International Journal of Forecasting 31 (2015) 815–833

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

# Copula modelling of dependence in multivariate time series

Michael Stanley Smith *

*Melbourne Business School, University of Melbourne, Australia*

ARTICLE INFO

*Keywords:*
Copula model
Nonlinear multivariate time series
Bayesian model averaging
Multivariate stationarity

ABSTRACT

Almost all existing nonlinear multivariate time series models remain linear, conditional on a point in time or latent regime. Here, an alternative is proposed, where nonlinear serial and cross-sectional dependence is captured by a copula model. The copula defines a multivariate time series on the unit cube. A drawable vine copula is employed, along with a factorization which allows the marginal and transitional densities of the time series to be expressed analytically. The factorization also provides for simple conditions under which the series is stationary and/or Markov, as well as being parsimonious. A parallel algorithm for computing the likelihood is proposed, along with a Bayesian approach for computing inference based on model averages over parsimonious representations of the vine copula. The model average estimates are shown to be more accurate in a simulation study. Two five-dimensional time series from the Australian electricity market are examined. In both examples, the fitted copula captures a substantial level of asymmetric tail dependence, both over time and between elements in the series.


## 1. Introduction

Capturing cross-sectional and serial dependence jointly in multiple time series can improve empirical analyses in many applications. These include problems in areas as diverse as political science (Enders & Sandler, 1993), neuroimaging (Goebel, Roebroeck, Kim, & Formisano, 2003) and particularly macroeconomics (Sims, 1980). While the linear vector autoregression (VAR) is the most widely used model for such data, linearity can prove a restrictive assumption, and, increasingly, nonlinear models are preferred. The most popular of these are models with time-varying parameters (e.g. Cogley & Sargent, 2005; Koop, Leon-Gonzalez, & Strachen, 2009; Primiceri, 2005) and/or regime-switching (e.g. Hamilton, 1990; Sola & Driffill, 1994). However, such models remain linear, conditional upon a point in time or latent state. An alternative is to use a copula model. This allows for nonlinear and asymmetric cross-sectional and serial dependence, using a copula function with constant parameters and no latent regimes.

Copula models (Joe, 1997; Nelsen, 2006) are used widely for modelling dependence in cross-sectional data. They are attractive because they separate the modelling of the location and scale from any dependence, the latter of which is modelled on the unit cube. Patton (forthcoming) provides a recent overview of the literature on multivariate time series copula models. However, almost all previous work has used copulas to capture cross-sectional dependence only. For example, Patton (forthcoming) suggests using a copula to capture cross-sectional dependence conditional on the past. For univariate time series, Joe (1997), Beare (2010), Chen and Fan (2006), Domma, Giordano, and Perri (2009), Ibragimov (2009) and Smith, Min, Almeida, and Czado (2010) all consider using copulas to model serial dependence. The aim of this paper is to extend their use to enable the modelling of both cross-sectional and serial dependence in multivariate time series data.

* Correspondence to: Melbourne Business School, 200 Leicester Street, Carlton, VIC, 3053, Australia.
*E-mail address:* mike.smith@mbs.edu.

http://dx.doi.org/10.1016/j.ijforecast.2014.04.003

A common approach to modelling multivariate time series is to select the form of the marginal distribution of the series at a point in time, along with a transitional density, from which the joint distribution of the series is derived. However, it is shown here that the analogous approach of selecting a copula to account for cross-sectional dependence, along with a transitional density on the unit cube, does not provide a well-defined copula model in general. Instead, the reverse approach is proposed, where a copula is selected to account for the dependence in the entire series, then the marginal distribution of the series and the transitional density are derived from this. In doing so, one difficulty is to select a flexible copula that provides closed form expressions for both of these. A copula called a drawable vine (or 'D-vine', see Aas, Czado, Frigessi, & Bakken, 2009; Bedford & Cooke, 2002) is used here to solve this problem.

Vines are copulas for higher dimensions that are constructed from a sequence of nested bivariate copula components called 'pair-copulas'. They are flexible because any combination of bivariate copulas can be used for the pair-copulas; see Haff, Aas, and Frigessi (2010) and Kurowicka and Cooke (2006) for recent overviews. The D-vine also has the property that the copula of any contiguous subvector remains a D-vine, with component pair-copulas that are subsets of those of the original vine. It is shown here how this 'closure' property makes the D-vine attractive for modelling the dependence in multivariate time series. Block functionals, which are equal to the product of subsets of the pair-copulas, are introduced to allow the marginal and transitional densities of the multivariate time series to be expressed concisely.

The block functionals extend the idea of traditional partial correlation matrices to the case of nonlinear time series, and also provide simple conditions under which the series is strongly stationary and/or Markovian. In these latter cases, it is shown that the D-vine is also a parsimonious representation of dependence, with equality between many of the block functionals. Parametric copulas that allow for bi-directional dependence (i.e., both positive and negative associations) should be employed for the pair-copulas. Here, a mixture of rotated Archimedean copulas that allows for bi-directional dependence is proposed. When Gaussian pair-copulas are used, the D-vine can be shown to be a Gaussian copula (Haff et al., 2010). In this special case, the model nests those employed for multivariate time series by Biller and Nelson (2003), Lambert and Vandenhende (2002) and Smith and Vahey (2013).

The dimension of the copula is large at $N = Tm$, where $m$ is the dimension of the multivariate vector and $T$ is the number of observations. For the D-vine copula model, Aas et al. (2009) and Smith et al. (2010) compute the likelihood using an $O(N^2)$ recursive serial algorithm. This can involve too many computations to be employed in practice here. However, we show here how the algorithm can be re-ordered to allow the computations to be undertaken efficiently in parallel, thus greatly speeding execution of the algorithm in real time. When the series is Markov of order $p$, it is also shown how the absolute number of computations can be reduced to $O(Tpm^2)$. In addition, in independent work, Brechmann and Czado (2012) propose the use of D-vines for the modeling of serial dependence in multivariate time series. However, without exploiting these computational insights, they restrict their attention to the bivariate case where $m = 2$. Moreover, they do not employ or exploit the block notation, nor do they consider the properties of the time series, such as stationarity. Rémillard, Papageorgiou, and Soustra (2012) also consider using a copula to model the dependence of a multivariate series on the unit cube. They suggest using either an elliptical copula or a $N$-dimensional Archimedean copula. However, the latter is an unrealistic choice beyond the bivariate case, because it characterises all dependence using a single parameter, nor allows for bi-directional dependence. In comparison, it is shown here that a D-vine copula for a stationary Markov series has a similar number of pair-copulas as there are parameters in a VAR.

A Bayesian approach for estimating the copula model is proposed for stationary series. Bayesian inference has become increasingly popular for multivariate time series models, where point priors and model averaging can improve estimates and predictions; see Garrat, Koop, Mise, and Vahey (2009), George, Sun, and Ni (2008), Jochmann, Koop, Leon-Gonzalez, and Stachan (2013) and Korobilis (2013) for examples. Here, a two-level prior is suggested that places point masses on the blocks for Markov order selection, and also on each pair-copula, so as to be equal to the independence copula for parsimony. The posterior distribution is evaluated using a computationally efficient Markov chain Monte Carlo (MCMC) sampling scheme. Posterior inference on the dependence structure is obtained via simulation from the fitted copula. This includes estimates of both partial and marginal measures of pairwise serial and cross-sectional dependence, predictive distributions and estimates of generalized impulse-response functions.

The effectiveness of the approach is demonstrated using a simulation study. This shows that the selection method can improve the accuracy of the estimates of both the conditional and unconditional dependence structures. To illustrate that the methodology is effective in practice, two five-dimensional time series are examined and a multivariate analysis is found to be beneficial. One is the daily maxima of electricity demand in the five regions of the Australian National Electricity Market (NEM), and the other is daily spot prices of electricity in the same five regions. In both examples, the copula model captures significant levels of asymmetric and heavy-tailed dependence. It is well known that accounting for such properties in cross-sectional dependence can be beneficial (Patton, 2006), and the examples here demonstrate that this is also the case for multivariate serial dependence. The first example extends copula models for multivariate extremes (e.g. Favre, El Adlouni, Perreault, Theiémonge, & Bobée, 2004) to also account for serial dependence. The second demonstrates the flexibility of the copula model when applied to a series that economic theory suggests may exhibit extensive nonlinear dependence.

The rest of the paper is organized as follows. Section 2 outlines the copula model for multivariate time series, including block notation and the conditions for the series to be stationary and/or Markov. Estimation using fast parallel algorithms for computing the likelihood, priors for model averaging and the MCMC scheme are discussed in Section 3. Section 4 contains the simulation study, and Section 5 the empirical applications, after which Section 6 concludes. The Appendix provides an algorithm for simulating iterates from the multivariate time series efficiently.

## 2. The copula model

### 2.1. The basic idea

Consider a continuous-valued random vector $\boldsymbol{Y}_t = (Y_{1,t}, \ldots, Y_{m,t})'$, observed at times $t = 1, \ldots, T$, with the corresponding realisation $\boldsymbol{y}_t = (y_{1,t}, \ldots, y_{m,t})'$. Re-order the elements into the univariate series $\boldsymbol{X} = (X_1, \ldots, X_N) = (\boldsymbol{Y}_1', \ldots, \boldsymbol{Y}_T')$, with $N = Tm$, and realisations $x_i = y_{j,t}$, such that

$$\boldsymbol{y}_t' = (x_{a(t)}, \ldots, x_{b(t)}), \tag{1}$$

where $a(t) = (t-1)m + 1$ and $b(t) = tm$. Then, by Sklar's Theorem (Nelsen, 2006, p. 45), there always exists a copula function $C : [0, 1]^N \to [0, 1]$ such that the joint density of $\boldsymbol{Y} = (\boldsymbol{Y}_1', \ldots, \boldsymbol{Y}_T')'$ is

$$f(\boldsymbol{y}) = c(\boldsymbol{u}) \prod_{t=1}^{T} \prod_{j=1}^{m} f(y_{j,t}).$$

Here, $\boldsymbol{u}' = (F(x_1), \ldots, F(x_N)) = (u_1, \ldots, u_N), f(y_{j,t})$ denotes the marginal density of $y_{j,t}$, $F(x_i) = F(y_{j,t})$ is the corresponding distribution function, and $c(\boldsymbol{u}) = \frac{\partial}{\partial \boldsymbol{u}} C(\boldsymbol{u})$ is the density function of $\boldsymbol{U} = (U_1, \ldots, U_N)' = (F(X_1), \ldots, F(X_N))'$ on $[0, 1]^N$ with uniform margins, often called the copula density. All dependence is captured by the copula and the problem of modelling the dependence of $\boldsymbol{Y}$ is therefore one of modelling the dependence of the random vector $\boldsymbol{U}$.

If $\boldsymbol{U}_t = (U_{a(t)}, \ldots, U_{b(t)})'$, then for $s < t$, the joint density of $\boldsymbol{U}_s, \ldots, \boldsymbol{U}_t$ is

$$c_{(s,\ldots,t)}(\boldsymbol{u}_s, \ldots, \boldsymbol{u}_t) = \int \cdots \int c(\boldsymbol{u}) \mathrm{d}\boldsymbol{u}_1, \ldots, \mathrm{d}\boldsymbol{u}_{s-1}, \mathrm{d}\boldsymbol{u}_{t+1}, \ldots, \mathrm{d}\boldsymbol{u}_T,$$

and is that of a copula because the margin of each element $U_{a(s)}, \ldots, U_{b(t)}$ is uniform. Throughout this paper, '$c$' and '$C$' exclusively denote densities and distribution functions for copulas. The densities of the marginal distribution of $\boldsymbol{Y}_t$, and the conditional distributions $\boldsymbol{Y}_t | \boldsymbol{Y}_{t-1}, \ldots, \boldsymbol{Y}_s$ and $\boldsymbol{U}_t | \boldsymbol{U}_{t-1}, \ldots, \boldsymbol{U}_s$, are

$$f(\boldsymbol{y}_t) = c_{(t)}(\boldsymbol{u}_t) \prod_{j=1}^{m} f(y_{j,t}),$$

$$f(\boldsymbol{y}_t | \boldsymbol{y}_{t-1}, \ldots, \boldsymbol{y}_s) = f(\boldsymbol{u}_t | \boldsymbol{u}_{t-1}, \ldots, \boldsymbol{u}_s) \prod_{j=1}^{m} f(y_{j,t}),$$

and

$$f(\boldsymbol{u}_t | \boldsymbol{u}_{t-1}, \ldots, \boldsymbol{u}_s) = c_{(s,\ldots,t)}(\boldsymbol{u}_s, \ldots, \boldsymbol{u}_t) / c_{(s,\ldots,t-1)}(\boldsymbol{u}_s, \ldots, \boldsymbol{u}_{t-1}). \tag{2}$$

The latter density is the transitional density of the series $\{\boldsymbol{U}_t\}_{t=1}^T$, and, while it is on the unit cube, it is not that of a copula, because the marginals are not uniform; see also Rémillard et al. (2012, Eqn. 4).

When modelling multivariate time series, it is common to begin by selecting a form for the transitional distribution of the series, then the joint and marginal distributions of the observations in the series are derived on the basis of this. However, an analogous approach is difficult here because of the constraint that the transition density in Eq. (2) must factor into the ratio of $c_{(s,\ldots,t)}$ and $c_{(s,\ldots,t-1)}$. This means that the initially appealing approach of selecting a copula $c_{(t)}$ to capture the cross-sectional dependence, along with a transitional density on the unit cube, does not generally produce a valid copula model. For this reason, it is necessary to approach the problem in the reverse direction: first select the copula of the joint distribution of the entire series, then derive the terms at Eq. (2) from this. Note that this differs from the approach of Patton (forthcoming), who adopts a copula decomposition of cross-sectional dependence, conditional on the past. In his approach, serial dependence is introduced by separate univariate transitional densities in each series which are unrelated to the copula.

In the copula modelling literature, a parametric copula function $C(\boldsymbol{u}; \boldsymbol{\phi})$, with parameters $\boldsymbol{\phi}$, is usually selected. This determines the contemporaneous and serial dependence structure of the multivariate time series. Biller (2009) suggests using a Gaussian copula (Song, 2000) or a canonical vine (Kurowicka & Cooke, 2006) for $C$, and Rémillard et al. (2012) an elliptical or an Archimedean copula. Here, a D-vine copula is used for $C$, which has a number of advantages. First, it provides closed form expressions for the densities in Eq. (2), each of which have dependence parameters that are subsets of $\boldsymbol{\phi}$. This allows cross-sectional and serial dependence to be both isolated and modelled separately. Second, the D-vine provides a parsimonious representation for both stationary and Markov multivariate time series. Third, any mix of bivariate copulas can be used for the pair-copulas, resulting in a flexible copula $C$ that can capture a wide range of dependence structures. Last, as is shown in Section 3, even the high-dimensional D-vine copulas employed here can be estimated efficiently using Bayesian methods, and predictive inference obtained. As such, the copula model allows for nonlinear dependence in a flexible, but amenable, fashion.

### 2.2. The D-vine copula

Vine copulas have proven popular choices for the flexible modelling of multivariate dependence in numerous situations. Important early work includes that of Aas et al. (2009), Bedford and Cooke (2002), Joe (1996), Haff et al. (2010) and Min and Czado (2010), while a recent overview was provided by Kurowicka and Joe (2010). Vine copulas have densities that are a product of bivariate copula densities on $[0, 1]^2$, known as 'pair' or 'linking' copulas. Parametric copulas are usually used for these, which are written as $c_{i,j}(v_1, v_2; \phi_{i,j})$, with $\phi_{i,j}$ being the parameters. These are usually scalars, but can be vectors for some choices of bivariate copulas, such as the BB series of Joe (1997), the $t$ copula (Demarta & McNeil, 2005), or the bidirectional Archimedean copulas discussed later. The density of the drawable (or 'D') vine copula is given by the product of $N(N-1)/2$ pair-copulas

$$c_{DV}(\boldsymbol{u};\boldsymbol{\phi}) = \prod_{i=2}^{N} f(u_i|u_{i-1},\ldots,u_1) = \prod_{i=2}^{N}\prod_{j=1}^{i-1} c_{i,j}(u_{i|j+1}, u_{j|i-1};\phi_{i,j}), \quad (3)$$

where the parameter vector $\boldsymbol{\phi} = \{\phi_{i,j};\ i=2,\ldots,N,\ j<i\}$. Here, $u_{i|i} \equiv u_i$, and for $i > j$, $u_{i|j+1} \equiv F(u_i|u_{i-1},\ldots,u_{j+1})$ and $u_{j|i-1} \equiv F(u_j|u_{i-1},\ldots,u_{j+1})$ are the distribution functions of $U_i$ and $U_j$, given the intervening variables $\{U_{i-1},\ldots,U_{j+1}\}$, respectively. These arguments can be computed using the $O(N^2)$ recursive given in Aas et al. (2009) and Smith et al. (2010).

Crucially, any contiguous sub-vector $(U_d,\ldots,U_e)$ of $\boldsymbol{U}$ has a density that is also a D-vine with pair-copula components that are a subset of those in Eq. (3), so that

$$c_{DV}(u_d,\ldots,u_e;\boldsymbol{\phi}^\star) = \prod_{i=d+1}^{e}\prod_{j=d}^{i-1} c_{i,j}(u_{i|j+1},u_{j|i-1};\phi_{i,j}), \quad (4)$$

with $\boldsymbol{\phi}^\star \subset \boldsymbol{\phi}$. As discussed below, it is this closure property that makes the D-vine particularly attractive for modelling the dependence in multivariate time series. When $X_1,\ldots,X_N$ is a univariate time series, Smith et al. (2010) point out that a D-vine is an attractive representation of the serial dependence structure. One aim of this paper is show that it also provides an attractive representation of dependence when $X_1,\ldots,X_N$ are the values of a multivariate time series, re-ordered as in Section 2.1.

### 2.3. Block copula representation and Markov order

The component pair-copula densities in the D-vine density in Eq. (3) can be grouped together into 'blocks' of pair-copulas. These can then be used to isolate cross-sectional and serial dependence in the multivariate time series. For $k_1 > k_2$, define the functionals

$$K_{k_1,k_2}(\boldsymbol{u}_{k_2},\ldots,\boldsymbol{u}_{k_1};\boldsymbol{\phi}_{(k_1,k_2)}) = \prod_{i=a(k_1)}^{b(k_1)}\prod_{j=a(k_2)}^{b(k_2)} c_{i,j}(u_{i|j+1},u_{j|i-1};\phi_{i,j}),$$

and

$$K_{k,k}(\boldsymbol{u}_k;\boldsymbol{\phi}_{(k)}) = \prod_{i=a(k)+1}^{b(k)}\prod_{j=a(k)}^{i-1} c_{i,j}(u_{i|j+1},u_{j|i-1};\phi_{i,j}), \quad (5)$$

where the pair-copula parameters are grouped together similarly as

$$\boldsymbol{\phi}_{(k_1,k_2)} = \{\phi_{i,j};\ a(k_1)\le i\le b(k_1),\ a(k_2)\le j\le b(k_2)\},$$

and

$$\boldsymbol{\phi}_{(k)} = \{\phi_{i,j};\ a(k) < i \le b(k),\ a(k)\le j < i\}.$$

The block $K_{k_1,k_2}$ is a function of $(\boldsymbol{u}_{k_2},\ldots,\boldsymbol{u}_{k_1})$, and $K_{k,k}$ a function of $\boldsymbol{u}_k$, through the arguments of the component pair-copulas, as is discussed in Section 3.1.

Using this block notation, the marginal copula density at time $t$ is $c_{(t)}(\boldsymbol{u}_t;\boldsymbol{\phi}_{(t)}) = K_{t,t}(\boldsymbol{u}_t;\boldsymbol{\phi}_{(t)})$, which is also a density of a D-vine copula with $m(m-1)/2$ pair-copula components. For $t > s$, the copula density

$$c_{(s,\ldots,t)}(\boldsymbol{u}_s,\ldots,\boldsymbol{u}_t) = \prod_{k_1=s}^{t}\prod_{k_2=s}^{k_1} K_{k_1,k_2}(\boldsymbol{u}_{k_2},\ldots,\boldsymbol{u}_{k_1};\boldsymbol{\phi}_{(k_1,k_2)}), \quad (6)$$

which can be seen by matching the pair-copula terms of the blocks with those in Eq. (4) when $d = a(s)$ and $e = b(t)$. Again, this is a density of a D-vine copula, the component blocks of which can be arranged in a lower triangular matrix as in Table 1. When $s = 1$ and $t = T$, Eq. (6) provides the block representation of the copula density $c_{DV}(\boldsymbol{u})$ of the entire series. For $t > s$, the conditional density

$$f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1},\ldots,\boldsymbol{u}_s) = \prod_{k=s}^{t} K_{t,k}(\boldsymbol{u}_k,\ldots,\boldsymbol{u}_t;\boldsymbol{\phi}_{(t,k)}), \quad (7)$$

which is obtained by considering the representation in Eq. (6) for the numerator and denominator in the ratio $c_{(s,\ldots,t)}/c_{(s,\ldots,t-1)}$. Using this expression, the multivariate series is Markov of order $p$ at time $t$, if and only if $K_{t,1} = \cdots = K_{t,t-p-1} = 1$; that is, when the pair-copula densities $c_{i,j}$ that make up these block functionals are all equal to 1. Therefore, the blocks determine the Markov order of the multivariate time series.

Last, the blocks generalize the partial autocorrelation matrices discussed by Tiao and Box (1981) in linear multivariate time series. For example, for $s < t$, consider the density of $(\boldsymbol{Y}_s,\boldsymbol{Y}_t)$ given the intervening values:

$$f(\boldsymbol{y}_s,\boldsymbol{y}_t|\boldsymbol{y}_{s+1},\ldots,\boldsymbol{y}_{t-1}) = \frac{c_{(s,\ldots,t)}(\boldsymbol{u}_s,\ldots,\boldsymbol{u}_t)\prod_{k=s}^{t}\prod_{j=1}^{m} f(y_{j,k})}{c_{(s+1,\ldots,t-1)}(\boldsymbol{u}_{s+1},\ldots,\boldsymbol{u}_{t-1})\prod_{k=s+1}^{t-1}\prod_{j=1}^{m} f(y_{j,k})} = K_{t,s}\left(\prod_{k=s}^{t-1}K_{k,s}\right)\left(\prod_{k=s+1}^{t}K_{t,k}\right)\prod_{j=1}^{m} f(y_{j,s})f(y_{j,t}),$$

which is derived from Eq. (6). If the Markov order is $p < t-s$, then $K_{t,s} = 1$ and

$$f(\boldsymbol{u}_s,\boldsymbol{u}_t|\boldsymbol{u}_{s+1},\ldots,\boldsymbol{u}_{t-1}) = \left(\prod_{k=s}^{t-1}K_{k,s}\right)\left(\prod_{k=s+1}^{t}K_{t,k}\right)$$

is separable with respect to $\boldsymbol{u}_s$ and $\boldsymbol{u}_t$. Therefore, the density above is also separable with respect to $\boldsymbol{y}_s$ and $\boldsymbol{y}_t$, so that $\boldsymbol{Y}_s$ and $\boldsymbol{Y}_t$ are independent, conditional on the intervening values $\boldsymbol{Y}_{s+1},\ldots,\boldsymbol{Y}_{t-1}$.

### 2.4. Block representation of stationary series

Multivariate stationary series can be represented using the block notation in a simple manner. Stationarity here refers to *strong stationarity*, where $f(\boldsymbol{y}_{t-k},\ldots,\boldsymbol{y}_t) = f(\boldsymbol{y}_{t'-k},\ldots,\boldsymbol{y}_{t'})$ for all $t, t'$ and $k \ge 0$, rather than the covariance (or weak) stationarity that is common in linear time series modelling. Note also that this is not the same as assuming that the univariate series $\{X_1,\ldots,X_N\}$ is stationary.

**Table 1**
Blocks of pair-copula terms that occur in the copula $c_{(s,\dots,t)}$ of the joint distribution of $(\boldsymbol{U}_s,\dots,\boldsymbol{U}_t)$, where each random vector $\boldsymbol{U}_t=(F(X_{a(t)}),\dots,F(X_{b(t)}))$.

| $k_1$ | $k_2$ | | | | | |
|---|---|---|---|---|---|---|
| | $s$ | $s+1$ | $s+2$ | $\cdots$ | $t-1$ | $t$ |
| $s$ | $K_{s,s}$ | | | | | |
| $s+1$ | $K_{s+1,s}$ | $K_{s+1,s+1}$ | | | | |
| $s+2$ | $K_{s+2,s}$ | $K_{s+2,s+1}$ | $K_{s+2,s+2}$ | | | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | | |
| $t-1$ | $K_{t-1,s}$ | $K_{t-1,s+1}$ | $K_{t-1,s+2}$ | $\cdots$ | $K_{t-1,t-1}$ | |
| $t$ | $K_{t,s}$ | $K_{t,s+1}$ | $K_{t,s+2}$ | $\cdots$ | $K_{t,t-1}$ | $K_{t,t}$ |

Note: if at time $t$ the process is also Markov of order $p$, then $K_{t,k}=1$ for $t-k>p$.

**Theorem** (*Strong Stationarity*). *If the series $\{\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_T\}$ has a copula decomposition $f(\boldsymbol{y})=c_{DV}(\boldsymbol{u})\prod_{t=1}^{T}\prod_{j=1}^{m}f(y_{j,t})$, with $c_{DV}$ being the D-vine copula denoted in block form as in Section 2.3, then:*

(i) *The series $\{\boldsymbol{U}_1,\dots,\boldsymbol{U}_T\}$, where $\boldsymbol{U}_t=(F(Y_{1,t}),\dots,F(Y_{m,t}))'$, is strongly stationary, with density $c_{(t-k,\dots,t)}(\boldsymbol{u}_{t-k},\dots,\boldsymbol{u}_t)=c_{(t'-k,\dots,t')}(\boldsymbol{u}_{t'-k},\dots,\boldsymbol{u}_{t'})$, if for all $t,t'$ and $k\ge 0$, the functionals*

$$K_{t,t-k}(\cdot;\boldsymbol{\phi}_{(t,t-k)})=K_{t',t'-k}(\cdot;\boldsymbol{\phi}_{(t',t'-k)}).$$

(ii) *The series $\{\boldsymbol{Y}_1,\dots,\boldsymbol{Y}_T\}$ is strongly stationary, if the series $\{\boldsymbol{U}_1,\dots,\boldsymbol{U}_T\}$ is strongly stationary, and also the densities $f(y_{j,t})=f(y_{j,t'})$, for each $j=1,\dots,m$, and all $t,t'$.*

**Proof.** Part (i) can be seen from the block copula representation of $c_{(t-k,\dots,t)}$ in Eq. (6). Part (ii) follows immediately from the copula decomposition

$$f(\boldsymbol{y}_{t-k},\dots,\boldsymbol{y}_t)=c_{(t-k,\dots,t)}(\boldsymbol{u}_{t-k},\dots,\boldsymbol{u}_t)\times\prod_{s=t-k}^{t}\prod_{j=1}^{m}f(y_{j,s}).\quad\square$$

In part (i) above, the block functionals are equal if the component pair-copulas of the blocks are also equal, thus greatly reducing the number of unique pair-copulas in the vine. In light of this theorem, the notation can be simplified for a stationary series as follows. Denote the unique block functionals $K_{t,t-k}$ as $K_k$, because the functionals do not vary with $t$. Similarly, the associated unique parameters $\boldsymbol{\phi}_{(t,t-k)}$ are denoted as $\boldsymbol{\phi}(k)=\{\phi_{l_1,l_2}(k)\}$, and the densities of the component pair-copulas $c_{i,j}$ as $c^{(k)}_{l_1,l_2}$. The superscript of the pair-copula density denotes the block with which it is associated, while the indices $l_1=i-m(t-1)$ and $l_2=j-m(t-k-1)$. Using this notation for the pair-copulas, the blocks are

$$K_k(\boldsymbol{u}_{t-k},\dots,\boldsymbol{u}_t;\boldsymbol{\phi}(k))=\begin{cases}\prod_{l_1=1}^{m}\prod_{l_2=1}^{l_1-1}c^{(0)}_{l_1,l_2}\left(u_{i|j+1},u_{j|i-1};\phi_{l_1,l_2}(0)\right) & \text{if } k=0\\ \prod_{l_1=1}^{m}\prod_{l_2=1}^{m}c^{(k)}_{l_1,l_2}\left(u_{i|j+1},u_{j|i-1};\phi_{l_1,l_2}(k)\right) & \text{if } k>0.\end{cases}\quad(8)$$

A D-vine copula model of a stationary series has the following simple representations for the densities in Eq. (2). The marginal copula densities $c_{(t)}(\boldsymbol{u}_t)=K_0(\boldsymbol{u}_t;\boldsymbol{\phi}(0))$, for $t=1,\dots,T$, so that the parameters $\boldsymbol{\phi}(0)$ determine the cross-sectional dependence in $\boldsymbol{U}_t$. Therefore, the marginal density of $\boldsymbol{Y}_t$ is $f(\boldsymbol{y}_t)=K_0(\boldsymbol{u}_t;\boldsymbol{\phi}(0))\prod_{j=1}^{m}f_j(y_{j,t})$, where the subscript for $f_j$ indicates that the density varies over dimension $j=1,\dots,m$, but not over time $t$. From Eq. (7), for $s<t$, the transitional density

$$f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1},\dots,\boldsymbol{u}_s)=K_0(\boldsymbol{u}_t;\boldsymbol{\phi}(0))\prod_{k=1}^{t-s}K_k(\boldsymbol{u}_{t-k},\dots,\boldsymbol{u}_t;\boldsymbol{\phi}(k)).\quad(9)$$

It follows that if the Markov order of the series is $p$, then $K_k=1$ for all $k>p$, and that $K_p\neq 1$. To illustrate, Table 2 arranges, in lower triangular form, the unique pair-copulas for a stationary series of dimension $m=3$, length $T=4$ and Markov order $p=1$.

Thus, D-vine copula representations of stationary and Markov multivariate series have much smaller numbers of unique pair-copulas and associated parameters. The following formulae give the numbers of unique pair-copulas in each case, and Table 3 gives the numbers for some common circumstances.

*Unrestricted vine copula process:* $(N-1)N/2=(Tm-1)Tm/2$.
*Stationary vine copula process:* $(T-1)m^2+m(m-1)/2$.
*Stationary vine copula process Markov order $p<T-1$:* $pm^2+m(m-1)/2$.

Lastly, note that, while the numbers of block functionals and pair-copula densities are reduced in these cases, they still have to be evaluated at the arguments $u_{i|j}$ and $u_{j|i}$, for $i>j$, in Eq. (8). Section 3 shows how this can be done efficiently.

## 2.5. Which pair-copulas?

While any mix of bivariate copulas can be employed for the pair-copulas, because the direction of dependence of each pair-copula is unknown, only those that allow for both positive and negative dependence should be employed. For problems where the dependence is likely to be symmetric and heavy-tailed, the bivariate $t$-copula can be used (Demarta & McNeil, 2005). To allow for bi-directional asymmetric dependence, the following mixture of rotated Archimedean copulas is suggested.

Let $c_A(u,v;\tau)$ be a bivariate Archimedean copula density parameterised in terms of Kendall's tau, $0<\tau<\tau^A_{max}$; then, the copula densities $c^{++}(u,v;\tau)=c_A(u,v;\tau)$, $c^{+-}(u,v;\tau)=c_A(u,1-v;\tau)$, $c^{-+}(u,v;\tau)=c_A(1-u,v;\tau)$ and $c^{--}(u,v;\tau)=c_A(1-u,1-v;\tau)$ are rotations of $c_A$ in the four quadrants. Negative dependence

**Table 2**
Pair-copulas $\{c_{i,j};\, i = 1, \ldots, N,\, j < i\}$ for a stationary multivariate time series of dimension $m = 3$, length $T = 4$ and Markov order $p = 1$.

| $i$ | $j$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 2 | $c_{2,1}^{(0)}$ | | | | | | | | | | |
| 3 | $c_{3,1}^{(0)}$ | $c_{3,2}^{(0)}$ | | | | | | | | | |
| 4 | $c_{1,1}^{(1)}$ | $c_{1,2}^{(1)}$ | $c_{1,3}^{(1)}$ | | | | | | | | |
| 5 | $c_{2,1}^{(1)}$ | $c_{2,2}^{(1)}$ | $c_{2,3}^{(1)}$ | $c_{2,1}^{(0)}$ | | | | | | | |
| 6 | $c_{3,1}^{(1)}$ | $c_{3,2}^{(1)}$ | $c_{3,3}^{(1)}$ | $c_{3,1}^{(0)}$ | $c_{3,2}^{(0)}$ | | | | | | |
| 7 | 1 | 1 | 1 | $c_{1,1}^{(1)}$ | $c_{1,2}^{(1)}$ | $c_{1,3}^{(1)}$ | | | | | |
| 8 | 1 | 1 | 1 | $c_{2,1}^{(1)}$ | $c_{2,2}^{(1)}$ | $c_{2,3}^{(1)}$ | $c_{2,1}^{(0)}$ | | | | |
| 9 | 1 | 1 | 1 | $c_{3,1}^{(1)}$ | $c_{3,2}^{(1)}$ | $c_{3,3}^{(1)}$ | $c_{3,1}^{(0)}$ | $c_{3,2}^{(0)}$ | | | |
| 10 | 1 | 1 | 1 | 1 | 1 | 1 | $c_{1,1}^{(1)}$ | $c_{1,2}^{(1)}$ | $c_{1,3}^{(1)}$ | | |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | $c_{2,1}^{(1)}$ | $c_{2,2}^{(1)}$ | $c_{2,3}^{(1)}$ | $c_{2,1}^{(0)}$ | |
| 12 | 1 | 1 | 1 | 1 | 1 | 1 | $c_{3,1}^{(1)}$ | $c_{3,2}^{(1)}$ | $c_{3,3}^{(1)}$ | $c_{3,1}^{(0)}$ | $c_{3,2}^{(0)}$ |

Notes: Ones appear in cells where $c_{i,j} = 1$, and the non-unity pair-copula densities are given by their block notation $c_{l_1,l_2}^{(k)}$. There are three unique pair-copulas in the blocks $K_0(\boldsymbol{u}_t; \boldsymbol{\phi}(0))$, and nine unique pair-copulas in the blocks $K_1(\boldsymbol{u}_{t-1}, \boldsymbol{u}_t; \boldsymbol{\phi}(1))$, giving a total of 12 unique pair-copulas in the model.

**Table 3**
Number of unique pair-copulas in the D-vine model for various combinations of the vector dimension $m$, and number of observations $T$.

| $T$ | $m$ | | | | |
|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 10 |
| | (i) *Unconstrained D-vine* | | | | |
| 40 | 3,160 | 7,140 | 12,720 | 19,900 | 79,800 |
| 80 | 12,720 | 28,680 | 51,040 | 79,800 | 319,600 |
| 120 | 28,680 | 64,620 | 114,960 | 179,700 | 719,400 |
| 500 | 499,500 | 1,124,250 | 1,999,000 | 3,123,750 | 12,497,500 |
| | (ii) *Stationary process* | | | | |
| 40 | 157 | 354 | 630 | 985 | 3,945 |
| 80 | 317 | 714 | 1,270 | 1,985 | 7,945 |
| 120 | 477 | 1,074 | 1,910 | 2,985 | 11,945 |
| 500 | 1,997 | 4,494 | 7,990 | 12,485 | 49,945 |
| | (iii) *Stationary process with Markov order* $p = 4$ | | | | |
| 40 | 17 | 39 | 70 | 110 | 445 |
| 80 | 17 | 39 | 70 | 110 | 445 |
| 120 | 17 | 39 | 70 | 110 | 445 |
| 500 | 17 | 39 | 70 | 110 | 445 |
| | (iv) *Serially independent process* | | | | |
| 40 | 1 | 3 | 6 | 10 | 45 |
| 80 | 1 | 3 | 6 | 10 | 45 |
| 120 | 1 | 3 | 6 | 10 | 45 |
| 500 | 1 | 3 | 6 | 10 | 45 |

Notes: If the parameter of each pair-copula is a scalar, then this is the same as the number of parameters.

is captured by $c^{+-}$ and $c^{-+}$, and positive by $c^{++}$ and $c^{--}$. The copula with density

$$c_{BA}(u, v; \tau, \alpha) = \begin{cases} \alpha c^{++}(u, v; \tau) + (1-\alpha)c^{--}(u, v; \tau) \\ \quad \text{if } \tau > 0 \\ 1 \text{ if } \tau = 0 \\ \alpha c^{-+}(u, v; -\tau) + (1-\alpha)c^{+-}(u, v; -\tau) \\ \quad \text{if } \tau < 0, \end{cases}$$

and parameters $0 \leq \alpha \leq 1$, $-\tau_{max}^A < \tau < \tau_{max}^A$, is an extension of the convex combination of Archimedean and survival copulas suggested by Junker and May (2005). Dependence is bi-directional and is determined smoothly by $\tau$, while $\alpha$ determines the level of tail dependence. The properties of this copula are given in Table 4. Bivariate copulas based upon rotations and mixtures of copulas are popular in empirical analysis; see Fortin and Kuzmics (2002) and Rodriguez (2007) for examples.

**Table 4**
Properties of the rotated and bi-directional Archimedean copulas, including: (i) copula function, (ii) conditional copula function, (iii) inverse conditional copula function, and (iv) lower and upper tail dependencies.

*Rotated clockwise* $(-\tau^A_{max} < \tau < 0)$
(i) $C^{+-}(u, v; \tau) = u - C_A(u, 1 - v; -\tau)$
(ii) $h^{+-}(u|v; \tau) = h_A(u|1 - v; -\tau)$
(iii) inverse $(h^{+-})^{-1}(u|v; \tau) = h_A^{-1}(u|1 - v; -\tau)$
(iv) Lower and upper tail dependence $\lambda_L^{+-} = \lambda^A_{Q2}$ and $\lambda_U^{+-} = \lambda^A_{Q4}$.

*Rotated anti-clockwise* $(-\tau^A_{max} < \tau < 0)$
(i) $C^{-+}(u, v; \tau) = v - C_A(1 - u, v; -\tau)$
(ii) $h^{-+}(u|v; \tau) = 1 - h_A(1 - u|v; -\tau)$
(iii) inverse$(h^{-+})^{-1}(u|v; \tau) = 1 - h_A^{-1}(1 - u|v; -\tau)$
(iv) Lower and upper tail dependence $\lambda_L^{-+} = \lambda^A_{Q4}$ and $\lambda_U^{-+} = \lambda^A_{Q2}$.

*Rotated 180 degrees, or survival copula* $(0 < \tau < \tau^A_{max})$
(i) $C^{--}(u, v; \tau) = u + v - 1 + C_A(1 - u, 1 - v; \tau)$
(ii) $h^{--}(u|v; \tau) = 1 - h_A(1 - u|1 - v; \tau)$
(iii) $(h^{--})^{-1}(u|v; \tau) = 1 - h_A^{-1}(1 - u|1 - v; \tau)$
(iv) Lower and upper tail dependence $\lambda_L^{--} = \lambda^A_U$ and $\lambda_U^{-+} = \lambda^A_L$.

*Bi-directional Archimedean copula* $(-\tau^A_{max} < \tau < \tau^A_{max})$
(i) $C_{BA}(u, v; \tau, \alpha) = \mathcal{I}(\tau = 0)uv + \mathcal{I}(\tau > 0)[\alpha C_A(u, v; \tau) + (1 - \alpha)C^{--}(u, v; \tau)] + \mathcal{I}(\tau < 0)[\alpha C^{+-}(u, v; \tau) + (1 - \alpha)C^{-+}(u, v; \tau)]$
(ii) $h_{BA}(u|v; \tau, \alpha) = \mathcal{I}(\tau = 0)u + \mathcal{I}(\tau > 0)[\alpha h_A(u|v; \tau) + (1 - \alpha)h^{--}(u|v; \tau)] + \mathcal{I}(\tau < 0)[\alpha h^{+-}(u|v; \tau) + (1 - \alpha)h^{-+}(u|v; \tau)]$
(iii) $h_{BA}^{-1}(u|v; \tau, \alpha)$ obtained numerically
(iv) Lower tail dependence $\lambda_L^{BA} = \mathcal{I}(\tau > 0)[\alpha\lambda^A_L + (1 - \alpha)\lambda^A_U] + \mathcal{I}(\tau < 0)[\alpha\lambda^A_{Q2} + (1 - \alpha)\lambda^A_{Q4}]$
Upper tail dependence $\lambda_U^{BA} = \mathcal{I}(\tau > 0)[\alpha\lambda^A_U + (1 - \alpha)\lambda^A_L] + \mathcal{I}(\tau < 0)[\alpha\lambda^A_{Q4} + (1 - \alpha)\lambda^A_{Q2}]$

Notes: These are based on any Archimedean copula with the corresponding distribution function $C_A(u, v; \tau)$, Kendall's tau $0 < \tau < \tau^A_{max}$, conditional copula function $h_A(u|v; \tau) = \frac{\partial}{\partial v} C_A(u, v; \tau)$, lower and upper tail dependences $\lambda^A_L$ and $\lambda^A_U$, and off-diagonal tail dependencies $\lambda^A_{Q2} = 1 - \lim_{q\downarrow 0} C_A(q, 1 - q; \tau)/q$ and $\lambda^A_{Q4} = 1 - \lim_{q\downarrow 0} C_A(1 - q, q; \tau)/q$. Note that for both the Gumbel and Clayton copulas, the off-diagonal tail dependencies are $\lambda^A_{Q2} = \lambda^A_{Q4} = 0$, while the other properties are given by Joe (1997) and Nelsen (2006) for these and other Archimedean copulas. The function $\mathcal{I}(X)$ denotes an indicator function which is equal to 1 if $X$ is true, and 0 otherwise.

In the special case where all of the pair-copulas are Gaussian copulas, Bedford and Cooke (2002) note that $c_{DV}(\boldsymbol{u}; \boldsymbol{\phi})$ is the density of a $N$-dimensional Gaussian copula

$$c_{\mathrm{Ga}}(\boldsymbol{u}; \Omega) = |\Omega|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{w}'(\Omega^{-1} - I_N)\boldsymbol{w}\right\},$$

with $\boldsymbol{w} = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_N))'$. Denote $W_i = \Phi^{-1}(U_i)$, then $(W_1, \ldots, W_N) \sim N(0, \Omega)$, and the pair-copula parameters

$$\phi_{i,j} = \begin{cases} \mathrm{corr}(W_i, W_j|W_{i-1}, \ldots, W_{j+1}) & \text{if } i > j + 1 \\ \mathrm{corr}(W_i, W_j) & \text{if } i = j + 1, \end{cases} \tag{10}$$

are partial correlations corresponding to $\Omega$. The advantages of representating a correlation matrix by these partial correlations are discussed by Bedford and Cooke (2002), Daniels and Pourahmadi (2009) and Joe (2006), and there is a one-to-one mapping between $\boldsymbol{\phi}$ and $\Omega$; see Anderson (2003, pp. 40–41). In this case, the copula $c_{(s,\ldots,t)}(\cdot) = c_{Ga}(\cdot; \Omega_{(s,\ldots,t)})$, where $\Omega_{(s,\ldots,t)}$ is the $(t-s)m \times (t-s)m$ submatrix of $\Omega = \{\omega_{i,j}\}$ comprising the elements $\{\omega_{i,j}; a(s) \le i, j \le b(t)\}$. The conditional density of $\boldsymbol{U}_t|\boldsymbol{U}_{t-1}, \ldots, \boldsymbol{U}_s$ can be expressed directly in terms of $\Omega_{(s,\ldots,t)}$ by considering the ratio

$$f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1}, \ldots, \boldsymbol{u}_s) = c_{Ga}\left((\boldsymbol{u}'_s, \ldots, \boldsymbol{u}'_t)'; \Omega_{(s,\ldots,t)}\right) / c_{Ga}\left((\boldsymbol{u}'_s, \ldots, \boldsymbol{u}'_{t-1})'; \Omega_{(s,\ldots,t-1)}\right).$$

However, the parameterization of the conditional density using the vine in Eq. (7) is both more flexible and more efficient computationally, in the same manner as the parameterization of a correlation matrix by its partial correlations.

## 3. Estimation and inference

This section discusses how to estimate a stationary D-vine copula time series model. It shows how the likelihood can be evaluated efficiently using a faster adaptation of the algorithm of Aas et al. (2009) and Smith et al. (2010). A Bayesian prior that shrinks both individual pair-copula densities and block functionals towards unity is also proposed, which corresponds to shrinking towards conditional independence. A MCMC algorithm is proposed to evaluate the posterior distribution, which is a Bayesian model average. It is discussed how dependence can be assessed using the fitted copula model, as well as how predictive inference can be constructed, based on Monte Carlo methods. To facilitate this, a computationally efficient algorithm for drawing samples from the fitted copula model is also given in the Appendix.

### 3.1. Likelihood

For a stationary series, the likelihood is $\mathcal{L}(\boldsymbol{\phi}, \boldsymbol{\theta}|\boldsymbol{y}) = c_{DV}(\boldsymbol{u}; \boldsymbol{\phi}) \prod_{t=1}^T \prod_{j=1}^m f_j(y_{j,t}; \boldsymbol{\theta}_j)$, where the marginal densities of $y_{j,t}$ have parameters $\boldsymbol{\theta}_j$, and $\boldsymbol{\theta} = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_m\}$. If the Markov order is $p$, then, from Eq. (9), the copula density can be written in terms of the block functionals

$$c_{DV}(\boldsymbol{u}; \boldsymbol{\phi}) = K_0(\boldsymbol{u}_1; \boldsymbol{\phi}(0)) \times \prod_{t=2}^{T} \left\{ K_0(\boldsymbol{u}_t; \boldsymbol{\phi}(0)) \prod_{k=1}^{\min(t-1,p)} K_k(\boldsymbol{u}_{t-k}, \ldots, \boldsymbol{u}_t; \boldsymbol{\phi}(k)) \right\}. \tag{11}$$

If $p = 0$, the inner most product is equal to unity and the time series is serially independent.

The main burden in computing the likelihood above is the evaluation of the arguments $u_{i|j}$ and $u_{j|i}$, for $i > j$, of the component pair-copulas of the blocks. These are first computed, then stored in an $N \times N$ matrix $\mathcal{M}$, where the value $u_{i|j}$ is stored as the $i, j$th element in the lower triangle of $\mathcal{M}$, and $u_{j|i}$ is stored as the $j, i$th element in the upper triangle. To outline the algorithm, first note that there is a one-to-one relationship between the indices of the pair-copulas in Eqs. (3) and (8):

$$\underline{(i,j) \to (t,k,l_1,l_2)} : t = \lceil i/m \rceil, s = \lceil j/m \rceil,$$
$$k = t - s, l_1 = i - m(t-1), l_2 = j - m(s-1);$$
$$\underline{(t,k,l_1,l_2) \to (i,j)} :$$
$$i = l_1 + m(t-1), j = l_2 + m(s-1), s = t - k,$$

where $\lceil \cdot \rceil$ is the ceiling function. The arguments are computed using an algorithm based on the following two recursive relationships:

$$u_{i|j} \equiv F(u_i|u_{i-1}, \ldots, u_j) = h^{(k)}_{l_1,l_2}(u_{i|j+1}|u_{j|i-1}) \quad \text{and}$$
$$u_{j|i} \equiv F(u_j|u_i, \ldots, u_{j+1}) = h^{(k)}_{l_1,l_2}(u_{j|i-1}|u_{i|j+1}). \tag{12}$$

The function $h^{(k)}_{l_1,l_2}(u|v)$ is the conditional distribution function of the pair-copula $c^{(k)}_{l_1,l_2}(u, v)$, and can be computed in closed form for many choices of bivariate copulas; for example, see Table 4 for that of the bi-directional Archimedean copula. Note that this function is also denoted $h_{i,j}(u|v)$ whenever the pair-copula is indexed as $c_{i,j}$.

The following two observations can be exploited to derive a computationally efficient algorithm. First, for Markov orders $p < T-1$, only a subset of these pair-copula arguments are needed to evaluate the blocks in Eq. (11). Second, $h^{(k)}_{l_1,l_2}(u|v) = u$ if $c^{(k)}_{l_1,l_2} = 1$, in which case the relationships in Eq. (12) are simplified to the two identities $u_{i|j} = u_{i|j+1}$ and $u_{j|i} = u_{j|i-1}$. This occurs whenever the block index is greater than the Markov order, i.e., $k > p$. In terms of the indices in Eq. (3), this implies that the pair-copula $c_{i,j} = 1$ whenever $i - j > (p+1)m - 1$. For example, in Table 2, $c_{i,j} = 1$ when $i - j > 5$, and also in the case of a small number of additional pair-copulas.

**Algorithm 1** (*Fast Evaluation of Pair-Copula Arguments*).
For $r = 1, \ldots, (p+1)m - 1$:
For $i = r + 1, \ldots, N$ (compute loop in parallel):
  Step 1. Set $j = i - r$, $s = \lceil j/m \rceil$, $t = \lceil i/m \rceil$,
    $k = t - s$, $l_1 = i - m(t-1)$, $l_2 = j - m(s-1)$.
  Step 2. If $k > p$ or $c^{(k)}_{l_1,l_2} = 1$ then
    Step 2.1. $u_{i|j} = u_{i|j+1}$ and $u_{j|i} = u_{j|i-1}$,
  else
    Step 2.2. $u_{i|j} = h^{(k)}_{l_1,l_2}(u_{i|j+1}|u_{j|i-1})$ and $u_{j|i} = h^{(k)}_{l_1,l_2}(u_{j|i-1}|u_{i|j+1})$.

In Algorithm 1, the outer loop fills in the elements of matrix $\mathcal{M}$ in the upper and lower diagonal bands, from band $r = 1$ to band $r = (p+1)m - 1$. The computations in the inner loop in $i$ can be undertaken in parallel, thus greatly decreasing the implementation time in real time.[1] In comparison, the algorithm of Smith et al. (2010) evaluates the elements of $\mathcal{M}$ column-by-column, and the computations cannot be undertaken in parallel. In addition, the computational burden of Algorithm 1 occurs in the repeated evaluation of the functions $h^{(k)}_{l_1,l_2}$ in Step 2.2. There are at most $Tpm^2$ of these here, whereas a direct application of the algorithm of Smith et al. (2010) results in $T^2m^2$ evaluations, which is a substantial reduction. Last, after running Algorithm 1, $c_{DV}$ is evaluated directly from Eq. (11) using the expressions for the blocks in Eq. (8).

[1] Parallel computing is now popular, with contemporary PCs having multi-core CPUs and GPUs. Loops can be evaluated in parallel using languages such as Fortran 95 and C using the OpenMP or other protocols, as well as in recent editions of Matlab.

### 3.2. Priors for Bayesian model averaging

Recall that if $K_0 = 1$, then there is contemporaneous independence between the elements of $\boldsymbol{Y}_t$, while the terms $K_1, \ldots, K_p$ determine serial dependence. The following two-level hierarchical prior shrinks these block functionals towards unity in order to provide a parsimonious representation. For the first level, latent binary indicators $\boldsymbol{\delta} = (\delta_0, \delta_1, \ldots, \delta_p)$ are introduced, where $K_k = 1$ iff $\delta_k = 0$, and $\delta_k = 1$ otherwise. The prior $\pi(\boldsymbol{\delta})$ can be chosen by the user, but in the empirical work here, $\pi(\delta_k = 1) = 1/(2+k)$, so that the probability of contemporaneous dependence is $\pi(\delta_0 = 1) = 1/2$, and lower order serial dependence is encouraged, similarly to the work of Barnett, Kohn, and Sheather (1996) for a stationary univariate autoregression.

The total numbers of pair-copulas in block $k$ are $N_k = m^2$ for $k > 0$, and $N_0 = m(m-1)/2$ for $k = 0$. To allow for further parsimony, in the second level of the hierarchical prior, each pair-copula has a point mass probability of being the independence copula, so that the pair-copula density is equal to unity. For example, when all of the pair-copulas are Gaussian, so that $c_{DV}$ is a Gaussian copula, the prior puts a point mass on each partial correlation in Eq. (10) being zero. Indicator variables $\gamma_{l_1,l_2}(k)$ are introduced, where $c^{(k)}_{l_1,l_2} = 1$ iff $\gamma_{l_1,l_2}(k) = 0$, and $\gamma_{l_1,l_2}(k) = 1$ otherwise. For a given Markov order $p$, denote the set of all indicators by $\boldsymbol{\gamma}$, so that $\boldsymbol{\gamma} = \{\boldsymbol{\gamma}(k); k = 0, \ldots, p\}$, $\boldsymbol{\gamma}(k) = \{\gamma_{l_1,l_2}(k); (l_1, l_2) \in L(k)\}$, with

$$L(k) = \begin{cases} \{(l_1, l_2);\ l_1 = 2, \ldots, m,\ l_2 = 1, \ldots, l_1\} \\ \quad \text{if } k = 0 \\ \{(l_1, l_2);\ l_1 = 1, \ldots, m,\ l_2 = 1, \ldots, m\} \\ \quad \text{if } 1 \le k \le p. \end{cases}$$

The indicators for the individual pair-copulas $\boldsymbol{\gamma}(k) = \mathbf{0}$ iff $\delta_k = 0$, so that the conditional prior $\pi(\boldsymbol{\gamma}(k)|\delta_k = 0) = \mathcal{I}(\boldsymbol{\gamma}(k) = \mathbf{0})$, where the function $\mathcal{I}(A) = 1$ if $A$ is true, and zero otherwise, and $\mathbf{0}$ is the zero vector. When $\delta_k = 1$, Scott and Berger (2010) discuss the importance of employing a prior that corrects for the multiplicity of the $2^{N_k} - 1$ possible values of $\boldsymbol{\gamma}(k)$, and the following prior is used to account for this:

$$\pi(\boldsymbol{\gamma}(k)|\delta_k = 1) \propto \mathrm{Beta}(S_k + 1, N_k - S_k + 1) \times (1 - \mathcal{I}(\boldsymbol{\gamma}(k) = \mathbf{0})).$$

Here, $S_k = \sum_{(l_1,l_2)\in L(k)} \gamma_{l_1,l_2}(k)$ is the number of pair-copulas in block $k$ which are not independence copulas, and $\mathrm{Beta}(\cdot,\cdot)$ is the beta function. The implied prior for the total number of non-zero elements of $\boldsymbol{\gamma}(k)$ is uniform with $\pi(S_k|\delta_k = 1) = 1/N_k$ for $S_k \ge 1$, $\pi(S_k = 0) = \pi(\delta_k = 0)$, and the priors $\pi(\gamma_{l_1,l_2}(k)|\delta_k = 1)$ are all equal for $(l_1, l_2) \in L(k)$. This prior has been used successfully for Bayesian variable selection (Kohn, Smith, & Chan, 2001) and elsewhere. Finally, the prior $\pi(\boldsymbol{\gamma})$ is obtained analytically by marginalising over $\boldsymbol{\delta}$ as follows:

$$\pi(\boldsymbol{\gamma}) = \sum_{\boldsymbol{\delta}} \pi(\boldsymbol{\gamma}|\boldsymbol{\delta})\pi(\boldsymbol{\delta}) = \prod_{k=1}^{m}\left(\sum_{\delta_k=0,1} \pi(\boldsymbol{\gamma}(k)|\delta_k)\pi(\delta_k)\right) = \prod_{k=1}^{m} \pi(\boldsymbol{\gamma}(k)),$$

where $\pi(\boldsymbol{\gamma}(k)) = \mathcal{I}(\boldsymbol{\gamma}(k) = \mathbf{0})\pi(\delta_k = 0) + \text{Beta}(S_k + 1, N_k - S_k + 1)(1 - \mathcal{I}(\boldsymbol{\gamma}(k) = \mathbf{0}))\pi(\delta_k = 1)$. In the MCMC scheme below, only $\boldsymbol{\gamma}$ is generated directly, not $\boldsymbol{\delta}$.

The prior for $\phi_{l_1,l_2}(k)$ varies depending on the choice of the pair-copula. For example, in the empirical work, uniform priors are used for the correlation parameters of the elliptical copulas, and for Kendall's tau and the mixture weights of the bi-directional Archimedean copulas.

### 3.3. Sampling schemes

The posterior distribution $f(\boldsymbol{\phi}, \boldsymbol{\gamma}, \boldsymbol{\theta}|\boldsymbol{y})$ can be evaluated using MCMC methods. For many parametric models, the marginal parameters $\boldsymbol{\theta}_j$ can be generated using Metropolis–Hastings with a normal approximation, as per Pitt, Chan, and Kohn (2006) and Smith et al. (2010), or similar proposals. It is also popular to employ nonparametric methods for the marginal models and estimate the dependence parameters conditional upon these, so that there are no marginal parameters $\boldsymbol{\theta}$ to estimate; see Shih and Louis (1995) for an early example. To generate the copula parameters, an extension of the approach suggested by Smith et al. (2010) is used here. The method involves an increased blocking of the parameters for an improved mixing of the Markov chain, and is outlined below.

Additional latent variables $\boldsymbol{\zeta} = \{\zeta_{l_1,l_2}(k); (l_1, l_2) \in L(k)\}$ are introduced, such that $\phi_{l_1,l_2}(k) = \zeta_{l_1,l_2}(k)$ iff $\gamma_{l_1,l_2}(k) = 1$. Note that $(\phi_{l_1,l_2}(k), \gamma_{l_1,l_2}(k))$ is known exactly from $(\zeta_{l_1,l_2}(k), \gamma_{l_1,l_2}(k))$. The prior $\pi(\zeta_{l_1,l_2}(k)|\gamma_{l_1,l_2}(k) = 1) \propto \pi(\phi_{l_1,l_2}(k)|\gamma_{l_1,l_2}(k) = 1)$, the domains of $\boldsymbol{\zeta}$ and $\boldsymbol{\phi}$ are the same, and $\pi(\boldsymbol{\zeta}, \boldsymbol{\gamma}) = \pi(\boldsymbol{\gamma}) \prod_k \prod_{(l_1,l_2)\in L(k)} \pi(\zeta_{l_1,l_2}(k))$. With these priors, generating from $f(\boldsymbol{\zeta}, \boldsymbol{\gamma}|\boldsymbol{\theta}, \boldsymbol{y})$ is equivalent to generating from $f(\boldsymbol{\phi}, \boldsymbol{\gamma}|\boldsymbol{\theta}, \boldsymbol{y})$.

The sampling scheme consists of Metropolis–Hastings (MH) multivariate random walk steps as follows. Randomly select two pair-copulas, indexed by '$B$', with corresponding values $\boldsymbol{\zeta}_B \subset \boldsymbol{\zeta}$ and $\boldsymbol{\gamma}_B \subset \boldsymbol{\gamma}$. Generate new values $\boldsymbol{\gamma}_B^{\text{new}}$ from the proposal mass function $q(\boldsymbol{\gamma}_B) = 1/4$, and $\boldsymbol{\zeta}_B^{\text{new}}$ from a random walk proposal, constrained to the domain of $\boldsymbol{\zeta}_B$, with a multivariate $t$ innovation with scale matrix $\sigma^2 I$ and degrees of freedom $d$. In the empirical work, we set $d = 5$ and $\sigma^2$, either adaptively or manually, in order to reduce the number of iterations needed for convergence and to increase the mixing of the Markov chain. Using common metrics, all Markov chains appear to converge to the posterior distribution within 5000 iterations in our empirical work.

The new iterate $(\boldsymbol{\zeta}_B^{\text{new}}, \boldsymbol{\gamma}_B^{\text{new}})$ is accepted over the old $(\boldsymbol{\zeta}_B^{\text{old}}, \boldsymbol{\gamma}_B^{\text{old}})$ with probability $\min(1, \alpha R)$, where $R$ is an adjustment due to any bounds on the domain of $\boldsymbol{\zeta}_B$. If the likelihood, $\mathcal{L}$, is denoted as a function of $(\boldsymbol{\phi}_B, \boldsymbol{\gamma}_B)$, then

$$\alpha = \frac{\mathcal{L}(\boldsymbol{\zeta}_B^{\text{new}}, \boldsymbol{\gamma}_B^{\text{new}})\pi(\boldsymbol{\zeta}_B^{\text{new}})\pi(\boldsymbol{\gamma}_B^{\text{new}}|\{\boldsymbol{\gamma}\backslash\boldsymbol{\gamma}_B^{\text{new}}\})}{\mathcal{L}(\boldsymbol{\zeta}_B^{\text{old}}, \boldsymbol{\gamma}_B^{\text{old}})\pi(\boldsymbol{\zeta}_B^{\text{old}})\pi(\boldsymbol{\gamma}_B^{\text{old}}|\{\boldsymbol{\gamma}\backslash\boldsymbol{\gamma}_B^{\text{old}}\})}.$$

Here, $\{\boldsymbol{\gamma}\backslash\boldsymbol{\gamma}_B\}$ denotes $\boldsymbol{\gamma}$ without elements $\boldsymbol{\gamma}_B$, and the conditional prior $\pi(\boldsymbol{\gamma}_B|\{\boldsymbol{\gamma}\backslash\boldsymbol{\gamma}_B\})$ can be computed analytically from $\pi(\boldsymbol{\gamma})$. Note that $\mathcal{L}$ is not a function of $\boldsymbol{\zeta}_B$ when $\boldsymbol{\gamma}_B = \mathbf{0}$, while $\mathcal{L}(\boldsymbol{\zeta}_B, \boldsymbol{\gamma}_B) = \mathcal{L}(\boldsymbol{\phi}_B, \boldsymbol{\gamma}_B)$ for other values of $\boldsymbol{\gamma}_B$. Also, when $\boldsymbol{\gamma}_B^{\text{old}} = \boldsymbol{\gamma}_B^{\text{new}} = (0, 0)'$, evaluation of $\alpha$ does not involve the computation of $\mathcal{L}$, so that the sampling scheme is faster the more frequently this case arises.

Sampling schemes that generate each $(\gamma_{l_1,l_2}(k), \zeta_{l_1,l_2}(k))$ individually can also be used. However, by generating the block $(\boldsymbol{\zeta}_B, \boldsymbol{\gamma}_B)$ associated with two randomly selected pair-copulas, a 'switch' in pair-copula indicators from $(0, 1)$ to $(1, 0)$ is possible in each generation; something that helps to improve the mixing of the Markov chain in the same manner as with reversible jump.

### 3.4. Estimates

After a burn-in period, $J$ iterates $\{\boldsymbol{\theta}^{[1]}, \boldsymbol{\gamma}^{[1]}, \boldsymbol{\phi}^{[1]}, \ldots, \boldsymbol{\theta}^{[J]}, \boldsymbol{\gamma}^{[J]}, \boldsymbol{\phi}^{[J]}\} \sim f(\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\phi}|\boldsymbol{y})$ are collected for constructing estimates, including posterior means, which are used as point estimates. Of particular interest are the posterior probabilities $\Pr(\gamma_{l_1,l_2}^{(k)} = 1|\boldsymbol{y}) \approx \frac{1}{J}\sum_{j=1}^{J} \gamma_{l_1,l_2}^{(k)[j]}$ and $\Pr(\delta_j = 0|\boldsymbol{y}) \approx \frac{1}{J}\sum_{j=1}^{J} \mathcal{I}(\boldsymbol{\gamma}(k)^{[j]} = \mathbf{0})$, which measure the level of parsimony and the Markov order, respectively. Posterior means of individual dependence metrics for each bivariate pair-copula, such as Kendall's tau or tail dependencies, can also be computed.

For any value $d > 0$, the fitted distribution with density

$$\hat{c}_{(t-d,\ldots,t)}(\boldsymbol{u}_{t-d}, \ldots, \boldsymbol{u}_t) = \int c_{(t-d,\ldots,t)}(\boldsymbol{u}_{t-d}, \ldots, \boldsymbol{u}_t; \boldsymbol{\phi}) f(\boldsymbol{\phi}|\boldsymbol{y}) d\boldsymbol{\phi},$$

is used to assess dependence. For a stationary series, $c_{(t-d,\ldots,t)}$ is invariant with respect to $t$, so that $\hat{c}_{(t-d,\ldots,t)}$ is likewise. It is straightforward to show that $\hat{c}_{(t-d,\ldots,t)}$ is a copula density, although no longer that of a D-vine. Nevertheless, in the empirical work in Section 5, fitted copula densities are often found, in practice, to be close to the D-vine density $c_{(t-d,\ldots,t)}(\cdot; \hat{\boldsymbol{\phi}})$, where $\hat{\boldsymbol{\phi}}$ is the posterior mean.

The level of cross-sectional and serial dependence can be measured by the unconditional pairwise Spearman correlations $\rho(Y_{i,t}, Y_{k,t-s}) = 12E(U_{i,t}U_{k,t-s}) - 3$, for $s = 0, 1, \ldots, d$. To estimate these, the expectation can be computed with respect to $\hat{c}_{(t-d,\ldots,t)}$ as follows. At each sweep $j = 1, \ldots, J$, generate iterates

$$(\boldsymbol{u}_{t-d}^{[j]}, \ldots, \boldsymbol{u}_t^{[j]}) \sim c_{(t-d,\ldots,t)}(\boldsymbol{u}_{t-d}, \ldots, \boldsymbol{u}_t; \boldsymbol{\phi}^{[j]}), \quad (13)$$

using Algorithm 2 in the Appendix. Then, because $\boldsymbol{\phi}^{[j]} \sim f(\boldsymbol{\phi}|\boldsymbol{y})$,

$$E_{\hat{c}_{(t-d,\ldots,t)}}(U_{i,t}U_{k,t-s}) = \int u_{i,t}u_{k,t-s}\hat{c}_{(t-d,\ldots,t)}(\boldsymbol{u}_{t-d}, \ldots, \boldsymbol{u}_t)\mathrm{d}\boldsymbol{u}_{t-d} \ldots \mathrm{d}\boldsymbol{u}_t \approx \frac{1}{J}\sum_{j=1}^{J} u_{i,t}^{[j]}u_{k,t-s}^{[j]}.$$

The estimate of the symmetric matrix $R_0 = \{\rho(Y_{i,t}, Y_{k,t})\}_{i=1,\dots,m;k=1,\dots,m}$ is a measure of cross-sectional dependence, while the estimate of the matrix $R_s = \{\rho(Y_{i,t}, Y_{k,t-s})\}_{i=1,\dots,m;k=1,\dots,m}$ measures serial dependence at lag $s$ and is asymmetric. The iterates in Eq. (13) can also be used to evaluate other aspects of the fitted distribution. These include measures of tail dependence, conditional moments, or multivariate margins of the fitted density, all of which are useful in assessing any nonlinearity in the dependence structure, as is illustrated in the empirical work.

Last, because the likelihood can be evaluated readily, it is easy to compute a number of information criteria for judging between models. Moreover, while this is not discussed here, it is straightforward to extend the model in order to allow for mixing over different pair-copula types. For example, setting $\gamma_{l_1,l_2}(k)$ to be multinomial would allow $c^{(k)}_{l_1,l_2}$ to take the form of two or more bi-directional Archimedean copulas, with estimates obtained using an extension of the sampling scheme discussed here.

### 3.5. Prediction

For a Markov order $p$, the predictive distribution $(\boldsymbol{Y}_{T+T'}, \dots, \boldsymbol{Y}_{T+1}|\boldsymbol{Y}_T, \dots, \boldsymbol{Y}_1, \boldsymbol{\phi})$ has density

$$f(\boldsymbol{y}_{T+T'}, \dots, \boldsymbol{y}_{T+1}|\boldsymbol{y}_T, \dots, \boldsymbol{y}_1, \boldsymbol{\phi}) = \prod_{t=T+1}^{T+T'} f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1}, \dots, \boldsymbol{u}_{t-p^*}, \boldsymbol{\phi}) \prod_{t=T+1}^{T+T'} \prod_{j=1}^{m} f_j(y_{j,t}), \quad (14)$$

where $p^* = \min(p, t-1)$, and $f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1}, \dots, \boldsymbol{u}_{t-p^*}, \boldsymbol{\phi})$ can be expressed as in Section 2.4. The evaluation of moments and other summaries of the predictive distribution via numerical integration is difficult for even moderate dimensions $m$ or forecast horizons $T'$. However, iterates can be generated quickly from Eq. (14) using Algorithm 2 outlined in the Appendix. Moreover, $\boldsymbol{\phi}$ is integrated out of the predictive density by appending Algorithm 2 to each sweep of the sampling scheme.[2]

The resulting iterates are used to construct Monte Carlo estimates of the forecast densities, and generalisations of the impulse-response functions (Koop, Pesaran, & Potter, 1996) which are appropriate for nonlinear multivariate time series. The latter are based on the distribution of the difference between the predictive distributions with and without a shock in each variable. That is, the distribution of

$$Z(t,k,j) = (Y_{k,T+t}|\{\boldsymbol{Y}\backslash Y_{j,T}\} = \{\boldsymbol{y}\backslash y_{j,T}\}, Y_{j,T} = y_{j,T} + \sigma_{T,j}) - (Y_{k,T+t}|\boldsymbol{Y} = \boldsymbol{y}). \quad (15)$$

Here, $\{\boldsymbol{Y}\backslash Y_{j,T}\}$ are the past values omitting $Y_{j,T}$, and $\sigma_{T,j}$ is a shock to $Y_{j,T}$ that is usually set equal to one or two standard deviations of this variable. Plots of the median or mean of $Z(t,k,j)$ against $t = 1, \dots, T'$ provide a generalisation of the impulse-response function for input dimension $j$ and output dimension $k$. These, and other quantiles of $Z$, are computed via simulation using Algorithm 2.

[2] Because Algorithm 2 is relatively fast to implement for lower Markov orders, 2–10 replicates of $\boldsymbol{y}_{T+1}, \dots, \boldsymbol{y}_{T+T'}$ are simulated at the end of each sweep in practice, to reduce the Monte Carlo variance of the estimators. Moreover, in a parallel processing implementation, each of these replicates can be evaluated separately on a different thread, further speeding implementation in real time.

## 4. Simulation study

The effectiveness of the approach in estimating the dependence structure on the unit cube is studied using a simulation study. One hundred datasets were simulated, with $m = 3$ and $T = 150$, from each of the following four cases:

A. *Serial independence:* All pair-copulas in block $K_0$ are bi-directional Gumbels, while blocks $K_k = 1$ for $k \geq 1$.
B. *Parsimonious order* 1 *serial dependence:* Two of three pair-copulas in block $K_0$, and four of nine in block $K_1$, are Gaussian copulas; the rest are independence copulas. Blocks $K_k = 1$ for $k \geq 2$.
C. *Parsimonious order* 2 *serial dependence:* All pair-copulas in block $K_0$, six of nine in block $K_1$, and three of nine in block $K_2$, are bi-directional Clayton copulas; the rest are independence copulas. Block $K_3 = 1$.
D. *Cross-sectional independence:* $K_0 = 1$, and $c_{l,l}(k)$ are bi-directional Gumbels for $k = 1, 2, 3$ and $l = 1, 2, 3$; all other pair-copulas are independence copulas.

The parameter values are reported in Table 5. Nine vine copula estimators with Markov order $p = 3$ are fitted to each case. These use Gaussian, bi-directional Clayton or bi-directional Gumbel pair-copulas, which are denoted E1, E2 and E3, respectively. A suffix 's' denotes an estimator where pair-copula selection is undertaken, and 'f' indicates that it is not (i.e., it is full). The suffix 'i' is used to denote an estimator with $\gamma_{l_1,l_2}(k) = 0$ for $l_1 \neq l_2, k \geq 0$, and $\gamma_{l,l}(k) = 1$ for $k > 0$. (The latter case corresponds to fitting three cross-sectionally independent univariate time series.) For example, 'E2s' is a D-vine copula with bi-directional Clayton pair-copula components, estimated with Bayesian selection. Estimators E3s/E3f, E1s/E1f, E2s/E2f and E1s/E1f/E3i are correct parametric models for cases A, B, C and D, respectively. Note that the ratio of the number of parameters to observations in these examples is challenging; for example, for the bi-directional Archimedean copulas, there are 60 continuous parameters.

The accuracy of the selection method is illustrated by classifying each pair-copula as an independence copula when the estimate of $\Pr(\gamma_{l_1,l_2}(k) = 0|\boldsymbol{y}) > 0.5$, and as dependent otherwise. Table 6 reports the percentage of correctly classified pair-copulas across simulation replicates, for all combinations of selection estimator and simulation cases. The results are broken down by the true values of $\gamma_{l_1,l_2}(k)$, and the selection method proves to be highly accurate. This is true even for estimators with different pair-copula types than that used in the simulation, a point which does not necessarily follow because the model is mis-specified. Table 7 reports the median, 1st and 9th deciles over the simulation of the estimates of $\Pr(\delta_k = 1|\boldsymbol{y})$, for $k = 0, 1, 2, 3$. Relative to the true values, the method does an impressive job of identifying the correct Markov order; or, for case D, the parsimony with $K_0 = 1$. The accuracy of the Bayesian selection methodology is in line with that documented previously for covariance selection (Smith & Kohn, 2002).

**Table 5**
The pair-copula parameters in the four cases of the simulation study.

| | | | | | |
|---|---|---|---|---|---|
| **Case A**: Bi-directional Gumbel pair-copulas with parameters $\tau, \alpha$ ($K_2 = K_3 = 1$) | | | | | |
| $K_0$ | | | $K_1$ | | |
| – | – | – | I | I | I |
| 0.7467, 0.0971 | – | – | I | I | I |
| −0.7900, 0.8235 | −0.2032, 0.1869 | – | I | I | I |
| **Case B**: Gaussian pair-copulas with parameter $\tau$ ($K_2 = K_3 = 1$) | | | | | |
| $K_0$ | | | $K_1$ | | |
| – | – | – | 0.2286 | 0.3990 | I |
| 0.3554 | – | – | −0.4111 | I | I |
| −0.2900 | I | – | 1 | 0.7028 | I |
| **Case C**: Bi-directional Clayton pair-copulas with parameters $\tau, \alpha$ | | | | | |
| $K_0$ | | | $K_1$ | | |
| – | – | – | 0.6491, 0.3500 | 0.3948, 0.7537 | I |
| 0.3467, 0.2435 | – | – | −0.7878, 0.1966 | 0.4284, 0.0135 | 0.5647, 0.1622 |
| 0.6517, 0.9600 | 0.5380, 0.7572 | – | I | I | 0.5637, 0.5678 |
| $K_2$ | | | $K_3$ | | |
| I | I | I | I | I | I |
| 0.4865, 0.4733 | I | −0.5391, 0.5285 | I | I | I |
| 0.0690, 0.3517 | I | I | I | I | I |
| **Case D**: Bi-directional Gumbel pair-copulas with parameters $\tau, \alpha$ | | | | | |
| $K_0$ | | | $K_1$ | | |
| – | – | – | 0.7249, 0.1455 | I | I |
| I | – | – | I | 0.4006, 0.9961 | I |
| I | I | – | I | I | 0.4164, 0.8693 |
| $K_2$ | | | $K_3$ | | |
| −0.1681, 0.8530 | I | I | 0.2354 | I | I |
| I | −0.1489, 0.6221 | I | I | −0.6998 | I |
| I | I | 0.1106, 0.3510 | I | I | 0.2045 |

Notes: The parameters are reported as four (3 × 3) blocks for $K_0, \ldots, K_3$, except in Cases A and B, where $K_2 = K_3 = 1$, and these blocks are therefore not reported. Independence pair-copulas are denoted by 'I'. Each block has nine pair-copulas, except for $K_0$, which only has three pair-copula components, reported as a lower triangle. In Cases A, B and D, the pair-copulas are bi-directional Archimedeans, so that two parameters are reported, whereas in Case C there is only one parameter per pair-copula.

**Table 6**
Correct pair-copula classification rates (in percentages) for the selection estimators E1s, E2s and E3s for the four simulation cases.

| | E1s | E2s | E3s | E1s | E2s | E3s | E1s | E2s | E3s | E1s | E2s | E3s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Case A. | | | Case B. | | | Case C. | | | Case D. | | |
| Independence | 99.2 | 96.9 | **97.7** | **90.5** | 93.0 | 96.0 | 78.6 | **93.3** | 90.3 | 97.4 | 98.1 | **99.1** |
| Dependent | 92.3 | 92.0 | **91.3** | **96.7** | 92.0 | 96.0 | 81.4 | **90.3** | 79.4 | 80.1 | 78.6 | **83.7** |

Notes: The results are broken down by correctly classified independent pair-copulas (with $\gamma_{l_1,l_2}(k) = 0$) and dependent pair-copulas (with $\gamma_{l_1,l_2}(k) = 1$). The results for the estimator that has the correct pair-copula family for each case are highlighted in bold.

To assess the accuracy of the approach for estimating the dependence structure, the posterior mean estimates of Kendall's tau for each pair-copula, and the Spearman correlations $R_0, \ldots, R_3$, are computed. Table 8 contains summaries of the absolute differences of these two coefficients from their true values. They measure the accuracy of the estimated conditional and unconditional dependence structures, respectively. Results are given for all combinations of simulation cases and estimators, with results for estimators that are based on the correct model in each case given in bold. A number of conclusions can be drawn. First, selection can improve the estimation of the conditional dependence structure substantially, even when an incorrect pair-copula type is chosen. Second, this generally extends to the unconditional dependence structure as well. Third, employing the incorrect pair-copula type does not always degrade the performance in terms of these two measures of dependence. For example, E1s and E3s are comparable in cases A, B and D. Last, selection appears to identify the cross-sectional independence reliably in case D.

## 5. Empirical applications

To illustrate the approach, two multivariate time series from the Australian National Electricity Market (NEM) are examined. Both series have the dimension $m = 5$, corresponding to the five regions in this market, which, in the order of the vector, are: New South Wales (NSW), Victoria (VIC), Queensland (QLD), South Australia (SA) and Tasmania (TAS).

### 5.1. Example 1: Daily electricity demand maxima

The first series is the daily maximum electricity demand in each region between 1 January 2010 and 31 August 2012 ($T = 974$). Forecasts of demand maxima are important to system managers for generation scheduling, as well as to electricity utilities for determining optimal bid schedules; for example, see Amjady (2001). Copula models have been used previously to model the cross-sectional dependence between multivariate extrema in hydrology (Favre et al.,

**Table 7**
Median values over the 100 simulation replicates of the estimates of the posterior probabilities that block $K_k \neq 1$, for $k = 0, \ldots, 3$ (i.e., the median of the values of $\Pr(\delta_k = 1|\boldsymbol{y})$).

| Block | E1s | E2s | E3s | True | E1s | E2s | E3s | True |
|---|---|---|---|---|---|---|---|---|
| | Case A. | | | | Case B. | | | |
| $k = 0$ | 1.00 | 1.00 | **1.00** | 1 | **1.00** | 1.00 | 1.00 | 1 |
| | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | |
| $k = 1$ | 0.13 | 0.07 | **0.06** | 0 | **1.00** | 1.00 | 1.00 | 1 |
| | (0.02, 0.46) | (0.01, 0.43) | (0.01, 0.22) | | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | |
| $k = 2$ | 0.14 | 0.07 | **0.07** | 0 | **0.33** | 0.22 | 0.19 | 0 |
| | (0.04, 0.40) | (0.01, 0.75) | (0.01, 0.35) | | (0.13, 0.94) | (0.09, 0.97) | (0.08, 0.74) | |
| $k = 3$ | 0.13 | 0.10 | **0.08** | 0 | **0.29** | 0.19 | 0.21 | 0 |
| | (0.02, 0.45) | (0.01, 0.47) | (0.01, 0.34) | | (0.16, 0.85) | (0.08, 0.59) | (0.08, 0.59) | |
| | Case C. | | | | Case D. | | | |
| $k = 0$ | 1.00 | **1.00** | 1.00 | 1 | 0.09 | 0.03 | **0.03** | 0 |
| | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | | (0.03, 0.27) | (0.00, 0.18) | (0.01, 0.10) | |
| $k = 1$ | 1.00 | **1.00** | 1.00 | 1 | 1.00 | 1.00 | **1.00** | 1 |
| | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | |
| $k = 2$ | 1.00 | **1.00** | 0.72 | 1 | 1.00 | 1.00 | **1.00** | 1 |
| | (1.00, 1.00) | (1.00, 1.00) | (0.09, 1.00) | | (0.62, 1.00) | (0.59, 1.00) | (0.46, 1.00) | |
| $k = 3$ | 0.51 | **0.18** | 0.54 | 0 | 1.00 | 1.00 | **1.00** | 1 |
| | (0.17, 0.99) | (0.08, 0.67) | (0.10, 0.99) | | (1.00, 1.00) | (1.00, 1.00) | (1.00, 1.00) | |

Notes: Results are reported for all combinations of selection estimator and case. The results for the estimator that has the correct pair-copula family for each case are highlighted in bold. Also reported in parentheses are the lower first and upper ninth deciles of these values over the 100 replicates, along with the true value of $\delta_k$.

**Table 8**
Summary of the accuracy of the estimated conditional and unconditional dependence structures, for all combinations of estimator and simulation case.

| Estimator: | E1s | E1f | E1i | E2s | E2f | E2i | E3s | E3f | E3i |
|---|---|---|---|---|---|---|---|---|---|
| | Absolute deviations for pair-copula Kendall's tau | | | | | | | | |
| Decile | Case A. | | | | | | | | |
| 1st | 0.000 | 0.007 | 0.000 | 0.000 | 0.007 | 0.000 | **0.000** | **0.005** | 0.000 |
| 5th | 0.000 | 0.034 | 0.002 | 0.000 | 0.006 | 0.001 | **0.000** | **0.027** | 0.001 |
| 9th | 0.015 | 0.088 | 0.747 | 0.037 | 0.609 | 0.747 | **0.011** | **0.076** | 0.747 |
| | Case B. | | | | | | | | |
| 1st | **0.000** | **0.006** | 0.000 | 0.000 | 0.007 | 0.000 | 0.000 | 0.006 | 0.000 |
| 5th | **0.002** | **0.032** | 0.012 | 0.001 | 0.036 | 0.002 | 0.001 | 0.031 | 0.002 |
| 9th | **0.043** | **0.082** | 0.290 | 0.063 | 0.089 | 0.290 | 0.036 | 0.080 | 0.290 |
| | Case C. | | | | | | | | |
| 1st | 0.000 | 0.011 | 0.000 | **0.000** | **0.002** | 0.000 | 0.000 | 0.006 | 0.000 |
| 5th | 0.024 | 0.067 | 0.198 | **0.002** | **0.016** | 0.198 | 0.005 | 0.038 | 0.203 |
| 9th | 0.390 | 0.357 | 0.636 | **0.043** | **0.065** | 0.594 | 0.183 | 0.183 | 0.614 |
| | Case D. | | | | | | | | |
| 1st | 0.000 | 0.007 | 0.000 | 0.000 | 0.006 | 0.000 | **0.000** | **0.005** | **0.000** |
| 5th | 0.002 | 0.040 | 0.037 | 0.001 | 0.033 | 0.038 | **0.001** | **0.029** | **0.028** |
| 9th | 0.098 | 0.116 | 0.168 | 0.092 | 0.095 | 0.149 | **0.070** | **0.086** | **0.148** |
| | Absolute deviations for unconditional Spearman correlations | | | | | | | | |
| Decile | Case A. | | | | | | | | |
| 1st | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | **0.00** | **0.00** | 0.00 |
| 5th | 0.01 | 0.04 | 0.01 | 0.01 | 0.04 | 0.01 | **0.01** | **0.03** | 0.01 |
| 9th | 0.03 | 0.11 | 0.61 | 0.06 | 0.11 | 0.62 | **0.03** | **0.10** | 0.61 |
| | Case B. | | | | | | | | |
| 1st | **0.00** | **0.00** | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 |
| 5th | **0.02** | **0.04** | 0.07 | 0.03 | 0.04 | 0.07 | 0.02 | 0.04 | 0.07 |
| 9th | **0.08** | **0.11** | 0.42 | 0.13 | 0.12 | 0.42 | 0.10 | 0.11 | 0.42 |
| | Case C. | | | | | | | | |
| 1st | 0.01 | 0.01 | 0.01 | **0.00** | **0.00** | 0.09 | 0.00 | 0.00 | 0.01 |
| 5th | 0.04 | 0.05 | 0.22 | **0.01** | **0.01** | 0.22 | 0.03 | 0.03 | 0.22 |
| 9th | 0.14 | 0.13 | 0.72 | **0.03** | **0.04** | 0.72 | 0.11 | 0.10 | 0.72 |
| | Case D. | | | | | | | | |
| 1st | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 | 0.00 | **0.00** | **0.01** | **0.00** |
| 5th | 0.01 | 0.06 | 0.01 | 0.01 | 0.04 | 0.01 | **0.01** | **0.04** | **0.01** |
| 9th | 0.08 | 0.18 | 0.07 | 0.09 | 0.13 | 0.09 | **0.06** | **0.13** | **0.06** |

Notes: The conditional dependence structure is measured by the absolute differences between the posterior mean estimates of Kendall's tau and their true values, for the pair-copulas, while the unconditional is measured by the absolute difference between the posterior mean estimates of the pairwise Spearman correlations and their true values. The summaries are the 1st, 5th and 9th deciles of all deviations across the simulation. Figures in bold correspond to estimators that are the correct vine copula for that case.

**Table 9**
Summaries of the three fitted copula multivariate time series models for real data Examples 1 and 2.

| | | | | | |
|---|---|---|---|---|---|
| *Example* 1. | | | | | |
| Vine | $E(S_0/N_0\|\boldsymbol{y})$ | $E(S_1/N_1\|\boldsymbol{y})$ | $E(S_2/N_2\|\boldsymbol{y})$ | $E(S_3/N_3\|\boldsymbol{y})$ | DIC |
| B.Clayton | 0.69 | 0.56 | 0.06 | 0.16 | −3784.9 |
| B.Gumbel | 0.50 | 0.62 | 0.09 | 0.17 | −3891.5 |
| Gaussian | 0.65 | 0.74 | 0.13 | 0.24 | −3519.8 |
| *Example* 2. | | | | | |
| Vine | $E(S_0/N_0\|\boldsymbol{y})$ | $E(S_1/N_1\|\boldsymbol{y})$ | $E(S_2/N_2\|\boldsymbol{y})$ | $E(S_3/N_3\|\boldsymbol{y})$ | DIC |
| B.Clayton | 0.745 | 0.808 | 0.008 | 0.021 | −7078.3 |
| B.Gumbel | 0.900 | 0.717 | 0.051 | 0.032 | −7166.1 |
| Gaussian | 0.987 | 0.834 | 0.005 | 0.003 | −6309.8 |
| Vine | $\Pr(\delta_0 = 1\|\boldsymbol{y})$ | $\Pr(\delta_1 = 1\|\boldsymbol{y})$ | $\Pr(\delta_2 = 1\|\boldsymbol{y})$ | $\Pr(\delta_3 = 1\|\boldsymbol{y})$ | |
| B.Clayton | 1.00 | 1.00 | 0.10 | 0.22 | |
| B.Gumbel | 1.00 | 1.00 | 0.92 | 0.43 | |
| Gaussian | 1.00 | 1.00 | 0.08 | 0.05 | |

Notes: The first four columns report the posterior means of the proportions of dependent pair-copulas in each block $S_k/N_k$, for $k = 0, 1, 2, 3$. Values close to zero correspond to parsimonious blocks. For Example 2, the posterior distributions of the lag indicators $\delta_k$ are also reported. The last column contains the Deviance Information Criterion (DIC) value for each vine copula.

2004; Renard & Lang, 2007). This example extends these models to account for serial dependence as well.

The daily maxima are mean-adjusted for day type and seasonality, both of which are known calendar effects (Smith, 2000). This is undertaken using separate regressions for each series, which employ eight dummy variables for the days of the week and public holidays, eight Fourier terms and a linear trend. The marginal distributions are then modelled using the five empirical distribution functions. Vine copula multivariate time series models are fitted using Gaussian and bi-directional Gumbel and Clayton pair-copulas, with a maximum Markov order of $p = 5$. In all three fits, $\Pr(\delta_k = 1|\boldsymbol{y}) > 0.95$ for $k = 0, \ldots, 3$, and $\Pr(\delta_k = 1|\boldsymbol{y}) < 0.02$ for $k = 4, 5$, indicating a lag length of three days. Table 9 reports the expected proportion of dependent pair-copulas, $E(S_k/N_k|\boldsymbol{y})$, for the first four block functionals, and shows that the fitted copulas are parsimonious, in terms of dependent pair-copulas. The table also reports the Deviance Information Criteria (DIC), computed as per Spiegelhalter, Best, Carlin, and Van Der Linde (2002), and according to this measure, the bi-directional Gumbel copulas are preferred.

Fig. 1 summarises the fitted vine with Gumbel pair-copulas, although the results for the other two pair-copula types are similar. Panels (a)–(d) provide the estimates of the posterior probabilities $\Pr(\gamma^{(k)}_{l_1,l_2} = 1|\boldsymbol{y})$, which show the parsimonious form of the block functionals $K_0, \ldots, K_3$, with many component pair-copulas being independence copulas. Panels (e)–(h) provide estimates of the posterior means of Kendall's tau for each pair-copula. Twenty-four of these were negative, underlining the importance of allowing for bi-directional dependence in the pair-copulas. However, for lags $k = 2, 3$, the absolute values of Kendall's tau were less than 0.1 for all pair-copulas, suggesting that serial dependence is captured largely by the first lag.

Estimates of the Spearman correlation matrices $R_0, \ldots, R_3$ are provided in Fig. 1(i)–(l). Positive pairwise dependence is seen throughout, probably due to shared weather conditions and human activity variables (such as television programming), which are known to affect electricity demand (Smith, 2000). Two other interesting features can also be seen. First, the maximum demand in TAS, which is a separate island with a distinct climate, has almost zero dependence with those in the four mainland regions. Second, the first order Spearman autocorrelations reflect the movement of weather systems eastwards across the Australian continent following the jetstream. For example, $\rho(Y_{1,t}, Y_{2,t-1}) = 0.538 > \rho(Y_{2,t}, Y_{1,t-1}) = 0.227$, which is because weather fronts tend to pass through VIC one day prior to NSW, not the other way around. Similarly, $\rho(Y_{2,t}, Y_{4,t-1}) = 0.584 > \rho(Y_{2,t}, Y_{4,t-1}) = 0.411$ (VIC is east of SA), $\rho(Y_{1,t}, Y_{4,t-1}) = 0.537 > \rho(Y_{4,t}, Y_{1,t-1}) = 0.190$ and $\rho(Y_{1,t}, Y_{4,t-2}) = 0.482 > \rho(Y_{4,t}, Y_{1,t-2}) = 0.094$ (NSW is east of SA).

To illustrate the asymmetry in the serial dependence that is captured by the copula model, the bivariate margin of the (mean-adjusted) maximum daily demand in NSW, and that in VIC one day prior, are examined further. Fig. 2 provides a scatterplot of 100,000 iterates of $(Y_{1,t}, Y_{2,t-1})$ in light gray simulated from the fitted copula model. The positive dependence is seen easily, along with a heavy tail in the upper right hand quadrant. The horizontal and vertical stripy visual effect is due to the use of the empirical distribution function for the marginal models, and would disappear with smoother margins, as was suggested by Chen and Fan (2006). A scatterplot of the $T - 1$ observations is superimposed, showing it to be consistent with the fitted bivariate margin. The expectation $E(Y_{1,t}|Y_{2,t-1})$ from the fitted copula model is also plotted. It is estimated using a cubic smoothing spline fitted to the iterates simulated from the fitted copula (not the data points), and is nonlinear, due to the asymmetric dependence between $Y_{1,t}$ and $Y_{2,t-1}$.

Plots of the predictive distributions of $\boldsymbol{U}_t$, for $t > T$ (not shown here), reveal a substantial degree of asymmetry in the distributions themselves, especially one to three days ahead. The same is true of the predictive distributions of $\boldsymbol{Y}_t$. To assess the impact of this, the predictive distributions one to three days ahead were computed using the copula model with Gumbel pair-copulas for the last 101 days in the series. Their accuracies are measured using the cumulative rank probability score (CRPS) of Gneiting and Raftery (2007), where lower values correspond to more accurate predictive distributions. A VAR(3) with Gaussian disturbances, where the Markov order was identified using Akaike's criterion, is employed as a benchmark. The mean

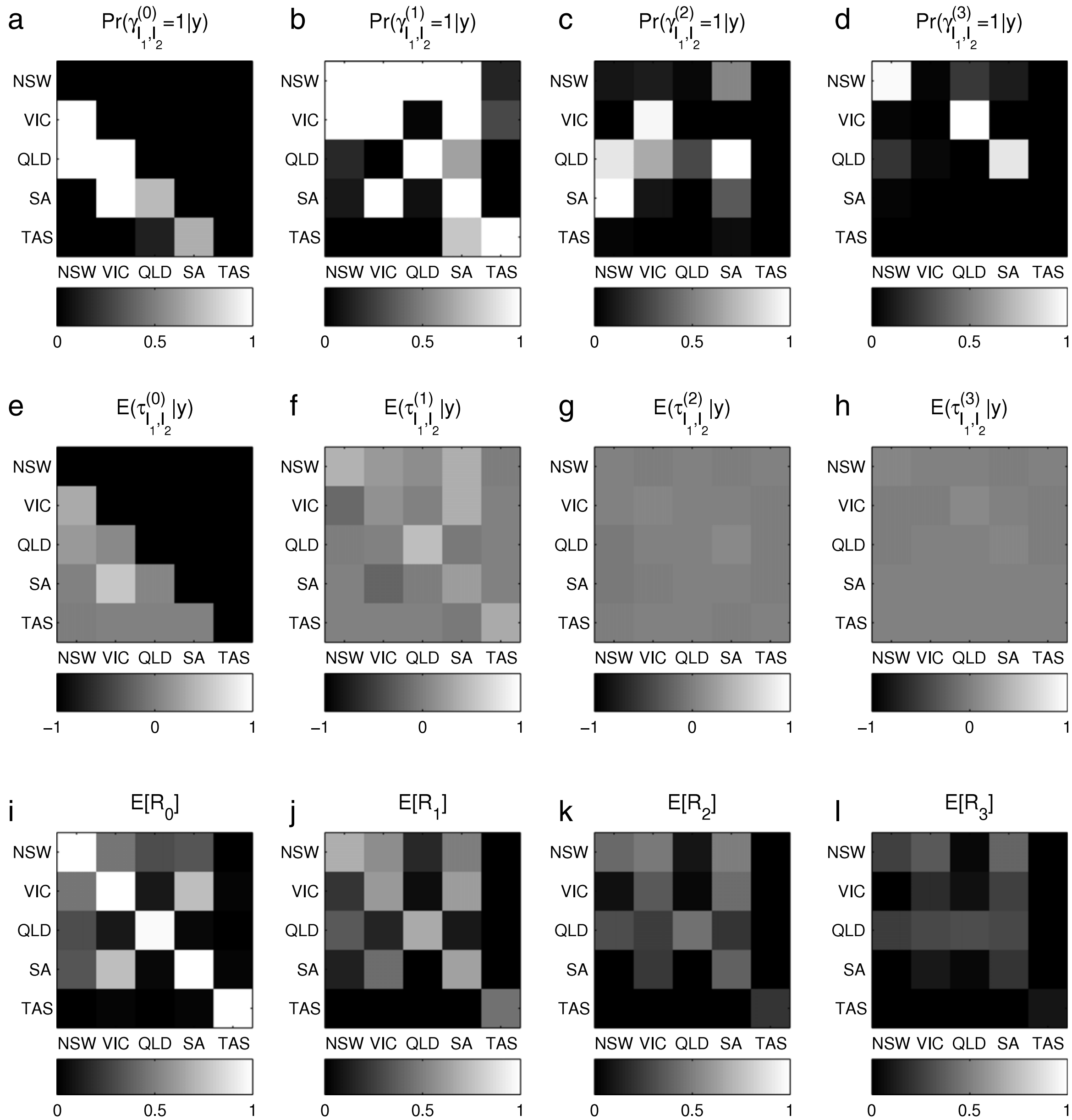


**Fig. 1.** Summary of fitted vine copula models with Gumbel pair-copulas for Example 1. In the top row, the estimated posterior probabilities $\Pr(\gamma^{(k)}_{l_1,l_2} = 1|\boldsymbol{y})$ are plotted for blocks $k = 0, \ldots, 3$. In the middle row, the estimated posterior means of Kendall's tau for each pair-copula, $\tau^{(k)}_{l_1,l_2}$, are plotted for blocks $k = 0, \ldots, 3$. In the bottom row, the estimates of the pairwise Spearman correlation matrices $R_0, \ldots, R_3$ from the fitted copula are plotted. Note that $R_0$ is symmetric, but $R_1$, $R_2$ and $R_3$ are not.

CRPS values taken across regions and samples, for one to three days ahead, were 105.2, 118.1 and 123.7 for the copula model, and 108.9, 122.0 and 128.2 for the VAR(3). This improvement in accuracy of the predictive distributions highlights the usefulness of a nonlinear time series model here.

### 5.2. Example 2: Daily electricity spot prices

To illustrate the flexibility of the approach further, a second series of the logarithm of daily electricity spot prices in each region ($m = 5$) between 1 January 2008 and 31 January 2010 ($T = 762$) is examined. Unlike many other assets, the spot price of electricity typically has a significant forecastable component; see Karakatsani and Bunn (2008) for a discussion. Of the approaches suggested, univariate nonlinear stationary time series models have proven particularly successful; see, for example, Janczura and Weron (2010) and Koopman, Ooms, and Carnero (2007). While each region has its own price setting mechanism, there is substantial inter-regional trade, and the prices exhibit a strong cross-sectional dependence. Panagiotelis and Smith (2008) show that exploiting this dependence using a multivariate time series model

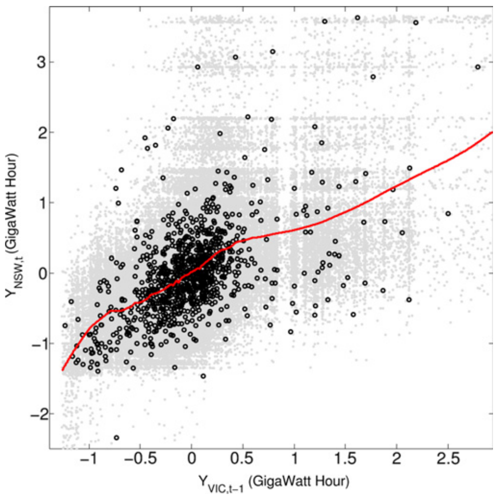


**Fig. 2.** Summary of the fitted bivariate margin of $(Y_{2,t-1}, Y_{1,t})$ in Example 1. 100,000 iterates simulated from the bivariate margin of the fitted copula model with Gumbel pair-copulas are plotted in light grey, while the superimposed dark circles form a scatterplot of the $T-1$ observations of these variables. The line is an estimate of $E(Y_{1,t}|Y_{2,t-1})$ over the range of $Y_{2,t-1}$. This is obtained using a cubic spline fit to the simulated iterates, and illustrates the nonlinear dependence between the two variables.

can improve the short-term density forecasts of prices. Moreover, many utilities are exposed to prices in multiple regions, in which case a multivariate analysis is essential.

As with the demand maxima, prices are mean-adjusted using a regression for a known day type effect. Standard linear time series tests provide strong evidence that these series are covariance stationary. Copula multivariate time series models are fitted to the data, with the marginal distributions modelled using the five empirical distribution functions. The maximal Markov order is $p = 3$, and Table 9 gives the estimates of $\Pr(\delta_k = 1|\boldsymbol{y})$ and $E(S_k/N_k|\boldsymbol{y})$. These suggest that the series has a Markov order of one or two. The DIC values are also reported, and the vine copula with Gumbel pair-copulas is optimal by this measure. The estimated Spearman correlation matrices $R_0, \ldots, R_3$ contain only positive entries, with regions that are connected directly by high voltage interconnectors (NSW, QLD), (NSW, VIC), (VIC, TAS) and (VIC, SA) having the highest price dependence. This is consistent with economic theory, where inter-regional trade produces increased price synchronisation.

To illustrate the flexibility of the copula model, Fig. 3(a) plots contours of the bivariate margin in $Y_{5,t-1}, Y_{2,t}$ computed via simulation from the fitted vine using Algorithm 2. The region VIC ($j = 2$) imports sizable amounts of electricity from TAS ($j = 5$), which has a low marginal cost of supply. The margin is bimodal, a feature that is also evident in the raw data (also plotted as a scatterplot). Bimodality is consistent with economic theory, which suggests that the market has two equilibrium price distributions, corresponding to two different cases. The first (labelled Mode A) is when the inter-regional trade in electricity is unconstrained, and the second (labelled Mode B) is when trade is constrained because the high voltage interconnectors are at their maximum capacity; see Janczura and Weron (2010) and Karakatsani and Bunn (2008). These two states also produce an asymmetric dependence structure. To highlight this, linear estimates of $E(Y_{2,t}|Y_{5,t-1})$ are also plotted for lower ($-0.75 < Y_{5,t-1} < 0.2$) and higher ($0.2 < Y_{5,t-1} < 0.8$) TAS prices. The dependence is stronger for higher prices, which is when electricity imports from TAS to VIC peak.

Two benchmark models are fitted for the comparison: a Gaussian VAR(2) and a two-state regime switching VAR(1). Both models were fit using maximum likelihood estimation, with lag lengths selected using Akaike's criterion. The equivalent bivariate margins are plotted in Fig. 3(b) and (c) for comparison, and two conclusions can be drawn. First, the elliptical contoured distributions from the fitted linear VAR(2) model cannot capture the bimodality or nonlinear dependence. Second, the copula model has bimodal fitted distributions that are similar to those identified using the regime switching model, but without the use of latent regime variables, emphasising its flexibility.

Last, Fig. 4 plots the generalized impulse-response functions, over a forecast horizon of 7 days from 31 January 2010. The impulse employed was two standard deviations of each of the five price series, and the plot is of the median and upper and lower quantiles of the distribution of $Z(t, k, j)$, as defined in Eq. (15). The results illustrate the structural role of each region in the NEM. For example, NSW is the largest importer of electricity, and the first row of Fig. 4 shows that an impulse in NSW barely affects median prices in the other regions. In comparison, VIC is the major net exporter of electricity, and the second row shows that an impulse in VIC has substantial effects on median prices in other regions. Note also that the distribution of $Z(t, k, j)$ is often asymmetric. For example, in the second row, the upper and lower quantiles are not equidistant around the median. This is due to the use of a nonlinear multivariate time series model.

## 6. Discussion

This paper suggests the employment of a high-dimensional copula for capturing both cross-sectional and serial dependence in multivariate time series data. This transforms the problem into one of modelling dependence on the unit cube. An important insight is that, because of the constraint in Eq. (2), it is not straightforward to construct the copula from the marginal copulas $c_{(t)}$ and arbitrary transitional densities $f(\boldsymbol{u}_t|\boldsymbol{u}_{t-1}, \ldots, \boldsymbol{u}_1)$. However, by employing a D-vine copula for dependence in the entire series, the marginal copulas and transitional densities can be derived in closed form instead. The D-vine nests popular elliptical copulas (Haff et al., 2010), and their use here generalises the Gaussian copula model of Biller and Nelson (2003). The bi-directional Archimedean copulas suggested here are promising pair-copula components for such a D-vine copula model when the data are likely to exhibit asymmetric dependencies.

The proposed block functional notation allows the transitional and marginal densities to be expressed succinctly. These block functionals generalise the partial correlation matrices of Tiao and Box (1981) to this nonlinear time series case, and allow for the derivation of the simple conditions required for the series to be stationary and/or

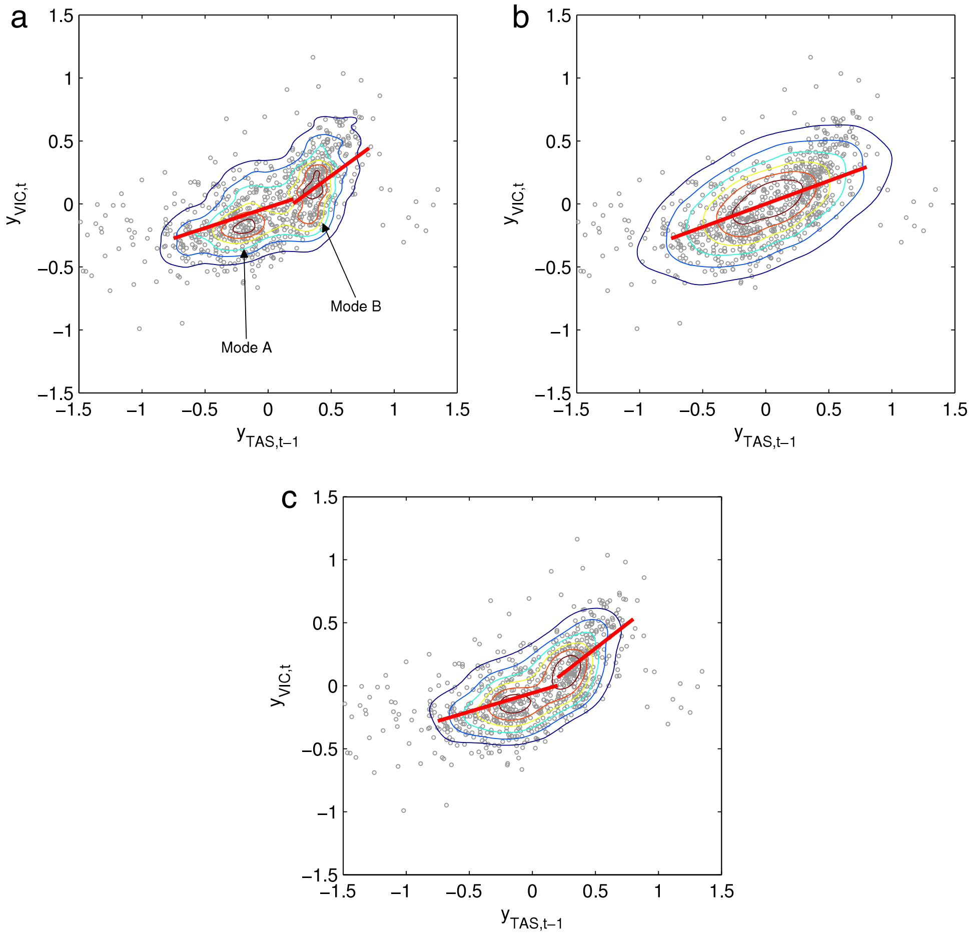


**Fig. 3.** Contour plots of the estimated bivariate distribution of $(Y_{5,t-1}, Y_{2,t})$ in Example 2, obtained from (a) the copula model; (b) a VAR; and (c) the regime-switching VAR. The contours are from bivariate kernel density estimates constructed using large Monte Carlo samples simulated from each model. Scatterplots of the 762 observations on the pairs are also included in each panel. The distributions are bi-modal in panels (a) and (c). The thick red lines denote estimates of $E(Y_{2,t}|Y_{5,t-1})$ for the two ranges $-0.75 < Y_{5,t-1} < 0.2$ and $0.2 < Y_{5,t-1} < 0.8$.

Markovian. This is important because stationary Markov series have a parsimonious D-vine representation, which is easy to estimate.

Another challenge is to develop an efficient likelihood-based estimation method. The parallel implementation of Algorithm 1 is important for computing the likelihood in practice, because of the high dimension of the copula. This insight can also be used to compute the likelihood of any D-vine copula model in parallel, including in other contexts (Aas et al., 2009). The proposed Bayesian estimation method allows for the selection of independence pair-copulas and blocks. Not only does this provide a parsimonious representation with a direct interpretation, but also Algorithms 1 and 2 are faster for sparse vines. This method extends the sampling scheme suggested previously by Smith et al. (2010). The simulation study demonstrates that the proposed model selection prior can improve the estimates of the dependence structure, and is effective at order selection. This is consistent with previous observations on Bayesian model averaging for linear VAR models by George et al. (2008) and Korobilis (2013), and the proposed method extends the work of these authors.

The first empirical application demonstrates the usefulness of the approach for modeling multivariate extrema, thereby extending recent work on the use of copulas for modelling the cross-sectional dependence of extrema in hydrology. The second empirical application highlights the nonlinear aspect of the model. Here, bimodal bivariate marginal distributions, which correspond to two economic equilibria, are captured without the aid of latent regimes. In both examples, the proposed approach identifies substantial asymmetric and heavy-tailed cross-sectional and serial dependence, which can only be captured by a nonlinear time series model. Both applications are important modeling and forecasting problems that are faced in the contemporary electricity industry, and the empirical work shows that the proposed model and estimation method are promising new tools for applied analysis.

## Acknowledgments

Michael Smith's work was supported by Australian Research Council Grant FT110100729. The author thanks Professor Shaun Vahey of Warwick University for helpful

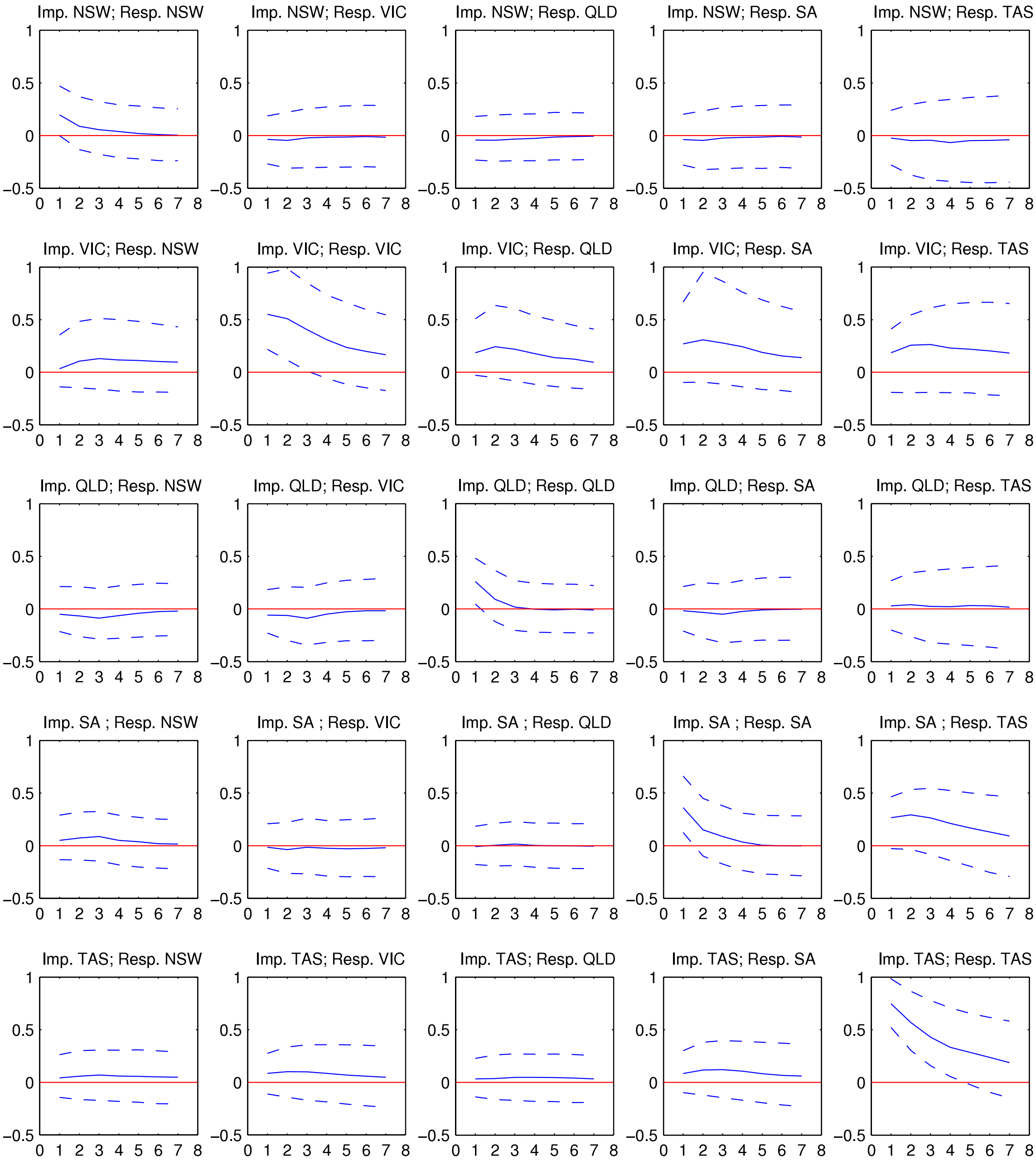


**Fig. 4.** Generalized impulse-response functions for the electricity spot prices on 31 January 2010 over a seven-day forecast horizon. The rows correspond to the region in which an impulse is made, while the columns correspond to the response in each region. In each panel, the solid line is the median of the forecast response distribution, while the dashed lines are the upper and lower quartiles of the forecast response distribution.

discussions on time series modeling in macroeconomics, Doctor Mohamad Khaled of the University of Queensland for helpful discussions on copulas, and two referees for helpful comments that improved the paper.

## Appendix

This appendix outlines an algorithm for simulating iterates from the distributions $(\boldsymbol{Y}_{t_1}, \ldots, \boldsymbol{Y}_{t_0+1} | \boldsymbol{Y}_{t_0}, \ldots, \boldsymbol{Y}_1, \boldsymbol{\phi}, \boldsymbol{\theta})$ and $(\boldsymbol{U}_{t_1}, \ldots, \boldsymbol{U}_{t_0+1} | \boldsymbol{U}_{t_0}, \ldots, \boldsymbol{U}_1, \boldsymbol{\phi})$, for $t_1 > t_0$. When $t_1 = d + 1$ and $t_0 = 0$, iterates are obtained from the distributions with densities $c_{(t-d,\ldots,t)}$ in Eq. (13) and $f(\boldsymbol{y}_{t-d}, \ldots, \boldsymbol{y}_t | \boldsymbol{\phi}, \boldsymbol{\theta})$. When $t_1 = T + T'$ and $t_0 = T$, iterates are obtained from the predictive distribution with the density in Eq. (14).

The algorithm first generates $(\boldsymbol{u}_{t_0+1}, \ldots, \boldsymbol{u}_{t_1})$ via composition by simulating the elements $u_i$ sequentially from $(U_i | U_{i-1}, \ldots, U_1, \boldsymbol{\phi})$ for $i = a(t_0 + 1), \ldots, b(t_1)$, using the inverse distribution function. Smith et al. (2010) show that, by repeated use of the recursions at Eq. (12), and by noting that $u_{i|i} \equiv u_i$, the conditional distribution function can be written as

$$F(u_i|u_{i-1},\ldots,u_1) = h_{i,1}\circ h_{i,2}\circ\cdots\circ h_{i,i-1}(u_i|u_{i-1}),$$

so that the inverse distribution function is

$$\begin{aligned} u_i &= F^{-1}(w|u_{i-1},\ldots,u_1)\\ &= h^{-1}_{i,i-1}\circ h^{-1}_{i,i-2}\circ\cdots\circ h^{-1}_{i,1}(w|u_{1|i-1}). \end{aligned} \tag{16}$$

To evaluate this efficiently, note first that $h^{-1}_{i,j}(w|u) = w$ whenever the pair-copula $c_{i,j} = 1$, while for dependent pair-copulas, $h_{i,j}$ can be inverted either analytically or quickly numerically; see Table 4 for the bi-directional Archimedean and Smith et al. (2010) for other examples. Whenever the series is Markov of order $p$, the distribution function can be simplified further as $F(u_i|u_i,\ldots,u_1) = F(u_i|u_i,\ldots,u_{i-q})$, where $q = \min((p+1)m-1, i-1)$ as in Section 3.1, so that the inverse is

$$\begin{aligned} u_i &= F^{-1}(w|u_{i-1},\ldots,u_{i-q})\\ &= h^{-1}_{i,i-1}\circ h^{-1}_{i,i-2}\circ\cdots\circ h^{-1}_{i,i-q}(w|u_{i-q|i-1}). \end{aligned}$$

To evaluate $h^{-1}_{i,j}(\cdot|u_{j|i-1})$, for $j = i-q,\ldots,i-1$, the values $u_{1|i-1},\ldots,u_{i-1|i-1}$ also need computing via the recursions in Eq. (12) in a similar manner to that in Algorithm 1. Last, iterates for $\boldsymbol{Y}_{t_0+1},\ldots,\boldsymbol{Y}_{t_1}$ are obtained by simply setting $y_{j,t} = F_j^{-1}(u_{t,j})$.

**Algorithm 2** (*Simulation*).
For $t = t_0+1,\ldots,t_1$:
For $i = a(t),\ldots,b(t)$:
Step 1. Generate $w_i \sim$ Uniform(0,1).
Step 2. Set $q = \min((p+1)m-1, i-1)$ and evaluate

$$u_i = h^{-1}_{i,i-1}\circ h^{-1}_{i,i-2}\circ\cdots\circ h^{-1}_{i,i-q}(w_i|u_{i-q|i-1}).$$

Step 3. Update the conditional distribution functions
(a) For $r = i-1,\ldots,1$, evaluate indices $(l_1,l_2,k)$ from $(i,r)$ and:
If $r < i-q+1$ and $\gamma^{(k)}_{l_1,l_2} = 1$, then set $u_{i|r} = h^{(k)}_{l_1,l_2}(u_{i|r+1}|u_{r|i-1})$, else set $u_{i|r} = u_{i|r+1}$.
(b) For $r = 1,\ldots,i-1$, evaluate indices $(l_1,l_2,k)$ from $(i,r)$ and:
If $r > i-q-1$ and $\gamma^{(k)}_{l_1,l_2} = 1$, then set $u_{r|i} = h^{(k)}_{l_1,l_2}(u_{r|i-1}|u_{i|r+1})$, else set $u_{r|i} = u_{r|i-1}$.
Step 4. Set $y_{j,t} = F_j^{-1}(u_i)$ with $j = i-a(t)+1$.

In Steps 3(a) and (b), the indices are obtained using the one-to-one transformation in Section 3.1. Similarly to Algorithm 1, a matrix $\mathcal{M}$ is needed to store the values of $u_{i|j}$ and $u_{i|j}$ that are computed in Algorithm 2, although here, this matrix is of size $(t_1 m)\times(t_1 m)$, which is larger when $t_1 > T$.

Last, when implementing this algorithm it is important to compute each $h^{-1}_{i,j}$ to a high level of numerical accuracy. In the empirical work, inversions are computed to between 8 and 12 decimal places of accuracy. For the bi-directional Archimedean copulas, inversion is numerical, and a hybrid of the Newton and Secant methods is used.

## References


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

Amjady, N. (2001). Short-term hourly load forecasting using time-series modeling with peak load estimation capability. *IEEE Transactions on Power Systems*, *16*(3), 498–505.

Anderson, T. W. (2003). *An introductions to multivariate statistical analysis* (3rd ed.). NY: Wiley.

Barnett, G., Kohn, R., & Sheather, S. (1996). Bayesian estimation of an autoregressive model using Markov chain Monte Carlo. *Journal of Econometrics*, *74*, 237–254.

Beare, B. K. (2010). Copulas and temporal dependence. *Econometrica*, *78*, 395–410.

Bedford, T., & Cooke, R. (2002). Vines — a new graphical model for dependent random variables. *Annals of Statistics*, *30*, 1031–1068.

Biller, B. (2009). Copula-based multivariate input models for stochastic simulation. *Operations Research*, *57*(4), 878–892.

Biller, B., & Nelson, B. (2003). Modeling and generating multivariate time-series input processes using a vector autoregressive technique. *ACM Transactions on Modeling and Computer Simulation*, *13*(3), 211–237.

Brechmann, E. C., & Czado, C. (2012). *COPAR — multivariate time series modeling using the COPula AutoRegressive model*. Unpublished working paper.

Chen, X., & Fan, Y. (2006). Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, *130*, 307–335.

Cogley, T., & Sargent, T. (2005). Drift and volatilities: monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics*, *8*, 262–302.

Daniels, M., & Pourahmadi, M. (2009). Modeling covariance matrices via partial autocorrelations. *Journal of Multivariate Analysis*, *100*, 2352–2363.

Demarta, S., & McNeil, A. (2005). The *t*-copula and related copulas. *International Statistical Review*, *73*, 111–129.

Domma, F., Giordano, S., & Perri, P. (2009). Statistical modeling of temporal dependence in financial data via a copula function. *Communications in Statistics*, *38*, 703–728.

Enders, W., & Sandler, T. (1993). The effectiveness of antiterrorism policies: a vector-autoregression-intervention analysis. *American Political Science Review*, *87*(4), 829–844.

Favre, A. C., El Adlouni, S., Perreault, L., Theiémonge, N., & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, *40*(12), W01101.

Fortin, I., & Kuzmics, C. (2002). Tail-dependence in stock-return pairs. *Intelligent Systems in Accounting, Finance and Management*, *11*(2), 89–107.

Garrat, A., Koop, G., Mise, E., & Vahey, S. P. (2009). Real-time prediction with U.K. monetary aggregates in the presence of model uncertainty. *Journal of Business and Economic Statistics*, *27*, 480–491.

George, E., Sun, D., & Ni, S. (2008). Bayesian stochastic search for VAR model restrictions. *Journal of Econometrics*, *142*, 553–580.

Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction and estimation. *Journal of the American Statistical Association*, *102*, 359–378.

Goebel, R., Roebroeck, A., Kim, D.-S., & Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magnetic Resonance Imaging*, *21*, 1251–1261.

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

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, *45*, 39–70.

Ibragimov, R. (2009). Copula-based characterizations for higher-order Markov processes. *Econometric Theory*, *25*, 819–846.

Janczura, J., & Weron, R. (2010). An empirical comparison of alternative regime-switching models for electricity spot prices. *Energy Economics*, *32*, 1059–1073.

Jochmann, M., Koop, G., Leon-Gonzalez, R., & Stachan, R. W. (2013). Stochastic search variable selection in vector error correction models with an application to a model of the UK macroeconomy. *Journal of Applied Econometrics*, *28*(1), 62–81.

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.

Junker, M., & May, A. (2005). Measurement of aggregate risk with copulas. *Econometrics Journal*, *8*, 428–454.

Karakatsani, N., & Bunn, D. (2008). Intra-day and regime-switching dynamics in electricity price formation. *Energy Economics*, *30*, 1776–1797.

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

Koop, G., Leon-Gonzalez, R., & Strachen, R. (2009). On the evolution of the monetary policy transmission mechanism. *Journal of Economic Dynamics and Control*, *33*, 997–1017.

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics*, *74*, 119–147.
Koopman, S. J., Ooms, M., & Carnero, M. A. (2007). Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices. *Journal of the American Statistical Association*, *102*, 16–27.
Korobilis, D. (2013). VAR forecasting using Bayesian variable selection. *Journal of Applied Econometrics*, *28*(2), 204–230.
Kurowicka, D., & Cooke, R. (2006). *Uncertainty analysis with high dimensional dependence modelling*. Wiley: New York.
Kurowicka, D., & Joe, H. (2010). *Dependence modeling: vine copula handbook*. World Scientific.
Lambert, P., & Vandenhende, F. (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.
Min, A., & Czado, C. (2010). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics*, *8*, 511–546.
Nelsen, R. (2006). *An introduction to copulas* (2nd ed.). New York: NY: Springer.
Panagiotelis, A., & Smith, M. (2008). Bayesian density forecasting of intraday electricity prices using multivariate skew *t* distributions. *International Journal of Forecasting*, *24*, 710–725.
Patton, A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, *47*, 527–556.
Patton, A. (2012). Copula methods for forecasting multivariate time series. In *Handbook of economic forecasting, vol. 2*. Springer Verlag, forthcoming.
Pitt, M., Chan, D., & Kohn, R. (2006). Efficient Bayesian inference for Gaussian copula regression models. *Biometrika*, *93*, 537–554.
Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, *72*, 821–852.
Rémillard, B., Papageorgiou, N., & Soustra, F. (2012). Copula-based semiparametric models for multivariate time series. *Journal of Multivariate Analysis*, *110*, 30–42.
Renard, B., & Lang, M. (2007). Use of a Gaussian copula for multivariate extreme value analysis: some case studies in hydrology. *Advances in Water Resources*, *30*, 897–912.
Rodriguez, J. (2007). Measuring financial contagion: a copula approach. *Journal of Empirical Finance*, *14*, 401–423.
Scott, J. G., & Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. *Annals of Statistics*, *38*(5), 2587–2619.
Shih, J. H., & Louis, T. A. (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, *51*(4), 1384–1399.
Sims, C. A. (1980). Macroeconometrics and reality. *Econometrica*, *48*, 1–48.
Smith, M. (2000). Short-term forecasting of New South Wales electricity system load. *Journal of Business and Economic Statistics*, *18*, 465–478.
Smith, M., & Kohn, R. (2002). Parsimonious covariance matrix estimation for longitudinal data. *Journal of the American Statistical Association*, *91*(460), 1141–1153.
Smith, M., Min, A., Almeida, C., & Czado, C. (2010). Modeling longitudinal data using a pair-copula decomposition of serial dependence. *Journal of the American Statistical Association*, *105*(492), 1467–1479.
Smith, M., & Vahey, S. (2013). *Asymmetric density forecasting of U.S. macroeconomic variables using a Gaussian copula model of cross-sectional and serial dependence*. Working Paper. Melbourne Business School.
Sola, M., & Driffill, J. (1994). Testing the term structure of interest rates using a vector autoregression with regime switching. *Journal of Economic Dynamics and Control*, *18*, 601–628.
Song, P. (2000). Multivariate dispersion models generated from Gaussian copula. *Scandinavian Journal of Statistics*, *27*, 305–320.
Spiegelhalter, D., Best, N., Carlin, B., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, *64*, 583–639.
Tiao, G. C., & Box, G. E. (1981). Modeling multiple time series with applications. *Journal of the American Statistical Association*, *76*, 802–816.

**Michael Stanley Smith** is a Chaired Professor of Management (Econometrics) at the University of Melbourne, and an Australian Research Council Future Fellow. His research interests include Bayesian econometric and statistical methodology, and its application to problems arising in marketing, imaging, transport science, the energy markets and elsewhere. His recent work has focused on the development and application of copula models for nonlinear dependence. His research has been published widely, including in the *Journal of the American Statistical Association, Journal of the Royal Statistical Society B, Journal of Econometrics, Journal of Business and Economic Statistics, Marketing Science, Journal of Computational and Graphical Statistics, International Journal of Forecasting, Statistics and Computing, Transportation Research Part B* and the *Journal of Applied Econometrics*.