Chapter 4
applied copulas

Seminar *Spatio-temporal dependence*,
09.02.2011

Benedikt Gräler
Institute for Geoinformatics
University of Muenster

Outline

1 Copulas and Indicator Kriging

- The Link
- The procedure
- Practical

ifgi
Institute for Geoinformatics
University of Münster

Outline

1 **Copulas and Indicator Kriging**
- The Link
- The procedure
- Practical

2 **Full copula approach**
- Multivariate Copulas
- The procedure
- The v-transform
- Practical

Outline

1. **Copulas and Indicator Kriging**
   - The Link
   - The procedure
   - Practical
2. **Full copula approach**
   - Multivariate Copulas
   - The procedure
   - The v-transform
   - Practical
3. **Pair-Copula Construction**
   - The Idea
   - Practical

Outline

Outline

1. **Copulas and Indicator Kriging**
   - The Link
   - The procedure
   - Practical
2. **Full copula approach**
   - Multivariate Copulas
   - The procedure
   - The v-transform
   - Practical
3. **Pair-Copula Construction**
   - The Idea
   - Practical
4. **temporal aspects**
   - lag-approach
   - full approach
5. **References & further readings**

Reminder: Indicator Kriging I

Instead of merely grounding the estimate $\hat{z}_1(s_0)$ on the observations of the variable of interest Z_1 we introduce additional variable(s) $Z_2, \ldots, Z_k$ which exhibit some correlation with the primary variable Z_1. The estimator based on a sample $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \ldots, \mathbf{Z}_k)$ is given by:

$$\hat{z}_0(s_0) = \sum_{i=1}^{n} \sum_{j=1}^{k} \lambda_{ij} z_j(s_i)$$

The weights are chosen to minimize the variance of the estimation error and to fulfill

$$\sum_{i=1}^{n} \lambda_{i0} = 1 \text{ and } \sum_{i=1}^{n} \lambda_{ij} = 0,\ 1 \leq j \leq k$$

Reminder: Indicator Kriging II

The weights λ_{ij} can be evaluated by

$$\begin{pmatrix} \lambda_{11} \\ \vdots \\ \lambda_{n1} \\ \lambda_{12} \\ \vdots \\ \lambda_{nk} \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix} = \begin{pmatrix} \mathbf{K}_{11} & \dots & \mathbf{K}_{1k} & \mathbf{1} & 0 & \dots & 0 \\ \mathbf{K}_{21} & \dots & \mathbf{K}_{2k} & 0 & \mathbf{1} & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ \mathbf{K}_{k1} & \dots & \mathbf{K}_{kk} & 0 & 0 & \dots & \mathbf{1} \\ \mathbf{1}^t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathbf{1}^t & 0 & 0 & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{1}^t & 0 & 0 & \dots & 0 \end{pmatrix}^{-1} \begin{pmatrix} K_{11}(s_0, s_1) \\ \vdots \\ K_{11}(s_0, s_n) \\ K_{12}(s_0, s_1) \\ \vdots \\ K_{1k}(s_0, s_n) \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$[k(n+1) \times 1] = [k(n+1) \times k(n+1)] \cdot [k(n+1) \times 1]$$

where $\mathbf{1}^t := (1, \dots, 1) \in \mathbb{R}^n$ and each $\mathbf{K}_{ij}$ is a matrix of dimension $(n \times n)$.

Reminder: Indicator Kriging III

In Indicator Kriging we used to have the matrices $\mathbf{K}_{ij}$ with $1 \leq i, j \leq k$ defined as

$$\mathbf{K}_{ij} := \begin{pmatrix} K_{ij}(s_1, s_1) & \dots & K_{ij}(s_1, s_n) \\ \vdots & \ddots & \vdots \\ K_{ij}(s_n, s_1) & \dots & K_{ij}(s_n, s_n) \end{pmatrix}$$

where $K_{ij}(s_u, s_v)$ denotes the auto-/cross-covariance of the indicators for the i-th quantile at the location s_u and the j-th quantile at the location s_v:

$$K_{ij}(s_u, s_v) := \mathrm{Cov}\big(\mathcal{I}(Z(s_u) \leq q_i), \mathcal{I}(Z(s_v) \leq q_j)\big)$$

for a set of quantiles $\{q_1, \dots, q_k\}$.

Introducing Copulas

The cross-/auto-covariances can be derived from a spatial copula through

$$K_{ij}(s_u, s_v) = C_{uv}\big(\hat{F}_u(q_i), \hat{F}_v(q_j)\big) - \hat{F}_u(q_i)\hat{F}_v(q_j)$$

where $\hat{F}_u$ and $\hat{F}_v$ are estimates of the marginal cdf at s_u and s_v respectively. C_{uv} is the spatial copula describing the dependence "between" location s_u and s_v.
In case of isotropic spatial random fields this identity simplifies to:

$$K_{ij}(h) = C_h\big(\hat{F}(q_i), \hat{F}(q_j)\big) - \hat{F}(q_i)\hat{F}(q_j)$$

(see e.g. [Bárdossy 2006])

Indicator Kriging with Copulas I

Assuming we have a sample $\mathbf{Z} = \{z_1, \ldots, z_n\}$ at n locations $\{s_1, \ldots, s_n\}$ of some staionary spatial random field $\mathcal{Z}$.

1. use $\mathbf{Z}$ to estimate the distribution function $\hat{F}_Z$ of Z
2. transform the sample $\mathbf{Z}$ to a uniform distributed sample (i.e. $\hat{F}_Z(\mathbf{Z})$, or a rank-order transformation)
3. generate a set of bivariate samples by building lag classes
4. estimate a spatial copula over all lag classes

Indicator Kriging with Copulas II

5. decide on a set of quantiles $\{q_1, \ldots, q_k\}$
6. apply an indicator transformation to the sample in order to generate a sample $\mathbf{Z}_k$ of k binary samples:

$$\mathbf{Z}_k := \{\mathcal{I}(\mathbf{Z} \leq q_j) | 1 \leq j \leq k\}$$

7. co-krige the k-variate variable with

$$K_{ij}(h) = C_h\big(\hat{F}_Z(q_i), \hat{F}_Z(q_j)\big) - \hat{F}_Z(q_i)\hat{F}_Z(q_j)$$

8. take care of some *order relation violations*

Benefits of the Copula Approach

Even though the estimation of a copula might be cumbersome this approach has benefits:

- covariances do not have to be calculated for all indicator combinations separately
- the copula is adjusted to fit the complete distribution and not just "slides"

However, the central co-Kriging matrix of dimension $[k(n+1)] \times [k(n+1)]$ consisting of all the $\mathbf{K}_{ij}$ still needs to be inverted which might be computationally expensive (GPUs might provide tools).

TASK

Realize steps 1 - 4 in order to prepare copula-krige for the meuse data set.

In general I

Instead of considering a spatial random filed $\mathcal{Z}$ as a collection of random variables $Z(s)$ over a vast set of locations $s \in S$ we restrict ourselves to a local neighborhood of the η closest locations (usually $4 \leq \eta \leq 24$).

This subset can then be modeled as a $(\eta + 1)$-dimensional distribution $H(z_1, \ldots z_{\eta+1})$.

In general II

This distribution H ideally has parameters which are sensitive to the configuration of the η closest locations as the configuration might alter within the spatial region S and thus the strength of dependence.

Configuration might include distances to the central location, or direction to the central location. An additional parameter like the elevation might be also possible.

The estimation relies on the conditional density.

conditional densities I

ifgi
Institute for Geoinformatics
University of Münster

We consider the 2-dim case first. We assume to know the bivariate distribution for *wind speed* and *temperature* $H(w, t)$ at some location in space.

Knowing the temperature, we would like to infer on the wind speed. The conditional density of the wind speed w given some fixed temperature t_0 is denoted as $h(w|t_0)$ and can be calculated through:

$$h(w|t_0) = \frac{h(w, t_0)}{h(t_0)}$$

The correction factor of the density $h(t)$ can be calculated by $h(t_0) = \int h(w, t_0) dw$

conditional densities II

ifgi
Institute for Geoinformatics
University of Münster

Conditioning the $(\eta + 1)$-dimensional distribution $H(z_1, \dots z_{\eta+1})$ on η-known variables $Z(s_2) = z_2, \dots, Z(s_{\eta+1}) = z_{\eta+1}$ gives a full conditional distribution estimate $\hat{Z}_\eta(s_0) = H(\cdot|z_2, \dots, z_{\eta+1})$ for the unknown location s_0:

$$f_Z(z_0) = \frac{h(z_0, z_2, \dots z_{\eta+1})}{h(z_2, \dots z_{\eta+1})}$$

where h denotes the density of the distribution H.

Why copulas?

Copulas let us drop the marginal distributions and enable us to merely model the dependence structure.

The fitting of the margins is in many cases quite well understood and refers to standard methods.

The copula glues the margins together and we get a full distribution in two *completely distinct* steps.

The problem

ifgi
Institute for Geoinformatics
University of Münster

Obviously, higher dimensional copulas allow for much higher flexibility in the dependence structure:

either the parameter space Θ of the copula grows in order to allow a huge variety of dependence structures

$\rightarrow$ *cumbersome estimation procedures*

or the parameter space Θ of the copula remains low dimensional but does not allow for many different kinds of dependence

$\rightarrow$ *easy to estimate, but inflexible copulas*

Some multivariate copulas I

An easy to estimate copula is the *multivariate normal copula*. It takes the correlation matrix of all pairs of variables as input.

The estimation relies on the estimation of the correlation matrix and can be done by applying the inverse estimators to all combinations of margins.

ifgi
Institute for Geoinformatics
University of Münster

Some multivariate copulas II

ifgi
Institute for Geoinformatics
University of Münster

(Some) *Archimedean Copulas* can easily be extend to higher dimensions. But, they only keep the bivariate parameter and reuse a single generator function φ_θ (that needs to be completely monotonic):

$$C_\theta(u_1, \dots, u_n) = \varphi_\theta^{[-]}(\varphi_\theta(u_1) + \dots + \varphi_\theta(u_n))$$

The `copula` package provides copulas up to a dimension of at least 6.

Some more copulas I

Two other approaches to construct multivariate copulas are

- *fully Nested Archimedean Copulas* with $n-1$ completely monotonic generator functions $\varphi_1, \ldots, \varphi_{n-1}$ with parameters $\theta_1 \geq \ldots \geq \theta_{n-1}$:

$$C_2(u_1, u_2) = \varphi_1^{[-]}\big(\varphi_1(u_1) + \varphi_1(u_2)\big)$$ and recursively defined:

$$C_n(u_1, \ldots, u_n) = \varphi_{n-1}^{[-]}\Big(\varphi_{n-1}\big(C_{n-1}(u_1, \ldots, u_{n-1})\big) + \varphi_{n-1}(u_n)\Big)$$

Some more copulas II

- *Hierarchical Archimedean Copulas* follow a binary tree structure where the child's degrees of dependence $\theta_{i,1}, \theta_{i,2}$ are strictly greater than the parent's degree of dependence $\theta_{i-1,\circ}$:

$$
\begin{aligned}
C_{n,1}(u_1, \ldots, u_n) \\
&= \varphi_{n,1}^{[-]}\Big[\varphi_{n,1}\big(C_{n-1,1}(u_1, \ldots, u_{n/2})\big) \\
&\quad + \varphi_{n,1}\big(C_{n-1,2}(u_{n/2+1}, \ldots, u_n)\big)\Big]
\end{aligned}
$$

see [Savu et al. 2006] for further details

Interpolation with full copulas I

Assuming we have a sample $\mathbf{Z} = \{z_1, \ldots, z_n\}$ at n locations $\{s_1, \ldots, s_n\}$ of some stationary, isotropic spatial random field $\mathcal{Z}$ with reasonable equally distributed samples (a regular sub-grid).

1. use $\mathbf{Z}$ to estimate the distribution function $\hat{F}_Z$ of Z
2. transform the sample $\mathbf{Z}$ to a uniform distributed sample (i.e. $\hat{F}_Z(\mathbf{Z})$, or a rank-order transformation)
3. choose the number of neighbors η to be considered and generate a set of $(\eta + 1)$-variate samples by grouping each location with its η closest neighbors
4. estimate a $(\eta + 1)$-variate copula $C_{\eta+1}(u_0, u_1, \ldots, u_\eta)$ and derive the η-variate copula $C_\eta(u_1, \ldots, u_\eta)$

ifgi
Institute for Geoinformatics
University of Münster

Interpolation with full copulas II

5. the full copula density at each unknown location s_0 is given by

$$c(u_0|u_1, \ldots, u_\eta) = \frac{c_{\eta+1}(u_0, u_1, \ldots, u_\eta)}{c_\eta(u_1, \ldots, u_\eta)}$$

6. an estimate can for instance be obtained by calculating the expected value:

$$\hat{z}_0 = \int_0^1 \hat{F}_Z^{-1}(x) c(x|u_1, \ldots, u_\eta) dx$$

   Or by the median:

$$\hat{z}_0 = \hat{F}_Z^{-1}\left(C^{-1}(0.5|u_1, \ldots, u_\eta)\right)$$

See for instance [Gräler et al. 2011].

the v-transform I

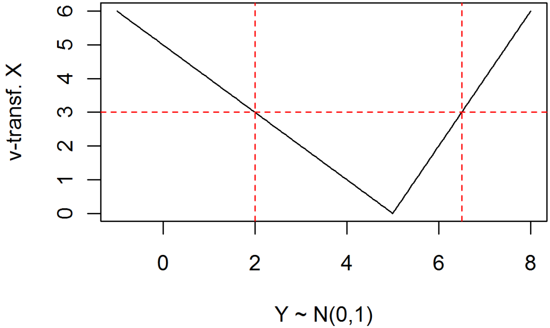
A simple transformation extending the multivariate normal copulas is the *v-transform* introduced in [Bárdossy et al. 2008].

The v-transformed n-variate variable $X = (X_1, \dots, X_n)$ is obtained from a normally distributed n-variate random variable $Y = (Y_1, \dots, Y_n)$ by:

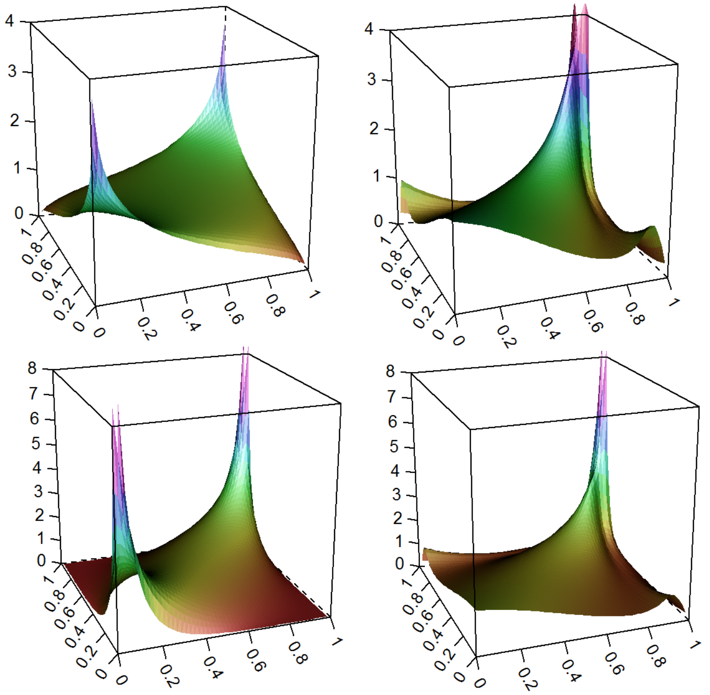
$$X_j := \begin{cases} k(Y_j - m) & \text{if } Y_j \geq m \\ m - Y_j & \text{if } Y_j < m \end{cases}$$

where $k \in (0, \infty)$ and $m \in \mathbb{R}$.

the v-transform II



$$F_X(x_i) = \mathbb{P}(X_i \leq x_i) = \mathbb{P}(m - x_i \leq Y \leq \frac{x_i}{k} - m)$$
$$= \mathbb{P}(Y \leq \frac{x_i}{k} + m) - \mathbb{P}(Y \leq m - x_i)$$
$$= F_Y(\frac{x_i}{k} + m) - F_Y(m - x_i)$$


0 1 2 3 4
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
0 1 2 3 4
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
0 1 2 3 4 5 6 7 8
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1
0 1 2 3 4 5 6 7 8
0 0.2 0.4 0.6 0.8 1
0 0.2 0.4 0.6 0.8 1

Tasks

- sample a v-transformed random variable
- can the v-transform be applied to other margins as well?
- estimate a copula for a local neighborhood of the `muese` data

The idea I

Any multivariate copula can be decomposed into a product of $n(n-1)/2$ bivariate copulas, e.g.:

$$
\begin{aligned}
& c\big(F(x_1),\dots,F(x_4)\big) \\
=& c_{12}\big(F_1(x_1),F_2(x_2)\big)\cdot c_{23}\big(F_2(x_2),F_3(x_3)\big)\cdot c_{34}\big(F_3(x_3),F_4(x_4)\big) \\
& \cdot c_{13|2}\big(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)\big)\cdot c_{24|3}\big(F_{2|3}(x_2|x_3),F_{4|3}(x_4|x_3)\big) \\
& \cdot c_{14|23}\big(F_{1|23}(x_1|x_2,x_3),F_{4|23}(x_4|x_2,x_3)\big)
\end{aligned}
$$

The idea II

ifgi
Institute for Geoinformatics
University of Münster

Where each of the conditional distribution functions $F_{i|j}$ or $F_{i|jk}$ can be expressed by

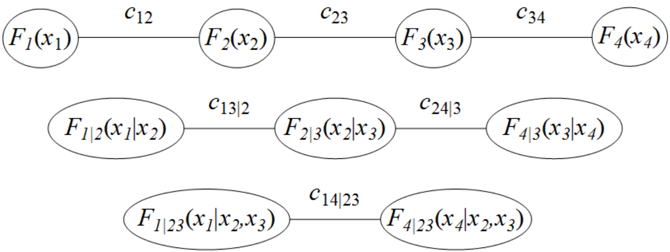
$$F_{i|j}(x_i|x_j) = \frac{\partial C_{ij}\big(F_i(x_i), F_j(x_j)\big)}{\partial F_j(x_j)}$$

$$F_{i|jk}(x_i|x_j, x_k) = \frac{\partial C_{ik|j}\big(F_{i|j}(x_i|x_j), F_{k|j}(x_j|x_k)\big)}{\partial F_{k|j}(x_k|x_j)}$$

The idea III

$$
\begin{aligned}
&c\big(F(x_1),\dots,F(x_4)\big)\\
=&c_{12}\big(F_1(x_1),F_2(x_2)\big)\cdot c_{23}\big(F_2(x_2),F_3(x_3)\big)\cdot c_{34}\big(F_3(x_3),F_4(x_4)\big)\\
&\cdot c_{13|2}\big(F_{1|2}(x_1|x_2),F_{3|2}(x_3|x_2)\big)\cdot c_{24|3}\big(F_{2|3}(x_2|x_3),F_{4|3}(x_4|x_3)\big)\\
&\cdot c_{14|23}\big(F_{1|23}(x_1|x_2,x_3),F_{4|23}(x_4|x_2,x_3)\big)
\end{aligned}
$$

The complex function simplifies if seen as a schematic plot (this structure is called a *d-vine*):



The idea IV

These patterns extend to the n-dimensional case as well.

In general, each copula $c_{ij|k}(F_{i|j}(x_i|x_k), F_{j|k}(x_j|x_k))$ may be dependent on the value of the k-th variable (x_k). That is the copula parameter/family changes for different values of x_k. This full approach would make the model infeasible.

Thus, the immediate influence of the conditional k-th variable on the copula $c_{ij|k}(F_{i|j}(x_i|x_k), F_{j|k}(x_j|x_k))$ is neglected and only implicitly captured through $F_{i|k}$ and $F_{j|k}$[1].

Drawback: Different kinds of decompositions may give different results!

[1] This method was introduced in [Aas et al. 2009] and the effect of the simplification is discussed in [Hobæk Haff et al. 2010].

estimate a pair-copula

- How does the conditional density of the d-vine look like?
- How can the conditional distribution functions be calculated?

in general

So far, we considered time only as a repetition of our random process.

Looking at the temporal development can with copulas be treated as an increase in dimension.

In the *lag approach* we group out data not only into spatial lags, but as well into temporal lags.

This results in a grid of lag classes with one spatial and one temporal axis (be aware of autocorrelation!)

Temporal dependence often appears to be asymmetric!

ifgi
Institute for Geoinformatics
University of Münster

two-dimensional convex combination

For each *knot* in the lag-grid we estimate a single copula.The spatio-temporal copula is then designed as a convex combination of four copulas, or by some interpolated parameter function [Gräler 2009].

The following tables are taken from [Gräler 2009] as well.

spatio-temporal copula for particular matter measurements in NRW I

observations per lag

	Temporal				
Sp.	0	1	2	3	4
0	743	372	361	361	361
1	1137	571	554	556	549
2	524	262	256	254	254
3	897	449	431	434	436
4	773	388	374	374	375
5	883	443	433	429	426
6	1006	507	488	489	486
7	767	381	371	374	375
8	748	375	361	365	363
9	881	441	432	431	426
10	885	440	427	431	433
11	889	442	431	435	433
12	894	445	434	435	435
13	894	445	433	434	435
14	636	316	310	310	310

spatio-temporal copula for particular matter measurements in NRW II

the best fitting families

	Temporal				
Sp.	0	1	2	3	4
0	Gaussian	Clayton	ASC	Clayton	Gaussian
1	Frank	Frank	Frank	Gaussian	Gaussian
2	Gaussian	Frank	ASC	Gaussian	Gaussian
3	Gaussian	ASC	Gaussian	ASC	Gaussian
4	Gaussian	Frank	ASC	Gaussian	Gaussian
5	Gaussian	Frank	ASC	ASC	Gaussian
6	Frank	Clayton	Clayton	Gaussian	Gaussian
7	Frank	ASC	Clayton	Clayton	Gaussian
8	Gaussian	Clayton	ASC	ASC	Gaussian
9	Frank	Frank	ASC	ASC	Gaussian
10	Frank	ASC	Gaussian	Gaussian	Gaussian
11	Frank	ASC	Clayton	Gaussian	Gaussian
12	Frank	ASC	Gaussian	Gaussian	Gaussian
13	Frank	ASC	Gaussian	Gaussian	Gaussian
14	Frank	ASC	Clayton	Gaussian	Gaussian

full spatio-temporal copula

We restructure the design of our data for the full copula.

Instead of collecting only the four nearest neighbors in space we add a column containing the preceding value(s).

It might be useful to use as well the preceding neighbors (forecasting).

ifgi
Institute for Geoinformatics
University of Münster

References & further readings I

- Aas, Kjersti, Claudia Czado, Arnoldo Frigessi & Henrik Bakken (2009), 'Pair-copula constructions of multiple dependence', Insurance: Mathematics and Economics 44, 182 - 198.
- Bárdossy, A. & Jing Li (2008), 'Geostatistical interpolation using copulas', Water Resources Research 44.
- Bárodssy, A. (2006), 'Copula-based geostatistical models for groundwater quality parameters'. Water Resources Research 42.
- Gräler, Benedikt (2009), 'Copulas for Spatio-Temporal Random Fields', Diploma thesis at the Institute of Mathematical Statistics and Institue for Geoinformatics, University of Muenster.

References & further readings II

- Gräler, Benedikt & Edzer Pebesma (2011): The pair-copula construction for spatial data: a new approach to model spatial dependency. Poster at: Spatial Statistics 2011 - Mapping global change. Enschede, The Netherlands, 23-25 March 2011 (accepted for prersentation and publications in *Procedia Environmental Sciences* by Elsevier).
- Hobæk Haff, Ingrid, Kjersti Aas & Arnoldo Frigessi (2010), 'On the simplified pair-copula construction – Simply useful or too simplistic?', Journal of Multivariate Analysis, 101(5), 1296 - 1310.
- Nelsen R. B. (2006), An Introduction to Copulas, 2nd Edition, Springer Science+Buisness, New York.

References & further readings III

Savu, C. & M. Trede (2006). 'Hierarchical Archimedean copulas', In: International Conference on High Frequency Finance. May 2006, Konstanz, Germany.