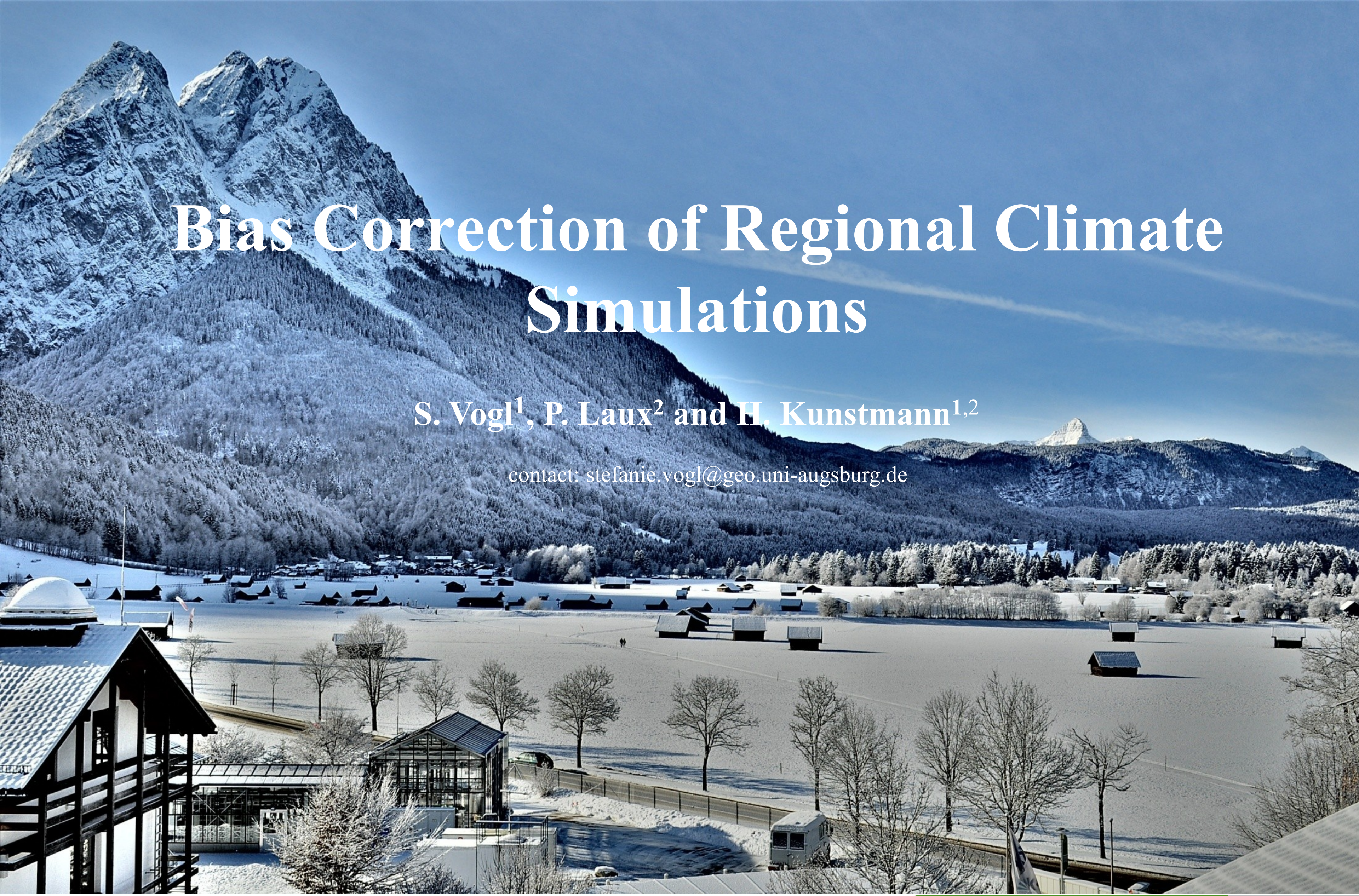


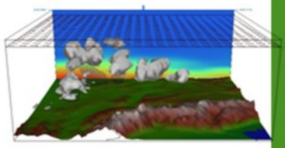
Bias Correction of Regional Climate Simulations

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 (IMK-IFU), Garmisch-Partenkirchen



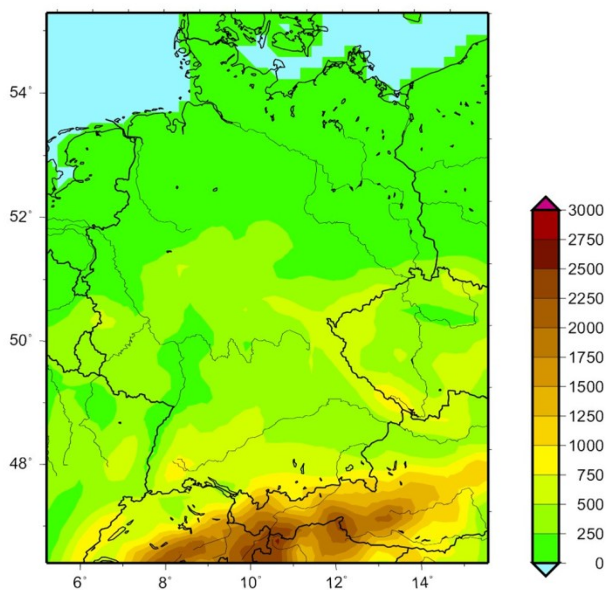
Institut für Geographie

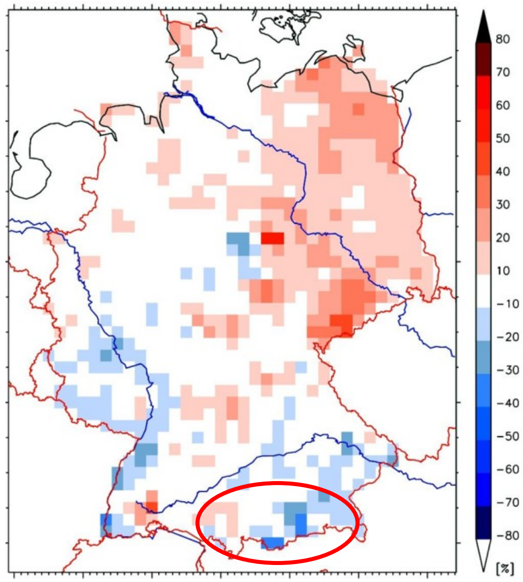
Lehrstuhl für Regionales Klima und Hydrologie
Prof. Dr. Harald Kunstmann

Universität Augsburg University

Application:

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)

- Rainfall is overestimated by MM5 for the whole eastern part of Germany, and strongly underestimated for the Rhine valley and the Alpine region of Germany.
- The underestimation in the Alpine region is possibly due to the complex terrain with very steep gradients of altitude.

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

Theory:

random sample from (X,Y):

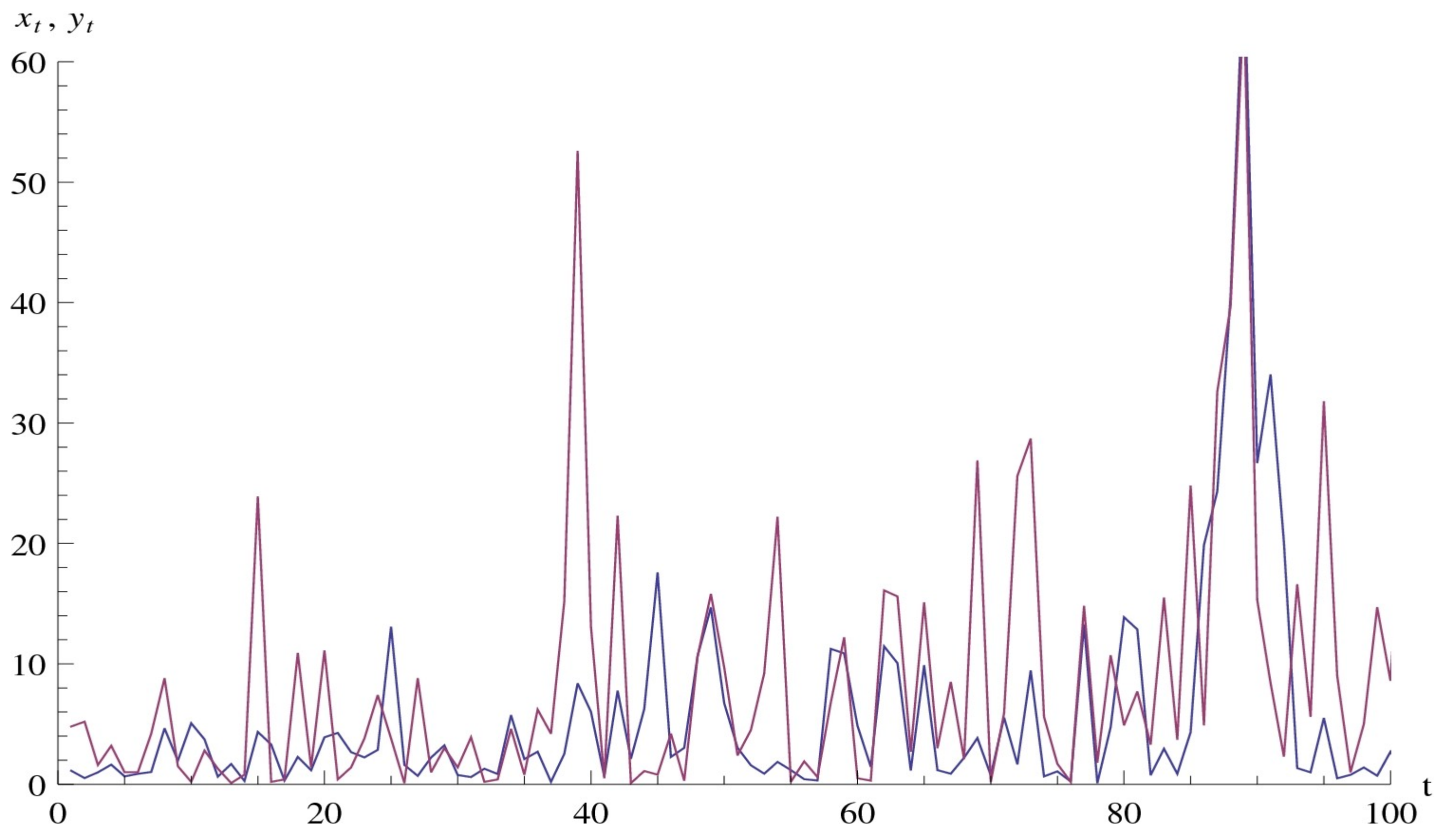
$$(X_1, Y_1), \ldots, (X_n, Y_n)$$

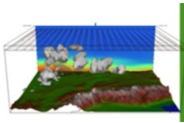
In practice we have two time series: e.g.

•modeled temperature and measured temperature

•modeled precip. and measured precip.

x_t model (blue), y_t measurement (red), station GAP





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

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

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

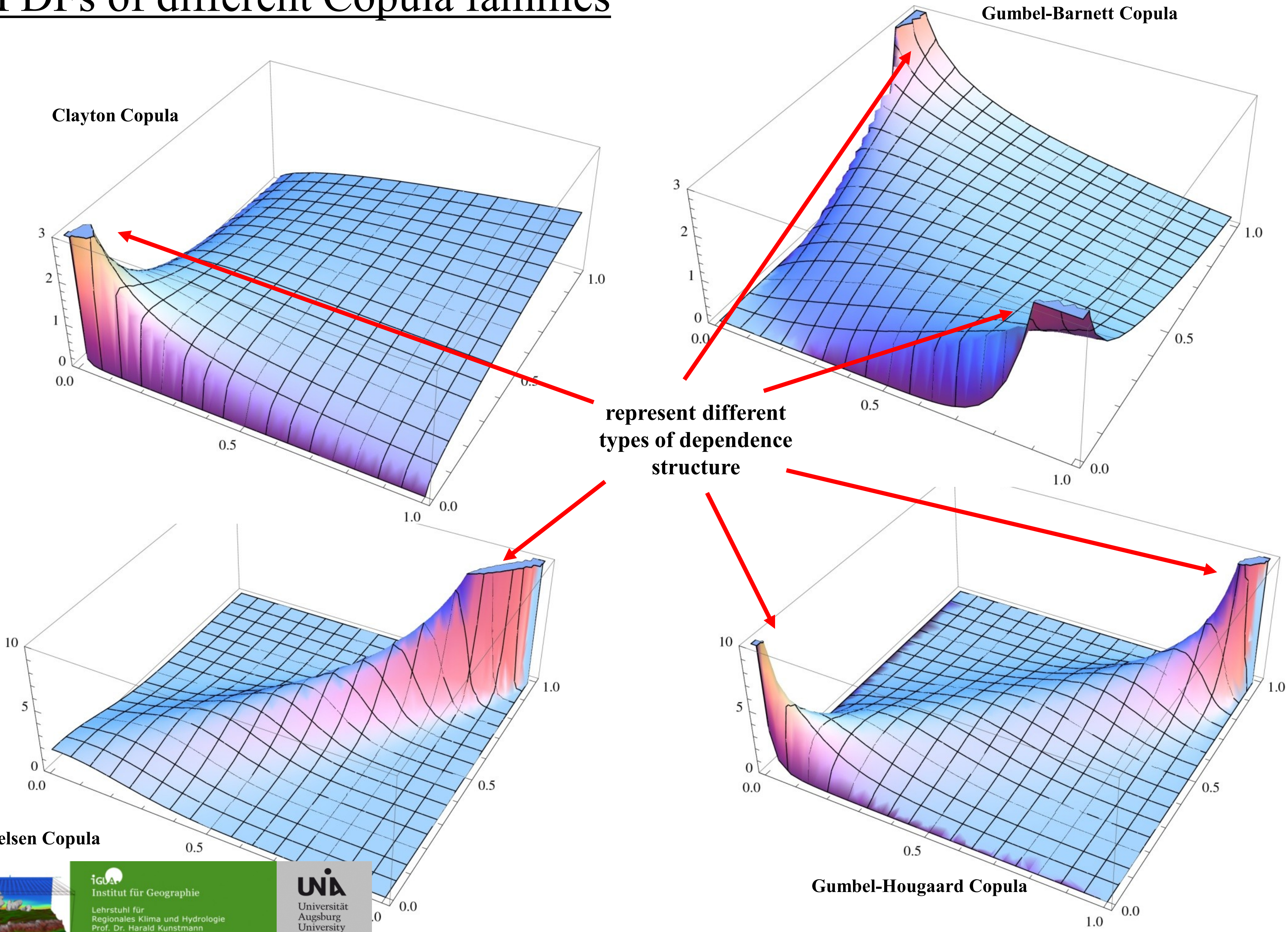
- by using a Copula it is possible to derive a bi- or multivariate PDF f(x,y) just by knowing the single marginal distributions $F_X(x)$ and $F_Y(y)$

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

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

the copula density c(u,v) is often called"dependence function"

PDFs of different Copula families



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

(+) $\tau = \quad n \; n - \quad - \iint \quad -$

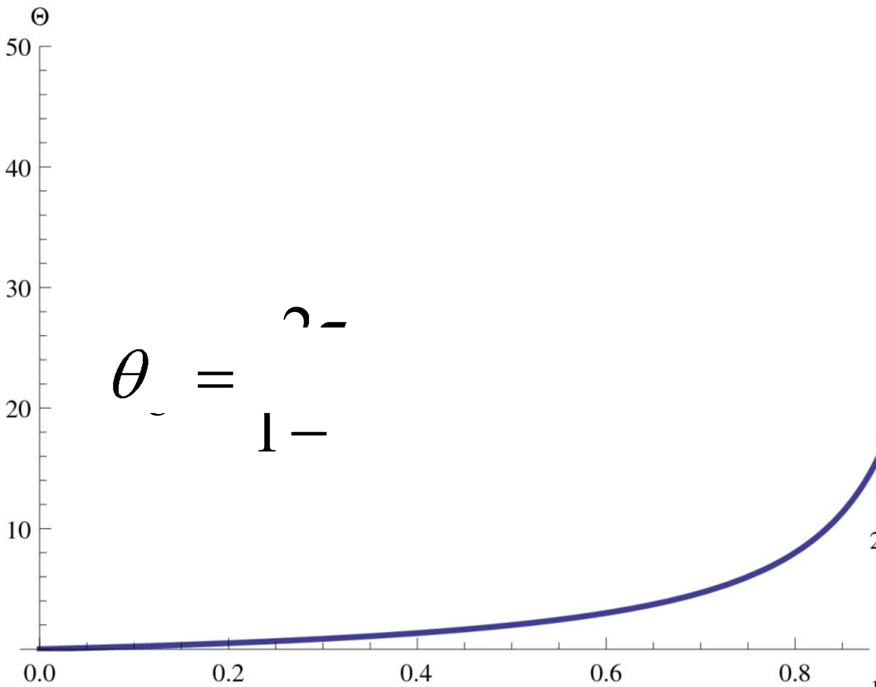
tau is a measure of dependence based on ranks

$C_\theta(\cdot) = \left(\quad \right)$ Clayton Copula

$C_\theta(\cdot) = \left(\quad \right)$ Gumbel-Hougaard Copula

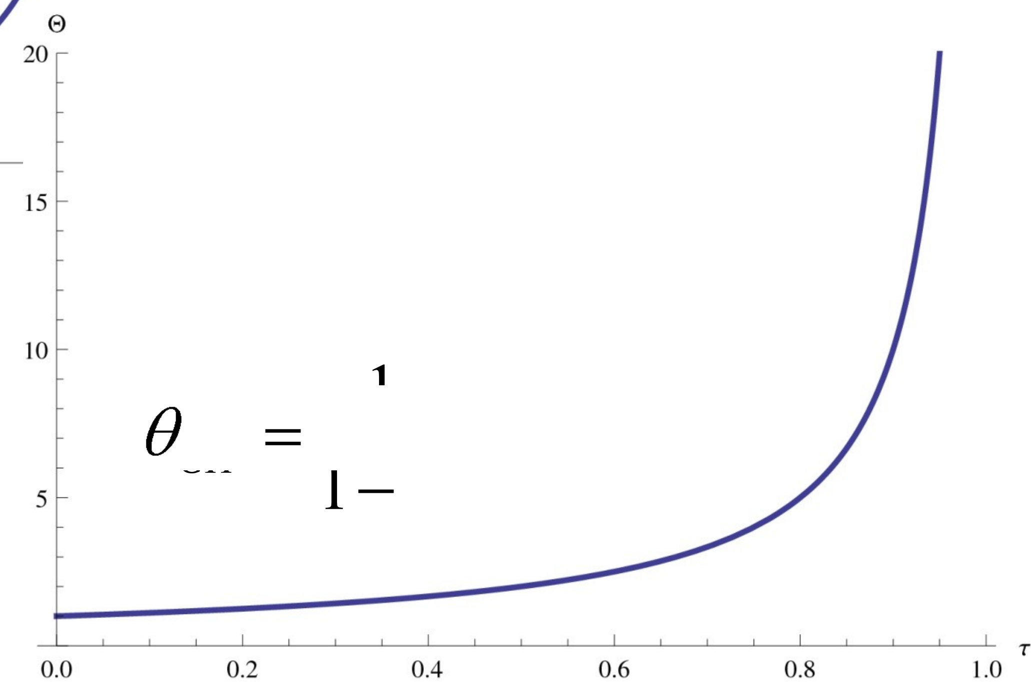
There is a relationship between Kendalls tau and the Copula parameter theta via (+)

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



Clayton Copula

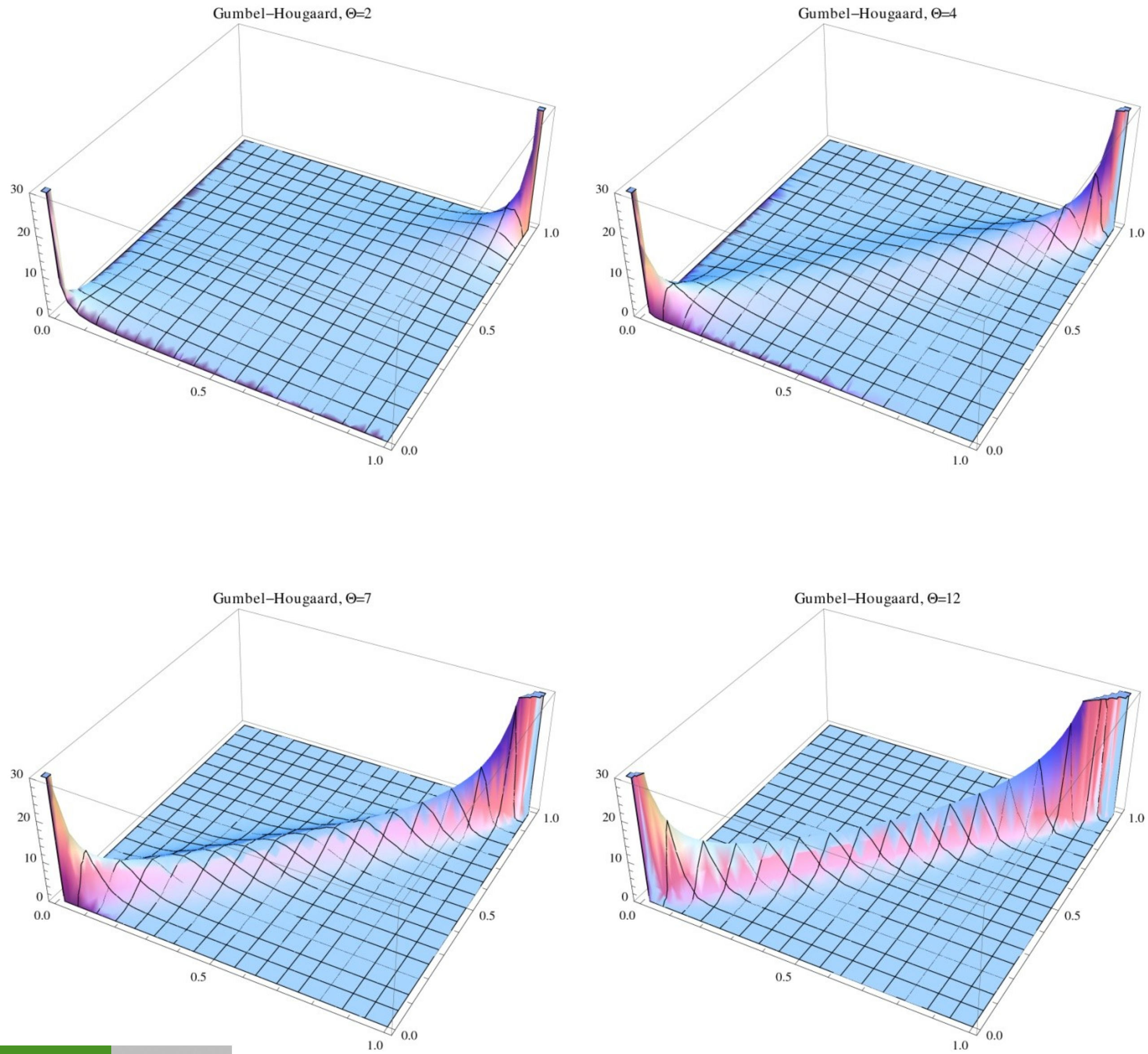
theta is a monotonically increasing function of tau



Gumbel-Hougaard Copula

high Copula parameters indicate strong correlation

larger Copula parameter - higher dependence

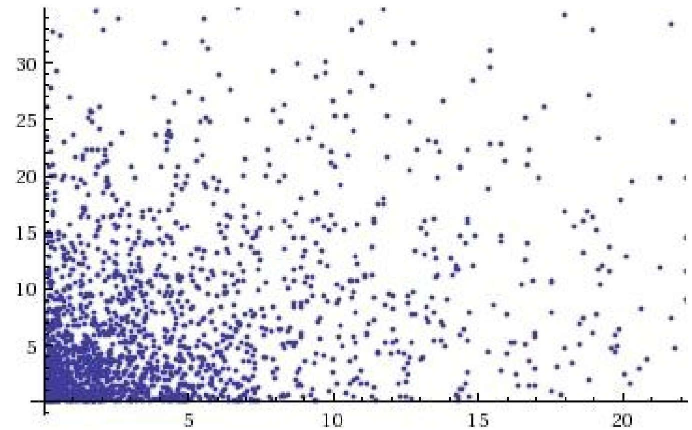


PDF of GH Copula for 4 different values of theta

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

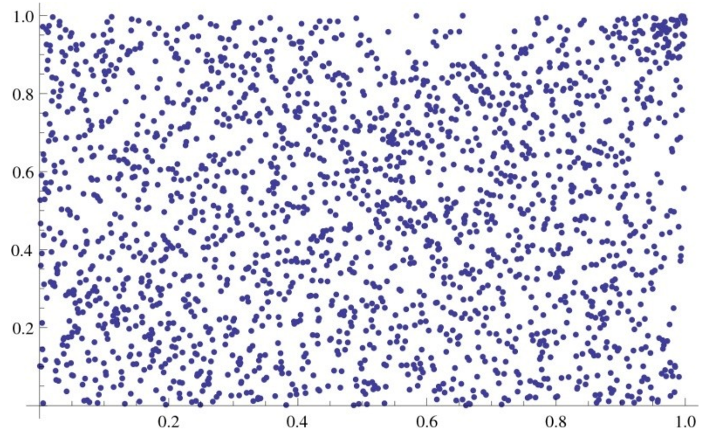
Find a theoretical Copula to model the dependency between model output and measurement

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

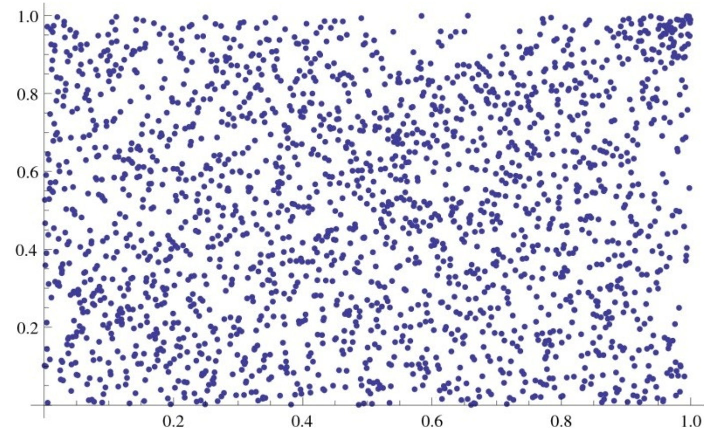
↓

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

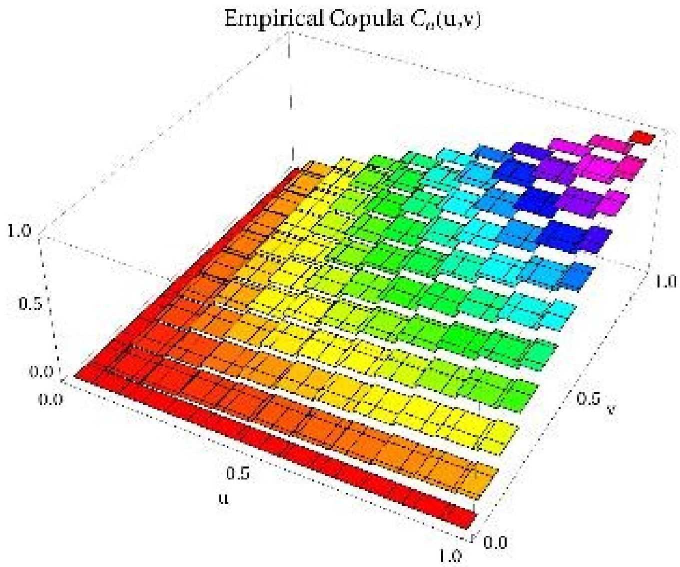
ranks have the same dependence structure as original data

➡ same Copula

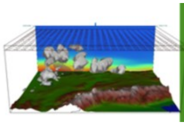
↓

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

↓



CDF of the empirical Copula



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

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

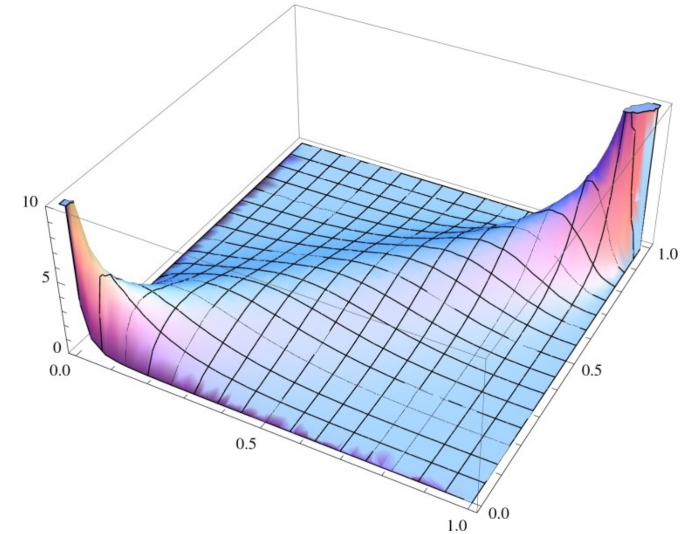
ranks (R_i, S_i) of the i.i.d. data

ranks have the same dependence structure as original data

➡ same Copula

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

decide which copula family is appropriate and estimate the copula parameter



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)

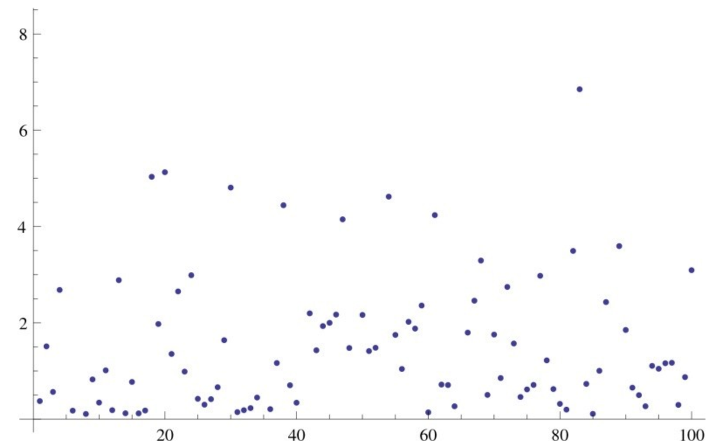
transform data to an i.i.d. data set (X_i^*, Y_i^*)

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



scatter plot of simulated data

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

estimate the Copula parameter and decide which copula family is appropriate

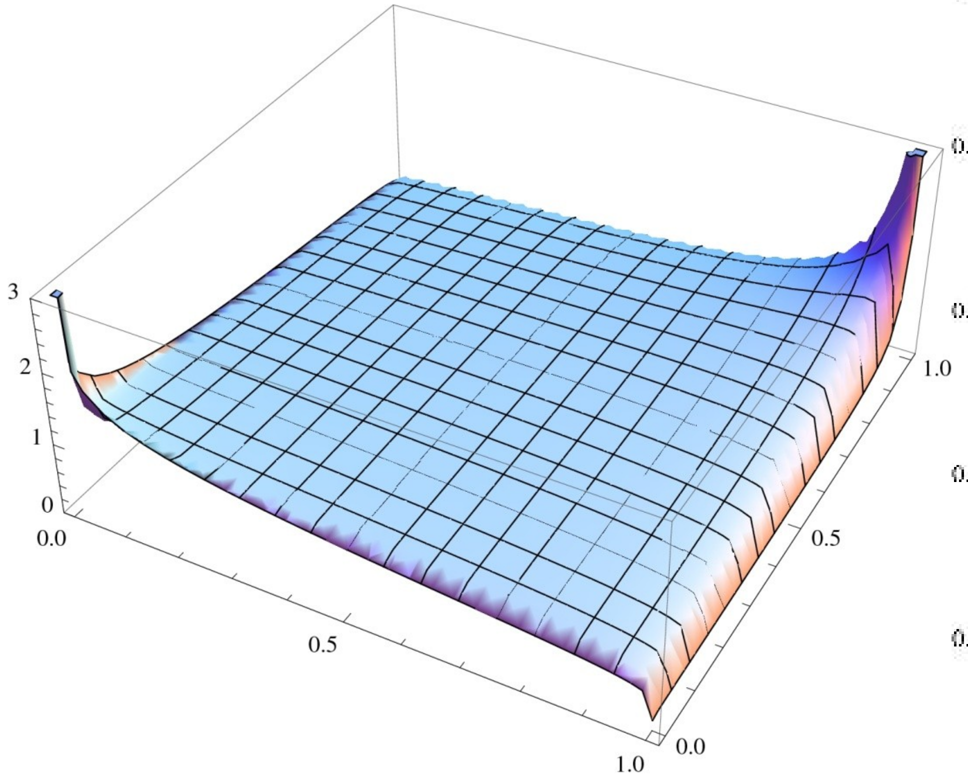
Simulate data from Copula

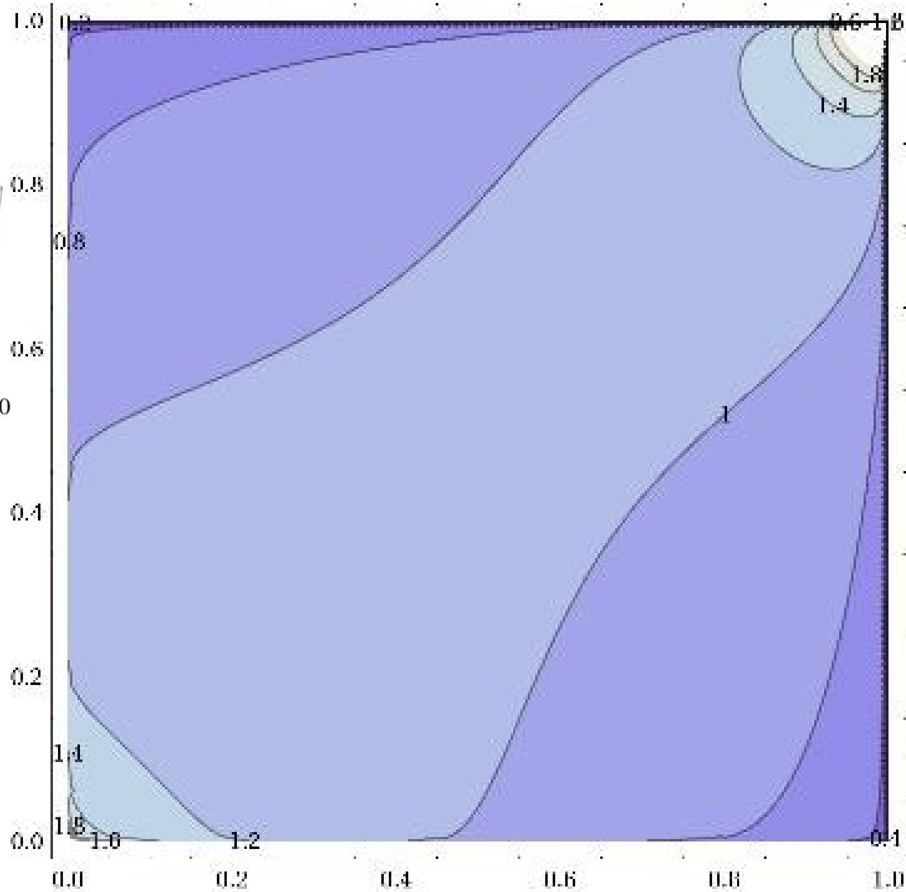
Back transformation of simulated data to a time series

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

$\theta =$

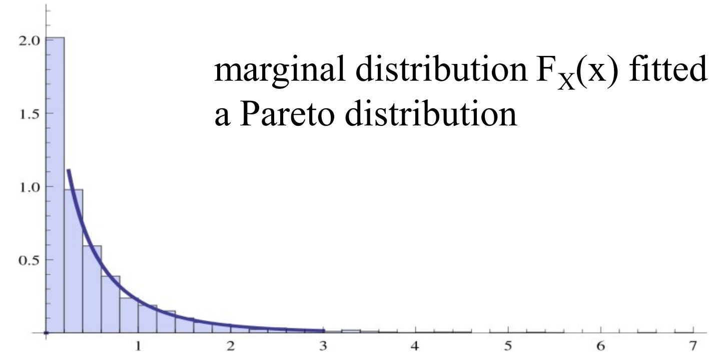




PDF of the Gumbel-Hougaard Copula with $\theta =$

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

Algorithm for a conditional simulation of (X,Y)



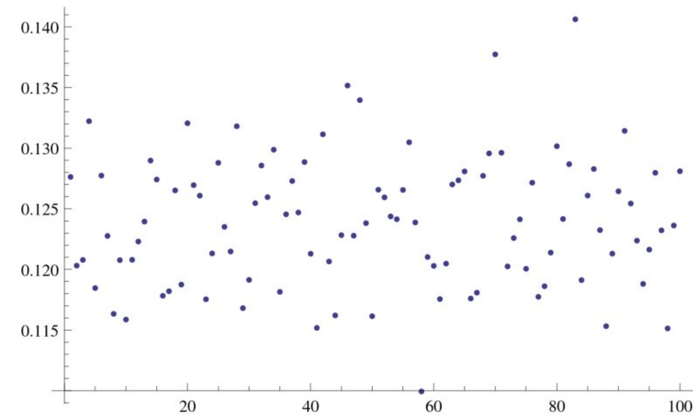
marginal distribution $F_X(x)$ fitted with a Pareto distribution

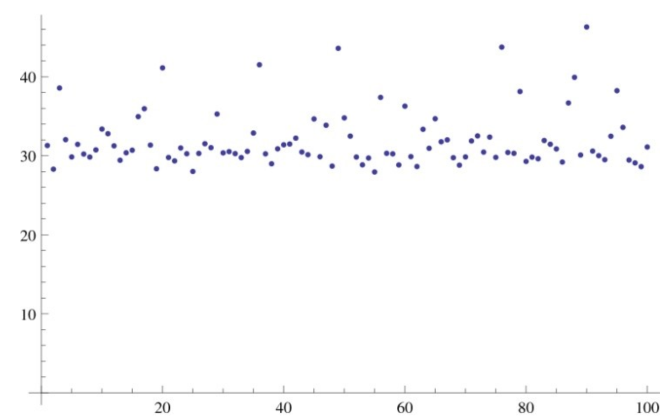
- calculate $u =$
- create random samples of v under the condition u using

$$\frac{\partial}{\partial} \quad = \qquad \leq \qquad =$$

- use $F_Y^{-}(v) =$ to calculate a sample of model values y

- based on the conditional CDF there is a **range of possible values for the variable y**

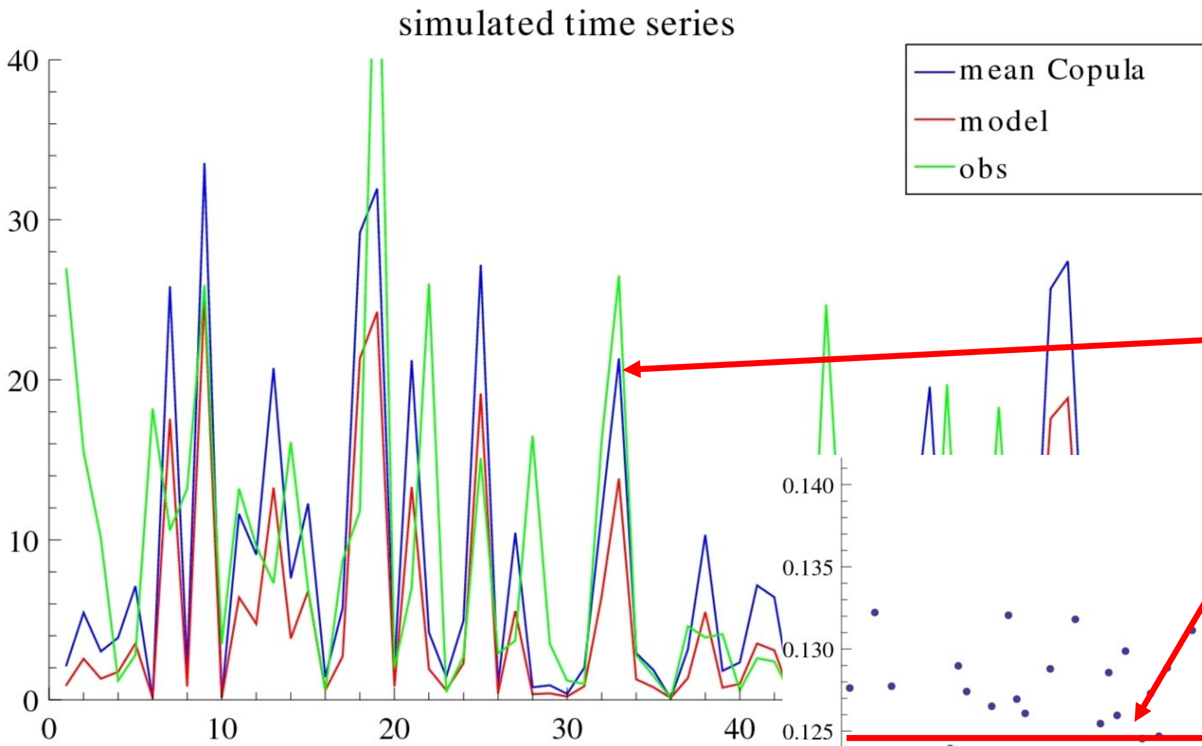


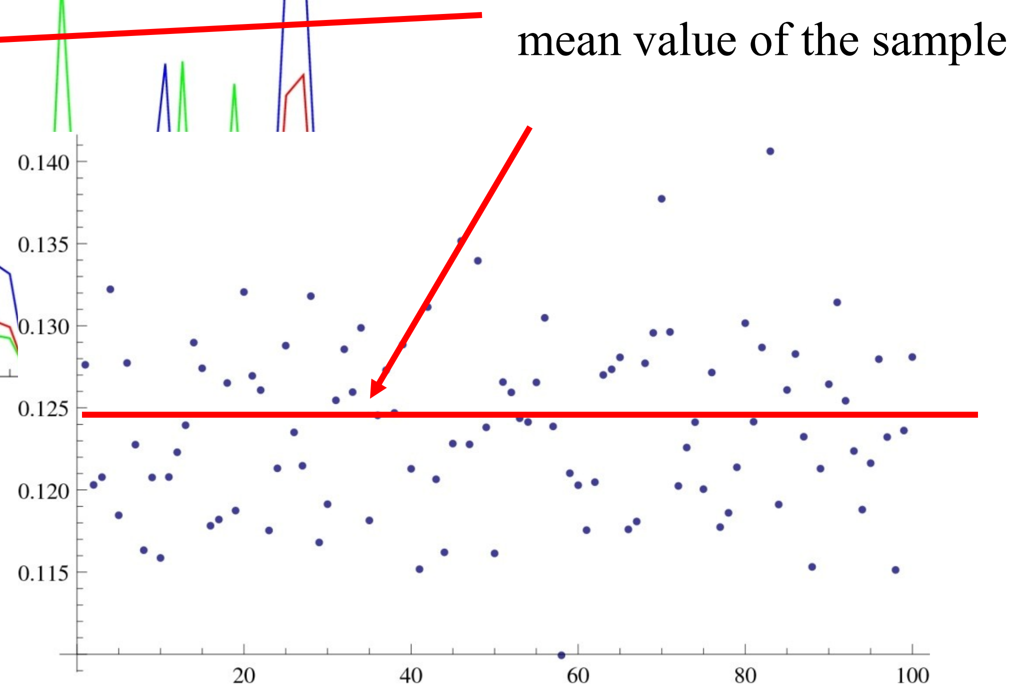


Samples of 100 pseudo-obs. y for two different values of u

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

First results

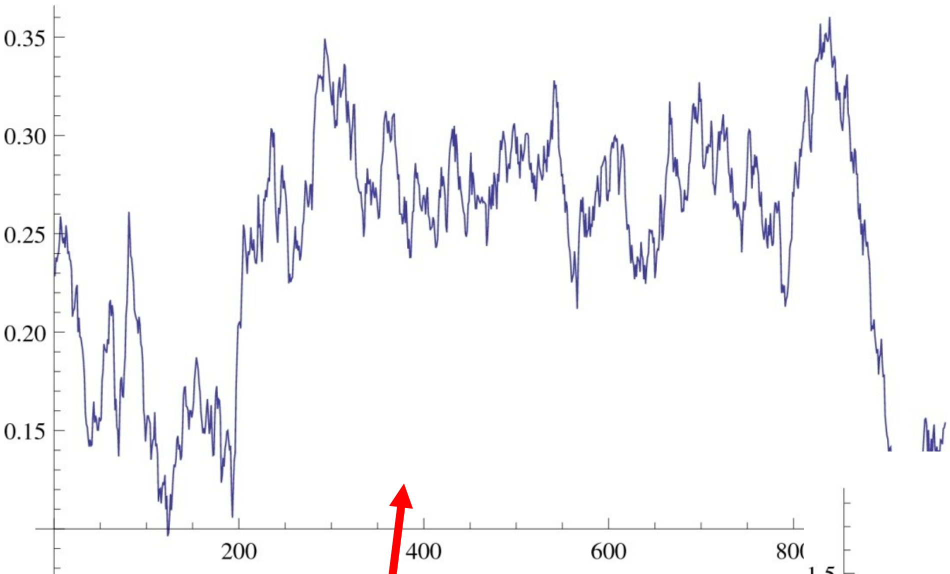




Is it possible to select a "good" value from the sample?

Possible improvement

$$\theta = \frac{1}{1 -}$$



tau is not constant over the range of data

theta is not constant over the range of data