Modelling Longitudinal Data using a Pair-Copula Decomposition of Serial Dependence

Michael Smith[1,⋆], Aleksey Min[2],

Carlos Almeida[2] & Claudia Czado[2]

1. Melbourne Business School, University of Melbourne

2. Zentrum Mathematik, Technische Universität München

First Version: September 2009

This Version: June 2010

Forthcoming in the *Journal of the American Statistical Association*, Theory and Methods Section.

(Supplementary Materials Also Available)

⋆ Corresponding author; address for correspondence: Professor Michael Smith, Melbourne Business School, 200 Leicester Street, Carlton, VIC, 3053, Australia. Email: mike.smith@mbs.edu.

Modelling Longitudinal Data using a Pair-Copula Decomposition of Serial Dependence

Abstract

Copulas have proven to be very successful tools for the flexible modelling of cross-sectional dependence. In this paper we express the dependence structure of continuous-valued time series data using a sequence of bivariate copulas. This corresponds to a type of decomposition recently called a 'vine' in the graphical models literature, where each copula is entitled a 'pair-copula'. We propose a Bayesian approach for the estimation of this dependence structure for longitudinal data. Bayesian selection ideas are used to identify any independence pair-copulas, with the end result being a parsimonious representation of a time-inhomogeneous Markov process of varying order. Estimates are Bayesian model averages over the distribution of the lag structure of the Markov process. Using a simulation study we show that the selection approach is reliable and can improve the estimates of both conditional and unconditional pairwise dependencies substantially. We also show that a vine with selection out-performs a Gaussian copula with a flexible correlation matrix. The advantage of the pair-copula formulation is further demonstrated using a longitudinal model of intraday electricity load. Using Gaussian, Gumbel and Clayton pair-copulas we identify parsimonious decompositions of intraday serial dependence, which improve the accuracy of intraday load forecasts. We also propose a new diagnostic for measuring the goodness of fit of high-dimensional multivariate copulas. Overall, the pair-copula model is very general and the Bayesian method generalizes many previous approaches for the analysis of longitudinal data. Supplemental materials for the article are also available online.

Keywords: Longitudinal Copulas; Covariance Selection; Inhomogeneous Markov Process; D-vine; Bayesian Model Selection; Goodness of Fit; Intraday Electricity Load

1 Introduction

Modelling multivariate distributions using copulas has proven to be highly popular. This is largely due to the flexibility that copula models provide, whereby the marginal distributions can be modelled arbitrarily, and any dependence captured by the copula. Major applications include survival analysis where much early work occurred (Clayton 1978; Oakes 1989), actuarial science (Frees & Valdez 1998) and finance (Cherubini et al. 2004; Patton 2006); Joe (1997) and Nelsen (2006) provide introductions to copula models and their properties. While there are many copulas from which to choose, only a few are readily applicable to high dimensional problems. Copula built from elliptical distributions, such as the Gaussian (Song 2000) or t (Demarta & McNeil 2005), are most popular in this case. However, these can prove restrictive and in the recent graphical models literature alternative copulas have been proposed that are constructed from series of bivariate copulas (Joe 1996). There are a large number of permutations in which this can be undertaken, and Bedford & Cooke (2002) organize the different decompositions in a systematic way. They label the resulting multivariate copulas 'vines', while Aas et al. (2009) label the component bivariate copulas 'pair-copulas'; see Kurowicka & Cooke (2006), Haff et al. (2010) and Czado (2010) for recent overviews.

To date copula models have been employed largely to account for cross-sectional dependence. Applications to serial dependence in time series and longitudinal data are rare, although the potential is large. Here, the marginal distribution of the process at each point in time can be modelled arbitrarily, while dependence over time is captured by a multivariate copula. Meester & MacKay (1994) and Joe (1997, pp.243-280) provide early examples, while more recent examples include Lambert & Vandenhende (2002), Frees & Wang (2006), Sun et al. (2008) and Domma et al. (2009). However, many of these authors employ multivariate copulas that do not fully exploit the time ordering of the margins. In this paper we aim to show that doing so results in a more flexible representation that is both more insightful and allows for improved estimates for continuous data.

We decompose the distribution of a continuous process at a point in time, conditional upon the past, into the product of a sequence of bivariate copula densities and the marginal density. This type of copula is called a D-vine by Bedford & Cooke (2002), where any mix of bivariate copulas can be used for the component pair-copulas, resulting in an extremely flexible modelling framework. When the process is Markovian, this can be accounted for by setting appropriate pair-copulas to the independence copula. For longitudinal data this results in a time-inhomogeneous Markov process with order that also varies over time. As we demonstrate here, not only does this produce greater insight into the underlying process, in high-dimensional longitudinal applications this parsimony can also lead to a substantial improvement in the quality of inference.

Bedford & Cooke (2002) give the theoretical construction of regular vines, however no estimation of pair-copula parameters is attempted. Kurowicka & Cooke (2006) estimate Gaussian vine copula parameters by minimising the determinant of the correlation matrix. Aas et al. (2009) estimate pair-copula using maximum likelihood for Gaussian and non-Gaussian pair-copulas in both C-vine and D-vines. Min & Czado (2010a) suggest a Bayesian method for the estimation of D-vines using Markov chain Monte Carlo (MCMC). Min & Czado (2010b) use vine copulas to model the dependency among foreign exchange rates using maximum likelihood. In all cases cross-sectional dependence is examined, where the determination of an appropriate ordering of the dimensions for the decomposition remains an open problem. However, for time-ordered data this issue does not arise, and one of the insights of this paper is that a pair-copula decomposition is arguably more appropriate.

We suggest a Bayesian approach for the estimation of a pair-copula decomposition for longitudinal data. Indicator variables are introduced to identify which pair-copulas are independence copulas. By doing so, we extend existing Gaussian covariance selection methods to a flexible non-Gaussian framework, both in the longitudinal case (Pourahmadi 1999; Smith & Kohn 2002; Huang et al. 2006; Liu et al. 2009) and more generally. We use a Metropolis-Hastings scheme to generate the indicator variable and dependence parameter(s) of a pair-copula jointly, where the proposal is based on a latent variable representation of the pair-copula parameter(s). The full spectrum of posterior inference is available, including measures of conditional and marginal pairwise dependence. All estimates are model averages over the distribution of the order of the Markov process. We also propose a diagnostic for the quality of fit of a multivariate copula using the distribution of the sum of the transformed uniform margins.

A simulation study using Gaussian, Gumbel and Clayton pair-copulas highlights the accuracy and reliability of the Bayesian procedure for both the selection and estimation of pair-copulas. The results show that selection can improve the estimated dependence structure, and that the vine copula provides a substantial improvement over the alternative of using a multivariate Gaussian copula; both when the Gaussian copula correlation matrix is unrestricted, or parsimonious in the fashion of Smith & Kohn (2002). We demonstrate the usefulness of the method using a longitudinal model for intraday electricity load in the Australian state of New South Wales. Here, marginal regressions with t disturbances are used with time and weather based covariates and load the previous day. Intraday dependence is captured flexibly using Gaussian, Gumbel and Clayton pair-copulas. A time-varying Markov structure is identified that is inline with that used in the energy forecasting literature (Cottet & Smith 2003; Soares & Medeiros 2008). We find that intraday dependence is nonlinear and better captured using the Gumbel pair-copula model. We demonstrate this using a forecasting study, where the choice of pair-copula type and selection are shown to improve the accuracy of out-of-sample intraday forecasts.

The rest of the paper is organised as follows. In Section 2 we outline the pair-copula decomposition of the joint distribution of time-ordered data, along with the resulting likelihood for longitudinal data. Section 3 discusses priors, Bayesian estimation and pair-copula selection. Also discussed are measures of conditional and unconditional pairwise dependence, and the diagnostic for the quality of fit. Section 4 contains the simulation study, Section 5 the electricity load example and Section 6 the conclusion. The supplementary materials contain an illustration of a Gaussian D-vine, additional figures and empirical results for the electricity example and a second empirical example.

2 The Model

2.1 Pair-copula Construction for Time Series

Consider a univariate time series $\boldsymbol{X} = \{X_1, \ldots, X_T\}$ of continuously distributed data observed at T possibly unequally-spaced points in time. If the underlying process is Markovian, then this can be exploited by selecting models for the conditionals in the decomposition of the joint density of $\boldsymbol{X}$:

$$f(\boldsymbol{x}) = \prod_{t=2}^{T} f(x_t|x_{t-1},\ldots,x_1)f(x_1)\,, \tag{2.1}$$

where $\boldsymbol{x} = (x_1,\ldots,x_T)$. Copulas can be used to construct a general representation for each conditional as follows. For $s < t$ there always exists a density $c_{t,s}$ on $[0,1]^2$, such that

$$\begin{aligned} f(x_t, x_s|x_{t-1},\ldots,x_{s+1}) &= c_{t,s}(F(x_t|x_{t-1},\ldots,x_{s+1}), F(x_s|x_{t-1},\ldots,x_{s+1}); x_{t-1},\ldots,x_{s+1}) \\ &\times \; f(x_t|x_{t-1},\ldots,x_{s+1})f(x_s|x_{t-1},\ldots,x_{s+1})\,. \end{aligned} \tag{2.2}$$

Here, $F(x_t|x_{t-1},\ldots,x_{s+1})$ and $F(x_s|x_{t-1},\ldots,x_{s+1})$ are the conditional distribution functions of X_t and X_s, respectively. This is the theorem of Sklar (1959) conditional upon $\{X_{t-1},\ldots,X_{s+1}\}$. In vine copula models $c_{t,s}$ is simplified by dropping dependence upon $(x_{t-1},\ldots,x_{s+1})$ and is called a 'pair-copula'; see Haff et al. (2010). We adopt this simplification throughout, and by setting $s = 1$, application of equation (2.2) gives:

$$f(x_t|x_{t-1},\ldots,x_1) = c_{t,1}(F(x_t|x_{t-1},\ldots,x_2), F(x_1|x_{t-1},\ldots,x_2))f(x_t|x_{t-1},\ldots,x_2)\,.$$

Repeated application with $s = 2, 3, \ldots, t-1$ leads to the following:

$$\begin{aligned} f(x_t|x_{t-1},\ldots,x_1) &= \prod_{j=1}^{t-2} \{c_{t,j}(F(x_t|x_{t-1},\ldots,x_{j+1}), F(x_j|x_{t-1},\ldots,x_{j+1}))\} \\ &\times \; c_{t,t-1}(F(x_t), F(x_{t-1}))f(x_t)\,, \end{aligned} \tag{2.3}$$

where $F(x_t)$ and $f(x_t)$ are the marginal distribution function and density of X_t, respectively.

We denote $u_{t|j} \equiv F(x_t|x_{t-1},\ldots,x_j)$ and $u_{j|t} \equiv F(x_j|x_t,\ldots,x_{j+1})$, where $j < t$. They correspond to projections backwards and forwards $t-j$ steps, respectively. By also denoting $u_{t|t} \equiv F(x_t)$, the joint density at equation (2.1) can be written as

$$f(\boldsymbol{x}) = \prod_{t=2}^{T} \left\{ \prod_{j=1}^{t-1} \left\{ c_{t,j}(u_{t|j+1}, u_{j|t-1}) \right\} f(x_t) \right\} f(x_1)\,, \tag{2.4}$$

which is a product of T marginal densities and $T(T-1)/2$ pair-copula densities.

Equation (2.4) can be recognised as a 'D-vine' and is one of a wider class of vine decompositions recently discussed in the context of graphical models by Bedford & Cooke (2002) and others. In this literature, the notation used makes the conditioning set explicit; for example, $c_{t,j|t-1,t-2,\ldots,j+1}$ would denote the copula density in equation (2.4). This is essential for differentiating between vine decompositions of general vectors $\boldsymbol{X}$. However, it is not necessary to uniquely identify the pair-copulas of the D-vine decomposition when the elements of $\boldsymbol{X}$ are time-ordered. Throughout the rest of the paper we employ parametric models for the pair-copulas and write each density as $c_{t,j}(u_1, u_2; \theta_{t,j})$, where $\theta_{t,j}$ are the parameters. If $\boldsymbol{u} = (u_{1|1}, \ldots, u_{T|T})$, then $c^\dagger(\boldsymbol{u}; \Theta) = \prod_{t=2}^{T} \prod_{j=1}^{t-1} c_{t,j}(u_{t|j+1}, u_{j|t-1}; \theta_{t,j})$ is the density of the multivariate D-vine copula with parameters $\Theta = \{\theta_{t,s}; (t,s) \in \mathcal{I}\}$ and $\mathcal{I} = \{(t,s); t = 2, \ldots, T, s < t\}$. In the special case where $c_{t,j}$ are bivariate Gaussian copula densities, Aas et al. (2009) show that $c^\dagger$ is the density of a T-dimensional Gaussian copula.

The most challenging aspect of the D-vine representation is the evaluation of $u_{t|j+1}$ and $u_{j|t-1}$ in equation (2.4). The following property (Joe 1996; p.125) proves useful in this regard:
Lemma: Let $u_1 = F(x_1|y)$ and $u_2 = F(x_2|y)$ be conditional distribution functions, and $F(x_1, x_2|y) = C(u_1, u_2; \theta)$, where C is a bivariate copula function with parameters θ, then

$$F(x_1|x_2, y) = h(u_1|u_2; \theta)\,, \text{ where } h(u_1|u_2; \theta) \equiv \frac{\partial C(u_1, u_2; \theta)}{\partial u_2}\,.$$

For $j < t$, application of the lemma to equation (2.2) with $y = \{x_{t-1}, \ldots, x_{j+1}\}$ gives the following recursive relationships:

$$\begin{aligned} u_{t|j} &= F(x_t|x_{t-1}, \ldots, x_j) = h_{t,j}(u_{t|j+1}|u_{j|t-1}; \theta_{t,j})\,, && (2.5) \\ u_{j|t} &= F(x_j|x_t, \ldots, x_{j+1}) = h_{t,j}(u_{j|t-1}|u_{t|j+1}; \theta_{t,j}), && (2.6) \end{aligned}$$

where $h_{t,j}(u_1|u_2; \theta_{t,j}) = \frac{\partial}{\partial u_2} C_{t,j}(u_1, u_2; \theta_{t,j})$ and $C_{t,j}$ is the distribution function corresponding to pair-copula density $c_{t,j}$. We label equation (2.6) a forwards recursion and equation (2.5) a backwards recursion, and from these it can be seen that $u_{t|j}$ and $u_{j|t}$ are functions not only of $\theta_{t,j}$, but also of the parameters of other pair-copulas. The recursions give the following algorithm for the evaluation of the values of $u_{t|j}$ and $u_{j|t}$ employed in equation (2.4):

Algorithm 1

Step (1): For $t = 1, \ldots, T$ set $u_{t|t} = F(x_t)$

Step (2): For $k = 1, \ldots, T-1$ and $i = k+1, \ldots, T$

Backwards Step: $u_{i|i-k} = h_{i,i-k}(u_{i|i-k+1}|u_{i-k|i-1}; \theta_{i,i-k})$

Forwards Step: $u_{i-k|i} = h_{i,i-k}(u_{i-k|i+1}|u_{i|i-k+1}; \theta_{i,i-k})$

Note that Step (2) involves the evaluation of the $T(T-1)/2$ functions $h_{t,j}$, for $j < t$, twice. Table 1 provides analytical expressions of h for some popular bivariate copulas. Figure S1 in the supplementary materials depicts the dependencies between $u_{t|j}, u_{j|t}$ resulting from the recursions in Algorithm 1. As an illustration, Part A of the supplementary materials shows how a Gaussian AR(2) can be decomposed into a D-vine using Gaussian pair-copulas and Gaussian margins.

2.2 Conditional Distributions and Simulation

From equation (2.5) $F(x_t|x_{t-1}, \ldots, x_1) = u_{t|1} = h_{t,1}(u_{t|2}|u_{1|t-1}; \theta_{t,1})$, where $u_{t|2} = F(x_t|x_{t-1}, \ldots, x_2)$ is a function of x_t, but $u_{1|t-1}$ is not. Repeated use of equation (2.5) provides expressions for $u_{t|2}, \ldots, u_{t|t-1}$, and by noting that $u_{t|t} = F(x_t)$, the conditional distribution function can be expressed as

$$F(x_t|x_{t-1}, \ldots, x_1) = h_{t,1} \circ h_{t,2} \circ \ldots h_{t,t-1} \circ F(x_t) \,. \tag{2.7}$$

To evaluate $h_{t,j}(\cdot|u_{j|t-1}, \theta_{t,j})$, for $j = t-1, \ldots, 1$, the values $u_{1|t-1}, \ldots, u_{t-1|t-1}$ also need computing, which can be obtained by running Algorithm 1, but with $T = t$. The expression at equation (2.7) can be used to provide the efficient algorithm below for simulating from D-vine via the method of composition. We simulate T independent uniforms $w_1, \ldots, w_T$, and compute $x_1 = F^{-1}(w_1)$ and $x_t = F^{-1}(w_t|x_{t-1}, \ldots, x_1)$ for $t = 2, \ldots, T$, so that $\boldsymbol{x}$ has the density at equation (2.4) and $\boldsymbol{u}$ that of $c^\dagger$.

Algorithm 2

For $t = 1, \ldots, T$:

Step (1): Generate $w_t \sim \text{Uniform}(0, 1)$

Step (2): If $t = 1$ set $x_1 = F^{-1}(w_1)$, otherwise set $x_t = F^{-1} \circ h^{-1}_{t,t-1} \circ \ldots \circ h^{-1}_{t,1}(w_t)$

Step (3): Set $u_{t|t} = F(x_t)$, and if $t > 1$ compute:

$$u_{t|j} = h_{t,j}(u_{t|j+1}|u_{j|t-1};\theta_{t,j}) \text{ for } j = t-1,\ldots,1$$

$$u_{j|t} = h_{t,j}(u_{j|t-1}|u_{t|j+1};\theta_{t,j}) \text{ for } j = 1,\ldots,t-1$$

The functions $h^{-1}_{t,j}$ are easily computed either analytically or numerically for commonly used copula; see Table 1. Moreover, Algorithm 2 can be adjusted to produce an iterate from the conditional distribution $F(x_T, x_{T-1},\ldots,x_{t_0+1}|x_{t_0},\ldots,x_1)$ simply by skipping Steps (1) and (2) for $t = 1,\ldots,t_0$, but not Step (3). This can be useful in computing forecasts, particularly when the vector is longitudinal as we demonstrate in Section 5. Both Kurowicka & Cooke (2007) and Aas et al. (2009) give algorithms that are equivalent to Algorithm 2, although the former do not provide an expression for the conditional distribution function, while that of the latter is less succinct.

2.3 Longitudinal Data and Pair-Copula Selection

While the decomposition at equation (2.4) provides a flexible representation for time series data generally, we focus here on the longitudinal case. That is, where there are n independent observations $\boldsymbol{x} = \{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n\}$ on a dependent time series vector $\boldsymbol{x}_i = (x_{i,1},\ldots,x_{i,T})$. In the case where the number of pair-copulas, $T(T-1)/2$, is large compared to the number of scalar observations nT, it can prove hard to obtain reliable estimates without imposing strong restrictions, and thus a data-driven method that allows for parsimony is useful. Following Smith & Kohn (2002) we introduce indicator variables $\Gamma = \{\gamma_{t,s}; (t,s) \in \mathcal{I}\}$, so that

$$\begin{aligned} c_{t,s}(u_1,u_2;\theta_{t,s}) &= 1 && \text{iff} \quad \gamma_{t,s} = 0 \\ c_{t,s}(u_1,u_2;\theta_{t,s}) &= c^\star_{t,s}(u_1,u_2;\theta_{t,s}) && \text{iff} \quad \gamma_{t,s} = 1. \end{aligned} \tag{2.8}$$

In the above $c^\star_{t,s}$ is a pre-specified bivariate copula density, such as a Gaussian, t, Gumbel or Clayton. While there is no reason why the pair-copula cannot vary with (t,s), for simplicity we assume $c^\star_{t,s}$ are all of the same form in our empirical work and therefore drop the subscripts of the copula density $c^\star$ and corresponding distribution function $C^\star$.

When $\gamma_{t,s} = 0$ the copula is the independence copula $C_{t,s}(u_1,u_2;\theta_{t,s}) = u_1u_2$, and implies that $h_{t,s}(u_1|u_2;\theta_{t,s}) = u_1$. Therefore, Γ determines the form of the time series dependency. For example, if $\gamma_{t,s} = 0$ for all $s \leq t-p$, then $f(x_t|x_{t-1},\ldots,x_1) = f(x_t|x_{t-1},\ldots,x_{t-p})$ and the process is Markov of order p. In general, Γ determines a parsimonious dependence structure that can vary with time t, extending antedependent models for longitudinal data (Gabriel 1962; Pourahmadi 1999; Smith & Kohn 2002) and covariance selection for Gaussian copulas (Pitt et al. 2006).

The likelihood $f(\boldsymbol{x}|\Theta,\Gamma) = \prod_{i=1}^n f(\boldsymbol{x}_i|\Theta,\Gamma)$, where

$$f(\boldsymbol{x}_i|\Theta,\Gamma) = \prod_{t=2}^{T}\left\{\prod_{j=1}^{t-1}\left\{\left(c^\star(u_{i,t|j+1},u_{i,j|t-1};\theta_{t,j})\right)^{\gamma_{t,j}}\right\} f(x_{i,t})\right\} f(x_{i,1})\,.$$

Here, the conditional copula data $u_{i,t|j+1} = F(x_{i,t}|x_{i,t-1},\ldots,x_{i,j+1})$ and $u_{i,j|t-1} = F(x_{i,j}|x_{i,t},\ldots,x_{i,j+1})$ are computed using Algorithm 1 applied separately to each observation $\boldsymbol{x}_i$. The following adjustment to Step (2) of Algorithm 1 can be employed:

$$h_{i,i-k}(u_1|u_2;\theta_{i,i-k}) = u_1 \quad \text{if} \quad \gamma_{i,i-k} = 0$$

$$h_{i,i-k}(u_1|u_2;\theta_{i,i-k}) = h^\star(u_1|u_2;\theta_{i,i-k}) \equiv \frac{\partial}{\partial u_2}C^\star(u_1,u_2;\theta_{i,i-k}) \quad \text{if} \quad \gamma_{i,i-k} = 1\,.$$

Exploiting this observation substantially increases execution speed when the proportion of zeros in Γ is high; something that is likely to be the case in many longitudinal studies.

When the marginal distributions are Gaussian, the framework nests a wide range of existing longitudinal models. If $c^\star$ is a bivariate Gaussian copula, then the longitudinal vector follows a Gaussian AR(p) when $\gamma_{t,j} = 0$ for $t > p$ and $j > t-p$, while $\gamma_{t,j} = 1$ otherwise. When the Gaussian pair-copula parameters $\{\theta_{t,j}|t>p, j>t-p\}$ vary with t, a time-varying parameter Gaussian autoregression is obtained. If the elements of γ vary, then the model is further extended to an antedependent model. However, by choosing non-Gaussian pair-copula densities $c^\star$, the approach allows for more complex models of nonlinear dependence. We show in our empirical work that this can have a considerable impact.

3 Bayesian Inference

3.1 Priors

The prior on Γ can be chosen to represent a preference for shorter Markov orders by setting the marginal priors $\pi(\gamma_{t,s} = 1) \propto \delta^{(t-s)}$, for $0 < \delta < 1$. Similarly, an informative prior can be used to ensure that $\gamma_{t,s} = 0$ if $\gamma_{t,s-1} = 0$. However, in our empirical work we do neither and place equal marginal prior weight upon each indicator. As observed by Kohn et al. (2001) such a prior can still prove highly informative when $N = T(T-1)/2$ is large. For example, if $K_\Gamma = \sum_{(t,s)\in\mathcal{I}} \gamma_{t,s}$ is the number of non-zero elements of Γ, then assuming the flat prior $\pi(\Gamma) = 2^{-N}$ puts very high prior weight on values of Γ which have $K_\Gamma \approx N/2$. To avoid this, beta priors can be employed (Kohn et al. 2001; Liu et al. 2009), although we adopt the prior

$$\pi(\Gamma) = \frac{1}{N+1} \begin{pmatrix} N \\ K_\Gamma \end{pmatrix}^{-1},$$

which has been used successfully in the component selection literature (Cripps et al., 2005; Panagiotelis & Smith 2008). It results in equal marginal priors, uniform prior weight on $\pi(K_\Gamma) = 1/(1+N)$, and the conditional prior

$$\pi(\gamma_{t,s}|\{\Gamma\backslash\gamma_{t,s}\}) \propto B(N - K_\Gamma + 1, K_\Gamma + 1)/(N+2)\,, \tag{3.1}$$

where $B(x, y)$ is the beta function. The priors of the dependence parameters $\theta_{t,s}$ vary according to choice of copula function $C^\star$. When Gaussian pair-copulas are employed, the $\theta_{t,s}$ are partial correlations, and independent beta priors can be adopted as suggested by Daniels & Pourahmadi (2009) or flat priors as in Pitt et al. (2006). When non-Gaussian pair-copulas are used, following equation (2.3), the parameters $\theta_{t,s}$ capture conditional dependence more generally. Unless mentioned otherwise, we employ independent flat priors on the domain of these dependence parameters. This extends the approaches suggested by Joe (2006) and Daniels & Pourahmadi (2009) for modelling covariance matrices for Gaussian data, and we show that this is an effective strategy for a range of copula functions in our empirical work.

3.2 Sampling Scheme

Given the margins, we generate iterates from the joint posterior $f(\Gamma, \Theta|\boldsymbol{x})$ by introducing latent variables $\tilde{\theta}_{t,s}$, for $(t,s) \in \mathcal{I}$, such that $\theta_{t,s} = \tilde{\theta}_{t,s}$ iff $\gamma_{t,s} = 1$. Following the definition of the indicator variables in equation (2.8), the pair-copula $c_{t,s}$ is known exactly from $(\tilde{\theta}_{t,s}, \gamma_{t,s})$. The prior $\pi(\tilde{\theta}_{t,s}|\gamma_{t,s} = 1) \propto \pi(\theta_{t,s}|\gamma_{t,s} = 1)$, and we assume prior independence between the latent and indicator variables, so that $\pi(\tilde{\Theta}, \Gamma) = \pi(\Gamma) \prod_{(t,s)\in\mathcal{I}} \pi(\tilde{\theta}_{t,s})$, with $\tilde{\Theta} = \{\tilde{\theta}_{t,s}; (t,s) \in \mathcal{I}\}$. We evaluate the posterior distribution using MCMC. The sampling scheme consists of Metropolis-Hastings (MH) steps that traverse the latent and indicator variable space by generating each pair $(\tilde{\theta}_{t,s}, \gamma_{t,s})$, one at a time for $t = 2, \ldots, T$ and $s = 1, \ldots, t-1$.

In the case where the pair-copula has a single dependence parameter, we adopt a MH proposal with density q that is independent in $\gamma_{t,s}$ and $\tilde{\theta}_{t,s}$, so that $q(\tilde{\theta}_{t,s}, \gamma_{t,s}) = q_1(\gamma_{t,s})q_2(\tilde{\theta}_{t,s})$. When there are multiple dependence parameters for a pair-copula, we simply generate each parameter independently in the same manner. Kohn et al. (2001) and Nott & Kohn (2005) compare the relative efficiency of a number of choices for q_1 in the regression variable selection problem. In this paper we consider two choices for q_1. The first corresponds to the simple proposal $q_1(\gamma_{t,s} = 1) = q_1(\gamma_{t,s} = 0) = 1/2$, while the second is Sampling Scheme 2 proposed by Kohn et al. (2001). This was the most computationally efficient scheme suggested by the authors, and employs the conditional prior at equation (3.1) as the proposal. For clarity, we label these two sampling schemes SS1 and SS2, respectively. In both cases we use a random walk proposal for q_2, with $\tilde{\theta}_{t,s}$ generated using a t-distribution with d degrees of freedom and scale σ^2.

Dropping the subscripts for notational convenience, the new iterate $(\tilde{\theta}^{\text{new}}, \gamma^{\text{new}})$ is accepted over the old $(\tilde{\theta}^{\text{old}}, \gamma^{\text{old}})$ with probability $\min(1, \alpha R)$, where R is an adjustment due to any bounds on the domain of θ. We denote the conditional prior at equation (3.1) for the case when $\gamma = 1$ as π_1, and $\pi_0 = 1 - \pi_1$. If the likelihood in Section 2.3 is denoted as a function of the element $(\tilde{\theta}, \gamma)$ as $L(\tilde{\theta}, \gamma)$, then α can be computed for the four different combinations of γ^{old} and γ^{new} as:

$$\alpha_{00} \equiv \alpha\left((\gamma^{\text{old}} = 0, \tilde{\theta}^{\text{old}}) \to (\gamma^{\text{new}} = 0, \tilde{\theta}^{\text{new}})\right) \quad = \quad \frac{\pi(\tilde{\theta}^{\text{new}})}{\pi(\tilde{\theta}^{\text{old}})}$$

$$
\begin{aligned}
\alpha_{01} \equiv \alpha\left((\gamma^{\text{old}}=0,\tilde{\theta}^{\text{old}}) \to (\gamma^{\text{new}}=1,\tilde{\theta}^{\text{new}})\right) &= \frac{L(\tilde{\theta}^{\text{new}},\gamma^{\text{new}}=1)\pi(\tilde{\theta}^{\text{new}})\pi_1}{L(\gamma^{\text{old}}=0)\pi(\tilde{\theta}^{\text{old}})\pi_0} \times \frac{q_1(0)}{q_1(1)} \\
\alpha_{10} \equiv \alpha\left((\gamma^{\text{old}}=1,\tilde{\theta}^{\text{old}}) \to (\gamma^{\text{new}}=0,\tilde{\theta}^{\text{new}})\right) &= \frac{L(\gamma^{\text{new}}=0)\pi(\tilde{\theta}^{\text{new}})\pi_0}{L(\tilde{\theta}^{\text{old}},\gamma^{\text{old}}=1)\pi(\tilde{\theta}^{\text{old}})\pi_1} \times \frac{q_1(1)}{q_1(0)} \\
\alpha_{11} \equiv \alpha\left((\gamma^{\text{old}}=1,\tilde{\theta}^{\text{old}}) \to (\gamma^{\text{new}}=1,\tilde{\theta}^{\text{new}})\right) &= \frac{L(\tilde{\theta}^{\text{new}},\gamma^{\text{new}}=1)\pi(\tilde{\theta}^{\text{new}})}{L(\tilde{\theta}^{\text{old}},\gamma^{\text{old}}=1)\pi(\tilde{\theta}^{\text{old}})}\,.
\end{aligned}
$$

The likelihood L is not a function of $\tilde{\theta}$ when $\gamma = 0$, while $L(\tilde{\theta}, \gamma = 1) = L(\theta, \gamma = 1)$. If the prior for $\tilde{\theta}$ is uniform, as is the case in much of our empirical work, $\pi(\tilde{\theta}^{\text{new}})/\pi(\tilde{\theta}^{\text{old}}) = 1$. If θ is constrained to the domain (a, b), so is $\tilde{\theta}$ and the factor

$$
R = \frac{T_d((b-\tilde{\theta}^{\text{old}})/\sigma) - T_d((a-\tilde{\theta}^{\text{old}})/\sigma)}{T_d((b-\tilde{\theta}^{\text{new}})/\sigma) - T_d((a-\tilde{\theta}^{\text{new}})/\sigma)}\,,
$$

where T_d is the distribution function of a t_d distribution. Note that the likelihood is not computed in the evaluation of α_{00}, and that with this proposal $\alpha_{00} = 1$. Therefore, the more frequently this case arises, the faster the estimation. For SS1, the choice of q_1 further simplifies $\alpha_{01} = L(\tilde{\theta}^{\text{new}}, \gamma^{\text{new}} = 1)\pi_1/L(\gamma^{\text{old}} = 0)\pi_0$, while $\alpha_{10} = L(\gamma^{\text{new}} = 0)\pi_0/L(\tilde{\theta}^{\text{old}}, \gamma^{\text{old}} = 1)\pi_1$. For the choice of q_1 in SS2, $\alpha_{01} = L(\tilde{\theta}^{\text{new}}, \gamma^{\text{new}} = 1)/L(\gamma^{\text{old}} = 0)$ and $\alpha_{10} = L(\gamma^{\text{new}} = 0)/L(\tilde{\theta}^{\text{old}}, \gamma^{\text{old}} = 1)$.

In our empirical work we condition on any marginal estimates and focus on studying inference for the serial dependence structure on $[0, 1]^T$. However, because the likelihood can be computed in closed form, the copula parameters can be estimated joint with any marginal parameters by appending additional MH steps as outlined by Pitt et al. (2006) for Gaussian copula. Nevertheless, joint estimation often does not affect the estimated dependence structure meaningfully; see Silva & Lopes (2008) for an empirical demonstration.

3.3 Posterior Inference

The sampling schemes are run for a burnin period, and then J iterates $\{\Theta^{[j]}, \Gamma^{[j]}\} \sim f(\Theta, \Gamma|\boldsymbol{x})$ collected. From these posterior inference is computed, including posterior means which we use in our empirical work as point estimates. Of particular interest here is $\Pr(\gamma_{t,s} = 0|\boldsymbol{x}) \approx \frac{1}{J}\sum_j (1-\gamma_{t,s}^{[j]})$, which is the estimate of the marginal probability that the (t, s)th pair-copula is the independence copula. We found little difference between SS1 and SS2, and all empirical results are from SS1. We assess convergence by checking our estimates are invariant to different initial conditions, and employ sample sizes between $J = 20,000$ and $J = 50,000$.

Because in our analysis we consider different pair-copula families it is important to measure dependence on a common metric. There is an extensive literature on measures of concordance, with a comprehensive summary given by Nelsen (2006; Chapter 5). We measure the level of dependence in each pair-copula using Kendall's tau. For each pair $(t,s) \in \mathcal{I}$ this can be expressed (Nelsen 2006; p.159) as

$$\tau_{t,s} = 4 \int_0^1 \int_0^1 C_{t,s}(u_1, u_2; \theta_{t,s}) \mathrm{d}C_{t,s}(u_1, u_2; \theta_{t,s}) - 1 \,.$$

For the independence copula $\tau_{t,s} = 0$, while Kendall's tau can be expressed as a function of $\theta_{t,s}$ for many common bivariate copula (Embrechts et al. 2003); for the Gaussian $\tau_{t,s} = \arcsin\left(\frac{2}{\pi}\theta_{t,s}\right)$, for the t $\tau_{t,s} = \arcsin\left(\frac{2}{\pi}\xi_{t,s}\right)$, for the Clayton $\tau_{t,s} = \theta_{t,s}/(\theta_{t,s}+2)$ and for the Gumbel $\tau_{t,s} = 1 - \theta_{t,s}^{-1}$. In these cases we can write Kendall's tau as $\tau_{t,s}(\theta_{t,s})$ and compute

$$E(\tau_{t,s}|\boldsymbol{x}) = \int \tau_{t,s}(\theta_{t,s}) f(\theta_{t,s}|\boldsymbol{x}) \mathrm{d}\theta_{t,s} = \int \tau_{t,s}(\tilde{\theta}_{t,s}) f(\tilde{\theta}_{t,s}, \gamma_{t,s} = 1|\boldsymbol{x}) \mathrm{d}\tilde{\theta}_{t,s} \approx \frac{1}{J} \sum_{j=1}^{J} \tau_{t,s}(\tilde{\theta}_{t,s}^{[j]}) \gamma_{t,s}^{[j]} \,.$$

This shows that the posterior mean is a model average over the indicator $\gamma_{t,s}$.

From equation (2.2), $\tau_{t,s}$ measures dependence between X_t and X_s, conditional upon intermediate values. To obtain a measure of marginal pairwise dependence, we use Spearman's rho $\rho_{s,t}(\Theta) = 12E(U_t U_s|\Theta) - 3$, where $U_t = F(X_t)$ and $U_s = F(X_s)$ (Nelsen 2006, p.170). We calculate its posterior mean by computing a Monte Carlo estimate of $m_{s,t}(\Theta) = E(U_t U_s|\Theta)$, based upon iterates $\{\boldsymbol{u}^{[1]}, \ldots, \boldsymbol{u}^{[J]}\}$ of $\boldsymbol{U} = (U_1, \ldots, U_T)$ simulated using Algorithm 2 appended to the end of each sweep of the sampling scheme. The estimate is

$$E(m_{s,t}|\boldsymbol{x}) = \int m_{s,t}(\Theta) f(\Theta|\boldsymbol{x}) d\Theta \approx \frac{1}{J} \sum_{j=1}^{J} \int u_s u_t f(\boldsymbol{u}|\Theta^{[j]}) \mathrm{d}\boldsymbol{u} \approx \frac{1}{J} \sum_{j=1}^{J} u_s^{[j]} u_t^{[j]} = \hat{m}_{t,s} \,,$$

where $\boldsymbol{u}^{[j]} \sim f(\boldsymbol{u}|\Theta^{[j]})$ and $\Theta^{[j]} \sim f(\Theta|\boldsymbol{x})$. Because Algorithm 2 is fast to implement, we actually generate 100 iterates of $\boldsymbol{U}$ at each sweep to make the estimate $\hat{m}_{t,s}$ more accurate.

3.4 Diagnostic

Currently, there are only a few approaches for judging the adequacy of the fit of a multivariate copula; for example, Glidden (2007) proposes diagnostics based upon pairwise dependence. We consider the sum $S(\Theta) = \sum_{j=1}^{T} \Phi^{-1}(U_j)$, where $\boldsymbol{U}$ has a multivariate copula as a distribution function. The sum S is both highly sensitive to the dependence structure of the copula and comparable across different copulas. From a Bayesian perspective, we consider the fitted distribution $f(S|\boldsymbol{x}) = \int f(S|\Theta)f(\Theta|\boldsymbol{x})\mathrm{d}\Theta$, where the parameters are integrated out with respect to their posterior distribution. For a Gaussian copula, $f(S|\Theta)$ can be shown to be a Gaussian density, and $f(S|\boldsymbol{x}) \approx \frac{1}{J}\sum_{j=1}^{J} f(S|\Theta^{[j]})$ is straightforward to compute. For non-Gaussian D-vines we select every twentieth iterate from $\{\boldsymbol{u}^{[1]}, \ldots, \boldsymbol{u}^{[J]}\}$, which are simulated as in Section 3.3, to obtain an approximately independent sample from the fitted vine. We then compute iterates $S^{[k]} = \sum_{j=1}^{T} \Phi^{-1}(u_j^{[k]})$, which form an approximately independent sample from $f(S|\boldsymbol{x})$ and can be used to construct a kernel density estimate (KDE). The fitted distribution $f(S|\boldsymbol{x})$ can be compared to two different benchmarks. The first is where the elements of $\boldsymbol{X}$ are assumed independent, so that $U_1, \ldots, U_T$ are independent uniforms and $S \sim N(0, T)$. The second is the empirically observed distribution of S. This is given as the KDE based on the sample $S_i = \sum_{t=1}^{T} \Phi^{-1}(u_{i,t}^{\text{obs}})$, for $i = 1, \ldots, n$, where $u_{i,t}^{\text{obs}} = \hat{F}_t(x_{it})$, and $\hat{F}_t(x_{it})$ is the empirical distribution function of the data $\{x_{1,t}, \ldots, x_{n,t}\}$. A parametric model that more adequately fits the observed dependence in the data will have $f(S|\boldsymbol{x})$ closer to this second benchmark distribution.

4 Simulation Study

We study the effectiveness of the approach in estimating the dependence structure on $[0,1]^T$ using a simulation study. We assume the marginal distributions are known, and consider the Gaussian, Clayton and Gumbel bivariate copulas. The latter two are popular Archimedean copulas; see Joe (1997; Chapter 5) for an introduction to their properties. We simulate 100 datasets from each of three models in two cases. The first case is with $T = 7$ margins and $n = 100$ observations on the longitudinal vector, and the second case is with dimension $T = 14$ and $n = 200$. The three models we consider are:

Model A: The dependence structure of a Gaussian AR(1) with autoregressive coefficient 0.85 and unit variance disturbances. In this case $\gamma_{t,t-1} = 1$, $\gamma_{t,s} = 0$ for $t-s > 1$, and $\theta_{t,t-1} = 0.85$. Figure 1(a) depicts the resulting values of $\tau_{t,s}$ for each pair-copula when $T = 7$.

Model B: Clayton pair-copula model with $\gamma_{t,s} = 1$ for all $t - s \leq 2$, and zero otherwise. This corresponds to a second order time-inhomogeneous Markov process, with dependence parameters set so that the values of $\tau_{t,s}$ for each pair-copula are as depicted in Figure 1(b) when $T = 7$.

Model C: Gumbel pair-copula model with dependence parameters set so that the values of $\tau_{t,s}$ for each pair-copula are as depicted in Figure 1(c) when $T = 7$. This specifies a time-inhomogeneous Markov process with varying order.

When $T = 14$ the designs are direct extensions, and Figure S2 in the supplementary materials depicts the values of $\tau_{t,s}$ for each pair-copula. We fit the following estimators to each dataset:

Estimators E1s/E1f: Estimation using the correctly specified pair-copula type and with/without selection, where the latter corresponds to assuming $\gamma_{t,s} = 1$ for $(t, s) \in \mathcal{I}$.

Estimator E2s/E2f: Estimation using an incorrectly specified pair-copula type (Clayton for Model A; Gumbel for Model B; and Gaussian for Model C) and with/without selection.

Estimator E3: Estimation with a Gaussian copula constructed from Gaussian pair-copulas without selection.

Estimator E4: Estimation with a Gaussian copula with flexible correlation matrix based upon the prior and method of Smith & Kohn (2002).

Estimator E3 employs flat priors on the parameters of the Gaussian pair-copulas. Estimator E4 corresponds to fitting a $N(0, \Sigma)$ distribution to the transformed copula data $x^{\star}_{i,j} = \Phi^{-1}(u_{i,j})$, where $u_{i,j} \in [0, 1]$ is the ith observation of the copula data from the jth margin. The covariance selection prior and method outlined in Smith & Kohn (2002) are used, but where the fitted Gaussian copula is based on the posterior mean of the correlation matrix $\text{diag}(\Sigma)^{-1}\,\Sigma\,\text{diag}(\Sigma)^{-1}$. We note that estimators E1f and E3 coincide for Model A, and E2f and E3 coincide for Model C.

Figure 1 provides a summary of the reliability of the pair-copula selection procedure of estimator E1s when $T = 7$. To quantify this, for each pair-copula we compute the mean posterior probability of being dependent over the simulation $\bar{P}_{t,s} = \frac{1}{100}\sum_{i=1}^{100} P_{t,s}(i)$, where $P_{t,s}(i)$ is the posterior probability $\Pr(\gamma_{t,s} = 1|\boldsymbol{x})$ in the ith dataset. Panels (d)-(f) plot these values for all pair-copulas and the three models, showing that the Bayesian selection approach is highly accurate. To confirm this, we also examine the performance of the approach for classification using a simple threshold. For each replicated dataset we classify each pair-copula as being dependent when $\Pr(\gamma_{t,s} = 1|\boldsymbol{x}) > 0.5$, or the independence copula otherwise. Over the three models, two cases, all pair-copulas and all simulation replicates, 99.8% of dependent pair-copulas and 99.5% of independence pair-copulas were correctly classified by estimator E1s. To also show that the method produces reliable estimates of the conditional dependence structure, for each pair-copula we estimate the bias $\hat{b}(\tau_{t,s}) = \frac{1}{100}\sum_{i=1}^{100}\left(\tau_{t,s}(i) - \tau^*_{t,s}\right)$, where $\tau^*_{t,s}$ is the true value and $\tau_{t,s}(i)$ the posterior mean for the ith dataset of Kendall's tau for pair-copula $c_{t,s}$. Figure 1(g)-(i) reports these estimated biases, with most being zero to two decimal places.

Table 2 compares the performance of all the estimators. The top and bottom half contains for the two combinations of dimension and sample size considered, although relative performance of the estimators is the same in both cases. Each column corresponds to a different combination of the six estimators and three models. Estimators E1 to E3 are D-vines, and for these summaries of the estimated Kendall's tau ($\tau_{t,s}$) for each pair-copula are provided. The first summary is the mean absolute bias (MAB), broken down by pair-copula type (dependent or independence) and also over all pair-copulas. Estimator E1 produces the best results throughout, highlighting the importance of the appropriate choice of pair-copula type for $c^\star$. Estimator E1s dominates E1f, suggesting that Bayesian identification of a parsimonious representation of a D-vine can substantially improve the estimation of the conditional dependence structure. Even when an incorrect pair-copula family is chosen, selection can enhance the estimated dependence structures, with E2s dominating E2f.

The second summary is the width $\hat{w}(\tau_{t,s})$ of the 90% posterior probability interval for $\tau_{t,s}$, defined for each dataset as follows. Order the iterates $\{\tau_{t,s}(\theta_{t,s}^{[1]}), \ldots, \tau_{t,s}(\theta_{t,s}^{[J]})\}$ from smallest to largest, and then compute the Monte Carlo estimate of the interval by counting off the lower and upper 5% of the iterates. The mean width, computed across pair-copulas and simulation replicates, is reported over all pair-copulas and also broken down by pair-copula type. Estimator E1s has substantially lower widths than the other estimators, and by at least an order of magnitude for the independence pair-copulas, again suggesting that the selection methodology is working well. For the selection estimators E1s and E2s, the posterior distribution of $\tau_{t,s}$ has a sizable point mass at 0 for many $(t, s) \in \mathcal{I}$. Consequently, posterior intervals of $\tau_{t,s}$ for different probabilities can be indistinguishable, inflating the coverage statistics and making them uninformative, so that we do not report them here.

To assess the accuracy of the estimation of the unconditional dependence structure we also compute the MAB for the marginal pairwise Spearman's correlations $\rho_{s,t}$. For Model A, estimators E1, E3 and E4 all fit Gaussian copulas, which are the correct parametric form for this model. Nevertheless, the flexible estimators E1s and E4 which employ Bayesian selection methodologies provide the best results. However, selection based on the D-vine decomposition is superior to that based on the Cholesky decomposition employed by Smith & Kohn (2002). Moreover, E1s dominates E4 in all three models. The standard errors in Table 2 show that the differences between estimator performance are significant. Overall, the simulation suggests that the selection method works well, and can improve both the estimated conditional and unconditional dependence structure. In every case, selection with the correct pair-copula family substantially out-performs all alternatives. Throughout, the pair-copula model substantially out-performs the common alternative of fitting a Gaussian copula, with or without a flexible correlation matrix.

5 New South Wales Intraday Electricity Load

Modelling and forecasting electricity load at an intraday resolution is an important problem faced by all electricity utilities; see Soares & Medeiros (2008) for a recent overview. When observed intraday, load has both strong periodic behaviour and meteorologically induced variation (Pardo et al. 2002). Numerous models have been proposed for intraday load, but some of the most successful are longitudinal (Cottet & Smith 2003) because they allow all aspects of the model to vary diurnally. We model electricity load in New South Wales (NSW) observed between 2 January 2002 and 2 January 2005 in MegaWatt hours (MWh). The data were used previously by Panagiotelis & Smith (2008), who employ a longitudinal model with multivariate Gaussian disturbances over the day. We also use a longitudinal model, but where the intraday dependence is captured by a more flexible pair-copula formulation.

For every hour ($t = 1, \ldots, 24$) load $L_{i,t}$ on day i is modelled with the marginal regression

$$L_{i,t} = \alpha_t^1 + \alpha_t^2 i + \boldsymbol{\beta}_t' \mathbf{z}_{i,t} + \alpha_t^3 |T_{i,t} - 18.3| + \delta_t L_{i-1,t} + \epsilon_{i,t} \,, \tag{5.1}$$

where $\{\epsilon_{1,t}, \ldots, \epsilon_{n,t}\}$ are t distributed with scale σ_t^2 and degrees of freedom ν_t. The coefficients α_t^1 and α_t^2 measure level and linear time trend, δ_t captures inter-day linear correlation, and $\mathbf{z}_{i,t}$ is a vector containing the 12 seasonal polynomials and 14 day type dummy variables listed in Panagiotelis & Smith (2008). The effect of air temperature[1] $T_{i,t}$ is nonlinear with a minimum at 18.3C (65F), which is a commonly employed functional form in the demand modelling literature (Pardo et al. 2002). Each of the $T = 24$ marginal models is estimated using maximum likelihood. Residual plots show that the regressions remove the strong signal in the load data, and quantile plots indicate that the marginal t distribution in equation (5.1) is appropriate. Figure S3 in the supplementary materials plots the estimates of σ_t and ν_t.

To account for the strong intraday dependence a pair-copula decomposition with $\boldsymbol{x}_i = (L_{i,1}, \ldots, L_{i,24})'$ is used, where the first element corresponds to load at 03:30, which is the approximate time of the overnight low in demand. We first employ Gaussian pair-copulas and selection, which produces strong positive dependencies between load at times t and $t-1$, with $\Pr(\gamma_{t,t-1} = 1|\boldsymbol{x}) \approx 1$ throughout, and $0.51 \leq E(\tau_{t,t-1}|\boldsymbol{x}) \leq 0.78$. The dependence structure is sparse, with $\Pr(\gamma_{t,s} = 1|\boldsymbol{x}) < 0.5$ for 212 of the $N = 276$ pair-copulas, and Figure S4 in the supplementary materials presents the estimated vine in full. We also employ Gumbel pair-copulas with selection. The Gumbel admits only positive dependence, but has proven

[1] The temperature $T_{i,t}$ is ambient air temperature in degrees centigrade at Bankstown airport in western Sydney, which is considered the centroid of demand in NSW by regulators.

particularly successful in modeling the bivariate dependence of financial returns. Figure 2 plots the estimates of $\Pr(\gamma_{t,s} = 1|\boldsymbol{x})$ in panel (a) and $E(\tau_{t,s}|\boldsymbol{x})$ in panel (b). As with the Gaussian pair-copula vine, strong dependencies between loads at time t and $t-1$ are captured, although the dependence structure is more sparse, with $\Pr(\gamma_{t,s} = 1|\boldsymbol{x}) < 0.5$ for 243 of the Gumbel pair-copulas.

To judge the adequacy of different copulas we employ the diagnostic based on the distribution of the sum discussed in Section 3.4. However, as our empirical benchmark we employ the copula data $u_{i,t} = T_\nu((L_{i,t} - \hat{L}_{i,t}; \hat{\sigma}_t)$, where $T_\nu(\cdot; \sigma)$ is a t distribution function with scale σ and ν degrees of freedom, computed over a 210 day long forecast period 3 January to 31 July 2005, so that $i = n, \ldots, (n + 210)$. These are the marginal predictive distributions with parameter values estimated from the in-sample data, but evaluated at the out-of-sample data points. Figure 3 plots KDEs constructed from the thinned Monte Carlo samples of S for vine copulas constructed from Gaussian, Gumbel and Clayton pair-copulas with and without selection. These are the Bayesian estimates of the fitted distribution $f(S|\boldsymbol{x})$. Also plotted is the distribution of S based on an assumption of independence, and that observed empirically over the forecast period. Ignoring the intraday dependence in the data leads to substantial under-statement of future variation in the sum. This translates directly into an under-statement in the variation of future daily total load, a quantity that is also important to electricity utilities. All three pair-copula models improve substantially on this benchmark; however, the Gaussian and Gumbel pair-copula models appear to be more inline with the observed load than the Clayton.

To further compare the three pair-copula models, we undertake a small intraday forecasting trial. We use a daily rolling window over the period 3 to 30 January 2005, and set $T = 12$, so that each margin corresponds observations at two hour intervals. Intraday forecasts are essential for effective system management by electricity utilities. Forecasts for peak periods are typically made at mid-morning, and are much more accurate than those made prior to 09:00 (Cottet & Smith 2003). The forecasts are constructed by evaluating the distribution $F(L_{i,12}, \ldots, L_{i,h+1}|L_{i,h}, \ldots, L_{i,1})$, where Θ is integrated out with respect to $f(\Theta|\boldsymbol{x})$ and $h = 4$ in our work, which corresponds to 09:30. To evaluate this conditional distribution we append Algorithm 2 to the end of the sampling scheme, but skip Steps (1) and (2) for $t \leq h$. Using the predictive means as forecasts, Table 3 reports the mean absolute deviation (MAD) and mean squared error (MSE) for all three vine copulas. The Gumbel pair-copula model produces the most accurate forecasts, and selection improves the forecasts throughout, showing the usefulness of a data-based parsimonious model here. Capturing intraday dependence is important, with all copula models out-performing the marginal models substantially.

The forecasts show there is a substantial difference between vines constructed from different pair-copulas. To further illustrate this, Figure 4 plots contours of the marginal bivariate distributions of load at 09:30 along with load at 11:30, 13:30, 15:30 and 17:30 on 3 January 2005. This is estimated by simulating iterates from the Gumbel pair-copula model with Θ integrated out with respect to $f(\Theta|\boldsymbol{x})$ in a Monte Carlo fashion. The four distributions are highly non-elliptical and pairwise dependence is nonlinear.

6 Discussion

We argue in this paper that pair-copula constructions, and in particular the D-vine, are suitable for the modeling of longitudinal data, where the time-ordering of the data is exploited. This is unlike the graphical models case, where establishing an ordering of the margins is often difficult. Bivariate Archimedean copulas (Genest & Rivest 1993), t copulas (Demarta & McNeil 2005) and skew t copulas (Smith et al. 2010) are promising choices of pair-copula, with the resulting D-vines capturing serial dependence in a more flexible manner than multivariate elliptical copula. Our approach extends the current literature on covariance modeling for longitudinal data from the Gaussian case (Smith & Kohn, 2002; Huang et al. 2006; Levina et al 2008) to a wide range of non-Gaussian situations. Ibragimov & Lentzas (2008) and Domma et al. (2009) both construct time series models using bivariate copulas to capture serial dependence. However, these are first order Markov models, where the resulting multivariate copula are not recognized as D-vines, and inference for higher order Markov models is not considered.

Our method extends the Bayesian selection approach of Pitt et al. (2006) for Gaussian copula, and also of Smith & Kohn (2002) when applied to the correlation matrix of a Gaussian copula. We show in our empirical work that the approach is highly reliable in identifying any parsimony in the conditional dependence structure, and provides more efficient estimates of both conditional and marginal pairwise dependencies when such parsimony exists. The practical benefits are demonstrated by the substantial improvement in forecasts that are obtained in the electricity load example. Bayesian selection is particularly appropriate for computing inference when the data have the potential to exhibit Markovian properties, by allowing exploration of the high dimensional model space. We mention here that because Algorithm 1 has to be run in its entirety to evaluate the likelihood, estimation of the D-vine for higher values of T is computationally burdensome. This is particularly the case when the pair-copula densities and corresponding h functions involve more computations to evaluate. However, if the dependence structure is parsimonious, so that $\gamma_{t,s} = 0$ for many $(t, s) \in \mathcal{I}$, estimation can be substantially faster. Alternative shrinkage methods (Huang et al. 2006; Levina et al. 2008) also have potential for the efficient estimation of vine copula models.

While modelling serial dependence is our objective, we note here that vine copula models have also been used to model cross-sectional dependence (Aas et al. 2009; Min & Czado 2010a; Czado et al. 2009). Recent work by Haff et al. (2010) suggests that vine copulas can account for a wide range of dependence, including that exhibited by elliptical copulas. Our Bayesian selection method can be used to estimate other vines as well. However, in the cross-sectional case the method identifies a parsimonious multivariate dependence structure, without the interpretation that is possible when using a D-vine to model serial dependence.

Acknowledgments

The work of Michael Smith was partially supported by Australian Research Council Grant DP0985505. Claudia Czado and Carlos Almeida gratefully acknowledge the financial support from the Deutsche Forschungsgemeinschaft (Cz 86/1-3: Statistical inference for high dimensional dependence models using pair-copulas). The authors thank three referees and associate editor, all of whom made comments that improved the paper.

7 Supplemental Materials

Part A An illustration of how a Gaussian AR(2) model can be represented as a D-vine.

Part B Second empirical example using cow liveweight data which has a low sample size and uses t pair-copulas.

Part C Additional figures and tables numbered with prefix 'S'.

Part D Data files for the two real data examples.

References

Aas, K., C. Czado, A. Frigessi & H. Bakken, (2009), 'Pair-copula constructions of multiple dependence', *Insurance: Mathematics and Economics*, 44, 2, 182-198.

Bedford, T. & R. Cooke, (2002), 'Vines - a new graphical model for dependent random variables', *Annals of Statistics*, 30, 1031-1068.

Cherubini, U. Luciano, E. & Vecchiato, W., (2004), *Copula methods in finance*, Wiley.

Clayton, D., (1978), 'A model for association in bivariate life tables and its application to epidemiological studies of family tendency in chronic disease incidence', *Biometrika*, 65, 141-151.

Cottet, R., & M. Smith, (2003), 'Bayesian modeling and forecasting of intraday electricity load', *Journal of the American Statistical Association*, 98, 839-849.

Cripps, E., C. Carter & R. Kohn, (2005), 'Variable Selection and Covariance Selection in Multivariate Regression Models', in D. Dey & C. Rao, (eds.), *Handbook of Statistics 25 Bayesian Thinking: Modeling and Computation*, Elsevier: North Holland, 519-552.

Czado, C., (2010), 'Pair-copula constructions of multivariate copulas' in F. Durante, W. Härdle, P. Jaworki & Rychlik, T. (Eds.) *Workshop on Copula Theory and its Applications*, Springer: Dortrecht, to appear.

Czado, C., F. Gärtner & A. Min (2009), 'Analysis of Australian electricity loads using joint Bayesian inference of D-vines with autoregressive margins' in Kurowicka, D. & H. Joe (eds.) *Dependence Modeling - Handbook on Vine Copulae*, to appear.

Daniels, M. & M. Pourahmadi, (2009), 'Modeling covariance matrices via partial autocorrelations', *Journal of Multivariate Analysis*, in press.

Demarta, S. & A. McNeil, (2005), 'The t-copula and related copulas', *International Statistical Review*, 73, 111-129.

Domma, F., S. Giordano, & P. Perri, (2009), 'Statistical Modeling of Temporal Dependence in Financial Data via a Copula Function', *Communications in Statistics*, 38, 703-728.

Embrechts, P., F. Lindskog, A. McNeil, (2003), 'Modelling dependence with copulas and applications to risk management', In: S. Rachev (ed.), *Handbook of Heavy Tailed Distributions in Finance*. North-Holland: Elsevier.

Frees, E. & P. Wang, (2006), 'Copula credibility for aggregate loss models', *Insurance Mathematics & Economics*, 38, 360-373.

Frees, E. & E. Valdez, (1998), 'Understanding Relationships Using Copulas', *North American Actuarial Journal*, 2, 1, 1-25.

Gabriel, K., (1962), 'Ante-dependent analysis of an ordered set of variables', *Ann. Math. Statist.*, 33, 201-212.

Genest, C. & L.-P. Rivest, (1993), 'Statistical Inference Procedures for Bivariate Archimedean Copulas', *Journal of the American Statistical Association*, 88, 1034-1043.

Glidden, D., (2007), 'Pairwise dependence diagnostics for clustered failure-time data', *Biometrika*, 94, 371-385.

Haff, I., K. Aas & A. Frigessi, (2010), 'On the simplified pair-copula construction- Simply useful or too simplistic?', *Journal of Multivariate Analysis*, 101, 1296-1310.

Huang, J., N. Liu, M. Pourahmadi & Liu, L. (2006), 'Covariance matrix selection and estimation via penalised normal likelihood' *Biometrika*, 93, 85-98.

Ibragimov, R. & G. Lentzas, (2008), 'Copulas and Long Memory', Harvard Institute of Economics Research Discussion Paper Number 2160.

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

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

Joe, H., (2006), Generating random correlation matrices based on partial correlations', *Journal of Multivariate Analysis*, 71, 2177-2189.

Kohn, R., M. Smith & C. Chan, (2001), 'Nonparametric regression using linear combinations of basis functions', *Statistics and Computing*, 11, 313-322.

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

Kurowicka, D. & R. Cooke, (2007), 'Sampling algorithms for generating joint uniform distributions using the vine-copula method', *Computational Statistics and Data Analysis*, 51, 2889-2906.

Lambert, P. & F. Vandenhende, (2002), 'A copula-based model for multivariate non-normal longitudinal data: analysis of a dose titration safety study on a new antidepressant', *Statistics in Medicine*, 21, 3197-3217.

Levina, E., A. Rothman & J. Zhu, (2008), 'Sparse Estimation of Large Covariance Matrices via a Nested Lasso Penalty', *Annals of Applied Statistics*, 1, 245-263.

Liu, X., M. Daniels & B. Marcus, (2009), 'Joint Models for the Association of Longitudinal Binary and Continuous Processes with Application to a Smoking Cessation Trial', *Journal of the Americian Statistical Association*, 104, 429-438.

Meester, S. & J. MacKay, (1994), 'A Parametric Model for Cluster Correlated Categorical Data', *Biometrics*, 50, 954-963.

Min, A., & C. Czado, (2010a), 'Bayesian inference for multivariate copulas using pair-copula constructions', *Journal of Financial Econometrics*, in press.

Min, A., & C. Czado, (2010b), 'SCOMDY Models based on Pair-copula Constructions with Application to Exchange Rates', working paper.

Nelsen, R., (2006), *An Introduction to Copulas*, 2nd ed., New York: NY: Springer.

Nott, D. & R. Kohn, (2005), 'Adaptive sampling for Bayesian variable selection', *Biometrika*, 92, 747-763.

Oakes, D., (1989), 'Bivariate Survival Models Induced by Frailties', *Journal of the American Statistical Association*, 84, 487-493.

Panagiotelis, A. & M. Smith, (2008), 'Bayesian identification, selection and estimation of semiparametric functions in high-dimensional additive models', *Journal of Econometrics*, 143, 291-316.

Patton, A., (2006), 'Modelling Asymmetric Exchange Rate Dependence', *International Economic Review*, 47, 527-556.

Pardo, A., V. Meneu & E. Valor, (2002). 'Temperature and seasonality influences on Spanish electricity load', *Energy Economics*, 24, 55-70.

Pitt, M., D. Chan & R. Kohn (2006), 'Efficient Bayesian Inference for Gaussian Copula Regression Models', *Biometrika*, 93, 537-554.

Pourahmadi, M., (1999), 'Joint mean-covariance models with applications to longitudinal data: Uncontrained parameterisation', *Biometrika*, 86, 677-69.

Silva, R. & H. Lopes, (2008), 'Copula, marginal distributions and model selection: a Bayesian note', *Statistics and Computing*, 18, 313-320.

Sklar, A. (1959), 'Fonctions de répartition à n dimensions et leurs marges', *Publications de l'Institut de Statistique de L'Universit de Paris*, 8, 229-231.

Smith, M. & R. Kohn, (2002), 'Parsimonious Covariance Matrix Estimation for Longitudinal Data', *Journal of the American Statistical Association*, 91, 460, 1141-1153.

Smith, M., Q. Gan & R. Kohn, (2010), 'Modeling Dependence using Skew t Copulas: Bayesian Inference and Applications', working paper.

Soares, L. & M. Medeiros, (2008), 'Modeling and forecasting short-term electricity load: A comparison of methods with an application to Brazilian data', *International Journal of Forecasting*, 24, 630-644.

Song, P. (2000), 'Multivariate Dispersion Models Generated from Gaussian Copula', *Scandinavian Journal of Statistics*, 27, 305-320.

Sun, J., E. Frees & M. Rosenberg, (2008), 'Heavy-tailed longitudinal data modeling using copulas', *Insurance Mathematics and Economics*, 42, 817-830.

Gaussian (or Normal) copula ($-1 \leq \theta \leq 1$)		
$C(u_1, u_2; \theta)$	$=$	$\Phi_2\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right)$
$c(u_1, u_2; \theta)$	$=$	$\frac{1}{\sqrt{1-\theta^2}} \exp\left\{-\frac{\theta^2(w_1^2+w_2^2)-2\theta\, w_1\, w_2}{2(1-\theta^2)}\right\}$, where $w_1 = \Phi^{-1}(u_1)$, $w_2 = \Phi^{-1}(u_2)$
$h(u_1 \mid u_2; \theta)$	$=$	$\Phi\left(\left[\Phi^{-1}(u_1) - \theta\, \Phi^{-1}(u_2)\right](1-\theta^2)^{-1/2}\right)$
$h^{-1}(u_1 \mid u_2; \theta)$	$=$	$\Phi\left(\Phi^{-1}(u_1)\,(1-\theta^2)^{1/2} + \theta\, \Phi^{-1}(u_2)\right)$
t copula ($\theta = \{\xi, \nu\}, -1 \leq \xi \leq 1,\ \nu > 0$)		
$C(u_1, u_2; \theta)$	$=$	$T_\nu\left(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2)\right)$
$c(u_1, u_2; \theta)$	$=$	$\frac{\Gamma\left(\frac{\nu+2}{2}\right)\Gamma\left(\frac{\nu}{2}\right)}{\sqrt{1-\xi^2}\left[\Gamma\left(\frac{\nu+1}{2}\right)\right]^2} \left(1+\frac{w_1^2}{\nu}\right)^{\frac{\nu+1}{2}} \left(1+\frac{w_2^2}{\nu}\right)^{\frac{\nu+1}{2}} \left(1+\frac{w_1^2+w_2^2-2\xi w_1 w_2}{\nu(1-\xi^2)}\right)^{-\frac{\nu+2}{2}}$, where $w_1 = t_\nu^{-1}(u_1)$, $w_2 = t_\nu^{-1}(u_2)$
$h(u_1 \mid u_2; \theta)$	$=$	$t_{\nu+1}\left(\left[t_\nu^{-1}(u_1) - \xi\, t_\nu^{-1}(u_2)\right]\left[\frac{\left(\nu+\left(t_\nu^{-1}(u_2)\right)^2\right)\left(1-\xi^2\right)}{\nu+1}\right]^{-1/2}\right)$
$h^{-1}(u_1 \mid u_2; \theta)$	$=$	$t_\nu\left(t_{\nu+1}^{-1}(u_1)\left[\frac{\left(\nu+\left(t_\nu^{-1}(u_2)\right)^2\right)\left(1-\xi^2\right)}{\nu+1}\right]^{1/2} + \xi\, t_\nu^{-1}(u_2)\right)$
Clayton copula ($\theta \in (-1, \infty)\backslash\{0\}$)		
$C(u_1, u_2; \theta)$	$=$	$\max\{\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta}, 0\}$
$c(u_1, u_2; \theta)$	$=$	$\max\left\{(1+\theta)(u_1 u_2)^{-1-\theta}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1/\theta-2}, 0\right\}$
$h(u_1 \mid u_2; \theta)$	$=$	$\max\left\{u_2^{-\theta-1}\left(u_1^{-\theta} + u_2^{-\theta} - 1\right)^{-1-1/\theta}, 0\right\}$
$h^{-1}(u_1 \mid u_2; \theta)$	$=$	$\left(\left[u_1 u_2^{\theta+1}\right]^{-\theta/(\theta+1)} + 1 - u_2^{-\theta}\right)^{-1/\theta}$
Gumbel copula ($\theta > 1$)		
$C(u_1, u_2; \theta)$	$=$	$\exp\left(-[(-\log u_1)^\theta + (-\log u_2)^\theta]^{1/\theta}\right)$
$c(u_1, u_2; \theta)$	$=$	$C(u_1, u_2; \theta)\,(u_1\, u_2)^{-1}((-\log u_1)^\theta + (-\log u_2)^\theta)^{-2+2/\theta}(\log u_1 \log u_2)^{\theta-1}$ $\times\left[1 + (\theta-1)\left((-\log u_1)^\theta + (-\log u_2)^\theta\right)^{-1/\theta}\right]$,
$h(u_1 \mid u_2; \theta)$	$=$	$C(u_1, u_2; \theta)\frac{1}{u_2}(-\log u_2)^{\theta-1}\left[(-\log u_1)^\theta + (-\log u_2)^\theta\right]^{1/\theta-1}$
$h^{-1}(u_1 \mid u_2; \theta)$		Obtained numerically using Newton's method.
Galambos copula ($\theta > 0$)		
$C(u_1, u_2; \theta)$	$=$	$u_1 u_2 \exp\left(-[(-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}]^{-1/\theta}\right)$
$c(u_1, u_2; \theta)$	$=$	$C(u_1, u_2; \theta) u_1^{-1} u_2^{-1}\Big[1 - \left((-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}\right)^{-1-1/\theta}$ $\times\left((-\log u_1)^{-\theta-1} + (-\log u_2)^{-\theta-1}\right)\Big]\left((-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}\right)^{-2-1/\theta}$ $\times(\log u_1 \log u_2)^{-\theta-1}\left(1 + \theta + \left((-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}\right)^{-1/\theta}\right)$
$h(u_1 \mid u_2; \theta)$	$=$	$C(u_1, u_2; \theta) u_1^{-1}\left[1 - \left((-\log u_1)^{-\theta} + (-\log u_2)^{-\theta}\right)^{-1-1/\theta}(-\log u_2)^{-1-\theta}\right]$
$h^{-1}(u_1 \mid u_2; \theta)$		Obtained numerically using Newton's method.

Table 1: Copula distribution functions, densities, $h-$functions and inverses for the Gaussian, t, Clayton, Gumbel and Galambos bivariate copulas. Function $\Phi(\cdot)$ denotes the standard normal distribution function and $t_\nu(\cdot)$ the t distribution function with ν degrees of freedom. The distribution function of a bivariate normal distribution with standard normal margins and correlation θ is denoted as Φ_2, while that of a bivariate t distribution with ν degrees of freedom, zero mean and correlation θ is denoted as T_ν.

Estimator	Model A					Model B						Model C				
Simulation Case 1: $T = 7, n = 100$																
Estimator	E1s	E1f/E3	E2s	E2f	E4	E1s	E1f	E2s	E2f	E3	E4	E1s	E1f	E2s	E2f/E3	E4
MAB($\tau_{t,s}$) $\times 10^2$																
Dependent	2.14	2.39	12.3	13.7	-	1.87	1.88	16.3	15.4	9.48	-	2.31	2.27	5.37	5.46	-
Independent	1.11	4.85	0.36	4.25	-	0.35	3.93	1.86	7.15	8.44	-	0.12	4.34	2.60	5.28	-
Overall	1.40	4.15	3.77	6.94	-	1.15	2.86	9.44	11.5	8.98	-	1.27	3.26	4.05	5.37	-
	(0.05)	*(0.12)*	*(0.53)*	*(0.44)*		*(0.12)*	*(0.14)*	*(0.94)*	*(0.75)*	*(0.46)*		*(0.17)*	*(0.18)*	*(0.39)*	*(0.37)*	
Mean $\hat{w}(\tau_{t,s})$ $\times 10^2$																
Dependent	8.72	9.00	10.9	11.3	-	7.23	7.29	13.3	13.5	10.9	-	8.64	8.70	10.1	10.2	-
Independent	6.35	20.8	1.12	14.9	-	1.46	16.1	3.71	12.8	21.2	-	0.34	9.80	12.4	19.3	-
Overall	7.03	17.4	3.92	13.9	-	4.48	11.5	8.72	13.2	15.8	-	4.69	9.22	11.2	14.5	-
	(0.12)	*(0.55)*	*(0.46)*	*(0.17)*		*(0.47)*	*(0.58)*	*(0.66)*	*(0.45)*	*(0.67)*		*(0.61)*	*(0.47)*	*(0.49)*	*(0.60)*	
MAB($\rho_{t,s}$) $\times 10^2$	2.56	3.84	15.7	14.0	3.24	0.95	1.04	13.6	11.2	3.86	4.23	1.55	2.31	3.02	4.30	5.50
	(0.09)	*(0.12)*	*(0.23)*	*(0.23)*	*(0.16)*	*(0.04)*	*(0.07)*	*(0.44)*	*(0.27)*	*(0.12)*	*(0.20)*	*(0.09)*	*(0.08)*	*(0.20)*	*(0.20)*	*(0.27)*
Simulation Case 2: $T = 14, n = 200$																
	Model A					Model B						Model C				
Estimator	E1s	E1f/E3	E2s	E2f	E4	E1s	E1f	E2s	E2f	E3	E4	E1s	E1f	E2s	E2f/E3	E4
MAB($\tau_{t,s}$) $\times 10^2$																
Dependent	1.42	1.70	12.2	13.8	-	1.43	1.58	15.5	15.5	9.49	-	1.53	1.67	4.73	4.96	-
Independent	0.34	3.60	0.13	2.95	-	0.08	2.78	1.28	4.30	6.07	-	0.03	3.03	0.66	3.73	-
Overall	0.49	3.33	1.85	4.50	-	0.45	2.45	5.20	7.37	7.01	-	0.41	2.68	1.69	4.04	-
	(0.04)	*(0.07)*	*(0.42)*	*(0.39)*		*(0.08)*	*(0.08)*	*(0.80)*	*(0.71)*	*(0.45)*		*(0.09)*	*(0.10)*	*(0.31)*	*(0.27)*	
Mean $\hat{w}(\tau_{t,s})$ $\times 10^2$																
Dependent	6.03	6.30	7.70	8.06	-	5.11	5.26	9.86	9.60	7.89	-	5.83	5.96	6.78	7.02	-
Independent	1.42	14.7	0.40	10.2	-	0.29	11.1	2.02	7.74	15.4	-	0.08	7.02	2.91	14.5	-
Overall	2.08	13.5	1.45	9.86	-	1.61	9.52	4.17	8.25	13.3	-	1.53	6.75	3.89	12.6	-
	(0.17)	*(0.30)*	*(0.26)*	*(0.08)*		*(0.28)*	*(0.32)*	*(0.48)*	*(0.29)*	*(0.40)*		*(0.23)*	*(0.24)*	*(0.30)*	*(0.41)*	
MAB($\rho_{t,s}$) $\times 10^2$	1.51	3.70	13.0	10.7	3.41	0.99	1.25	15.0	12.2	3.33	4.46	1.06	2.17	1.85	4.13	4.93
	(0.04)	*(0.09)*	*(0.13)*	*(0.06)*	*(0.14)*	*(0.09)*	*(0.11)*	*(0.36)*	*(0.22)*	*(0.08)*	*(0.18)*	*(0.03)*	*(0.04)*	*(0.07)*	*(0.08)*	*(0.16)*

Table 2: Summary of simulation results over all 100 replicates, with the first case in the top half of the table, and the second in the bottom half. The columns give the results for each combination of the 3 models and 6 estimators. For Model A, estimators E1f & E3 coincide, as do estimators E2f & E3 for Model C. For estimators E1 to E3 summaries of the estimates of Kendall's tau $\tau_{t,s}$ for the pair-copulas are given, broken down by type (dependent or independence) and overall. The summaries are the mean absolute bias (MAB) and mean posterior interval width $\hat{w}(\tau_{t,s})$. For all estimators the MAB of the estimates of the pairwise unconditional Spearman's rho values $\{\rho_{t,s}; (t,s) \in \mathcal{I}\}$ is given in the last row. Figures in parentheses are standard errors of the sample means immediately above. They are computed over the replicates as the standard error of the sample mean of 100 MAB values, each evaluated over the elements in $\mathcal{I}$.

t (hour)	Gaussian Full	Gaussian Select	Clayton Full	Clayton Select	Gumbel Full	Gumbel Select	Marginal
			Mean Absolute Deviation (MAD)				
5 (11:30)	0.276	0.280	0.337	0.345	0.249	**0.247**	0.424
6 (13:30)	0.255	0.260	0.364	0.369	0.229	**0.221**	0.459
7 (15:30)	0.232	0.233	0.380	0.382	0.220	**0.208**	0.556
8 (17:30)	0.223	0.211	0.361	0.353	0.199	**0.176**	0.812
9 (19:30)	0.152	0.133	0.269	0.263	0.142	**0.122**	0.918
10 (21:30)	0.097	0.091	0.200	0.196	0.107	**0.082**	0.479
11 (23:30)	0.078	0.074	0.138	0.135	0.079	**0.061**	0.374
12 (01:30)	0.077	0.077	0.123	0.121	0.082	**0.066**	0.407
			Mean Squared Error (MSE)				
5 (11:30)	0.125	0.131	0.195	0.200	0.107	**0.106**	0.265
6 (13:30)	0.112	0.116	0.226	0.231	0.098	**0.093**	0.371
7 (15:30)	0.095	0.097	0.265	0.267	0.098	**0.092**	0.479
8 (17:30)	0.081	0.072	0.231	0.224	0.073	**0.062**	0.867
9 (19:30)	0.049	**0.031**	0.138	0.132	0.0405	0.034	1.129
10 (21:30)	0.018	**0.016**	0.093	0.092	0.0275	0.021	0.343
11 (23:30)	0.011	**0.012**	0.043	0.042	0.0141	**0.011**	0.184
12 (01:30)	0.011	0.010	0.027	0.027	0.012	**0.008**	0.216

Table 3: Performance of intraday forecasts for the three pair-copula models with selection (Select) and without selection (Full). Also reported is the performance of forecasts made from the marginal regression models which do not account for any intraday dependence. Forecasts are made at 09:30 for periods ahead at two hour intervals. The best performing model at each forecast period is denoted in bold.

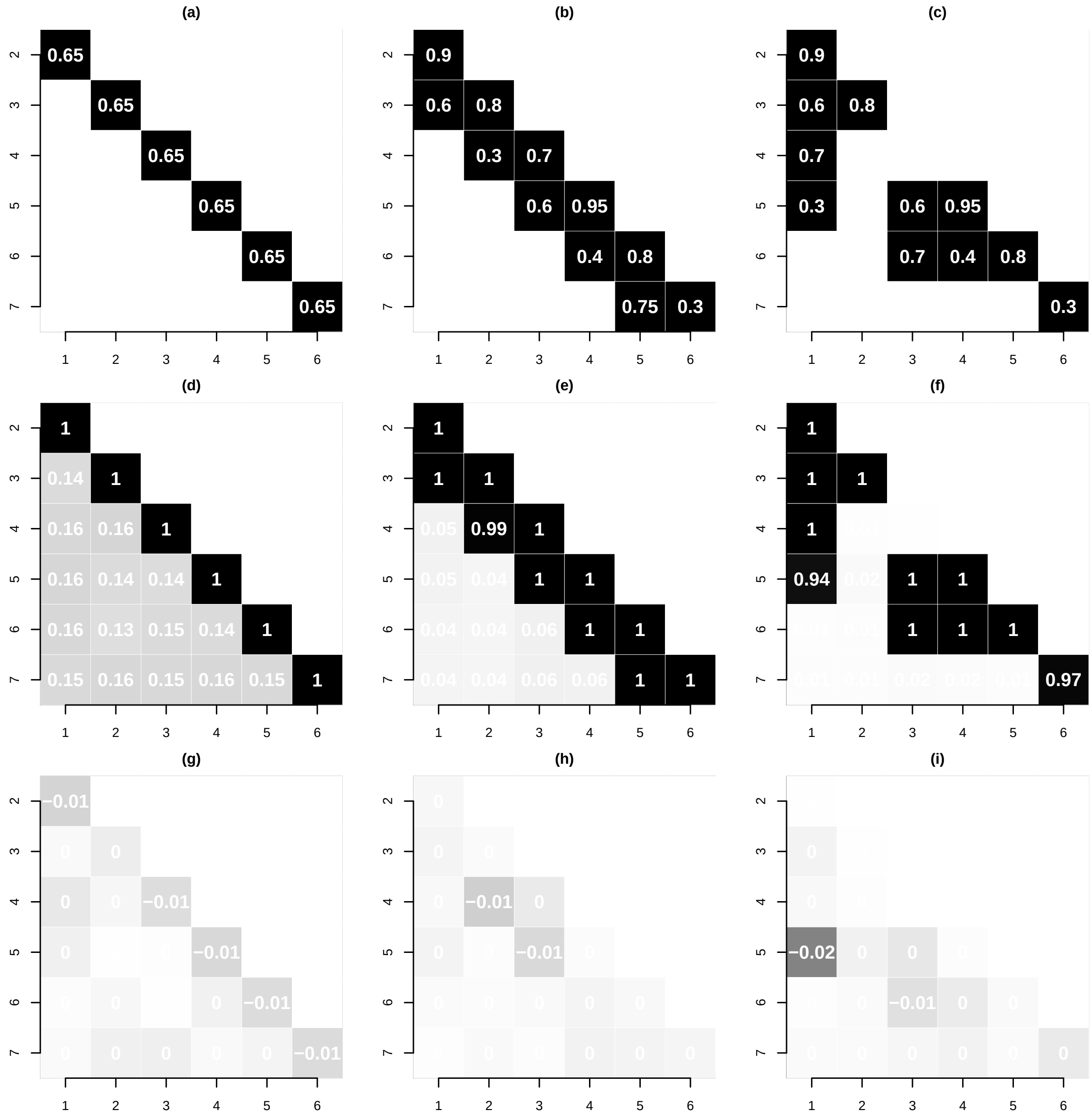

Figure 1: Simulation design when $T = 7$ and results for estimator E1s, with the three columns of panels corresponding to Models A, B and C, respectively. Panels (a)-(c) plot the true values of $\tau_{t,s}$ for the models in each row t and column s of each panel. Panels (d)-(f) plot the values of $\bar{P}_{t,s}$ defined in Section 4 in row t and column s of each panel. Panels (g)-(i) plot the estimated bias values $\hat{b}(\tau_{t,s})$ in row t and column s of each panel. Throughout, darker cells correspond to higher absolute values, lighter cells to lower absolute values and transparent cells to zero.

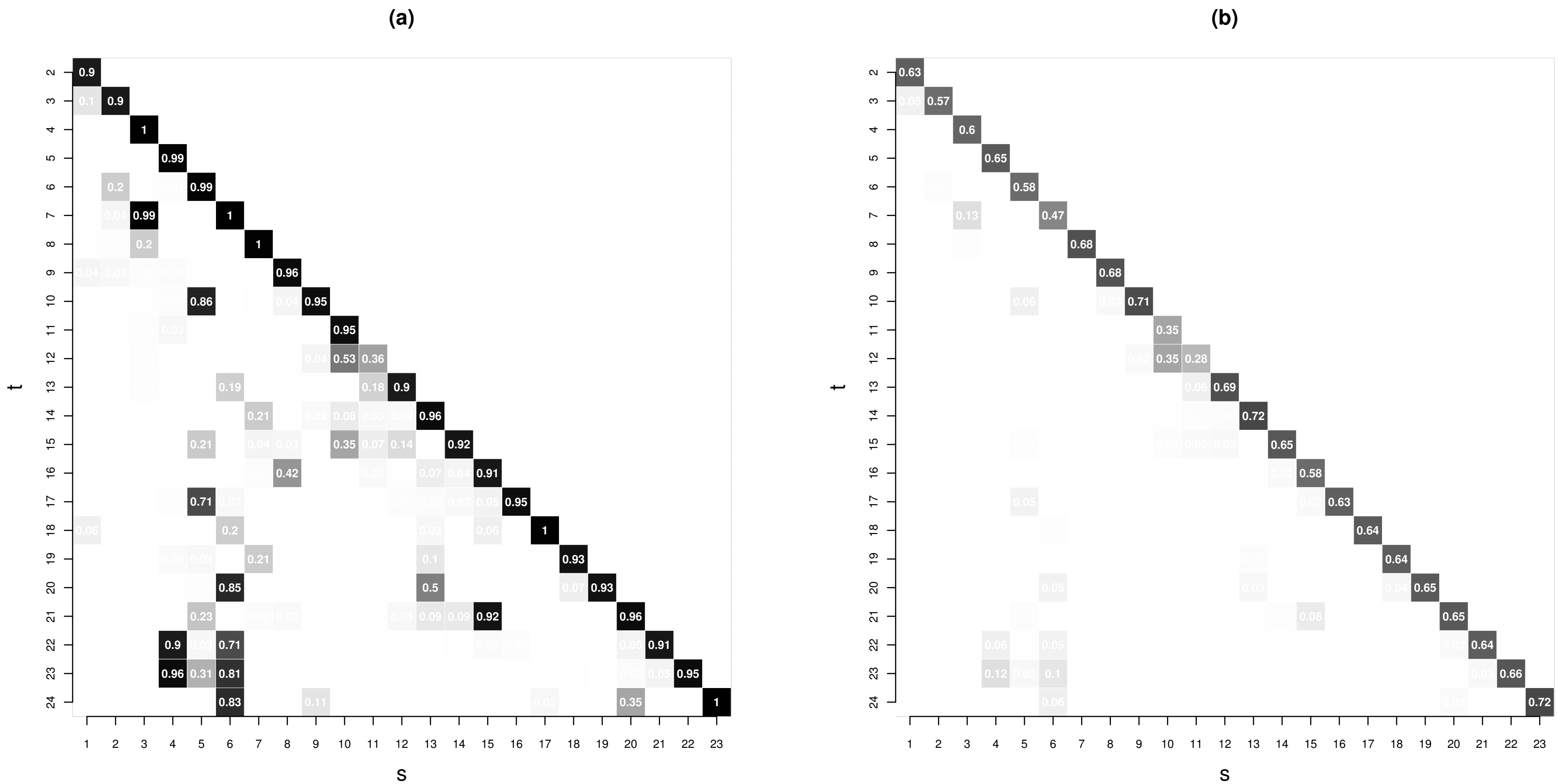

Figure 2: Gumbel pair-copula estimates for the NSW electricity example. Panel (a) depicts the probabilities $\Pr(\gamma_{t,s} = 1|\boldsymbol{x})$, and panel (b) the estimated posterior means $E(\tau_{t,s}|\boldsymbol{x})$, in row t and column s. In both panels, higher absolute values correspond to darker cells and lower absolute values correspond to lighter cells.

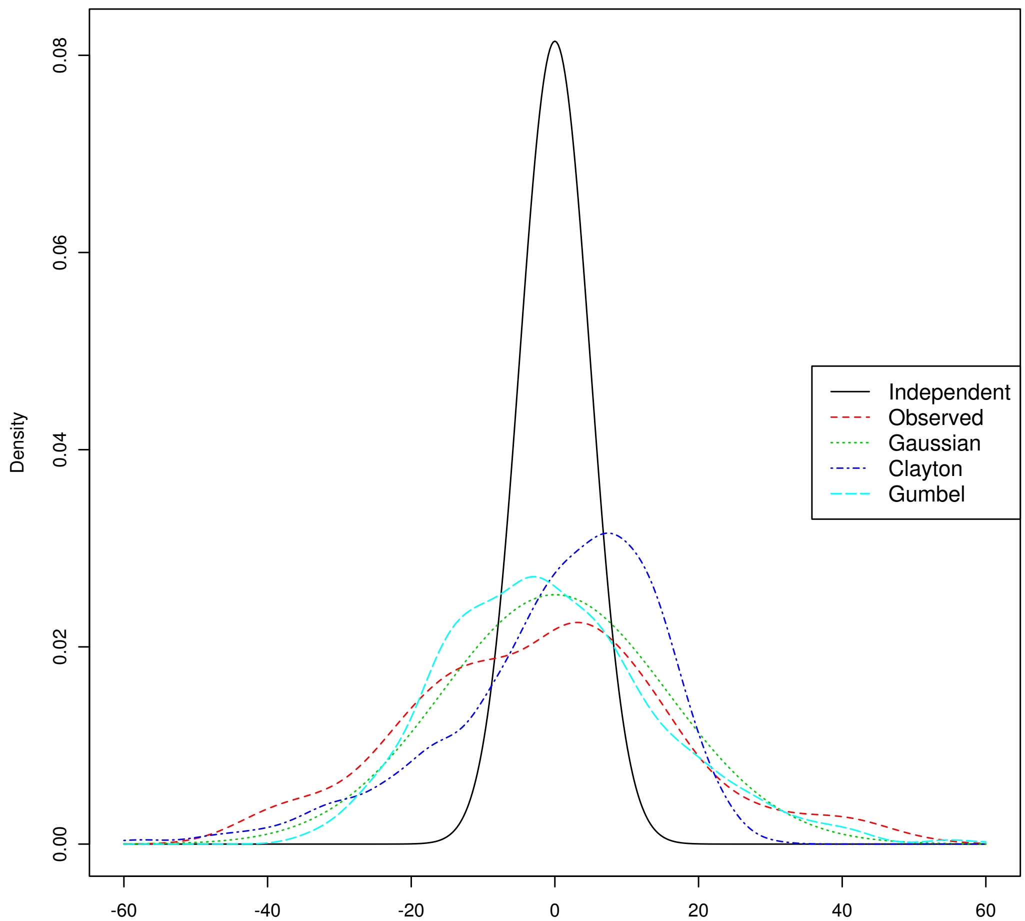

Figure 3: Distributions of diagnostic S for the NSW electricity load example. The solid line corresponds to an assumption of intraday independence, while the distribution of the empirically observed data is given by the dashed (red) line. The distributions corresponding to the three parametric pair-copula models are also shown with line types (colors) as indicated.

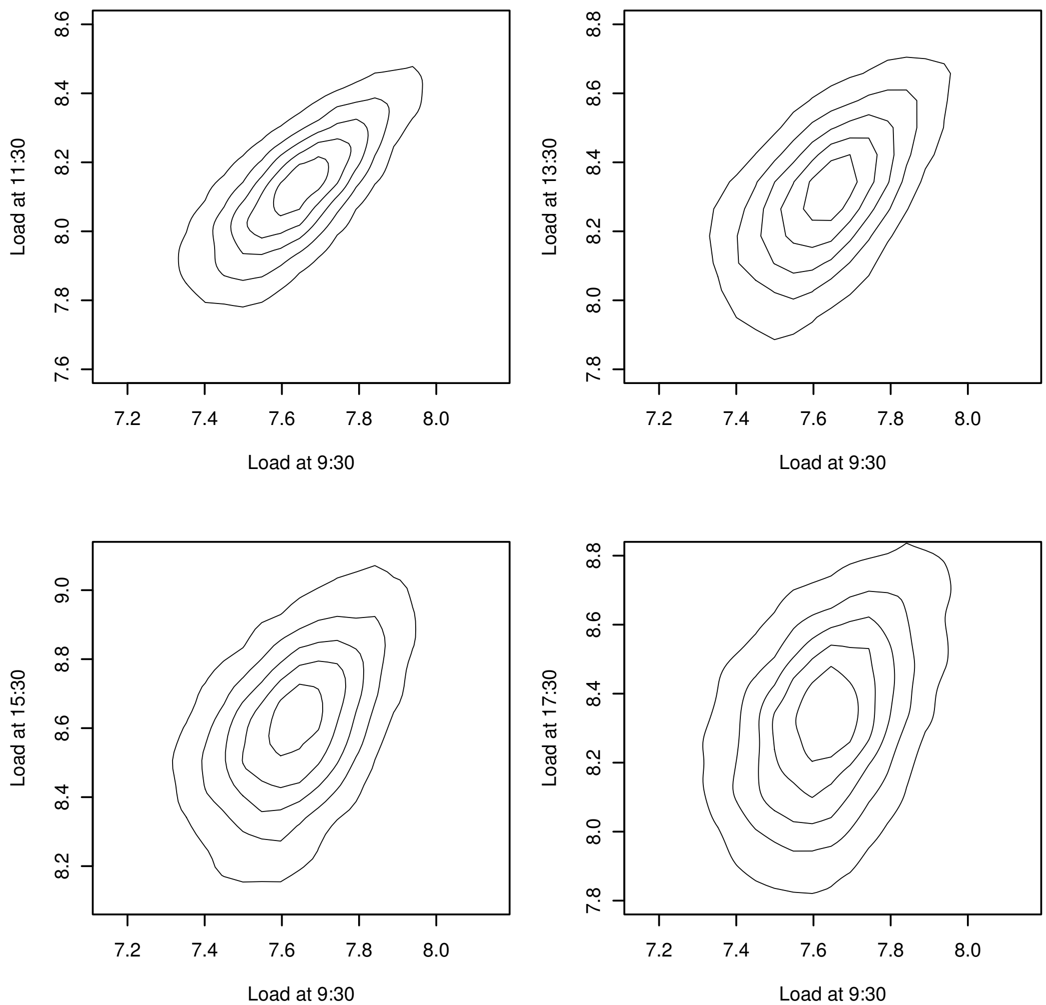

Figure 4: Contour plots of marginal bivariate densities of load (in GigaWatt hours) on 3 January 2005 from the fitted Gumbel pair-copula model. The four panels are for load at (09:30,11:30), (09:30,13:30), (09:30,15:30) and (09:30,17:30). The densities were constructed using KDEs of 19,500 iterates simulated by appending Algorithm 2 to the end of each sweep of the sampling scheme.