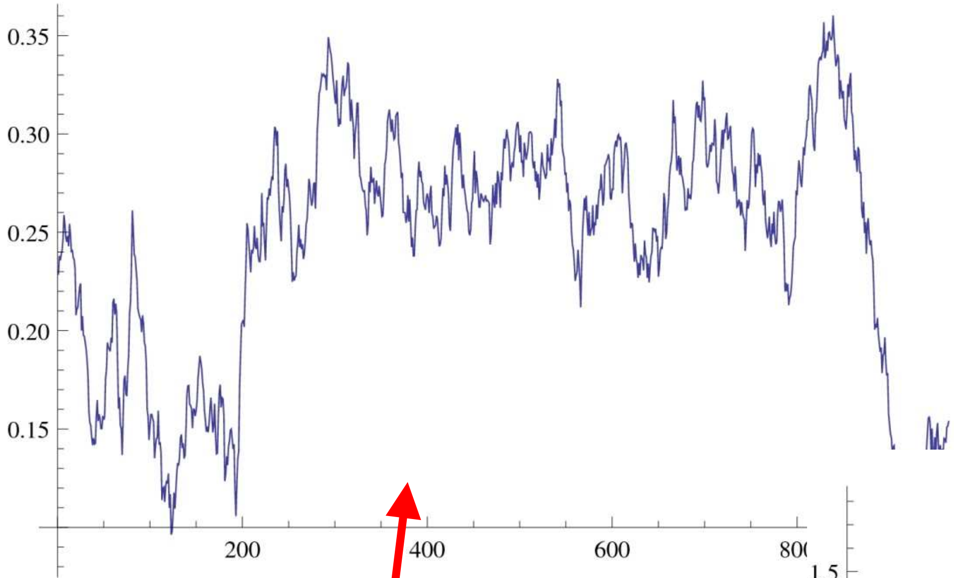
Conditioning on cyclonity gives the best result

Laux, P., Vogl, S., Qiu, W., Knoche, H. R., and Kunstmann, H. (2011): Copula-based statistical refinement of precipitation in RCM simulations over complex terrain, HESS D 8, 3001–3045, 2011.

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

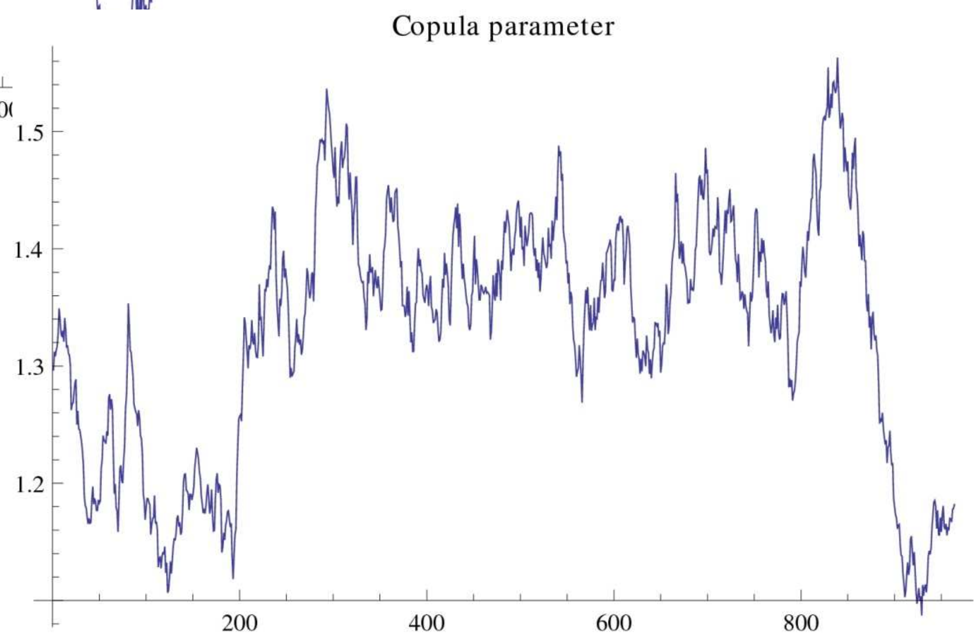
Possible improvement

$$\theta \approx \frac{1}{1-\tau}$$



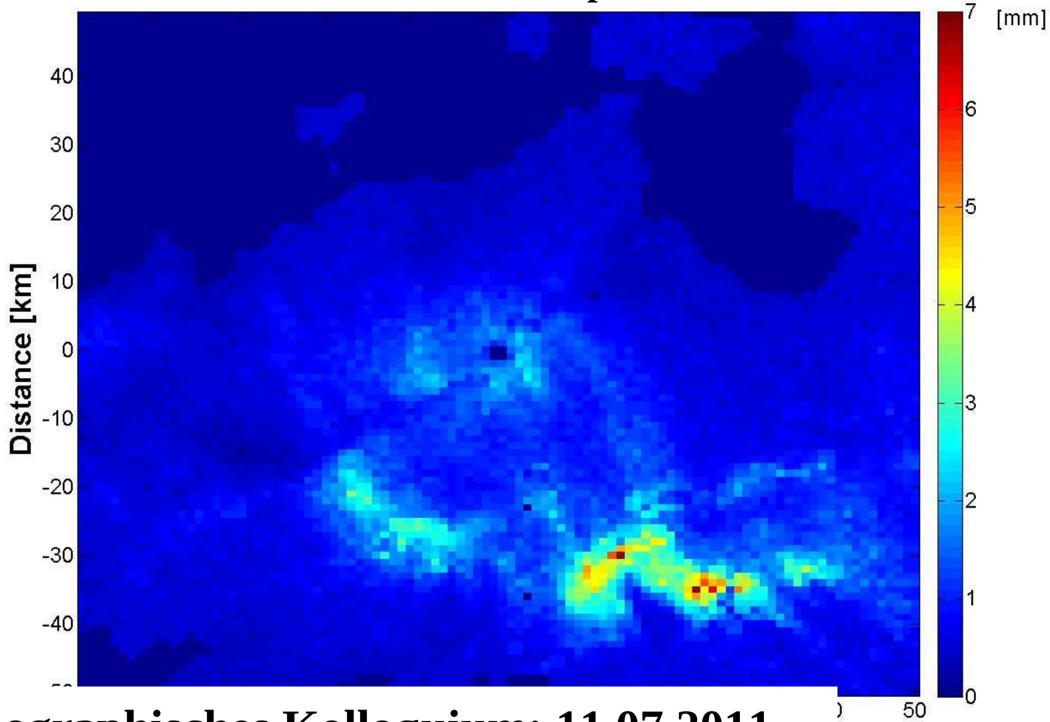
τ is not constant over the range of data

θ is not constant over the range of data



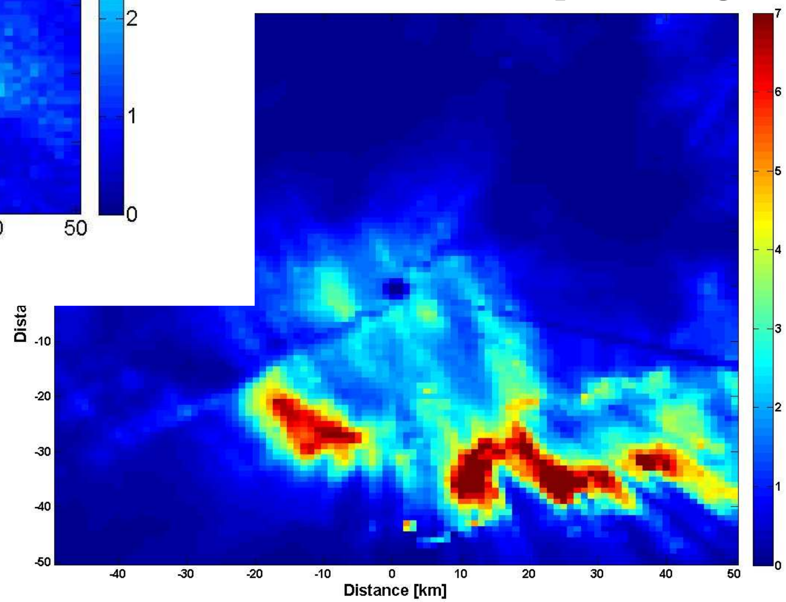
Spatial application

Field of pseudo-observations



Using radar fields for copula-based spatial interpolation of gauge-data

Radar Hohenpeißenberg



Geographisches Kolloquium: 11.07.2011

Wei Qiu

Copula based Precipitation Estimation by combining Gauge-, Radar, and Microwave link observations

Institut für Geographie
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann
Universität Augsburg University

Future work

- ➢Compare the results of the new approach with traditional methods of bias correction
- ➢Extension to (0,1) and (1,0), and (0,0) - case
- ➢Develop a copula that is tailored to the particular needs of hydrological/meteorological data such as precipitation, temperature etc.
- ➢Expand the bivariate to the multivariate case