Model selection of vine copulas with applications

Claudia Czado
Technische Universität München and University of Plymouth
cczado@ma.tum.de

Lehrstuhl für Mathematische Statistik

International Workshop on High-Dimensional Dependence and Copulas,
Jan 3-5, 2014, Beijing, China

Why are vine copulas useful?

- Multivariate data has often complex dependency patterns, such as asymmetry and dependence in the extremes
- Cannot be captured by the multivariate normal distribution.
- The copula approach allows for these dependency patterns
- Current classes of multivariate copulas such as Gaussian, Student t and Archimedean copulas are too restrictive
- They require often exchangeability and that the distribution of pairs are of same kind
- Vine copulas allow for flexible modeling of (conditional) pairs

Overview

What are copulas and how it all started ...

Consider d random variables $\mathbf{X} = (X_1, ..., X_d)$ with

	density function	distribution function
marginal	$f_i(x_i),\ i = 1, ..., d$	$F_i(x_i),\ i = 1, ..., d$
joint	$f(x_1, ..., x_d)$	$F(x_1, ..., x_d)$
conditional	$f_{i\|j}(x_i\|x_j),\ i \neq j$	$F_{i\|j}(x_i\|x_j),\ i \neq j$

Copula

A d-dimensional copula C is a multivariate distribution on $[0,1]^d$ with uniformly distributed marginals.

Copula density function: $c(u_1, ..., u_d) := \frac{\partial^d}{\partial u_1 ... \partial u_d} C(u_1, ..., u_d)$

Sklar's Theorem

Theorem (Sklar 1959)

$$F(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$$
$$f(x_1, ..., x_d) = c(F_1(x_1), ..., F_d(x_d))f_1(x_1)...f_d(x_d)$$

for some d-dimensional copula C.

$d = 2:$

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2)$$
$$f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2))f_2(x_2)$$

What are these vine copulas?

- Multivariate vine copulas are copulas built out of bivariate copulas.
- A pair copula construction (PCC) is possible through conditioning. Joe (1996) gave a first example.
- Many PCC's are feasible. Bedford and Cooke (2002) introduced a graphical structure to organize them.
- Gaussian vines were analyzed in Kurowicka and Cooke (2006) while ML estimation for Non Gaussian ones started with Aas et al. (2009).
- See also `vine-copula.org`

How does this work in 3 dimensions?

$$f(x_1, x_2, x_3) = f_{3|12}(x_3|x_1, x_2) f_{2|1}(x_2|x_1) f_1(x_1)$$

Using Sklar for $f(x_1, x_2), f(x_2, x_3)$ and $f_{13|2}(x_1, x_3|x_2)$ implies

$$\begin{aligned} f_{2|1}(x_2|x_1) &= c_{12}(F_1(x_1), F_2(x_2)) f_2(x_2) \\ f_{3|12}(x_3|x_1, x_2) &= c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) f_{3|2}(x_3|x_2) \\ &= c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) c_{23}(F_2(x_2), F_3(x_3)) f_3(x_3) \end{aligned}$$

$$\begin{aligned} f(x_1, x_2, x_3) &= c_{13;2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) c_{23}(F_2(x_2), F_3(x_3)) \\ &\times c_{12}(F_1(x_1), F_2(x_2)) \\ &\times f_3(x_3) f_2(x_2) f_1(x_1) \end{aligned}$$

The copula corresponding to the distribution of (X_1, X_3) given $X_2 = x_2$ is denoted by $c_{13;2}$. Only bivariate copulas and univariate conditional cdf's are used. This can be easily generalized to d dimensions.

What bivariate copula families are available?

Elliptical: Construction by inversion

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)), \quad u_1, u_2 \in (0, 1),$$

where F is elliptical.

Examples: Gaussian, Student's t

Archemedean: Construction through generator φ

$$C(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)), \quad u_1, u_2 \in (0, 1).$$

(McNeil and Nešlehová 2009)

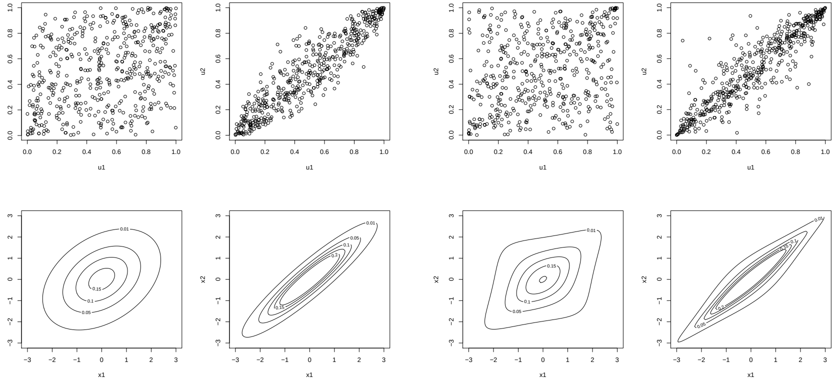
Examples: Clayton, Gumbel, Frank, Joe ...

Books: Joe (1997) and Nelsen (2006)
Extensions: Rotations by 90, 180 (survival) and 270 degree

Bivariate elliptical copula families

Gaussian copula
(left $\tau = .25$, right: $\tau = .75$)
symmetric dependence

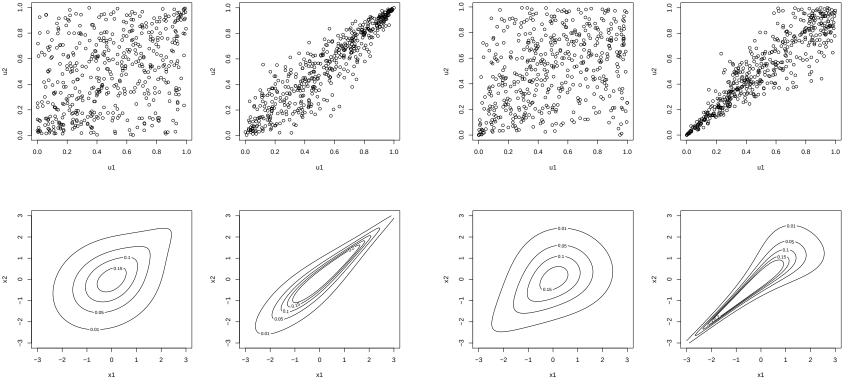
t-copula with $df = 3$
(left $\tau = .25$, right: $\tau = .75$)
symmetric dependence



Bivariate Archimedean copula families

Gumbel copula
(left $\tau = .25$, right: $\tau = .75$)
upper tail dependent

Clayton copula
(left $\tau = .25$, right: $\tau = .75$)
lower tail dependent



How do vines work in higher dimensions?

- **Which pairs of variables are needed?**
- **What are the conditioning variables?**

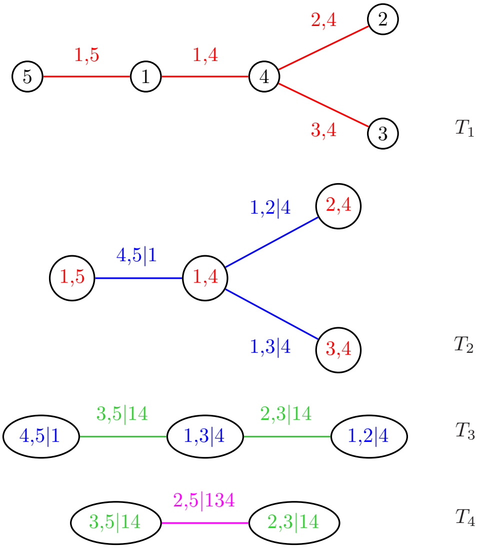
Components of a regular vine $R(\mathcal{V}, \mathcal{C}, \boldsymbol{\theta})$ distribution

1. Tree structure $\mathcal{V}$ of linked trees identifies the pairs of variables and conditioning variables.
2. Parametric bivariate copulas $\mathcal{C} = \mathcal{C}(\mathcal{V})$ for each edge in the tree structure
3. Corresponding parameter value $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{C}(\mathcal{V}))$

- Joe (1996) showed that conditional distribution functions can be computed recursively. For $\mathbf{v} = (v_j, \mathbf{v}_{-j})$ we have

$$F(x|\mathbf{v}) = \frac{\partial C_{xv_j;\mathbf{v}_{-j}}(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j}))}{\partial F(v_j|\mathbf{v}_{-j})}.$$

Can we see an example of a tree structure?



Density

$$f = f_1 \cdot f_2 \cdot f_3 \cdot f_4 \cdot f_5$$

$\cdot\ c_{14} \cdot c_{15} \cdot c_{24} \cdot c_{34}$

$\cdot\ c_{12;4} \cdot c_{13;4} \cdot c_{45;1}$

$\cdot\ c_{23;14} \cdot c_{35;14}$

$\cdot\ c_{25;134}$

How is a regular vine tree structure defined?

An d-dimensional vine tree structure $\mathcal{V} = \{T_1, \ldots, T_{d-1}\}$ is a sequence of $d-1$ linked trees with

Vine tree structure (Bedford and Cooke (2002))

- Tree T_1 is a tree on nodes 1 to d.
- Tree T_j has $d+1-j$ nodes and $d-j$ edges.
- Edges in tree T_j become nodes in tree T_{j+1}.
- **Proximity condition:** Two nodes in tree T_{j+1} can be joined by an edge only if the corresponding edges in tree T_j share a node.

Are there special cases?

- D-vines use only path like trees
- canonical (C)-vines use only star like tree

How do these C and D-vines look like?

C-vine: each tree has a unique node connected to $d - j$ edges

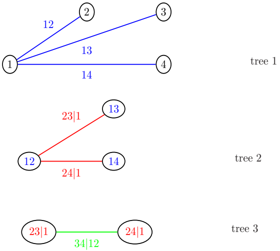
$$f_{1234} = \left[\prod_{i=1}^{4} f_i\right] \cdot c_{12} \cdot c_{13} \cdot c_{14} \cdot c_{23;1} \cdot c_{24;1} \cdot c_{34;12}$$

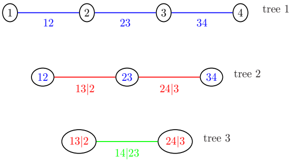
useful for ordering by importance

D-vine: no node is connected to more than 2 edges

$$f_{1234} = \left[\prod_{i=1}^{4} f_i\right] \cdot c_{12} \cdot c_{23} \cdot c_{34} \cdot c_{13;2} \cdot c_{24;3} \cdot c_{14;23}$$

useful for temporal ordering of variables





General density expressions

C-vine (Aas et al. 2009)

$$f(x_1, \ldots x_d) = \left[\prod_{k=1}^{d} f(x_k)\right] \times \left[\prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+i;1,\ldots,j-1}\right]$$

D-vine (Aas et al. 2009)

$$f(x_1, \ldots x_d) = \left[\prod_{k=1}^{d} f(x_k)\right] \times \left[\prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{i,i+j;i+1,\ldots,i+j-1}\right]$$

Regular vine (Dißmann et al. 2013)

$$f(x_1, \ldots, x_d) = \left[\prod_{k=1}^{d} f_k(x_k)\right] \times \left[\prod_{j=d-1}^{1} \prod_{i=d}^{j+1} c_{m_{j,j}, m_{i,j}; m_{i+1,j}, \ldots, m_{n,j}}\right]$$

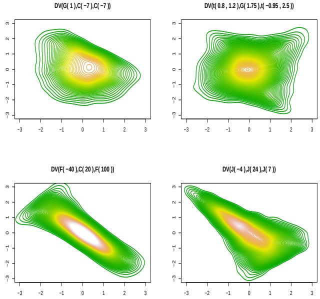
Here, $m_{i,j}$ refers to element (i, j) in the matrix representation of the R-vine.

What is the scope of the vine models?

Vine copula classes
(Stöber et al. (2013))

- multivariate Gaussian copula (pair copulas are Gauss and parameters are partial correlations)
- multivariate t copula (pair copulas are t and df increases by ℓ for trees $\ell \geq 2$)
- multivariate Clayton copula (Takahasi (1965))

Contours of bivariate (1,3) margins with standard normal margins



(C=Clayton, G=Gumbel, t=Student, F=Frank, J=Joe)

Number of R-vine tree structures and copulas

Dimension n	#R-vine tree structure [1]	#R-vine copulas [2]
2	1	7
3	3	1,029
4	24	2,823,576
5	480	1.3559e+11
6	23,040	1.0938e+17
7	2,580,480	1.4413e+24
8	660,602,880	3.0387e+32
9	3.8051e+11	1.0090e+42
10	4.8705e+14	5.2118e+52

Efficient estimation and model selection are crucial

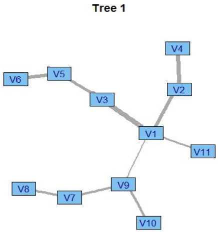
[1] see Morales-Nápoles et al. (2010) for details.

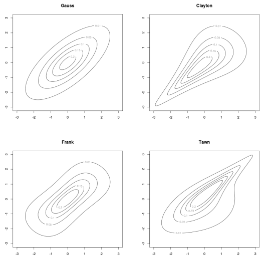
[2] This assumes 7 candidate pair copula families.

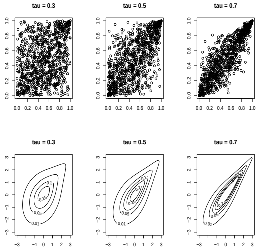
How can we estimate and select PCCs?

Three problems: (Czado et al. (2013))

1. How to estimate the pair copula parameters for a given vine tree structure and pair copula families for each edge?
2. How to choose the pair copula families and estimate the corresponding parameters for a given vine tree structure?
3. How to select and estimate all components of a regular vine?







Problem 1: Parameter estimation for given tree structure and copula families

- **Sequential estimation:**
  - Parameters are sequentially estimated starting from top tree until last (Aas et al. (2009), Czado et al. (2012)).
  - Asymptotic theory developed by Hobæk Haff (2013), Hobæk Haff (2012), however standard error estimates can only be bootstrapped.
  - starting values for maximum likelihood.
- **Maximum likelihood estimation:**
  - Asymptotically efficient and standard errors have been directly estimated in Stoeber and Schepsmeier (2013)
  - Uncertainty in value-at-risk (high quantiles) is difficult to assess.
- **Bayesian estimation:**
  - Posterior is tractable using Markov Chain Monte Carlo (Min and Czado (2011) for D-vines and Gruber (2011) for R-vines)
  - Prior beliefs can be incorporated and credible intervals allow to assess uncertainty.

How does sequential and ML estimation work?

Parameters: $\Theta = (\theta_{12}, \theta_{23}, \theta_{13;2})$

Observations: $\{(x_{1t}, x_{2t}, x_{3t}), t = 1, \cdots, T\}$

Sequential estimates:

- Estimate θ_{12} from $\{(x_{1,t}, x_{2,t}), t = 1, \cdots, T\}$
- Estimate θ_{23} from $\{(x_{2,t}, x_{3,t}), t = 1, \cdots, T\}$.
- Define pseudo observations

$$\hat{v}_{1|2t} := F(x_{1t}|x_{2t}, \hat{\theta}_{12}) \text{ and } \hat{v}_{3|2t} := F(x_{3t}|x_{2t}, \hat{\theta}_{23})$$

Finally estimate $\theta_{13;2}$ from $\{(\hat{v}_{1|2t}, \hat{v}_{3|2t}), t = 1, \cdots, T\}$.

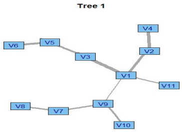
Maximum likelihood

$$\begin{aligned} L(\Theta|x) &= \sum_{t=1}^{T} [\log c_{12}(x_{1t}, x_{2t}|\theta_{12}) + \log c_{23}(x_{2t}, x_{3t}|\theta_{23}) \\ &+ \log c_{13;2}(F(x_{1t}|x_{2t}, \theta_{12}), F(x_{3t}|x_{2t}, \theta_{23})|\theta_{13;2})] \end{aligned}$$

Problem 2: Joint estimation of pair copula families and parameters

- **Classical approach:**
  - Restrict to a set of bivariate pair copula families and use AIC or Vuong test to select family
  - Check for truncation possibilities (Brechmann et al. (2012)) by using independence copulas in higher trees
- **Bayesian approach:**
  - Reversible jump (RJ) MCMC (Min and Czado (2011))
  - MCMC with model indicators (Smith et al. (2010)) choosing between an independence copula and a fixed copula family.

Only one more problem to go ...



- **sequential treewise approach** (see Dißmann et al. (2013))
- **Bayesian sequential and joint approaches** (see Gruber and Czado (2012), Gruber and Czado (2013))

Problem 3: How does this treewise selection work?

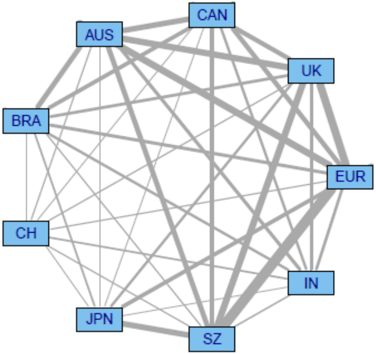
Idea: Capture strong pairwise dependencies first

For Tree $\ell = 1, \ldots, d-1$

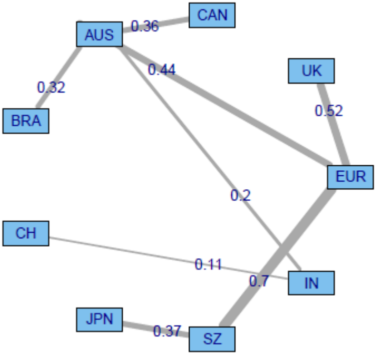
1. Calculate an empirical dependence measure $\hat{\delta}_{jk|D}$ for all variable pairs $\{jk|D\}$ ($\rightarrow$ **edge weights**: Kendall's τ, tail dependence coefficients) allowed by proximity (D is empty for Tree 1).
2. Select the tree on all nodes that maximizes the sum of absolute empirical dependencies ($\rightarrow$ **maximum spanning tree**). Choose independence copula if possible.
3. For each selected edge $\{j,k\}$ ($\{j,k\}|D$) in Tree 1 (in Tree $\ell > 1$), select copula family and estimate the corresponding parameter(s).
4. Transform to pseudo observations: $F_{j|k \cup D}(u_{ij}|\mathbf{u}_{i,k\cup D}, \hat{\theta}_{j,k;D})$ and $F_{k|j\cup D}(u_{ik}|\mathbf{u}_{i,j\cup D}, \hat{\theta}_{j,k;D})$, $i = 1, \ldots, n$.

What does this look like for Tree 1?

(1) Pairwise dependencies.



(2) Maximum dependence tree.



Czado, Jeske, and Hofmann (2013) compare sequential selection strategies

Problem 3: Three approaches to full R-vine selection

- $d_{\mathcal{V}}$ number of parameters in R-vine $(\mathcal{V}, \mathcal{B}(\mathcal{V}), \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})))$
- T_ℓ tree ℓ of vine tree structure $\mathcal{V}$
- $\mathcal{B}_\ell$ set of pair copula families for all edges in T_ℓ
- d_ℓ number of parameters in Tree T_ℓ

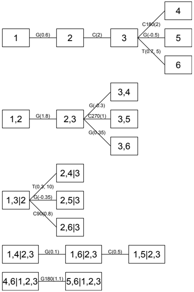
	Dissmann et. al 2013	Gruber/Czado 2012	Gruber/Czado 2013
Appr.	Frequentist level-by-level	Bayesian level-by-level	Bayesian all levels jointly
Priors		$T_\ell \sim Unif(\cdot)$ $\mathcal{B}_\ell \mid T_\ell \sim \exp(-\lambda d_\ell)$ $\boldsymbol{\theta}_\ell \mid T_\ell, \mathcal{B}_\ell \sim Unif(\cdot)$	$\mathcal{V} \sim Unif(\cdot)$ $\mathcal{B}_{\mathcal{V}} \mid \mathcal{V} \sim \exp(-\lambda d_{\mathcal{V}})$ $\boldsymbol{\theta}_{\mathcal{V}} \mid \mathcal{V}, \mathcal{B}_{\mathcal{V}} \sim Unif(\cdot)$
Method	Select MST with weights	RJ MCMC	RJ MCMC

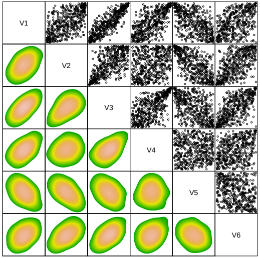
Proposal strategies (Gruber and Czado 2012; Gruber and Czado 2013)

- Use a mixture of two mutually exclusive, collectively exhaustive algorithms for the between models move:
  - FAM only updates the pair copula families;
  - TREE updates the tree structure and the pair copula families and guarantees that the current tree is not proposed.
- Draw proposal trees from a uniform distribution over all trees allowed by the proximity condition (only TREE).
- Compute the MLEs of the parameters for all candidate pair copula families.
- Draw the proposal pair copulas from a discrete distribution with weights proportional to the copulas' maximum likelihoods.
- Draw the proposal parameters from a mixture of truncated normal distributions with varying variances, centered at the MLE.

Simulation study for R-vine copula selection

- 4 R-vine copula models in 6 dimensions were chosen.
- Model 1 and 2 have the same R-vine tree structure and pair copula families but different parameter values, stronger dependencies in Model 2
- Model 3 is a C-vine and Model 4 is Gauss copula
- **Model 1:**





Average percentage of true likelihood recovered

Procedure	Model 1	Model 2	Model 3	Model 4	Mean
Gruber/Czado 2013	**98.4**	**98.6**	**96.5**	**99.2**	**98.2**
Gruber/Czado 2012	88.6	80.9	84.9	99.9	88.6
Dißmann et.al. 2013	86.6	75.1	78.9	99.7	85.1

- Results are from Gruber and Czado (2013) based on 10 data sets of size 500
- The joint Bayesian procedure performs best
- Model 4 is the multivariate Gaussian copula, which can be expressed as any Gaussian vine
- Because the selection of the regular vine tree structure does not matter for Model 4, all model selection procedures perform uniformly well.

How does copula based estimation work?

- **Original scale**: $\mathbf{x}_i = (x_{i1}, \ldots, x_{id}) \in \mathbb{R}^d$ i.i.d. sample
- **Copula scale**:
  - Known margins: $\mathbf{u}_i := (F_1(x_{i1}), \ldots, F_d(x_{id})) \in [0,1]^d$ (Probability integral transform)
  - Unknown margins: Estimate margins F_j either parametrically or non-parametrically and then transform (two step procedure)
- **Marginal structures**: If each margin has time series or regression structure, then a copula model will be applied to the fitted standardized residuals.

Application: Euro Stoxx 50:

- 50 large Eurozone companies.
- Major market indicator for the Eurozone.
- Brechmann and Czado (2013) consider 46 members from 5 countries (Germany, France, Italy, Spain and the Netherlands) together with their national indices.
- Daily log returns: May 2006 to April 2010 (985 obs.)

Questions

- **How do stock returns depend on the European and the national indices? Is dependence on the national index dominant?**
- **Which dependencies are most important? Are they asymmetric and/or heavy-tailed?**

Copula based models for Euro Stoxx 50 returns:

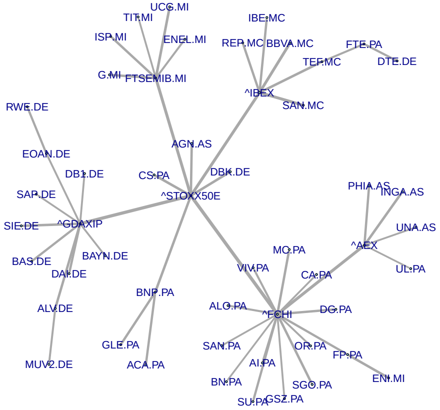
- Fit appropriate (ARMA-)GARCH models for each return time series.
- Fit copula model such as R- and C-vine copulas as well as multivariate Student-t copula for comparison to copula data based on standardized residuals

Results

Copula	Log likelihood	No. of param.	BIC
R-vine	30879.60	596	-57651.19
C-vine	30839.68	685	-56957.90
Student-t	30691.36	1327	-52236.18

R-vine $>$ C-vine $>$ Student-t

First tree of R-vine and C-vine order



Order	Root nodes
1^{st}	^STOXX50E
2^{nd}	GLE.PA
3^{rd}	^FCHI
4^{th}	^GDAXIP
5^{th}	^IBEX
6^{th}	INGA.AS
7^{th}	FTSEMIB.MI
⋮	⋮
13^{th}	^AEX
⋮	⋮

Application: Stock and volatility indices (Beil 2013)

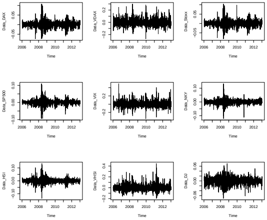
- DAX: German stock index
- VDAX: volatility index to the DAX
- STOXX: Dow Jones Euro Stoxx 50
- VSTOXX: volatility index to the STOXX
- SP500: Standard and Poor's 500 index
- VIX: volatility index to SP500
- DJ: Dow Jones UBS commodity index ex-agriculture and live stock
- NKY: Nikkei -225 stock average
- VNKY: volatility index to NKY
- HSI: Hong Kong Hang Seng index
- VHSI: volatility index to HSI

Daily values from June 2006 until June 2013 considered (1786 obs)

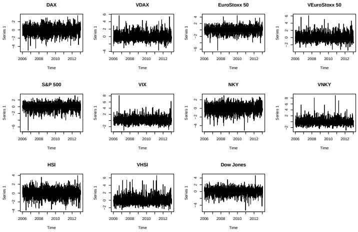
Marginal models

- Value of volatility index is the implied volatility of a 30 day option on its underlying asset by the Black Scholes model
- ARMA-(e)GARCH models with generalized hyperbolic innovation distribution are to each time series

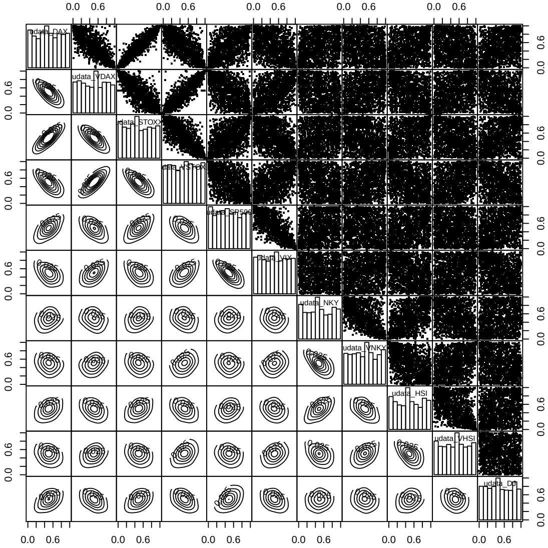
original time series



residual time series



Copula data and normalized contour plots



Using AIC/BIC to compare models not reduced by independence tests

	loglik	par	AIC	BIC
R-vine	8663.53	71	-17185.05	-16795.42
D-vine	8652.68	70	-17165.35	-16781.21
C-vine	8541.96	73	-16937.91	-16537.31
Gauss-copula	8033.01	55	-15956.03	-15654.20
T-copula	8308.15	56	-15755.92	-15448.60
T-vine	8525.14	80	-16890.28	-16451.26

If $df > 30$ for pair in T-vine then Gauss copula is used

- R-vines are selected using the Dissmann algorithm with Gauss, T, Gumbel, Clayton, Frank, Joe pair copulas and rotations are allowed
- For T-vine only pair t-copulas are allowed
- R-vine performs best compared to Gauss and T-copula

Using AIC/BIC to compare models reduced by independence tests comparison

More parsimonious models can be achieved by independence tests using asymptotic theory of $\hat{\tau}$.

	loglik	par	AIC	BIC
R-vine-ind	8606.27	51	-17110.54	-16830.67
D-vine-ind	8692.77	71	-17235.54	-16823.96
C-vine-ind	8524.54	58	-16933.07	-16614.78
Gauss-vine-ind	8013.50	41	-15945.00	-15720.01
T-vine-ind	8454.55	68	-16773.10	-16399.93
T-vine-fixed-ind	8470.37	59	-16822.74	-16498.96

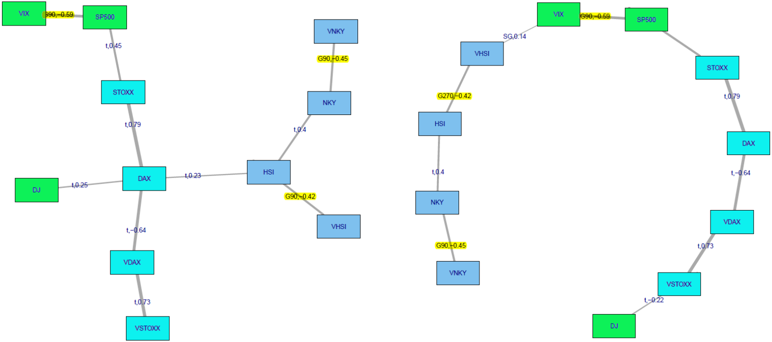
- T-vine-fixed-ind has the same tree structure as R-vine-ind but only T-pair copulas are allowed
- T-vine-ind might not have the same tree structure as R-vine-ind and only T- pair copulas are allowed

Model comparison using the Vuong test

statistic	p.value	decision	alpha
0.16	0.87	R-vine equivalent D-vine	0.10
4.54	< 0.01	R-vine better C-vine	0.10
4.05	< 0.01	D-vine better C-vine	0.10
7.79	< 0.01	R-vine better T-vine	0.10
7.50	< 0.01	R-vine better T-vine-fixed	0.10

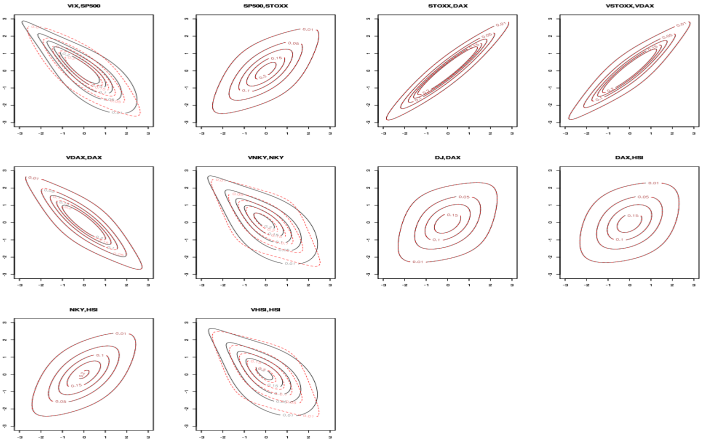
- Test of Vuong (1989) is suited for comparing non nested models
- BIC correction is used to obtain parsimony
- R-vine and T-vine-fixed have same tree structure, while this is not true for R-vine and T-vine
- Everywhere models are reduced by independence tests

First tree of R-vine and D-vine model



- tree structures are quite different
- asymmetry between volatility and asset indices
- geographic clusters

Unconditional fitted contours of R-vine and T-vine with same R-vine structure



T-vine versus R-vine: Different fit especially for dependencies between VIX-SP500, VNKY-NKY and VHSI-HSI

Recent advances for vines

- Simplified and non simplified vines: Acar et al. (2012)
- Time varying/regime switching regular vines: Almeida et al. (2012), Stöber and Czado (2013)
- Discrete and discrete/continuous vines: Panagiotelis et al. (2012), Stöber (2013)
- Non Gaussian DAG's using pair copula constructions: Hanea et al. (2006), Bauer and Czado (2012)
- Vines with non parametric pair copulas: Kauermann and Schellhase (2013), Lopez-Paz et al. (2013), Hobaek Haff and Segers (2012)
- Acceleration of MCMC algorithms: Schmidl et al. (2013)

Selected Applications

- **Financial risk management:**
  - Systemic risk simulation: Brechmann, Hendrich, and Czado (2013)
  - Operational risk: Brechmann, Czado, and Paterlini (2013)
  - Multivariate options: Bernard and Czado (2013)
  - Realized volatility: Vaz de Melo Mendes and Accioly (2013)
  - Insurance: Erhardt and Czado (2012)
  - Portfolio management: Low, Alcock, Faff, and Brailsford (2013)
- **Hydrology:** Gräler, van den Berg, Vandenberghe, Petroselli, Grimaldi, De Baets, and Verhoest (2013)
- **Machine learning:** Lopez-Paz, Hernández-Lobato, and Ghahramani (2013),
- **Health:** comorbidity Stöber, Czado, Hong, and Ghosh (2012)
- **Environmental Science**: Gräler and Pebesma (2011), Pachali (2011), Erhardt (2013)

What have we learned?

- Standard multivariate copulas are less flexible, while PCCs such as C-, D- and R-vines are much more flexible.
- Sequential and MLE parameter estimation of C-, D- and R-vines are available in **R packages** `CDVine` and `VineCopula`.
- The vine tree and the pair copula familes matter in the selection of good fitting PCCs.
- The catalog of possible pair copula families should also include nonsymmetric pair copulas such as the Tawn copula (Eschenburg (2013))
- Sequential and full Bayesian estimation and Bayesian model selection of vine trees and copula families for regular vines available, but need further testing and development

What needs to be done?

- non parametric pair copulas, spatial vines, large non Gaussian Bayesian belief networks, vines for data mining, high dimensional GoF for vines (Schepsmeier 2013)
- more applications in finance, insurance, health, genetics ...

Vine resource page: `vine-copula.org`
Vine workshop book: Kurowicka and Joe (2011)

Spatial Copula Workshop

Host: Institute for Geoinformatics, Münster, Germany
Dates: 22nd and 23rd September 2014
Organizers: C. Czado and B. Gräler

Thanks to my collaborators





Aas, K., C. Czado, A. Frigessi, and H. Bakken (2009).
Pair-copula constructions of multiple dependence.
Insurance, Mathematics and Economics 44, 182–198.

Acar, E. F., C. Genest, and J. Nešlehová (2012).
Beyond simplified pair-copula constructions.
Journal of Multivariate Analysis 110, 74–90.

Almeida, C., C. Czado, and H. Manner (2012).
Modeling high dimensional time-varying dependence using D-vine scar models.
preprint.

Bauer, A. and C. Czado (2012).
Pair-copula Bayesian networks.
preprint.

Bedford, T. and R. M. Cooke (2002).
Vines - a new graphical model for dependent random variables.
Annals of Statistics 30(4), 1031–1068.

Beil, M. (2013).
Modeling dependencies among financial asset returns using copulas.
Master's thesis, Technische Universität München.

Bernard, C. and C. Czado (2013).
Multivariate option pricing using copulae.
Appl. Stochastic Models Bus. Ind. 29, 509–526.

Brechmann, E. and C. Czado (2013).
Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50.
to appear in Statistics & Risk Modeling.

Brechmann, E., C. Czado, and K. Aas (2012).
Truncated regular vines in high dimensions with application to financial data.
Canadian Journal of Statistics 40, 68–85.

Brechmann, E., C. Czado, and S. Paterlini (2013).
Flexible dependence modeling of operational risk losses and its impact on total capital requirements.
to appear in Journal of Banking and Finance.

Brechmann, E., K. Hendrich, and C. Czado (2013).
Conditional copula simulation for systemic risk stress testing.
Insurance: Mathematics and Economics 53(3), 829–839.

Czado, C., S. Jeske, and M. Hofmann (2013).
Selection strategies for regular vine copulae.
Journal de la Société Francaise de Statistique 154, 174–191.

Czado, C., U. Schepsmeier, and A. Min (2012).
Maximum likelihood estimation of mixed C-vine pair copula with application to exchange rates.
Statistical Modeling 12, 229–255.

Dißmann, J., E. Brechmann, C. Czado, and D. Kurowicka (2013).
Selecting and estimating regular vine copulae and application to financial returns.
Computational Statistics and Data Analysis 52(1), 52–59.

Erhardt, T. (2013).
Predicting temperature time series using spatial vine copulae.
Master's thesis, Technische Universität München.

Erhardt, V. and C. Czado (2012).
Modeling dependent yearly claim totals including zero claims in private health insurance.
Scandinavian Actuarial Journal 2012(2), 106–129.

Eschenburg, P. (2013).
Properties of extreme-value copulas.
Master's thesis, Technische Universität München.

Gräler, B. and E. Pebesma (2011).
The pair-copula construction for spatial data: a new approach to model spatial dependency.
Procedia Environmental Sciences 7(0), 206 – 211.

Gräler, B., M. J. van den Berg, S. Vandenberghe, A. Petroselli, S. Grimaldi, B. De Baets, and N. E. C. Verhoest (2013).
Multivariate return periods in hydrology: a critical and practical review focusing on synthetic design hydrograph estimation.
Hydrol. Earth Syst. Sci. 17, 1281–1296.

Gruber, L. F. (2011).
Bayesian analysis of R-vine copulas.
Master's thesis, Technische Universität München.

Gruber, L. F. and C. Czado (2012).
Sequential Bayesian Model Selection of Regular Vine Copulas.
Preprint.

Gruber, L. F. and C. Czado (2013).
Bayesian Model Selection of Regular Vine Copulas.
Preprint.

Hanea, A., D. Kurowicka, and R. Cooke (2006).
Hybrid method for quantifying and analyzing bayesian belief nets.
Quality and Reliability Engineering International 22, 709–729.

Hobæk Haff, I. (2012).
Comparison of estimators for pair-copula constructions.
Journal of Multivariate Analysis 110, 91–105.

Hobæk Haff, I. (2013).
Estimating the parameters of a pair copula construction.
Bernoulli 19, 462–491.

Hobaek Haff, I. and J. Segers (2012, January).
Nonparametric estimation of pair-copula constructions with the empirical pair-copula.
ArXiv e-prints.

Joe, H. (1996).
Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters.
In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics*.

Joe, H. (1997).
Multivariate Models and Dependence Concepts.
London: Chapman & Hall.

Kauermann, G. and C. Schellhase (2013).
Flexible pair-copula estimation in D-vines with penalized splines.
submitted.

Kurowicka, D. and R. Cooke (2006).
Uncertainty analysis with high dimensional dependence modelling.
Chichester: Wiley.

Kurowicka, D. and H. Joe (2011).
Dependence Modeling - Handbook on Vine Copulae.
Singapore: World Scientific Publishing Co.

Lopez-Paz, D., J. M. Hernández-Lobato, and Z. Ghahramani (2013).
Gaussian process vine copulas for multivariate dependence.
arXiv preprint arXiv:1302.3979.

Low, R. K. Y., J. Alcock, R. Faff, and T. Brailsford (2013).
Canonical vine copulas in the context of modern portfolio management: Are they worth it?
Journal of Banking and Finance 37(8), 3085 – 3099.

McNeil, A. J. and J. Nešlehová (2009).
Multivariate Archimedean copulas, *d*-monotone functions and ℓ_1-norm symmetric distributions.
Annals of Statistics 37(5B), 3059–3097.

Min, A. and C. Czado (2011).
Bayesian model selection for multivariate copulas using pair-copula constructions.
Canadian Journal of Statistics 39, 239–258.

Morales-Nápoles, O., R. Cooke, and D. Kurowicka (2010).
About the number of vines and regular vines on n nodes.
Submitted for publication.

Nelsen, R. (2006).
An Introduction to Copulas.
New York: Springer.

Pachali, M. (2011).
Modeling dependence among weather measurements and tree ring data.
Master's thesis, Technische Universität München.

Panagiotelis, A., C. Czado, and H. Joe (2012).
Pair copula constructions for multivariate discrete data.
Journal of the American Statististical Association 107, 1063–1072.

Schepsmeier, U. (2013).
A goodness-of-fit test for regular vine copula models.
ArXiv e-prints.

Schmidl, D., C. Czado, S. Hug, and F. Theis (2013).
A vine-copula based adaptive MCMC sampler for efficient inference of dynamical systems (with discussion).
Bayesian Analysis 8(1), 1–22.

Sklar, A. (1959).
Fonctions dé repartition á n dimensions et leurs marges.
Publ. Inst. Stat. Univ. Paris 8, 229–231.

Smith, M., A. Min, C. Almeida, and C. Czado (2010).
Modeling longitudinal data using a pair-copula construction decomposition of serial dependence.
Journal of the American Statistical Association 105, 1467–1479.

Stöber, J. (2013).
Regular vine copulas with the simplifying assumption, time-variation, and mixed discrete and continuous margins.
Ph.D. thesis, Technische Universität München.

Stöber, J. and C. Czado (2013).
Regime switches in the dependence structure of multidimensional financial data.
Computational Statistics and Data Analysis, in press, –.

Stöber, J., C. Czado, H. Hong, and P. Ghosh (2012).
Interrelation among chronic diseases in the elderly: Longitudinal patterns identified by copula design for mixed responses.
submitted.

Stöber, J., H. Joe, and C. Czado (2013).
Simplified pair copula constructions: Limitations and extensions.
Journal of Multivariate Analysis 119(0), 101 – 118.

Stoeber, J. and U. Schepsmeier (2013).
Estimating standard errors in regular vine copula models.
to appear in Computational Statistics.

Takahasi, K. (1965).
Note on the multivariate burr's distribution.
Annals of the Institute of Statistical Mathematics 17, 257–260.

Vaz de Melo Mendes, B. and V. B. Accioly (2013).
Robust pair-copula based forecasts of realized volatility.
Applied Stochastic Models in Business and Industry.

Vuong, Q. H. (1989).
Likelihood ratio tests for model selection and non-nested hypotheses.
Econometrica 57, 307–333.