# Analysis of Australian electricity loads using joint Bayesian inference of D-Vines with autoregressive margins

**Claudia CZADO**[1], **Florian GÄRTNER** and **Aleksey MIN**

Zentrum Mathematik
Technische Universität München
Boltzmannstr. 3
D-85747 Garching, Germany

**Abstract**

Sklar's theorem allows the construction of models for dependent components using a multivariate copula together with marginal distributions. For estimation of the copula and marginal parameters, a two step procedure is often used to avoid high dimensional optimization. Here, marginal parameters are estimated first, then used to transform to uniform margins and in a second step, the copula parameters are estimated. This procedure is not efficient. Therefore, we follow a joint estimation approach in a Bayesian framework using Markov Chain Monte Carlo (MCMC) methods. This allows also for the assessment of parameter uncertainty using credible intervals. D-Vine copulas are utilized and as marginal models we allow for autoregressive models of first order. Finally, we apply these methods to Australian electricity loads demonstrating the usefulness of this approach. Bayesian model selection is also discussed and applied using a method suggested by Congdon (2006).

*Keywords:* multivariate copulas, vines, AR(1) margins, Bayesian inference, MCMC

## 1 Introduction

The celebrated result of Sklar (1959) shows that dependence among random variables can be separated from the marginal distributions. This forms the basis for the construction of many multivariate models in statistics and finance (see for example the books by Joe (1997), Nelsen (2006) and Cherubini, Luciano, and Vecchiato (2004)). While there are many bivariate copulas available for modelling bivariate dependence, the catalogue of multivariate copulas is less rich. Joe (1997) used a decomposition into pair-copulas to construct multivariate distributions. Bedford and Cooke (2002) systemized these constructions using a graphical tree representation and the book by Kurowicka and Cooke (2006) contains an overview.

[1]Corresponding author.

Aas, Czado, Frigessi, and Bakken (2007) derived statistical inference methods based on these pair copula constructions (PCC) using bivariate t-copulas. While the maximum likelihood estimation is feasible in small dimensions, the number of copula pairs increases quadratically with the number of dimensions, therefore assessing the variability of the estimates is difficult requiring the numerical inversion of large dimensional Hessian matrices. Therefore, we prefer to use a Bayesian approach to estimation and inference. Min and Czado (2008) develop such a Bayesian approach using Markov Chain Monte Carlo (MCMC) methods to estimate the posterior distribution for PCC's with bivariate t-copulas as building blocks for multivariate copula data, i.e. for an i.i.d. multivariate sample with uniform margins.

For financial applications, one usually starts with multivariate time series and in a first step, one estimates for each marginal time series its structure as an ARMA or GARCH structure. In a second step, one determines standardized residuals, which are assumed to form an i.i.d marginal sample. Depending on whether the distribution of the residuals are known or unknown, one uses a parametric or empirical probability transform to transform to data with approximate uniform margins. This separates the marginal distribution from the dependence structure. In a final step, this dependency is modelled using a multivariate copula and copula parameter are estimated. The statistical properties of such two step estimation procedures are investigated by Joe (2005) for a known standardized residual distribution and by Chen and Fan (2006) for unknown standardized residual distribution, respectively. It is known that such two step procedures are not efficient. The loss in efficiency depends on the specific data structure and model.

The main contribution of this paper is to provide a Bayesian analysis which jointly estimates the marginal and copula parameters. For this, we start with a simple marginal time series structure such as the first order autoregressive structure. In Section 2 we introduce the model and in Section 3 we derive a MCMC algorithm which will be used to facilitate the Bayesian analysis. In Section 4 we apply our methods to Australian electricity load data. Here we remove the trend and seasonal effects. In Section 5 we consider Bayesian model selection when one wants to choose among a small number of alternative model specifications based on approach suggested by Congdon (2006). We close with a summary and discussion section.

# 2 Multivariate time series with D-vine dependency and marginal autoregressive structure

First we describe the marginal time series structure. For this let $\boldsymbol{Y}^t := (Y_{1t}, \cdots, Y_{mt})$ for $t = 1, \cdots, m$ denote $m$ dependent time series with marginal AR(1) time series structure and normal errors, i.e.

$$\begin{aligned} Y_{it} &= \gamma_i \cdot Y_{i(t-1)} + \epsilon_{it} \quad t = 1, \ldots, T, \text{ and } \epsilon_{it} \sim \mathcal{N}(0, \sigma_i^2) \text{ i.i.d. }, i = 1, \ldots, m \\ Y_{i0} &\sim \mathcal{N}(0, \frac{\sigma_i^2}{1-\gamma_i^2}) \text{ and } |\gamma_i| < 1. \end{aligned}$$

From this it follows that $\boldsymbol{Y_i} := (Y_{i1}, \cdots, Y_{iT})' \sim \mathcal{N}_T(\mathbf{0}, \Sigma_i)$ with the $(r, s)$ th element of $\Sigma_i$ given by $\Sigma_{i,r,s} = \frac{\sigma_i^2}{1-\gamma_i^2}\gamma_i^{|r-s|}$ for $i = 1, \cdots, m$. We first transform each marginal time series $\boldsymbol{Y_i}$ to achieve i.i.d. margins by defining

$$\boldsymbol{Z_i} := \Sigma_i^{-1/2}\boldsymbol{Y_i} \sim \mathcal{N}_T(\mathbf{0}, I_T), \tag{1}$$

where $I_T$ is the identity matrix of size $T$. Wise (1955) shows that $\Sigma_i^{-1}$ is a tridiagonal matrix and it is easy to determine the Cholesky factorization $\Sigma_i^{-1/2}$ as

$$\Sigma_i^{-1/2} = \sigma_i^{-1} \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ -\gamma_i & 1 & 0 & & \vdots \\ 0 & -\gamma_i & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\gamma_i & \sqrt{1-\gamma_i^2} \end{pmatrix}$$

and

$$\det \Sigma_i^{-1/2} = \sigma_i^{-T} \sqrt{1-\gamma_i^2}.$$

For the dependency structure between the $m$ marginal time series, we now impose a $m$ dimensional D-vine (see Bedford and Cooke (2001) and Kurowicka and Cooke (2006)) structure on the i.i.d. random vectors $\boldsymbol{Z}^t := (Z_{1t}, \cdots, Z_{mt})$ for $t = 1, \cdots, m$. In particular, each random vector $\boldsymbol{Z}^t$ has a D-vine density $f(z_1, \cdots, z_m)$ given by

$$\prod_{k=1}^{m} f(z_k) \prod_{j=1}^{m-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\dots,i+j-1} \left\{ F(z_i|z_{i+1}, \dots, z_{i+j-1}), F(z_{i+j}|z_{i+1}, \dots, z_{i+j-1}) \right\}, \tag{2}$$

where $c_{i,i+j|i+1,\dots,i+j-1}(\cdot,\cdot)$ are arbitrary bivariate copula densities depending on parameters $\boldsymbol{\theta}_{i,i+j|i+1,\dots,i+j-1}$. Here $F(\cdot|\cdot)$ denote conditional cdf's. Utiliziling (1), we can use (2) together with the density transformation theorem to construct a multivariate density for $\boldsymbol{Y^t}$ for $t = 1, \cdots, T$. In the application, we will consider the case $m = 4$, where the D-vine density is given by

$$\begin{aligned} f(x_1, x_2, x_3, x_4) = & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \cdot \\ & \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{23}(F_2(x_2), F_3(x_3)) \cdot c_{34}(F_3(x_3), F_4(x_4)) \cdot \\ & \cdot c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \cdot c_{24|3}(F(x_2|x_3), F(x_4|x_3)) \cdot \\ & \cdot c_{14|23}(F(x_1|x_2, x_3), F(x_4|x_2, x_3)). \end{aligned} \tag{3}$$

For details and exact density expressions for $\boldsymbol{Y}^t$ in the case $m = 4$ see Chapter 5 of Gärtner (2008). In our application, we use bivariate t-copulas $c(u, v) = c(u, v|\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\nu, \rho)$, where $\nu$ is the degree of freedom (df) parameter and $\rho$ the correlation parameter. Correlation refers to the corresponding t-distribution and not to the copula.

# 3 Bayesian analysis of multivariate time series with D-vine dependency and marginal autoregressive structure

The multivariate time series models introduced in the previous section have many parameters to estimate, for example the model when $m = 4$ and bivariate t-copulas are used require 20 parameters. We could follow a two step approach by first estimating marginal parameters using the R function arima within the stats package. In a second step we transform to the Z-level using these estimated parameters and apply the D-vine R package of Daniel Berg (private communication) to determine maximum likelihood estimates of the D-vine parameters. However, at the moment, there is no efficient estimation of the standard errors of such estimates reflecting the uncertainty in the marginal parameters. Therefore, we follow a Bayesian approach where credible intervals can be determined naturally to assess the significance of the parameter.

First we have to specify prior distributions to complete the model specification. In particular we would like to use a uniform prior for the correlation matrix in the D-vine based on bivariate t-copulas. Using the results of Lewandowski, Kurowicka, and Joe (2007) and the calculations of Joe (2006), we assume a $Beta((4-k)/2,(4-k)/2)$ distribution on $(-1,1)$ for $\rho_{j|i_1,\cdots,i_k}$, where $k$ is the cardinality of the set of conditioning variables. This choice of conditional correlations results in an unconditional correlation matrix which is uniformly distributed over the space of correlation matrices. For the degree of freedom parameter, we choose a half Cauchy distribution as prior, i.e. $\pi(\nu) \propto \frac{1}{1+(\frac{\nu-1}{2})^2}, \nu \in (1,\infty)$ and for the marginal error variances $\sigma_i^2$ an inverse Gamma prior, given by $\pi(u) = \frac{0.001}{\Gamma(1)} u^{-2} \exp\left(-\frac{0.001}{u}\right)$ as suggested by Congdon (2003, p. 15). Finally for the autoregressive parameters $\gamma_i$, we choose a $N(0,10)$ prior on the transformed scale $a_i := \frac{1}{2}\log\frac{1+\gamma_i}{1-\gamma_i}$. In addition, we assume prior independence among all parameters.

The posterior distribution is however not analytically tractable, therefore we use Markov Chain Monte Carlo (MCMC) methods (see for example Gamerman and Lopes (2006)). In particular, we utilize the Metropolis-Hastings (MH) algorithm with independence proposal introduced by Tierney (1994). As proposal distribution, we take a normal distribution centered around the mode with covariance matrix a multiple of the inverse Hessian matrix evaluated at the mode as suggested by Gamerman and Lopes (2006, p. 83). For determining the mode and Hessian on the transformed scale of the parameters, the delta method is applied. To reduce computing time, we update the proposal distribution only every 20th iteration. The corresponding acceptance probabilities for the MH algorithm were developed in Gärtner (2008, Chapter 4). In particular, each parameter is updated individually.

# 4 Modeling Australian electricity loads

Initially the Australian electricity supply was organized as a vertically integrated monopoly with almost no trade or connection between different states. The liberalization started with the opening of the National Electricity Market (NEM) in December 1998. The first members were Victoria, Queensland, New South Wales, the Australian Capital Territory and then South Australia and Tasmania joined in 2005. It is a wholesale market for electricity to supply retailers and end-users.

The connection between the electricity producers and electricity consumers is facilitated by the establishment of the National Electricity Market Management Company (NEMMCO). This company manages a pool where the output of all generators is aggregated and scheduled to meet the forecasted demand.

Wholesale trading is done as a real-time market where supply and demand are instantaneously matched through a centrally dispatched process. As the offers are submitted by the generators every five minutes, NEMMCO determines the necessary plants and they are dispatched into production. So the market clearing price is determined every five minutes and is averaged for each trading interval (30 minutes).

Since the Australian electricity market is an energy-only market, there are a lot of price spikes. But the experience shows that these price spikes are enough incentives for the electricity companies to build new generation plants, for example as seen in South Australia in the period 1998 - 2003 where the generation capacities increased heavily after a series of price spikes to meet these peak demands. For more details on the Australian electricity market see Weron (2006).

The used load data consists of four time series of daily observations dating from May 16th, 2005 to June 30th, 2008, in total 1142 data points per time series. It describes the average daily load demand in Gigawatt (GW) for the regions Queensland, New South Wales, Victoria and South Australia calculated by averaging the half-hourly observed data for one day. This data is available at www.nemmco.com.au.

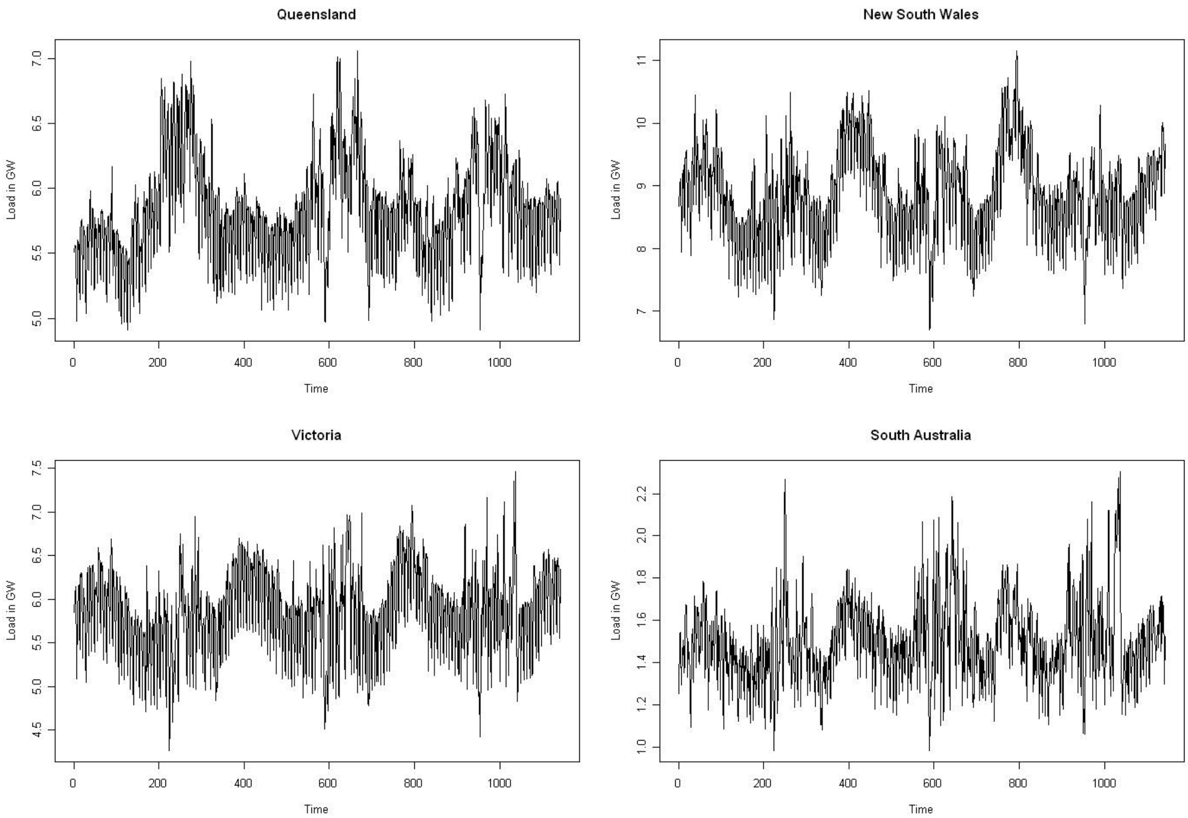


Figure 1: Time series plots of the observed data in the four different states

Electricity demand clearly shows seasonal fluctuations, mostly due to changing climate conditions like temperature or the number of daylight hours. We follow the classical technique of seasonal decomposition by thinking of a trend component $T_t$, a seasonal component $S_t$ and the remaining stochastic component $Y_t$, i.e we can represent the observed daily load data $\{x_1, \ldots, x_n\}$ as

$$x_t = T_t + S_t + y_t, \qquad t = 1, \ldots, n.$$

Using techniques of Weron (2006) and Brockwell and Davis (1991), we investigate the original time series to identify the trend and seasonal component. In a preprocessing step, these components will be removed before we analyse the marginal and across time series dependency.

Looking at the original data in Figure 1, we see that the trend component is negligible. Based on estimated (partial) autocorrelations and periodograms, we see evidence of a weekly and yearly cycle. Weron (2006) uses a rolling volatility technique to remove the annual seasonality. We apply this to our data. In a second step we remove the weekly cycle by fitting a moving average MA(7) model where only the seventh moving average coefficient is not equal to zero, i.e.

$$X_t = \epsilon_t + \theta_7 \epsilon_{t-7},\, \theta_7 \neq 0, \epsilon_t \sim WN(0, \sigma^2).$$

After the preprocessing, we want to test this data for stationarity or if there is a unit root. We use the augmented Dickey-Fuller test (ADF) with lag order 1, a Phillips-Perron test (PP) (both test for a unit root) and the KPSS test for stationarity (for details cf. Banerjee et al. (1993) and Kwiatkowski et al. (1992)). The KPSS test has the null hypothesis 'K0: The time series is stationary.' versus the alternative 'K1: The time series is not stationary.' and ADF and PP test have the null hypothesis 'H0: The time series has a unit root, i.e. the autoregressive coefficient has an absolute value of 1.' against the alternative 'H1: The absolute value of the autoregressive coefficient is smaller than 1.' The test results are given in Table 1. For each of the preprocessed

| | QLD | | NSW | | VIC | | SA | |
|---|---|---|---|---|---|---|---|---|
| | stat. | p-value | stat. | p-value | stat. | p-value | stat. | p-value |
| KPSS | 0.0671 | $> 0.1$ | 0.126 | $> 0.1$ | 0.099 | $> 0.1$ | 0.065 | $> 0.1$ |
| | accept K0 | | accept K0 | | accept K0 | | accept K0 | |
| ADF | -12.74 | $< 0.01$ | -13.70 | $< 0.01$ | -15.11 | $< 0.01$ | -15.04 | $< 0.01$ |
| | reject H0 | | reject H0 | | reject H0 | | reject H0 | |
| PP | -334.31 | $< 0.01$ | -347.94 | $< 0.01$ | -413.37 | $< 0.01$ | -421.71 | $< 0.01$ |
| | reject H0 | | reject H0 | | reject H0 | | reject H0 | |

Table 1: Tests for Stationarity and Unit Root for the preprocessed time series

time series, we cannot reject the KPSS test for stationarity. In addition, we have to reject the ADF and PP test for unit roots. So, we assume stationarity for all four time series. Therefore, we will model them marginally with an AR(1) process. The preprocessed time series are given in Figure 2. For more details on the preprocessing methods and alternative preprocessing see Gärtner (2008, Chapter 3).

Now we are ready to investigate the dependencies among the preprocessed time series. For this, we have preprocessed time series for Queensland, New South Wales, Victoria and South Australia, each consisting of 1134 observations available. We chose this order because of geographical reasons. These states are neighbored and connected with all its infrastructure in this way along the Eastern Coast of Australia beginning with Queensland in the North, then New South Wales, then Victoria and South Australia following in the South West.

Applying the joint MCMC (JMCMC) approach developed in Section 2, we ran 10000 MCMC iterations. Using trace plots and estimated autocorrelation (see Gärtner (2008, Section 5.3)), we determined appropriate burnin and thining parameters. A burnin of 1000 iterations for all but one parameter and a thining to every 20th iterations were sufficient.

For comparison, we also determined the two step MLE estimates, i.e. first marginal parameters are estimated (marg. MLE) and used to transform to the Z-level. For the copula parameters, the Z-level time series are further transformed using the standard normal cdf to time series with uniform margins. Finally, this data is used to determine the MLE's of the copula parameters (C-MLE). The resulting kernel density estimates of the posterior distribution for each parameter are given in Figures 3 and 4, while summary statistics are presented in Table 2.

The MLE's of the marginal autoregressive parameters are higher than the corresponding JMCMC estimates indicating the effect of ignoring the joint dependency. Some of these differences are credible since the marginal MLE values for the parameters $\gamma_2$, $\gamma_3$ and $\gamma_4$ lie outside of the estimated posterior 95% credibility interval. There is however good agreement for the remaining parameters. Some of the df parameters are quite high especially for the df parameters corresponding to conditional copula pairs. This leads to the question if one can reduce the

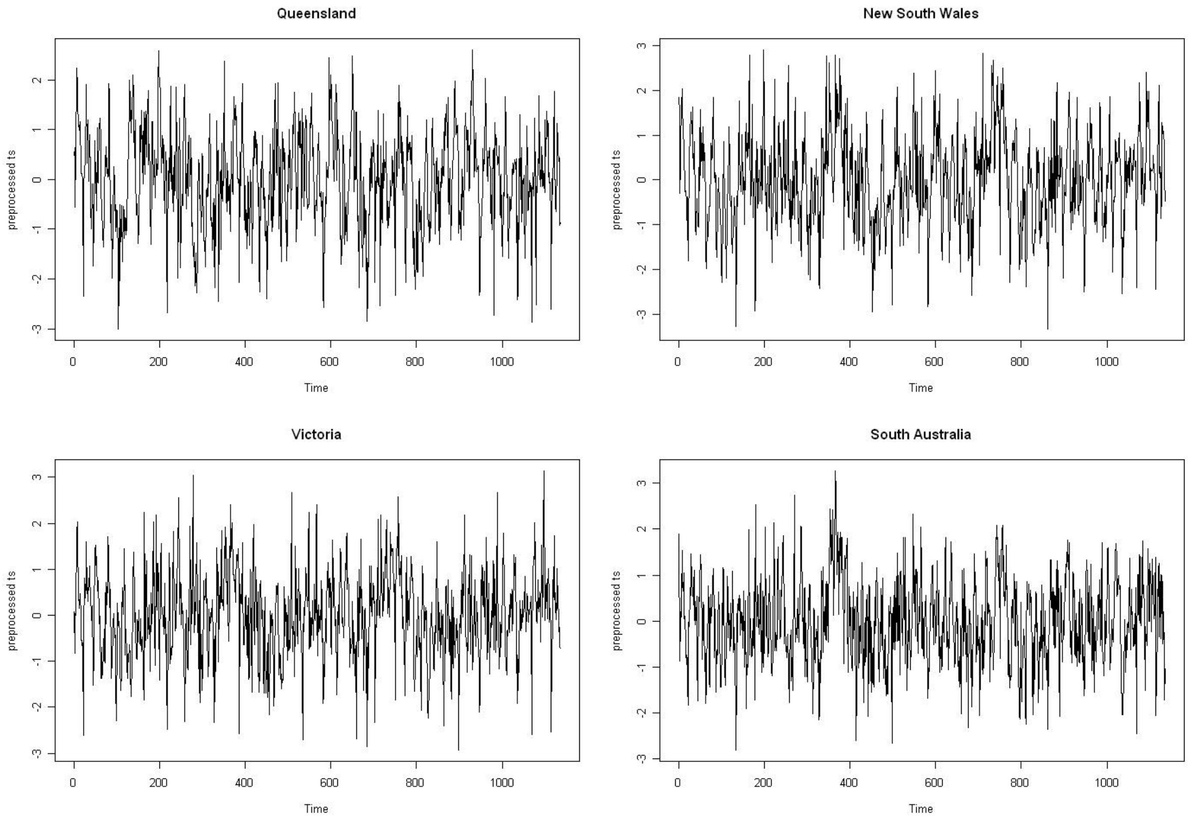


Figure 2: Time series plots of the preprocessed data in the four different states

copula dependency model to Gaussian bivariate copulas for those conditional copula pairs. We will investigate this question in the following section.

# 5 Bayesian model selection

As we have seen in the previous section, we might want to compare several model specifications in a Bayesian setup. For this we want to compare posterior model probabilities for models of interest. Assume that we have fitted models $M_1, \ldots, M_k$ with a MCMC method where model $M_k$ has parameters $\boldsymbol{\theta}_k$ and want to estimate

$$P(\text{Model } M_k|\text{data}), \qquad k = 1, \ldots, K.$$

Congdon (2006) gives the following estimation procedure under the assumption that the distribution of the data under Model $M_k$ is independent of $\{\boldsymbol{\theta}_{j\neq k}\}$ and he assumes independence among all $\boldsymbol{\theta}_k$ given Model $M$. Then he shows that the posterior distributions are independent and can be sampled individually. He uses the relationship

$$P(M = M_k|\text{data}, \boldsymbol{\theta}) \propto P(\text{data}|\boldsymbol{\theta}, M = M_k)P(\boldsymbol{\theta}|M = M_k)P(M = M_k). \tag{4}$$

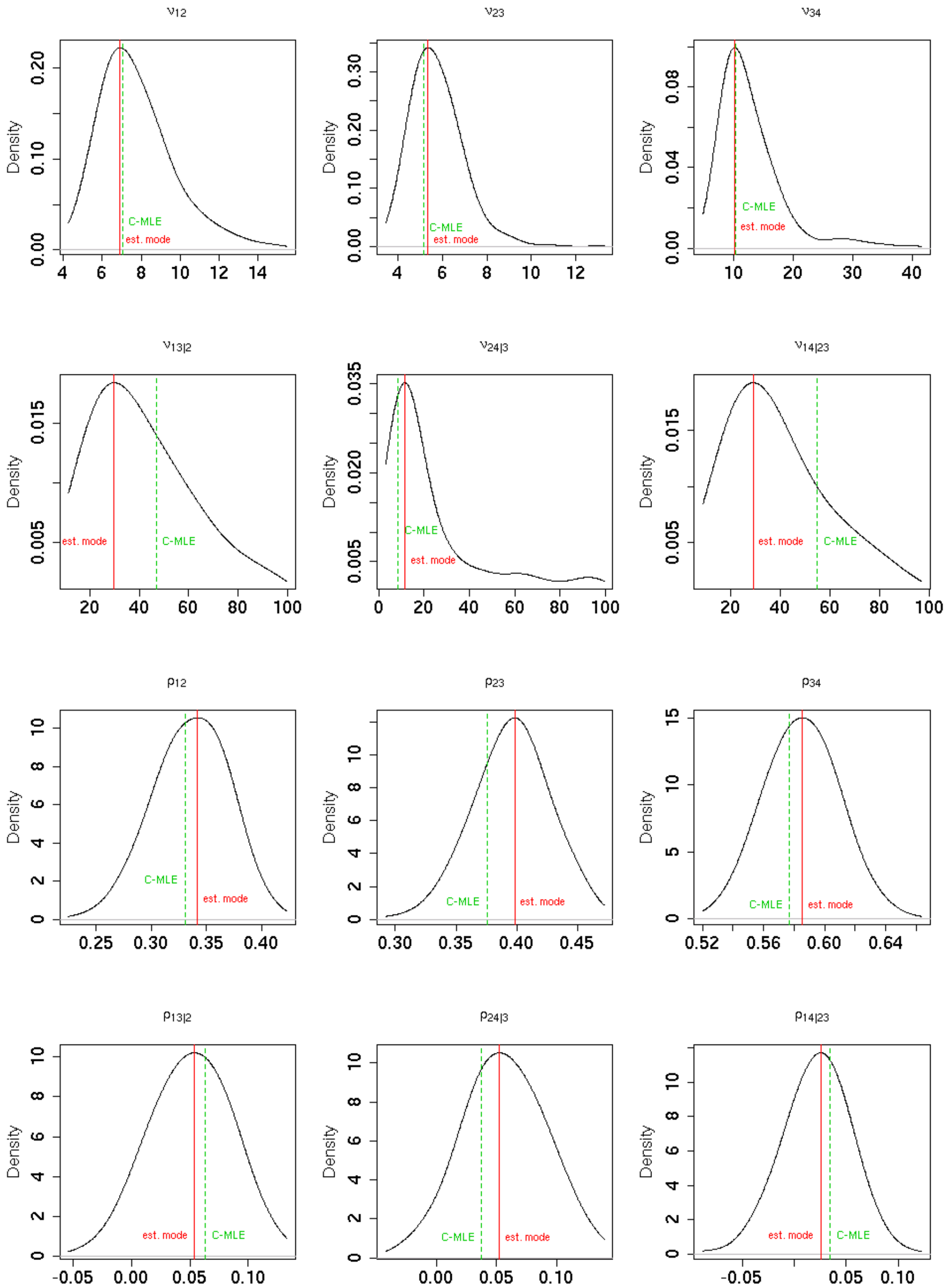


Figure 3: Plots of the estimated kernel density of $\nu$ and $\rho$ estimated of the observed real data based on the thinned out MCMC chain

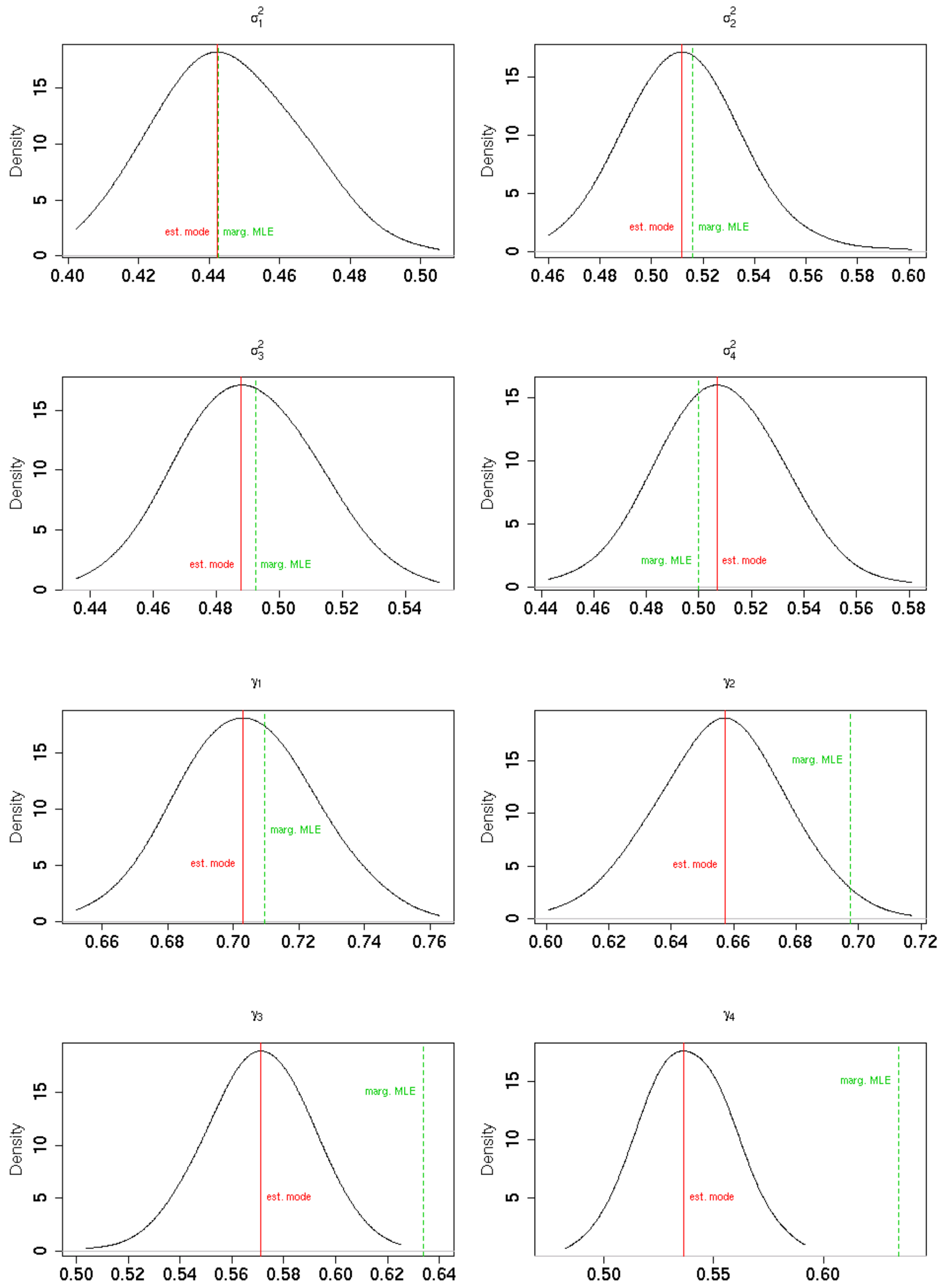


Figure 4: Plots of the estimated kernel density of $\sigma^2$ and $\gamma$ estimated of the observed Real Data based on the thinned out MCMC chain

| Copula | 2.5% | 5% | 50% | 95% | 97.5% | mean | mode | C-MLE |
|---|---|---|---|---|---|---|---|---|
| $\nu_{12}$ | 5.17 | 5.45 | 7.37 | 11.72 | 12.71 | 7.80 | 6.92 | 7.32 |
| $\nu_{23}$ | 4.11 | 4.26 | 5.57 | 7.65 | 8.60 | 5.76 | 5.36 | 5.52 |
| $\nu_{34}$ | 7.08 | 7.69 | 11.45 | 24.29 | 29.68 | 12.89 | 10.23 | 11.17 |
| $\nu_{13\|2}$ | 15.01 | 16.28 | 36.89 | 84.56 | 91.77 | 41.44 | 29.80 | 38.55 |
| $\nu_{24\|3}$ | 4.20 | 4.73 | 14.25 | 78.22 | 93.44 | 22.84 | 11.60 | 8.34 |
| $\nu_{14\|23}$ | 12.32 | 14.58 | 34.43 | 78.22 | 86.81 | 38.94 | 29.00 | 58.07 |
| $\rho_{12}$ | 0.27 | 0.28 | 0.34 | 0.38 | 0.39 | 0.34 | 0.34 | 0.33 |
| $\rho_{23}$ | 0.33 | 0.35 | 0.40 | 0.45 | 0.45 | 0.40 | 0.40 | 0.39 |
| $\rho_{34}$ | 0.54 | 0.55 | 0.59 | 0.62 | 0.63 | 0.59 | 0.59 | 0.57 |
| $\rho_{13\|2}$ | $-0.01$ | $-0.00$ | 0.05 | 0.10 | 0.11 | 0.05 | 0.05 | 0.06 |
| $\rho_{24\|3}$ | $-0.01$ | 0.00 | 0.05 | 0.11 | 0.12 | 0.06 | 0.05 | 0.04 |
| $\rho_{14\|23}$ | $-0.04$ | $-0.03$ | 0.02 | 0.07 | 0.08 | 0.02 | 0.03 | 0.03 |
| Marginal | 2.5% | 5% | 50% | 95% | 97.5% | mean | mode | marg. MLE |
| $\sigma_1^2$ | 0.41 | 0.41 | 0.44 | 0.48 | 0.48 | 0.45 | 0.44 | 0.44 |
| $\gamma_1$ | 0.67 | 0.67 | 0.70 | 0.74 | 0.74 | 0.70 | 0.70 | 0.71 |
| $\sigma_2^2$ | 0.47 | 0.48 | 0.51 | 0.55 | 0.56 | 0.51 | 0.51 | 0.52 |
| $\gamma_2$ | 0.62 | 0.62 | 0.66 | 0.69 | 0.69 | 0.66 | 0.66 | 0.70 |
| $\sigma_3^2$ | 0.45 | 0.46 | 0.49 | 0.52 | 0.53 | 0.49 | 0.49 | 0.49 |
| $\gamma_3$ | 0.53 | 0.54 | 0.57 | 0.60 | 0.61 | 0.57 | 0.57 | 0.63 |
| $\sigma_4^2$ | 0.47 | 0.47 | 0.51 | 0.54 | 0.55 | 0.51 | 0.51 | 0.50 |
| $\gamma_4$ | 0.50 | 0.51 | 0.54 | 0.57 | 0.58 | 0.54 | 0.54 | 0.63 |

Table 2: Estimated posterior mean, mode and quantiles of JMCMC as well as marginal MLE, starting values and C-MLE for the preprocessed Australian load data

We assume now that the K independent MCMC runs result in

$$\begin{array}{lcr} M_1 : \boldsymbol{\theta}_1^{(t)}, r = 1, \dots R & & p(\boldsymbol{\theta}_1|\text{data}) \\ \vdots & \text{which approximate} & \vdots \\ M_K : \boldsymbol{\theta}_K^{(t)}, r = 1, \dots R & & p(\boldsymbol{\theta}_K|\text{data}). \end{array}$$

We use $\{\boldsymbol{\theta}^{(r)} := (\boldsymbol{\theta}_1^{(r)}, \dots, \boldsymbol{\theta}_K^{(r)}), r = 1, \dots R\}$ and hence, we can approximate

$$P(M|\text{data}) = \int P(M|\boldsymbol{\theta}, \text{data}) p(\boldsymbol{\theta}|\text{data}) \mathrm{d}\boldsymbol{\theta}$$

by

$$\hat{P}(M|\text{data}) := \frac{1}{R} \sum_{r=1}^{R} P(M|\boldsymbol{\theta}^{(r)}, \text{data}).$$

Using Equation (4), we can estimate $P(M = M_k|\text{data}, \boldsymbol{\theta}^{(r)})$ by

$$w_k^{(r)} := \frac{G_k^{(r)}}{\sum_{j=1}^{K} G_k^{(r)}},$$

where

$$\begin{aligned} G_k^{(r)} &:= \exp(L_k^{(r)} - L_{\max}^{(r)}) \\ L_k^{(r)} &:= \log(P(\text{data}|\boldsymbol{\theta}^{(r)}, M = M_k) P(\boldsymbol{\theta}^{(r)}|M = M_k) P(M = M_k)) \\ L_{\max}^{(r)} &:= \max_{k=1,\dots,K} L_k^{(r)}. \end{aligned}$$

Therefore, we get

$$\hat{T}_k := \frac{1}{R}\sum_{r=1}^{R} w_k^{(r)}$$

as an estimator for $P(M = M_k|\text{data})$.

We investigated 6 models for the Australian load data, which are described in Table 3 together with their estimated posterior model probability. These results indicate clearly that the model with a marginal AR(1) structure and a four-dimensional t-copula where the conditional correlation parameters are fixed to 0 gives the best fit to the observed data.

| Model | Model | $\hat{T}_k$ |
|---|---|---|
| M1 | Joint Bayesian estimation with marginal AR(1) and D-Vine of pair t-copulas | $4.2689 \cdot 10^{-06}$ |
| M2 | Joint Bayesian estimation of reduced model M1: marginal AR(1), unconditional pair copulas as t-copulas, conditional pair copulas as t-copulas with correlation 0 and df = 100 | 0.0351 |
| M3 | Joint Bayesian estimation with marginal AR(1) and D-Vine of normal pair copulas (approximated by a t-copula with df=100) | $4.5999 \cdot 10^{-13}$ |
| M4 | Joint Bayesian estimation of reduced model M4: marginal AR(1), unconditional pair copulas as Gauss copulas (approximated by a t-copula with df=100), conditional pair copulas as pair t-copulas with correlation 0 | $3.2536 \cdot 10^{-14}$ |
| M5 | marginal AR(1) and four-dimensional t-copula with common df | 0.3245 |
| M6 | marginal AR(1) and four-dimensional t-copula with with common df and the conditional correlation parameters fixed to 0 | 0.6404 |

Table 3: Estimated posterior model probabilities for six models for the Australian load data

Finally, we present the parameter estimates for the Model M6 in Table 4. This shows that there are strong marginal autocorrelations present in the four time series. The strongest one is observed in Queensland, while the lowest one is in South Australia. The dependence between the time series on the Z variable level has a first order Markov structure determined by the unconditional bivariate t-copulas. Since conditional and partial correlations are the same for elliptical distributions (see Baba and Sibuya (2005)) and there exists a one-to-one relationship between partial correlations and unconditional correlations, the remaining unconditional correlations in M6 can be determined. In particular the posterior mode for $\rho_{13}, \rho_{24}$ and $\rho_{14}$ are estimated to be $.13, .23$ and $.075$ respectively. This shows that the strongest dependence is Victoria and South Australia followed by New South Wales and Victoria. This is reasonable since South Australia and New South Wales are neighbors of Victoria and Victoria is the most populated region among the 4 regions.

# 6 Summary and discussion

This paper developed a joint Bayesian analysis of a multivariate copula model with AR(1) time series margins. This avoids efficiency loss introduced by the usual two step estimation procedure. Model selection was facilitated by applying the approach by Congdon (2006). This approach

| | 2.5% | 5% | 50% | 95% | 97.5% | mean | mode | C-MLE |
|---|---|---|---|---|---|---|---|---|
| $\nu$ | 5.80 | 6.08 | 7.44 | 9.30 | 9.79 | 7.54 | 7.32 | 9.59 |
| $\rho_{12}$ | 0.27 | 0.28 | 0.33 | 0.38 | 0.39 | 0.33 | 0.33 | 0.30 |
| $\rho_{23}$ | 0.34 | 0.35 | 0.40 | 0.45 | 0.46 | 0.40 | 0.40 | 0.34 |
| $\rho_{34}$ | 0.54 | 0.55 | 0.59 | 0.62 | 0.63 | 0.59 | 0.59 | 0.58 |
| | 2.5% | 5% | 50% | 95% | 97.5% | mean | mode | marg. MLE |
| $\sigma_1^2$ | 0.40 | 0.41 | 0.44 | 0.48 | 0.48 | 0.44 | 0.44 | 0.44 |
| $\gamma_1$ | 0.67 | 0.67 | 0.71 | 0.74 | 0.74 | 0.71 | 0.71 | 0.71 |
| $\sigma_2^2$ | 0.46 | 0.47 | 0.50 | 0.54 | 0.54 | 0.50 | 0.50 | 0.52 |
| $\gamma_2$ | 0.62 | 0.63 | 0.66 | 0.69 | 0.70 | 0.66 | 0.66 | 0.70 |
| $\sigma_3^2$ | 0.45 | 0.46 | 0.49 | 0.52 | 0.53 | 0.49 | 0.49 | 0.49 |
| $\gamma_3$ | 0.52 | 0.53 | 0.56 | 0.59 | 0.60 | 0.56 | 0.56 | 0.63 |
| $\sigma_4^2$ | 0.47 | 0.48 | 0.51 | 0.55 | 0.56 | 0.51 | 0.51 | 0.50 |
| $\gamma_4$ | 0.50 | 0.50 | 0.54 | 0.57 | 0.57 | 0.54 | 0.54 | 0.63 |

Table 4: Estimated posterior mean, mode and quantiles of the joint MCMC as well as marginal MLE and C-MLE for marginal AR(1) and four-dimensional t-Copula with the conditional correlation parameters fixed to 0

however has been criticized by Robert and Marin (2008) in general. Alternatively, one can use reversible jump MCMC (RJMCMC) developed by Green (1995). However, it is our experience that Congdon's method is a close approximation to RJMCMC for copula models based on vines.

Several extensions are of interest, for example using higher order autoregressive models and especially to use GARCH margins for financial applications. For Bayesian approaches to univariate GARCH models see Bauwens and Lubrano (1998) and Ardia (2008). The joint Bayesian estimation of such marginal models together with copula models based on vines is subject of current research.

## Acknowledgement

Claudia Czado is supported by the German Research Foundation (DFG).

## References

Aas, K., C. Czado, A. Frigessi, and H. Bakken (2007). Pair-Copula Construtions of Multiple Dependence. *Insurance: Mathematics and Economics*. doi: 10.1016/j.insmatheco.2007.02.001.

Ardia, D. (2008). *Financial Risk Management with Bayesian Estimation of GARCH Models*. Springer Verlag.

Baba, K. and M. Sibuya (2005). Equivalence of partial and conditional correlation coefficients. *Journal of Japanese Statistical Society 35*, 1–19.

Banerjee, A., J. J. Dolado, J. W. Galbraith, and D. F. Hendry (1993). *Conintegration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford University Press.

Bauwens, L. and M. Lubrano (1998). Bayesian inference on GARCH models using the Gibbs sampler. *Econometrics Journal 1*(1), C23–C46.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a New Graphical Model for Dependent Random Variables. *The Annals of Statistics 30*(4), 1031–1068.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods* (2nd ed.). Springer Series in Statistics. New York: Springer.

Chen, X. and Y. Fan (2006). Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics 135*, 125–154.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. New York: John Wiley & Sons.

Congdon, P. (2003). *Applied Bayesian Modelling*. Wiley & Sons.

Congdon, P. (2006). Bayesian model choice based on Monte Carlo estimates of posterior model probabilities. *Computational Statistics & Data Analysis 50*, 346 – 357.

Gamerman, D. and H. F. Lopes (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* (2nd ed.). CRC Press.

Gärtner, F. (2008). *Bayesian analysis of multivariate time series models based on pair-copula construction*. Diploma thesis in the Zentrum Mathematik: Technische Universität München.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*(4), 711–732.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman Hall.

Joe, H. (2005). Asymptotic effiency of the two stage estimation method for copula-based models. *Journal of Multivariate Analysis 94*, 401–419.

Joe, H. (2006). Generating Random Correlation Matrices Based on Partial Correlations. *Journal of Multivariate Analysis 97*, 2177–2189.

Kurowicka, D. and R. M. Cooke (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. New York: Wiley & Sons.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root. *Journal of Econometrics 54*, 159 – 178.

Lewandowski, D., D. Kurowicka, and H. Joe (2007). Generating Random Correlation Matrices Based on Vines and Extended Onion Method. *Journal of Multivariate Analysis*, submitted.

Min, A. and C. Czado (2008). Bayesian inference for multivariate copulas using pair-copula constructions. Preprint, available under http://www-m4.ma.tum.de/Papers/index.html.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Springer Series in Statistics. Springer.

Robert, C. and J.-M. Marin (2008). On some difficulties with a posterior probability approximation technique. Preprint, available under http://arxiv.org/abs/0801.3513.

Sklar, A. (1959). Fonctions de répartitions à n dimensions et leur marges. *Publications de l'Institut de Statistique de l'Université de Paris 8*, 229 – 231.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics 22*(4), 1701 – 1762.

Weron, R. (2006). *Modeling and Forecasting Electricity Loads and Prices.* Chichester, GB: Wiley & Sons.

Wise, J. (1955). The Autocorrelation Function and the Spectral Density Function. *Biometrika 42*(1/2), 151 – 159.