Forecasting the Joint Distribution of Australian Electricity Prices using Dynamic Vine Copulae

Hans Manner[1], *Farzad Alavi Fard*[2], *Armin Pourkhanali,*[3] *Laleh Tafakori*[4]

June 1, 2017

Abstract

We consider the problem of modelling and forecasting the distribution of a vector of prices from interconnected electricity markets using a flexible class of drawable vine copula models, where we allow the dependence parameters of the constituting bivariate copulae to be time-varying. We undertake in-sample and out-of-sample tests using daily electricity prices, and evidence that our model provides accurate forecasts of the underlying distribution and outperforms a set of competing models in their abilities to forecast one-day-ahead conditional quantiles of a portfolio of electricity prices. Our study is conducted in the Australian National Electricity Market (NEM), which is the most efficient power auction in the world. Electricity prices exhibit highly stylised features such as extreme price spikes, price dependency between regional markets, correlation asymmetry and non-linear dependency. The developed approach can be used as a risk management tool in the electricity retail industry, which plays an integral role in the apparatus of modern energy markets. Electricity retailers are responsible for the efficient distribution of electricity, while being exposed to market risk with extreme magnitudes.

Key Words: Electricity prices, SCAR model, Dvine Copula, Back-testing, Nonlinear Dependence

[1]Institute of Economics, University of Graz, Email: hans.manner@uni-graz.at
[2]School Economics, Finance and Marketing RMIT University, Email: farzad.alavifard@rmit.edu.au
[3]Department of Banking and Finance, Monash University, Email: armin.pourkhanali@monash.edu
[4]School of Mathematics and Statistics, The University of Melbourne, Email: laleh.tafakori@unimelb.edu.au

1. Introduction

A multivariate analysis of electricity prices has become a critical issue for electricity retailers who simultaneously operate in multiple interconnected markets. Over the course of the past two decades, many countries have deregulated their electricity industries. The vertically integrated electricity industry has been restructured into three parties: power generators, transmission and distribution network service providers, and retailers. Reliance has been placed on the forces of competition to not only achieve efficient operational and investment decisions by generators, but also fairer energy prices for the end users. Further, to promote stability, local markets have formed interconnected power grids, so that electricity can be transmitted between different regions. Therefore, when the demand for electricity in a region surges, electricity can be displaced from adjoining regions, if the price is low enough. Concurrently, retail competition has been introduced through the market deregulation, such that retailers now operate in multiple markets and compete with each other over larger market shares. On the other hand, the retailers have to bid for electricity in the wholesale power auction and sell at a contractual fixed rate to the end-users. This creates an incredibly high exposure to the market risk, especially when considering the fact that extreme price movements of up to 1000% are observed frequently. Electricity, by the virtue of not being (economically) storable[5], exhibits extreme price movements (spikes) and volatility clustering. These stylised features, in tandem with the interconnection of regional markets, have turned the problem of modelling electricity prices to a challenging, yet intriguing, question.

Recent empirical studies evidence non-linear dynamic dependence in every interconnected electricity market, e.g. Nord Pool, Continental U.S. power transmission grid, and Australian National Electricity Market (NEM) (cf. Misiorek, Trück, and Weron (2006), Higgs (2009) and references therein). However, NEM is of particular interest, because electricity is traded in a constrained real time spot market, whereas in other markets electricity is traded one-day-ahead. In their seminal work, Maskin and Tirole (1988) note that day-ahead price-setting mechanisms, especially when the market consists of a few producers, endogenises the timing of production, which in turn allows the producers to collect monopoly rent. Therefore, this type of inefficiency is removed by real time trading in NEM, thus our results could be extended to other commodities (e.g. oil and gas).

The multivariate analysis of Australian electricity markets was first studied in Higgs (2009), where she employs the DCC-GARCH[6] model of Engle (2002). She found evidence for time-varying correlations between the electricity prices between Australia's different regional markets. However, empirical studies suggest that the electricity prices in Australia are extremely sensitive to small imbalances in the electricity demand and supply (cf. Higgs and Worthington 2008), often resulting in large changes in electricity prices (or spikes) and spillovers to other regional markets. This characteristic results in asymmetric dependence and violates the assumption of elliptical dependence that is the basis of classic econometric models, such as the DCC-GARCH. To address this issue, Smith, Gan, and Kohn (2012) construct a copula model from the skew t-distribution of Sujit K. Sahu and Branco (2003), and the copula approach was further explored in Ignatieva and Trück (2016). One of the important advantages of using copula based models is their ability to capture asymmetric tail dependencies. This, Ignatieva and Trück (2016) show, provides a flexible framework of modelling non-linearity in the dependence structure of electricity markets, which is particularly relevant when analysing heavy-tailed and skewed data sets (cf. Siburg, Stoimenov, and Weiß 2015). However, two major drawbacks in Ignatieva and Trück (2016) are that they focus on bivariate copulae only, and that the dependence parameter is assumed to be time-invariant. Particularly, when considering the multidimensional nature of power grids, it becomes of

[5]In recent years there has been a number of technological advancements in building electricity storage facilities, but they are not economically feasible to be used in large scales.

[6]Dynamic Conditional Correlation Generalised Autoregressive Conditional Heteroskedasticity

critical importance to consider the interconnection of markets in a system as a whole, as opposed to pairs of markets in isolation.

The literature on time-varying non-linear dependence modelling has gained a substantial momentum in the literature of financial econometrics (cf. Engle 2002, Kole, Koedijk, and Verbeek 2007, Koliai 2016). This has motivated a strand of research to construct flexible, non-standard multivariate models with the aid of *dynamic copulae*. The history of copula can be traced back to Sklar (1959), as a flexible framework, through which dependencies are modelled independently of the marginal distributions. It was then introduced by Wang (1998) in the insurance context and popularised by Li (1999) in finance literature. The early applications of copula models only considered constant copula (cf. Cherubini, Luciano, and Vecchiato 2004 for a complete review), until the seminal work of Patton (2006), where he offers an extension to the Sklar's theorem for conditional distributions forming the basis for dynamic copula modelling. This innovation has sparked a rapid growth in the literature of the theory and the applications of time-varying copulae (cf. Manner and Reznikova 2012 and references therein).

In parallel to the above advancements, the high demand for multivariate analysis of financial time-series has motivated another strand of research aiming at practically useful higher dimensional copulae, such as hierarchical Archimedean copulae by Savu and Trede (2010), factor copula models by Oh and Patton (2017a), or vine copula models by Aas, Czado, Frigessi, and Bakken (2009). In particular, the latter class has become extremely popular because of its structure that allows for the sequential estimation of a large number of parameters.[7] Vines are copulae constructed from a sequence of nested pair-copulae. Although they offer the flexibility of using any combination of pair-copulae, one challenge is to select a flexible structure that provides tractable analytical expressions. In this paper, we use 'drawable' vine, or Dvine, copulae (Aas, Czado, Frigessi, and Bakken 2009), which offer a tractable structure due to their 'closure' property. This property ensures that any adjacent sub-vector retains the Dvine structure, with bivariate copulae that are subsets of those of the initial Dvine. Most recently, Almeida, Czado, and Manner (2016) propose two copula models by combining Dvine with stochastic autoregressive copula (SCAR) and generalised autoregressive score (GAS) models (cf. Creal, Koopman, and Lucas 2013), to capture high-dimensional dependence which changes over time. The former model is termed SCAR-Dvine, which is used in this paper.

The contribution of this paper is to analyse the complex dependence structure in the NEM. We show that the SCAR-Dvine model is the most suitable model for forecasting extreme price movements, better than any other model applied to the literature before. Furthermore, the statistically determined structure of the model resembles the geographical locations of the markets, adding to the interpretability of the results. Our research design is based on the dynamic financial analysis approach, looking at an enterprise's risks holistically, as opposed to the traditional risk analysis which analyses risks individually. Modelling risk of the five markets as a complex interconnected system, as opposed to analysing markets individually, for the electricity retailers is more relevant than ever before. In late 1990s, when the Australian electricity was deregulated, the market consisted of numerous retailers operating in local markets. However, over the past 15 years, many of the small market players amalgamated into large enterprises operating in all five markets simultaneously. Many of the remaining regional retailers also went bankrupt due to poor risk management strategies, especially during the volatile market conditions between 2005 and 2011 (Australian Energy Regulator 2012).

From the financial risk management point of view, the central problem is the actual quantification of the risks. Since the adoption of Value-at-Risk (VaR) by Basel Committee in the second Basel Accord, the regulatory capital requirements of commercial banks with trading activities are based on (VaR) estimates. This important measure of market risk is defined as a pre-specified quantile of the conditional

[7]Examples of financial applications of vine copula models are Chollete, Heinen, and Valdesogo (2009), Low, Alcock, Faff, and Brailsford (2013), Pourkhanali, Kim, Tafakori, and Fard (2016), amongst many others.

distribution of portfolio returns and its estimates are routinely generated by the banks' internal models. Conventionally, VaR is defined as the quantile of the conditional distribution of the portfolio returns. However, research on electricity markets, predominantly, concentrate on log-price series (cf. Bennedsen 2017). That is because electricity is a non-storable commodity and returns cannot be defined in the traditional sense, as in other financial assets. Thus, to avoid confusion, we use the term *Conditional Quantile* (CQ), instead of VaR, which is a more general definition of the same concept.

Finally, we conduct a thorough out-of-sample back-testing exercise to evaluate the performance of the proposed SCAR-Dvine model, as well as a number of benchmark models, in forecasting CQs accurately. In order to have a conclusive analysis, we implement four statistical tests, namely, the test of Kupiec (1995), the autocorrelation test of Christoffersen (1998), the dynamic quantile test of Engle and Manganelli (2004), the duration-based test of Christoffersen and Pelletier (2004), as well as the Risk Map of Colletaz, Hurlin, and Pérignon (2013). Our study includes point estimation of eight CQs based on a set of five competing models. Further, the sort of back-testing procedures suggested here are the statistical diagnostic tests carried out on various aspects of the risk model. Therefore, it is quite natural to anticipate some mixed evidence across such a large amount of results. To assist concluding the presented evidence, we develop a simple scorecard system as a 'decision making' tool. Our scorecard system summarises the p-values from the omnibus back-testing exercise and aggregates the scores for each model to conclude the overall performance. We show that SCAR-Dvine provides appropriate and reliable out-of-sample CQ forecasts in the conducted study. This is a very significant finding, since many researchers have concluded due to the spiky and extremely volatile behaviour of electricity spot prices, no known statistical model can provide appropriate out-of-sample (conditional) quantile forecasts (cf. Bierbrauer, Menn, Rachev, and Trück (2007), Weron and Misiorek (2008), Ignatieva and Trück (2016) and reference therein)

The remainder of this paper is structured as follows. The model and its estimation are discussed in Section 2. Section 3 introduces the data and the in-sample estimation results. Section 4 discusses the out-of-sample backtesting of the conditional quantiles and Section 5 concludes.

2. The Model

2.1. Seasonal and Weekly Periodicity

A preparatory, but important, issue in fitting stochastic models to electricity spot prices is the estimation of a component to deal with trends and seasonality in the data. Seasonal fluctuations in electricity demand are well documented (cf. Pilipovic 2007 and references therein). As a result, electricity prices tend to follow a similar periodicity. The seasonal patterns in electricity prices are intrinsically different to those observed in other commodity markets, which has motivated the growth of a body of literature on different methods of capturing the seasonal periodicity. Early investigations in modelling seasonal patterns often lacked robustness (or elegance). For example, Garcia, Contreras, Van Akkeren, and Garcia (2005) use many AR terms (more than 500) to model the mean process and do not distinguish between long-term and short-term trends. The recent literature notes that the long-term seasonal pattern is often the more complex part of the mean process to model (cf Janczura, Trück, Weron, and Wolff 2013), so the distinction between the long- and short-term seasonal patterns is important.

Janczura, Trück, Weron, and Wolff (2013) provide a comprehensive review of the literature on different methodologies available to capture the long-term seasonal components of electricity prices. They classify the available models into three larger categories, namely 1) piecewise constant functions (cf. Higgs and Worthington 2008, Cartea and Villaplana 2008), 2) wavelet filters (cf. Nowotarski, Tomczyk, and Weron 2013), and 3) sinusoidal functions (cf. Bierbrauer, Menn, Rachev, and Trück 2007, Keles, Hartel, Möst, and Fichtner 2012).

The first class of models, typically, yields a non-smooth seasonal component which is not appropriate for complex empirical studies. Most of the recent literature either uses the wavelet filter, or a version of sinusoidal functions. Janczura, Trück, Weron, and Wolff (2013) show that a combination of an exponentially weighted moving average and a sinusoidal function (sin-EWMA) does accurately capture the seasonal pattern in the electricity price data. They further compare their results to that obtained from a wavelet filter and found that the ability of both models to capture the seasonal component of electricity prices are very similar.

Janczura, Trück, Weron, and Wolff (2013), also note that the estimation routines for the long-term and short-term seasonal patterns are usually quite sensitive to extreme observations, known as electricity price spikes. They recommend that the robustness of models can be improved by filtering the data with some reasonable procedure for outlier detection, before using estimation and testing procedures on the filtered data. In this study, we use the sin-EWMA model to capture the long-term periodicity in electricity log-price series, together with the recursive filtration procedure proposed in Janczura, Trück, Weron, and Wolff (2013) to treat outliers.

We denote by $P_t = (p_{1,t}, \ldots, p_{d,t})$ the vector of electricity prices at time $t = 1, \ldots, T$, for the $d = 5$ states NSW, Qld, SA, TAS, and VIC, receptively. Furthermore, $Y_t = (y_{1,t}, \ldots, y_{d,t})$ denotes the vector of log-price, such that $y_{i,t} := \ln(p_{i,t})$. The logarithmic transformation is especially suitable and commonly used for extremely spiky and non-negative spot prices (cf. Janczura, Trück, Weron, and Wolff 2013). Additionally, it makes the spot price distribution more symmetric and less leptokurtic, hence better suited for statistical analysis.

The vector of log-prices, Y_t. consists of two independent parts: a trend-seasonal component S_t and a stochastic component X_t, thus,

$$Y_t = S_t + X_t. \tag{1}$$

Further, we let S_t be composed of two components 1) a short-term (weekly) seasonal component ST_t, and 2) a long-term trend-seasonal component LT_t, i.e. $S_t = ST_t + LT_t$. The latter component represents long-term non-periodic fuel price levels, changing climate related consumption conditions throughout the years and variation in bidding practices on the strategic level (e.g., due to changes in the generation portfolio).

The model for LT_t used in this study is given by:

$$LT_t = b_1 \sin\left\{2\pi\left(\frac{t}{365} + b_2\right)\right\} + b_3 + b_4 EWMA^{\lambda}, \tag{2}$$

where λ denotes the decay factor λ, such that

$$EWMA_t^{\lambda} = (1-\lambda)\cdot P_t + \lambda \cdot EWMA_{t-1}^{\lambda}.$$

The parameters $b_i, i \in \{1,2,3,4\}$ and λ are estimated using maximum likelihood estimation.

The price series without the long-term trend is obtained by subtracting LT_t from Y_t. Next, the short-term (weekly) periodicity ST_t can simply be removed by subtracting weekly averages of the long-term deseasonalised log-prices corresponding to each day of the week. This approach is equivalent to having dummies for each day of the week, as in De Jong (2006). Note that the median or a truncated mean can be used instead of the mean as an alternative more robust to outliers; however, usually the differences are not substantial. Public holidays are treated as the eighth day of the week. It should be noted that electricity consumption profiles on working days directly preceding/following holidays or enclosed by a public holiday and the weekend may possibly resemble holiday consumption patterns more than working day profiles.

2.2. DVine Stochastic Autoregressive Copula (SCAR-Dvine) Model

We consider the following model for the joint (conditional) distribution of the deseasonalised log-prices $X_t = (x_{1,t}, \dots, x_{d,t})$ for $t = 1, \dots, T$. We assume that each variable $x_{i,t}$ follows an $AR(1) - GARCH(1,1)$ process, i.e.

$$\begin{aligned} x_{i,t} &= \mu_{i,t} + \sigma_{i,t}\varepsilon_{i,t} \\ \mu_{i,t} &= \beta_{i,0} + \beta_{i,1}x_{i,t-1} \\ \sigma_{i,t}^2 &= \alpha_{i,0} + \alpha_{i,1}\varepsilon_{i,t-1}^2 + \gamma_i\sigma_{i,t-1}^2. \end{aligned} \tag{3}$$

The usual stationarity conditions are assumed to hold. Denote the joint distribution of the standardised innovations $\varepsilon_{i,t}$ by $G(\varepsilon_{1,t}, \dots, \varepsilon_{d,t})$ and let their marginal distributions be $F_i(\varepsilon_{i,t})$ for $i = 1, \dots, d$. Note that the marginal distributions are specified as the skew-t distribution by Hansen (1994) parametrised by the degrees-of-freedom ν_i and the skewness parameter λ_i. By Sklars theorem there exists a copula C such that

$$G(\varepsilon_{1,t}, \dots, \varepsilon_{d,t}) = C\left(F_1(\varepsilon_{1,t}), \dots, F_d(\varepsilon_{d,t})\right) \tag{4}$$

Since all the marginal behaviour is captured by the marginal distributions, the copula captures the complete contemporaneous dependence of the distribution. Let $u_{i,t} = F_i(\varepsilon_{i,t})$ be the innovations transformed to $U(0,1)$ random variable and define $\boldsymbol{u}_t := (u_{1,t}, \dots, u_{d,t})$. Since the copula distribution is dynamic, then

$$\boldsymbol{u}_t \sim c(\boldsymbol{u}_t; \boldsymbol{\omega}, \mathcal{F}_{t-1}), \tag{5}$$

where $c(\cdot)$ is the copula density, $\mathcal{F}_{t-1}$ the information set at time $t-1$ and $\boldsymbol{\omega}$ the vector of time invariant parameters.

The challenge in specifying an appropriate parametric model for the copula is that this model must be able to capture time-varying dependence in a flexible way, that it should be able to capture different types of dependence such as (potentially asymmetric) tail dependence and that it must possible to handle such a model for higher dimensional data ($d = 5$ in our case). Such a model was proposed by Almeida, Czado, and Manner (2016) relying on drawable vine copula constructions (Dvine) and stochastic latent dependence parameters for the bivariate dynamic copula models entering the pair copula construction.

As in Czado (2010), we define a d-dimensional dynamic Dvine copula using the recursive decomposition of a multivariate density:

$$c(u_{1,t}, \dots, u_{d,t}; \boldsymbol{\theta}_t) := \prod_{j=1}^{d-1}\prod_{i=1}^{d-j} c_{l(i,j)}\left(F(u_{i,t}|\boldsymbol{u}_{i+1:i+j-1,t}; \boldsymbol{\theta}_t^{l(i,j)}), F(u_{i+j,t}|\boldsymbol{u}_{i+1:i+j-1,t}; \boldsymbol{\theta}_t^{l(i,j)})\right), \tag{6}$$

where $l(i,j) := i, i+j|i+1 : i+j-1$ and $\boldsymbol{\theta}_t := \{\boldsymbol{\theta}_t^{l(i,j)}; j = 1, \dots, d-1, i = 1, \dots, d-j\}$ is the time-varying copula parameter vector. Here $c_{l(i,j)}(\cdot, \cdot; \boldsymbol{\theta}_t^{l(i,j)})$ is the bivariate copula density corresponding to the bivariate dynamic copula.

Most copula models in the literature consider the copula dependence parameters to be fixed in time. To capture possible dynamics, Patton (2006) extends Sklar's theorem for conditional distributions and proposes the use of time-varying copulae, with parameters following a restricted nonlinear ARMA model. da Silva Filho, Ziegelmann, and Dueker (2012) allow the ARMA dynamics for the dependence parameters to be driven by a hidden Markov Chain. For higher dimensions, Oh and Patton (2017b) use the GAS model, proposed by Creal, Koopman, and Lucas (2013), to describe time dynamics of a factor copula. Heinen, Valdesogo Robles, et al. (2009) build on the DCC GARCH model for specifying dynamic dependence and adapt the approach to copula models. For elliptical copulas this approach is readily available in higher dimensions and below we consider this model based on the Gaussian and Student copulas as benchmark models. Different approaches for time-varying copula models are surveyed in Manner and Reznikova (2012). The choice of the most appropriate model requires a level of pragmatism, in particular, when considering high dimensional problems. As discussed in Oh and Patton (2017a), one should strike a balance between number of model parameters, flexibility, and computation time.

We assume that $\theta_t^{l(i,j)}$ is driven by an unobserved stochastic process $\lambda_t^{l(i,j)}$ such that $\theta_t^{l(i,j)} = \Gamma(\lambda_t^{l(i,j)})$, where $\Gamma : \mathbb{R} \to \Theta$ is an appropriate transformation to ensure that the copula parameter remains in its domain and whose functional form depends on the choice of copula. The underlying process $\lambda_t^{l(i,j)}$ driving the dependence parameter, which is unobserved, is assumed to follow a latent Gaussian $AR(1)$ process given by

$$\lambda_t^{l(i,j)} = \mu_{l(i,j)} + \phi_{l(i,j)}\lambda_{t-1}^{l(i,j)} + \eta_{l(i,j)} z_t^{l(i,j)}, \tag{7}$$

where $z_t^{l(i,j)}$ are independent standard normal innovations. We further assume $|\phi_{l(i,j)}| < 1$ for stationarity and $\eta_{l(i,j)} > 0$ for identification. This model specification is known as stochastic autoregressive copula (SCAR), which was first proposed in Hafner and Manner (2012).

For the parametric bivariate copulae we consider the Gaussian, Gumbel, and Clayton copulas and rotated versions of the latter two.[8] For the Gaussian copula we apply the inverse Fisher transform to ensure that correlations are in the interval $(-1, 1)$,

$$\Gamma(\lambda_t^{l(i,j)}) := \frac{\exp(2\lambda_t^{l(i,j)}) - 1}{\exp(2\lambda_t^{l(i,j)}) + 1}. \tag{8}$$

For the Gumbel and Clayton copulae, as well as their rotated versions, we first transform $\lambda_t^{l(i,j)}$ to Kendall's τ using the inverse Fisher transformation and then we use the relationship between Kendall's τ and the copula parameter $\theta = r(\tau)$ to obtain the implied parameter of these copulas [9].

$$\boldsymbol{\theta}_t^{l(i,j)} = r(\tau_t^{l(i,j)}) := r(\Gamma(\lambda_t^{l(i,j)})) \tag{9}$$

where, as above, $\Gamma(x)$ is the inverse Fisher transform that maps $\lambda_t^{l(i,j)}$ into $(-1, 1)$, the domain of $\tau_t^{l(i,j)}$. Whenever τ_t is negative we rotate the copula by 90° to allow the models to have negative dependence.

The use of the Gaussian SCAR model instead of tail dependent copulas such as the t-copula is justified by the result of Manner and Segers (2011). There it is shown that Gaussian copulas with stochastic correlations have the property of *near asymptotic dependence*, which implies that the dependence in the tails is basically indistinguishable from true tail dependence at any practically relevant quantiles. This translates into an excellent fit of such models for financial data and we expect this also to be the case for electricity prices.

[8] A rotation by 180°, e.g., transforms the lower tail dependent Clayton copula into an upper tail dependent copula. A rotation by 90°, on the other hand, transforms a copula that only allows for positive dependence into one that allows only for negative dependence. In fact, we use this approach to extend the Gumbel and Clayton copulas to allow the time-varying parameter to correspond to negative dependence as well.

[9] For the Gumbel family the relationship is $\theta = 1/(1-\tau)$ and for the Clayton copula $\theta = 2\tau/(1-\tau)$.

Finally, similar to tail properties of Dvines studied in Joe, Li, and Nikoloulopoulos (2010) the choice of time-varying pair copulae in the first tree propagates to the whole distribution, in particular all pairs of variables have an induced time-varying Kendall's τ if the copulas on the first tree are specified to be time-varying. In order to keep the models tractable and reduce the number of parameters to be estimated we also allow for static copulas as pair copulas and in the set of static models we include the t-copula, as well as the independence copula. The particular choice for each pair copula is made using the Bayesian information criterion (BIC).

2.3. Sequential Estimation of SCAR-Dvine models

We need to estimate the parameters of both the marginal models and the stochastic copula models. The joint density of our model is given by the product of the marginal and the copula densities. The estimation process involves two steps. First, the marginal parameters for the conditional mean and variance as specified in (3), as well as the parameters of the skewed-t distribution for the errors, are estimated separately and then the standardised residuals are formed, for which we employ a parametric probability integral transformation as in Joe (2005). The transformation must be applied with due care, as Kim, Silvapulle, and Silvapulle (2007) note that the gross misspecification of the marginal models impairs the accuracy of the approximation to the true copula data. We perform goodness-of-fit tests to ensure that this is not the case. Second, the parameters of the dependence model are estimated sequentially for each pair-copula. This subsection provides a brief discussion of the latter. Aas, Czado, Frigessi, and Bakken (2009) show that the form of the Dvine density given in (6) allows for a sequential parameter estimation approach starting from the first tree until the last tree. The technique was further developed and expanded for dynamic Dvine in Almeida, Czado, and Manner (2016).

We are interested in estimating the copula parameter vector $\boldsymbol{\omega} := (\mu, \phi, \sigma)$. For convenience, and with slight abuse of notations, we drop the indices i and j, wherever it does not cause ambiguity. Denote $\boldsymbol{u}_i = \{u_{i,t}\}_{t=1}^T$, $\boldsymbol{u}_j = \{u_{j,t}\}_{t=1}^T$ and $\Lambda = \{\lambda_t\}_{t=1}^T$, and let $f(\boldsymbol{u}_i, \boldsymbol{u}_j, \Lambda; \boldsymbol{\omega})$ be the joint density of the observable variables $(\boldsymbol{u}_i, \boldsymbol{u}_j)$ and the latent process Λ. Further, let $\Lambda_t := \{\lambda_\tau\}_{\tau=1}^t$, and similarly for $\boldsymbol{u}_{j,t}$ and $\boldsymbol{u}_{i,t}$. Then the likelihood function of the parameter vector $\boldsymbol{\Omega}$ can be obtained by integrating the latent process Λ out of the joint likelihood. Richard and Zhang (2007) propose an efficient importance sampling (EIS) procedure to approximate the likelihood function, taking advantage of an auxiliary sampler $\{m(\lambda_t|\Lambda_{t-1}, a_t)\}_{t=1}^T$, indexed by the auxiliary parameters a_t. Then, the likelihood can be rewritten as

$$L(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \int \prod_{t=1}^{T} \left[\frac{f(u_{i,t}, u_{j,t}, \lambda_t | \boldsymbol{u}_{i,t-1}, \boldsymbol{u}_{j,t-1}, \Lambda_{t-1}, \boldsymbol{\omega})}{m(\lambda_t|\Lambda_{t-1}, a_t)} \right] m(\lambda_t|\Lambda_{t-1}, a_t) d\Lambda.$$

This can be evaluated using N trajectories $\{\tilde{\lambda}_t^{(i)}(a_t)\}_{t=1}^T$ drawn from the importance sampler m by

$$\tilde{L}(\boldsymbol{\omega}; \boldsymbol{u}_i, \boldsymbol{u}_j) = \frac{1}{N} \sum_{s=1}^{N} \left(\prod_{t=1}^{T} \left[\frac{f(u_{i,t}, u_{j,t}, \tilde{\lambda}_t^{(s)}(a_t) | \boldsymbol{u}_{i,t-1}, \boldsymbol{u}_{j,t-1}, \tilde{\Lambda}_{t-1}^{(s)}(a_{t-1}), \boldsymbol{\omega})}{m(\tilde{\lambda}_t^{(s)} | \tilde{\Lambda}_{t-1}^{(s)}, a_t)} \right] \right).$$

$\{\tilde{\lambda}_t^{(i)}(a_t)\}_{t=1}^T$ is a natural sampler that are sampled independently of the observed variables $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$. For the details concerning the implementation of the EIS algorithm the interested reader is referred to Liesenfeld and Richard (2003). Dvine copula is the product of bivariate (conditional) copulae. Estimation of all copula parameters of our model in one step is computationally expensive, due to the large number of parameters. Thus, we utilise a sequential procedure of the copula parameters, similar to Almeida, Czado, and Manner (2016). To begin with, we need a fast recursive way to compute conditional cdf's which enter as arguments. Let $\varepsilon_D := (\varepsilon_{i_1}, \ldots, \varepsilon_{i_k})$, then by the virtue of margins being uniform in Dvine, we can express $F(\varepsilon_j|\varepsilon_D)$ as

$$F(\varepsilon_j|\varepsilon_D) = h\left(F(\varepsilon_j|\varepsilon_{D_{-i}})|F(\varepsilon_i|\varepsilon_{D_{-i}}), \boldsymbol{\theta}_{ij|D_{-i}}\right)$$

where

$$h(u_j|u_i, \boldsymbol{\theta}_{ij}) := \frac{\partial C_{i,j}(u_i, u_j; \boldsymbol{\theta}_{ij})}{\partial u_i} \tag{10}$$

The form of the Dvine density given in (6) allows for a sequential parameter estimation approach starting from the level 1 tree until the level $d-1$ tree. The algorithm begins by estimating the parameters corresponding to the pair-copulae in the first level. For the copula parameters in level 2, we first transform the data with the h function in (10) required for the appropriate conditional cdf using estimated parameters to determine pseudo realisations needed in the level 2 tree. Using these pseudo observations the parameters in the second level are estimated, the pseudo data is again transformed using the h function and so on.

Naturally, in the level 1 tree one would assume that time-varying dependence can be computed and inserted into the corresponding (time-dependent) h functions to compute the pseudo observations. Nevertheless, in the SCAR model, given that the bivariate models in the level 1 tree have been estimated, it is not possible to apply the h function directly to obtain the pseudo observations that are needed to obtain the parameters on the second tree. The problem can be circumvented by calculating the pseudo observations, using N simulated trajectories $\tilde{\theta}_t^{(s)}$ from the importance sampler:

$$u_{j|i,t} = \frac{1}{N}\sum_{s=1}^{N} h(u_{j,t}|u_{i,t}, \tilde{\theta}_t^{(s)}). \tag{11}$$

Even though the parameters of the underlying process are of interest themselves, ultimately we wish to get estimates of the latent process Λ and transformations thereof. In particular, we are interested in estimating $\tau_t = \Gamma(\lambda_t)$ for $t = 1\ldots, T$, where $\Gamma(\cdot)$ denotes the inverse Fisher transform given in (8). Smoothed estimates of $\Gamma(\lambda_t)$ given the entire history of the observable information $\boldsymbol{u}_i$ and $\boldsymbol{u}_j$ can be computed as

$$E[\Gamma(\lambda_t)|\boldsymbol{u}_i, \boldsymbol{u}_j] = \frac{\int \Gamma(\lambda_t) f(\boldsymbol{u}_i, \boldsymbol{u}_j, \Lambda; \boldsymbol{\omega}) d\Lambda}{\int f(\boldsymbol{u}_i, \boldsymbol{u}_j, \Lambda; \boldsymbol{\omega}) d\Lambda}. \tag{12}$$

Note that the denominator in (12) corresponds to the likelihood function with $t-1$ observations, which is $L(\boldsymbol{\omega}; \boldsymbol{u}_{i,t-1}, \boldsymbol{u}_{j,t-1})$. Then both integrals can be evaluated using draws from the importance sampler $m(\lambda_t|\Lambda_{t-1}, \boldsymbol{a}_t)$.

Finally, in order to obtain 1-step ahead predictive distributions we predict the conditional mean and variances, which is straightforward, as well as the copula parameters. The dependence parameters are predicted for each level as described in Hafner and Manner (2012). Using the parameters from the predictive distribution we simulate a large number of random draws from the predictive distribution using the algorithm form Aas, Czado, Frigessi, and Bakken (2009), which allow us to construct the predictive distribution of a portfolio.

3. Empirical Analysis: Dependence Modelling and In-Sample Estimation

In this section, we present the empirical results of fitting our model to the electricity prices from five Australian regional Markets, namely, New South Wales (NSW), Queensland (QLD), South Australia (SA), Tasmania (TAS), and Victoria (VIC). For benchmarking the forecasting performance of the SCAR-Dvine copula model, we use four competing models, namely,

1. the DCC-GARCH model of Engle (2002),
2. the DCC-Copula-GARCH model with Gaussian copula (cf. Heinen, Valdesogo Robles, et al. 2009),
3. the DCC-Copula-GARCH model with Student's t copula (cf. Heinen, Valdesogo Robles, et al. 2009), and
4. the static Dvine copula model (cf. Aas, Czado, Frigessi, and Bakken 2009, Brechmann, Czado, and Aas 2012).

The first model is studied as it was applied to Australian electricity prices by Higgs (2009). The DCC-Copula-GARCH models are fairly flexible time-varying specifications based on elliptical copulas that have the advantage of being relatively easy to implement at a low computational cost. The last benchmark model is included to assess whether time-varying parameters do indeed lead to improved distributional forecasts. For the copula models we use the same process of choosing marginal distributions as used for the SCAR-Dvine model. For simplicity, we label the the second and third models N-DCC-Copula and t-DCC-Copula, respectively. A natural question is whether more complex models lead to improved predictions of the joint density of electricity prices. The first three models allow for time-varying dependence, but rely on elliptical models that do not allow for asymmetries. The static, Dvine, on the other hand, does potentially capture asymmetric tail dependence, but as the name suggests, not the dynamic dependence. Thus we will be able to judge which of these characteristics, all of which are captured by the SCAR-Dvine model, are necessary for appropriate density forecasts of electricity prices.

We estimate the SCAR-Dvine model using a simulated maximum likelihood, programmed in MATLAB. The computation time for the estimation using the in-sample period is about 2 hours on a computer with an $i5$ core processor. The most computationally-expensive part of the algorithm is the out-of-sample back-testing, discussed in the next section. Here, we just mention that due to the large number of iterations in the problem, we performed the computations in parallel on a cluster with 32 core processors, taking close to 40 hours for all four models.

3.1. The Australian National Electricity Market (NEM)

As a wholesale market the NEM in Australia began operating in December 1998, forming a grid between five regional market jurisdictions (New South Wales, Queensland, South Australia, Tasmania, and Victoria)[10]. In the market, electricity load is displaced when the price of electricity in an adjoining state (or country) is low enough, or the demand in a particular region is significantly higher than the amount of electricity that can be provided by local generators. Electricity transmission involves three parties; namely, 1) power plants (or generators), 2) electricity retailers, 3) consumers (commercial or residential); but only the first two engage in trading electricity in NEM. Transmission between power plants

[10]Western Australia and Northern Territory are not connected to the NEM

Introduction	NSW	QLD	SA	TAS	VIC	Portfolio
Panel (a): Descriptive statistics of the full sample (2010-2015)						
Number of observations	2190	2190	2190	2190	2190	2190
Minimum	2.85	-1.11	0.11	1.29	1.94	2.1207
Maximum	7.16	7.54	7.76	6.69	7.15	5.6232
Mean	3.6158	3.6053	3.6731	3.5837	3.5397	3.6035
Std. Deviation	0.3870	0.5486	0.5535	0.3879	0.4365	0.3746
Skewness	1.647	.847	0.922	0.631	1.287	0.2160
Kurtosis	9.608	9.209	5.946	5.286	7.481	0.2267
Jarque–Bera	9370	7960.3	3518.2	2680	5684	21.566
(p-value)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Panel (b): Descriptive statistics of the in-sample period (2010-2012)						
Number of observations	1095	1095	1095	1095	1095	1095
Minimum	2.85	-1.11	0.11	1.29	1.94	2.1207
Maximum	7.16	6.97	7.76	6.69	7.15	5.6232
Mean	3.4601	3.3758	3.4602	3.4280	3.3939	3.4235
Std. Deviation	0.4244	0.4775	0.5218	0.3949	0.45918	0.3662
Skewness	2.968	0.4161	1.834	1.301	2.490	1.1797
Kurtosis	15.649	18.090	14.246	10.292	13.627	3.0188
Jarque–Bera	12666	14818	9780.1	5091.8	9516.4	663.83
(p-value)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Panel (c): Descriptive statistics of the out-of-sample period (2013-2015)						
Number of observations	1095	1095	1095	1095	1095	1095
Minimum	2.96	1.01	1.62	2.21	2.45	2.8362
Maximum	6.16	7.54	6.69	4.94	6.21	4.8377
Mean	3.7714	3.8348	3.8861	3.7394	3.6855	3.7834
Std. Deviation	0.2668	0.51840	0.4999	0.3108	0.3575	0.2863
Skewness	0.8140	1.4840	0.5350	0.7150	0.2120	-0.2983
Kurtosis	8.228	8.215	3.798	2.524	4.050	0.2592
Jarque–Bera	3177.3	3447.2	702.12	379.72	747.71	19.107
(p-value)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 1: Descriptive statistics of daily log-prices.

and consumers is facilitated through electricity retailers, who have contractual supply agreements, fixed in price, with the consumers, but purchase electricity from the spot market. NEM acts as an exchange between generators and retailers, where the output from all power plants are aggregated and instantaneously scheduled to meet demand through a centrally-coordinated dispatch process. The Australian Energy Market Operator (AEMO) coordinates the demand-dispatch process in compliance with the provisions of *Australian National Electricity Law* and *Australian National Electricity Rules*. Every five minutes, AEMO receives offers from scheduled generators, then determines the required generation capacity based on the prevailing demand in the most cost-efficient way. Subsequently, a dispatch price is determined and automatic generation control target signals are sent to the generators. For the purpose of the settlement of all financial transactions, spots or derivatives, AEMO uses the dispatch prices; thus, 'dispatch prices' and 'spot prices' are interchangeably used.

We use daily spot prices from five electricity markets in Australia: NSW, QLD, SA, TAS and VIC. For each market, the sample of 2190 daily observations covers the time period from 01.01.2010 to 31.12.2015. We split the data into an in-sample period covering the years 2010-2012 and an out-of-sample period covering 2013-2015. Table 1 provides the summary statistics for the log-prices and an equally weighted portfolio of all five markets, as well as the Jarque–Bera statistics showing that they all reject the normal distribution at the 1% level of significance. This documents the well accepted stylised features of electricity log-prices, which has been reported in the literature (c.f. Janczura, Trück, Weron, and Wolff 2013).

This research aims to model the dynamic and nonlinear dependence between different regional electricity markets. There now exists a wide-spread consensus in the empirical literature that the dependence between the returns of financial assets is asymmetric, nonlinear, and time-varying. In particular, there is a growing body of evidence that electricity prices in different interconnected grids exhibit such complex dependence between local markets (cf. Misiorek, Trück, and Weron 2006, Smith, Gan, and Kohn 2012, Lindström and Regland 2012, Ignatieva and Trück 2016, among others). The upper triangle of Figure **??**, plots the Kendall's τ of different pairs of electricity log-prices from different regional markets for the full sample implementing a rolling window approach with lags of 50 days. For all pairs, we can see significant variation of the dependence measure over time, which strongly supports our model assumptions. Despite the wide use of the Pearson correlation coefficient, there are several undesirable properties associated with its use (cf. Dietrich 1991), due to which we choose Kendall's τ as a better alternative for measuring dependence. Kendall's τ, is a well accepted measure of dependence, which assess statistical associations based on the ranks of the data. It provides, arguably, the best alternatives to the linear correlation coefficient as a measure of dependence for non-elliptical distributions, for which the linear correlation coefficient is inappropriate and often misleading. Further, Kendall's τ plays a special role in the parametrisation of copula functions (see Subsection 2.2).

$<<$**Insert Figure ??**$>>$

One of the fundamental challenges in analysing portfolios of financial assets is the presence of *asymmetric correlations* between them, which means assets move more often with the market when the market goes down than when it goes up (cf. Pourkhanali, Kim, Tafakori, and Fard 2016 and references therein). Asymmetric correlation, also, manifests itself in electricity price data in the form of the shocks' spill-over from one regional market to another. Intuitively, prolonged periods of excess demand are expected to impose more severe distress on the grid than periods of excess supply. This is because the price elasticity of electricity demand is extremely low, but the elasticity of supply is typically slightly greater than unit elasticity (cf. Moral-Carcedo and Vicens-Otero 2005, Chang, Kim, Miller, Park, and Park 2016). For example, often industrial consumers with mission critical support services (MCSS) directly compete with power retailers in the market. MCSSs guarantee uninterrupted power supply to their clients (e.g. data centres), thus, when a regional market experience a high load of excess demand, they instantly place bids in neighbouring markets. These bids are always limit-to-market, and the sheer load-per-hour required for these infrastructures catapult the prices, propagating shocks through the grid.

In the lower triangle of Figure **??** we present the *exceedance correlation plots* of different pairs of log-prices series, based on the method proposed in Ang and Chen (2002) and Hong, Tu, and Zhou (2007). Exceedance correlation captures three features of the dependence relationship in the joint distribution, namely, correlation levels, correlation asymmetry, and tail-dependence, all of which are evident in the data. The presence of such strong asymmetric correlations can cause problems in hedging effectiveness. Additionally, it hampers the effectiveness of diversification for the electricity retailers, through expanding operations into different states.

3.2. Seasonality and the Marginal Models

As discussed in Section 2.1, we apply the recursive filtration procedure proposed in Janczura, Trück, Weron, and Wolff (2013) to treat outliers, before estimating the seasonal trend. This is because the analysis will be more robust, as the estimation procedure can be sensitive to extreme observations. The parameter estimation of the sinusoidal pattern is performed on the full-sample using a maximum likelihood estimator. The results are reported in Table 2, Panel A. Figure **??** presents the time series of the log-prices, adjacent to their deseasonalised time series.

Panel A: Parameter Estimation for sin-EWMA model

	λ	b_1	b_2	b_3	b_4	MSE
NSW	0.3930	-0.0380	4.9476	-0.1654	1.0453	0.1475
Qld	0.2550	0.0671	7.4587	-0.2155	1.0593	0.2720
SA	0.0490	-0.0876	6.9011	-0.1147	1.0314	0.3991
Tas	0.2490	0.0214	4.4398	-0.1904	1.0531	0.1849
Vic	0.3150	0.0721	7.449	-0.1828	1.0511	0.1879

Panel B: Parameter Estimation for Marginal Models

Parameters	NSW	QLD	SA	TAS	VIC
	Conditional Mean Equation				
β_0	0.0085 (0.1577)	0.0077 (0.0278)	0.0036 (0.0386)	0.0074 (0.3206)	0.0067 (0.0529)
β_1	0.5185 (0.0878)	0.5547 (0.0813)	0.6455 (0.0260)	0.4773 (0.0258)	0.5701 (0.0488)
	Conditional Variance Equation				
α_0	0.8569 (0.0062)	0.0114 (0.0027)	0.0529 (0.0072)	0.0100 (0.0024)	0.0095 (0.0028)
α_1	0.9234 (0.1635)	0.8569 (0.0625)	0.9693 (0.1270)	0.7490 (0.0606)	0.6466 (0.0224)
γ	0.0766 (0.0470)	0.1431 (0.0275)	0.0307 (0.0547)	0.2510 (0.1383)	0.1049 (0.0400)
CS*	0.0807 (0.1284)	0.0241 (0.1410)	0.0008 (0.0342)	0.0801 (0.0376)	0.0275 (0.0349)
df^{**}	2.6082 (0.1086)	2.5446 (0.1319)	2.4826 (0.0685)	2.5321 (0.3287)	2.6677 (0.1881)
Log Likelihood	1473	836	148	847	1009
K-S p-value***	0.4201	0.3890	0.3708	0.3726	0.4110
NS p-value****	0.9867	0.9648	0.6988	0.5143	0.9594

$*$ CS denotes the coefficient of skewness.
$**$ df denotes the degree of freedom.
$***$ p-values for Kolmogorov–Smirnov test for residuals' goodness-of-fit to skewed Student's t.
$****$ p-values for Neyman smooth test for residuals' goodness-of-fit to skewed Student's t.

Table 2: This table reports the parameter estimation for the long-term seasonal component of the in-sample data using sin-EWMA model (Panel A), as well as the model for marginal distributions using a AR(1)-GARCH(1,1) (Panel B).

Next we proceed to the modelling of the marginal distributions using the $AR(1) - GARCH(1,1)$ specification, where the errors are assumed to follow a skewed Student-t distribution. We have also considered a range of different models and lag orders on the in-sample period, including $ARMA - GARCH$, $ARMA - EGARCH$, $ARMA - GAS$, and $ARMA - EGAS$. We did not find any significant evidence of out-performance by any other model, based on the BIC information criteria. Looking at the results in Table 2, Panel B, we note the following: (i) The log-prices are characterised by signification autocorrelation, (ii) the underlying GARCH processes appear to be integrated for four out of five markets for which $\alpha_{1,i} + \gamma_i = 1$, (iii) there almost no evidence of skewness in the error distribution, (iv) the error distribution has very heavy tails as indicated by the low estimates for the degrees-of-freedom, and (v) the Kolmogorov-Smirnov (KS) and Neyman smooth (NS) goodness-of-fits indicate an adequate fit of the marginal models for all markets.

<<**Insert Figure ??** >>

Parameters	NSW	QLD	SA	TAS	VIC
NSW	1.0000				
QLD	0.3807	1.0000			
SA	0.3307	0.1631	1.0000		
TAS	0.2349	0.1070	0.2188	1.0000	
VIC	0.5375	0.2668	0.4602	0.3062	1.0000

Table 3: Matrix of the in-sample Kendall's τs for residuals (full-sample).

3.3. Copula Estimation

After choosing the Dvine decomposition of the multivariate copula, we still have to decide which permutation of the data to choose, since there are $5!/2$ possible distinct permutations and the results are not invariant to the chosen permutation. A rule for selecting the best permutation for Dvines, according to Nikoloulopoulos, Joe, and Li (2012), consists of choosing and connecting the most dependent pairs in the first tree, which is the approach typically used in the literature. Using the sample Kendall's τs computed from $AR(1) - GARCH(1,1)$ residuals, reported in Table 3, we choose the following structure:

$<<$**Insert Figure ??**$>>$

It is noteworthy that this structure is chosen purely based on the strength of dependence of the standardised residuals. However, it perfectly matches the geographical location of the inter-connectors, since adjacent states presenting the strongest ranked correlation in the data set. The only exception is TAS, which we have to connect to SA, because the structure of Dvine trees does not allow three edges to be connected to VIC.

In Table 4, Panel A, we report the parameter estimates for the SCAR-Dvine model. We have examined different copula functions (i.e. Gaussian, Student's t, Gumbel, Clayton, and rotated versions thereof, described in Cherubini, Luciano, and Vecchiato 2004) for each pair, and chosen the best one based on their respective BIC criteria. At first, it might seem striking that the majority of the chosen copulae in the level 1 tree are Gaussian. However, a similar behaviour has been identified in Hafner and Manner (2012) and Manner and Segers (2011), which can be explained by the property of near asymptotic dependence discussed in Subsection 2.3. Additionally, it is also surprising that the symmetric Gaussian copula outperforms the two asymmetric models, even though upper-tail dependence behaviour is expected in electricity log-prices. It seems that to some extent this asymmetry is accounted for by the time-varying dependence parameter. Furthermore, the asymmetric models, e.g. the Clayton copula, may simply underestimate the dependence for large observations and this may outweigh the advantage of allowing for lower tail dependence (see Hafner and Manner 2012 for a similar situation).

To verify these arguments, we compare our results with the static Dvine model, the results of which are reported in Table 4, Panel B. In the static Dvine model, the dependence copula parameter θ is constant, as opposed to the stochastic process (9). Therefore, it is the Dvine extension of the copula model of Ignatieva and Trück (2016), applied to a similar dataset. We can see that the Student t-copula is the dominant copula for all pairs in the first tree, whereas different copulas have been selected on the higher trees. However, the results are not at odds with the dynamic model. For example, for the pairs "QLD-NSW", "NSW-VIC", and "VIC-SA" Student copula provides the best static fit, while SCAR Gaussian provides the best dynamic fit, both of which alluding to the existence of both (near) upper- and lower-tail dependence (recall that dynamic Gaussian copulas have the property of *near asymptotic dependence*, making their tails indistinguishable from asymptotically dependent one). Overall, the SCAR version of the Dvine model provides a better fit with a log-likelihood of 1337.37, compared to 1100.57 in the static

Panel A: SCAR-Dvine Model

Pairs states	Copula type	Prameters			
		μ	ϕ	σ	LL
		Tree Level 1			
QLD-NSW	SCAR Gaussian	0.2119	0.7460	0.3774	283.9477
NSW-VIC	SCAR Gaussian	0.2920	0.7363	0.4138	448.3619
VIC-SA	SCAR Gaussian	0.3130	0.6789	0.3811	382.7633
SA-TAS	SCAR Gumbel	0.1067	0.7675	0.2762	104.7408
		Tree Level 2			
QLD-VIC|NSW	Static Gumbel	1.0450	-	-	2.5947
NSW-SA|VIC	SCAR Gumbel	0.0045	0.6872	0.2540	16.2751
VIC-TAS|SA	SCAR Gaussian	0.2670	0.4250	0.3948	79.5261
		Tree Level 3			
QLD-SA|NSW VIC	Independence	0	-	-	0
NSW-TAS|VIC SA	SCAR Gaussian	-0.0149	0.7644	0.2581	8.7489
		Tree Level 4			
QLD-TAS|NSW VIC SA	SCAR Gumbel	0.0226	0.7447	0.1206	10.4091

Panel B: Static Dvine Model

Pairs states	Copula type	Prameters		
		θ	ν	LL
		Tree Level 1		
QLD-NSW	Student's t	0.52	5.15	208.0061
NSW-VIC	Student's t	0.77	3.46	193.1331
VIC-SA	Student's t	0.56	8.09	486.3622
SA-TAS	Student's t	0.60	20.76	101.2675
		Tree Level 2		
QLD-VIC|NSW	Student's t	0.14	6.63	1.7173
NSW-SA|VIC	Gaussian	-0.27	-	30.4260
VIC-TAS|SA	Gaussian	0.17	-	16.9282
		Tree Level 3		
QLD-SA|NSW VIC	Clayton	0.89	-	6.12062
NSW-TAS|VIC SA	Gaussian	0.39	-	52.0421
		Tree Level 4		
QLD-TAS|NSW VIC SA	Clayton	0.32	-	4.1845

Table 4: In-sample parameters estimation for the SCAR-DVine model (Panel A) and the static Dvive model (Panel B). Selected copula models and parameter estimates are provide for both models and for daily electricity log-prices of five regional markets in Australia, covering the period 01.01.2010 to 31.12.2012.

version. In comparison, the log-likelihood values of the Gaussian and Student DCC-copulas are 979 and 1136, respectively, indicating a clearly worse fit than the SCAR-Dvine. However, the dynamic Student copula fits better than the static Dvine.[11] This is also reflected in the predictive power of the models, a thorough discussion of which will follow in the next section.

The findings here are in line with the observation that regional energy markets exhibit significant tail dependence. Additionally, the choice of Gumbel indicate higher upper-tail dependence on the lower trees of the model. The survival Gumbel copula on the second tree, on the other hand, indicate (conditional) lower tail dependence. The estimates of the persistence parameters ϕ indicate that the time-varying dependence is positively autocorrelated. In the third and fourth levels the persistence is equally strong as in the lower trees. However, the conditional dependence is rather weak with μ being close to zero showing that most of the dependence has been captured by the first two levels. It is noteworthy that there is conditional independence between QLD and SA.

Important features of the joint dependence among different markets can be inferred from the preceding estimation results (cf Joe, Li, and Nikoloulopoulos 2010). Because the pair-copulae in the first level

[11] This ranking of the models is preserved when computing information criteria based on the log-likelihoods such as AIC or BIC that penalize highly parametrized models.

of the estimated Dvines have upper and lower tail dependence, the multivariate copulae also have upper and lower tail dependence. Moreover, since one of these pair-copulae are tail asymmetric (SA-TAS), the range of upper/lower tail dependence for the bivariate margins is quite flexible. These characteristics are in accordance with empirical evidence found in Ignatieva and Trück (2016) that electricity price data tends to exhibit tail dependence and asymmetries.

The dynamics of the dependencies over the in-sample period can be observed in Figures **??** and **??**. For the SCAR-Dvine model Kendall's τ has been computed based on 10000 Monte Carlo simulations, whereas for the N-DCC and t-DCC copulas the time-varying correlations have been transformed to Kendall's τ to make the results comparable. Time-variation clearly visible and it is more pronounced for the SCAR-Dvine model. Further evidence on the importance of modelling time-varying dependence is provided in the next section, where we compare the results of our model to the static Dvine model in an out-of-sample test.

<<Insert Figures ?? and ??>>

4. Out-of-Sample Back-testing of Conditional Quantiles

Assessment of risk models (or back-testing) is an essential part of the process of internal control for market risk management. The internal risk management teams in financial institutions are particularly interested in detecting clustering in *exceedances*[12]. Empirical studies have shown that large losses that occur in rapid succession are more likely to lead to disastrous events such as bankruptcy (Christoffersen and Pelletier 2004). In this section, we consider the problem of estimating the Conditional Quantile (CQ) of an equally weighted portfolio of different regional markets. Recall that our out-of-sample covers the period between 01.01.2013 to 31.12.2015, consisting of 1095 days.

4.1. Back-testing

We are interested in comparing the SCAR-Dvine copula model's ability to forecast the occurrence of extreme events with the alternative models. For this purpose, we conduct the following out-of-sample back-testing exercise. Given the estimation set of $\{1, \dots, T\}$ daily observations for the copula model and the testing set consisting of $T^* = 1095$ observations $\{T+1, \dots, T+T^*\}$, the exercise is done as follows:

1. For $h = 1, \dots, T^*$:

   (a) From the fitted copula model, we simulate 10,000 samples $u_{1,T+h}, \dots, u_{5,T+h}$ from the estimated copula using one-day ahead prediction the copula parameter, $\theta_{t+h|t+h-1}$.

   (b) For $j = 1, \dots, 5$, we convert $u_{j,T+h}$ to $\hat{\varepsilon}_{j,T+h}$, using the (estimated) inverse cdf's, i.e., $\hat{\varepsilon}_{j,T+h} = \hat{F}_j^{-1}(u_{j,T+h})$.

   (c) For $j = 1, \dots, 5$, we convert $\hat{\varepsilon}_{j,T+h}$ to the deseasonalised log-price using one-day ahead forecasts of the conditional mean and variance as

$$\hat{x}_{j,T+h} = \hat{\mu}_{j,T+h|T+h-1} + \hat{\sigma}_{j,T+h|T+h-1} \cdot \hat{\varepsilon}_{j,T+h}.$$

[12] Kupiec (1995) defines exceedance as an event where the ex post portfolio loss exceeds the ex ante risk measure.

($c^{\dagger}$) For $j = 1, \ldots, 5$, we add seasonal trend back to the deasonalised forecasts by[13]

$$\hat{y}_{j,T+h} = \hat{x}_{j,T+h} + s_{j,T+h}.$$

(d) Then we compute the portfolio forecasts as

$$\hat{P}_{T+h} = \frac{1}{5}\sum_{j=1}^{5} x_{j,T+h}.$$

($d^{\dagger}$) Then we compute the portfolio forecasts as

$$\hat{P}^{\dagger}_{T+h} = \frac{1}{5}\sum_{j=1}^{5} y_{j,T+h}.$$

2. For significance levels $\alpha = \{0.005, 0.01, 0.05, 0.1, 0.9, 0.95, 0.99, 0.995\}$, we compute the one-day ahead conditional quantile $Q_\alpha(T+h)$ forecast for the day $T+h$ based on the 10,000 simulated samples from the predictive distribution.

3. To implement backtesting of a sequence of CQ forecast, we follow Christoffersen (1998) in defining the *hit-sequence*, $\{I_t\}_{t=T}^{T+T^*}$, as follows:

$$I_t := \mathbb{I}(P_t < Q_\alpha(t|t-1).$$

   Where $\mathbb{I}(\cdot)$ is the indicator function. That is if the observed value of the portfolio log-prices for the day is less than Q_α, then a hit (i.e., violation or exceedance) is said to occur.

The portfolio CQs are shown graphically in Figure **??**. Visually it is difficult to distinguish the out-of-sample forecasting performance of different models. Therefore, in order to have a conclusive analysis, we implement four important statistical tests, namely, the test of Kupiec (1995), the autocorrelation test of Christoffersen (1998), the dynamic quantile test of Engle and Manganelli (2004), and the duration-based test of Christoffersen and Pelletier (2004). These tests check whether the proportion of CQ exceptions is consistent with the chosen confidence level α and whether the exceedances are independent.

<Place holder for Figure ?? >

By definition, the conditional probability of violating the VaR should always be

$$\mathrm{P}(I_{t+1} = 1|\mathcal{F}_t) = \alpha, \tag{13}$$

for every t. Equation (13) implies that no information available to the risk manager at the time the CQ was made is helpful in forecasting the probability of violations. Otherwise, this information would be incorporated into calculating a better CQ with unpredictable violations. We will refer to tests of this property as conditional coverage (CC) tests.

Kupiec (1995) developed an unconditional coverage (UC) test of whether $\Pr(I_{t+1} = 1) = \alpha$, under the assumption of independence for the hits. The test is built on the idea that if the actual fraction of CQ exceedences is statistically different than α, the UC test rejects the null. While Kupiec test

[13]The steps with superscript (†) are optional depending on the application. The vector of seasonal component is assumed to be deterministic, by construction (see Subection 2.1); thus, it has no effect on the analysis.

provide a useful benchmark for assessing the accuracy of a given CQ model, it focuses exclusively on the unconditional coverage property of an adequate CQ measure and do not examine the extent to which the independence property is satisfied.

According to Christoffersen (1998), I_t should be an i.i.d. sequence, following a Bernoulli distribution with first moment α. Christoffersen (1998) proposes a formal test, using a general first-order Markov process, where the one-step-ahead transition probabilities $\Pr(I_{t+1}|I_t)$ are given by

$$\begin{pmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{pmatrix}.$$

Here, $\pi_{ij} := \Pr(I_{t+1} = j | I_t = i)$. Under the null hypothesis, the hits have a constant conditional mean, i.e., $\pi_{01} = \pi_{11} = \alpha$.

Another test that we use to test the adequacy of the estimated CQs is the Dynamic Quantile test of Engle and Manganelli (2004). The null hypothesis of the test states that the model is correctly specified and that CQ is not under- or over-estimated. The test is based on F statistics and tests $H_0 : b_0 = b_1 = \cdots = b_6 = 0$ for the regression

$$I_t = a_0 + \sum_{k=1}^{5} b_k I_{t-k} + b_6 CQ_\alpha(t) + u_t.$$

In their estimation method, Engle and Manganelli (2004) assume that the error term u_t has a logistic distribution resulting in a logistic regression.

Finally, we apply the duration-based test of Christoffersen and Pelletier (2004). Under the null hypothesis of the test, hits should occur at random time intervals. Let D_i denote the duration between two time intervals; then, $D_i = t_i - t_{i-1}$, where t_i is the observation time of hit number i. Using the Bernoulli property, the probability of a violation in d periods is $\Pr(D_i = d) = \Pr(I_{t+1} = 0, I_{t+2} = 0, \ldots, I_{t+d} = 1)$. Consider the hazard rate

$$\lambda(D_i) = \frac{\Pr(D_i = d)}{1 - \Pr(D_i < d)}.$$

Under the null hypothesis, $\lambda(D_i)$ should be flat and equal to α. The only memory free (continuous) random distribution is the exponential, thus Christoffersen and Pelletier (2004) specify that under the null the distribution of the no-hit durations should be

$$f_{\exp}(D_i; \alpha) = \alpha \exp(-\alpha D_i).$$

Table 5 reports the p-values for all four statistical tests of interest[14], applied to SCAR-Dvine model and the four competing models, across all values of α under consideration. The results of the tests for the lower tail quantiles are provided in the left part of the table and the upper tail results are displayed on the right side. Our results indicate that the SCAR-Dvine provide a sufficient coverage and outperforms the competing models in sense that it has the fewest rejections for any given significance level.

To cross validate the results, we also use the *Risk Map*, which has been recently introduced in Colletaz, Hurlin, and Pérignon (2013). Risk Map is a graphical tool, which summarises all information about the performance of a risk model. It jointly considers the number and the magnitude of extreme

[14]For brevity, we have excluded the backtesting results for the UC autocorrelation test of Christoffersen (1998), the UC conditional autoregressive test of Engle and Manganelli (2004), and the UC duration-based test of Christoffersen and Pelletier (2004). The results of these test are in line with the tests presented in Table 5 and are available upon request.

Panel A: Kupiec's Unconditional Coverage Test of Kupiec (1995)

	Lower tail				Upper tail			
	0.005	0.01	0.05	0.1	0.90	0.95	0.99	0.995
SCAR-Dvine	0.8378	0.1311	0.075	0.0269	0.3915	0.2085	0.7719	0.2468
t-DCC-Copula	0.8215	0.1007	0.0002	0.0001	0.0342	0.0075	0.5433	0.8381
N-DCC-Copula	0.5080	0.1007	0.0109	0.0022	0.2921	0.0415	0.9855	0.5080
DCC-GARCH	0.1665	0.0728	0.3429	0.0006	0.0001	0.1656	0.0003	0.0001
Static Dvine Copula	0.8230	0.8979	0.0004	0.0003	0.5940	0.1996	0.7729	0.5274

Panel B: Autocorrelation Conditional Coverage Test of Christoffersen (1998)

	Lower tail				Upper tail			
	0.005	0.01	0.05	0.1	0.90	0.95	0.99	0.995
SCAR-Dvine	0.9567	0.2771	0.2035	0.0425	0.5292	0.4149	0.8744	0.5071
t-DCC-Copula	0.9433	0.2522	0.0006	0.0001	0.0838	0.026	0.711	0.9571
N-DCC-Copula	0.7916	0.2522	0.0345	0.0075	0.2847	0.1215	0.8441	0.7916
DCC-GARCH	0.3563	0.2103	0.6341	0.0022	0.0005	0.2747	0.0007	0.0001
Static Dvine Copula	0.5637	0.2497	0.0008	0.0009	0.8395	0.4468	0.8744	0.7916

Panel C: Dynamic Quantile Conditional Coverage Test of Engle and Manganelli (2004)

	Lower tail				Upper tail			
	0.005	0.01	0.05	0.1	0.90	0.95	0.99	0.995
SCAR-Dvine	0.9987	0.7017	0.1381	0.0540	0.7939	0.7147	0.9871	0.8899
t-DCC-Copula	0.9777	0.6806	0.0001	0.1177	0.1792	0.0704	0.9129	0.9987
N-DCC-Copula	0.9804	0.6806	0.0117	0.0147	0.5162	0.308	0.9864	0.9804
DCC-GARCH	0.6119	0.1741	0.7856	0.0265	0.0066	0.5383	0.0002	0.0001
Static Dvine Copula	0.8155	0.3212	0.0017	0.0024	0.3429	0.1815	0.3212	0.5696

Panel D: Duration-based Conditional Coverage Test of Christoffersen and Pelletier (2004)

	Lower tail				Upper tail			
	0.005	0.01	0.05	0.1	0.90	0.95	0.99	0.995
SCAR-Dvine	0.6438	0.1936	0.1687	0.0047	0.5788	0.5443	0.9245	0.1765
t-DCC-Copula	0.6211	0.1294	0.0001	0.0001	0.0524	0.0454	0.9043	0.4841
N-DCC-Copula	0.4750	0.1294	0.0308	0.0015	0.5671	0.1771	0.5561	0.5061
DCC-GARCH	0.4436	0.0868	0.5814	0.0045	0.0007	0.2723	0.0030	0.0001
Static Dvine Copula	0.4750	0.1294	0.0090	0.0076	0.4503	0.2741	0.1266	0.2234

Table 5: Comparison of CQ backtesting results obtained using alternative procedures. For each set of models, the table reports the p-values of A) UC Kupiec (1995) test, B) the CC autocorrelation test of Christoffersen (1998), C) the CC conditional autoregressive test of Engle and Manganelli (2004), and D) the CC duration-based test of Christoffersen and Pelletier (2004).

events, relying on the concept of *super exceptions*. Colletaz, Hurlin, and Pérignon (2013) define a supper exception as an extreme event which exceeds both the standard CQ_α and a $CQ_{\alpha'}$ defined at an extremely low coverage probability α'. Here we assume, $\alpha' = 0.2\%$ for the lower tail CQs and $\alpha' = 99.8\%$ for the upper tail CQs. Colletaz, Hurlin, and Pérignon (2013) name several advantages of Risk Map: 1) it incorporates the magnitude of the extreme events into the Kupiec test[15]; 2) it allows us to jointly test the null hypothesis that both the numbers of CQ exceptions and super exceptions are accurate; 3) it allows joint validation of the standard and stressed CQs that the risk manager must compute; and 4) the approach is general and no distributional assumptions is needed.

The hypothesis test underlying the Risk Map boils down to jointly test the number of CQs exceptions and super exceptions:

[15]Thus, the Risk Map approach is a three-dimensional generalisation of the "Traffic Light" system of on Banking Supervision (2011). The three colour "Traffic Light" system of on Banking Supervision (2011) is the reference back-test methodology for banking regulators.

$$H_0 : E[I_t(\alpha)] = \alpha, \quad \text{and} \quad E[I_t(\alpha')] = \alpha'.$$

This joint conditional coverage null hypothesis can be tested using the hit regression test of Engle and Manganelli (2004), under which the test statistic is asymptotically chi-square with two degrees of freedom[16].

Figure **??** and **??** presents the Risk Maps for different quantiles, constructed based on the rejection zones for different confidence levels. Note that the cells below the diagonal are not coloured as they correspond to situations in which the number of super exceptions exceeds the number of exceptions, which is of course impossible. If the (exception, super exception) pair corresponds to a green cell, we conclude that we cannot reject the null for either the exception or the super exception, at the 95% confidence level. If the pair falls in the orange zone, we can reject the null at the 95% but not at the 99% confidence level. Finally, a red cell implies that we can reject the null hypothesis at the 99% confidence level. The key take away from this figures is that the p-values remain remarkably stable for SCAR-Dvine models, which confirms the robustness of our results.

<Place holder for Figure ?? >

Overall, we have shown that SCAR-Dvine has performed very well for predicting the distribution of electricity log-prices; however, it is less than straightforward to draw a conclusion due to the following observations. The sheer quantity of information in Table 5 makes it hard to determine which models perform best overall, as the p-values are often close to each other, we are testing for numerous quantiles indicating a multiple testing problem and in some cases, although SCAR-Dvine has performed better than the competing models, the p-values still suggest rejection at conventional significance levels. For instance, in quantile $\alpha = 0.1$ in Panel D the duration based model rejects all models, including SCAR-Dvine.

Therefore, we develop a ‘decision making’ tool, to assist us in assessing the reliability of the CQ forecasts by summarising the p-values, using a simple ‘scorecard’ system. Thereby, we compare the quality of results from each test, including the Risk Map, then by aggregating the scores for each model we can choose the better risk model with more conclusive evidence. The idea of linear scoring of market-risk models has first appeared in So and Chan (2014), where they ranked the ratio of absolute and theoretical proportion of exceedances. However, the method proposed here is more robust, because instead of basing the scorecard on the sequence of violations, we construct it based on the p-values obtained from each backtest. Having the p-values already calculated, we can proceed to our scoring scheme. We score each p-value from 0 to 3, based on the strength of evidence against the null hypothesis. A small p-value of less than 0.01 indicates strong evidence against the null hypothesis, so the test in the quantile receives a Score of 0. A p-value larger than 0.01 indicates weak evidence against the null hypothesis, so we fail to reject the null hypothesis. If the p-value is between 0.01 and 0.05 we accept the null with ‘weak’ statistical evidence, and give it a Score of 1; if between 0.05 and 0.10, we accept the null with an ‘acceptable’ statistical evidence, and give it a Score of 2; if greater than 0.10 we accept the null with a ‘reliable’ statistical evidence, giving it a Score of 3. As for the Risk Map, the situation is simpler. If the (exception, super exception) pair corresponds to a green cell, we give the model a Score of 3, if it corresponds to orange we give a Score of 2, and if it corresponds to red we give a score of 1.

The scorecard is presented in Table 6. The total score for each model is presented in the last column, based on which the five risk models can be compared in each test. It is clear that SCAR-Dvine has obtained a higher total score in every instance. We must highlight that this decision making tool addresses

[16]The p-values of the test can be based on either the asymptotic distribution or the finite sample distribution, which can be generated by simulation. See Colletaz, Hurlin, and Pérignon (2013) for more discussion on finite sample properties.

Panel A: Kupiec's Unconditional Coverage Test of Kupiec (1995)									
	0.005	0.01	0.05	0.1	0.90	0.95	0.99	0.995	Total Score
SCAR-Dvine	3	3	2	1	3	3	3	3	21*
t-DCC-Copula	3	3	0	0	1	0	3	3	13
N-DCC-Copula	3	3	1	0	3	1	3	3	17
DCC-GARCH	3	2	3	0	0	3	0	0	11
Static Dvine Copula	3	3	0	0	3	2	3	3	17
Panel B: Autocorrelation Conditional Coverage Test of Christoffersen (1998)									
SCAR-Dvine	3	3	3	1	3	3	3	3	22*
t-DCC-Copula	3	3	0	0	2	1	3	3	15
N-DCC-Copula	3	3	1	0	3	3	3	3	19
DCC-GARCH	3	3	3	0	0	3	0	0	12
Static Dvine Copula	3	3	0	0	3	3	3	3	18
Panel C: Dynamic Quantile Conditional Coverage Test of Engle and Manganelli (2004)									
SCAR-Dvine	3	3	3	2	3	3	3	3	23*
t-DCC-Copula	3	3	0	3	3	2	3	3	20
N-DCC-Copula	3	3	1	1	3	3	3	3	20
DCC-GARCH	3	3	3	1	0	3	0	0	13
Static Dvine Copula	3	3	0	0	3	3	3	3	16
Panel D: Duration-based Conditional Coverage Test of Christoffersen and Pelletier (2004)									
SCAR-Dvine	3	3	3	2	3	3	3	3	23*
t-DCC-Copula	3	3	0	3	3	2	3	3	20
N-DCC-Copula	3	3	1	1	3	3	3	3	20
DCC-GARCH	3	3	3	1	0	3	0	0	13
Static Dvine Copula	3	3	0	0	3	3	3	3	18
Panel E: Risk Map of Colletaz, Hurlin, and Pérignon (2013)									
SCAR-Dvine	3	3	3	2	2	3	3	3	22*
t-DCC-Copula	3	3	1	1	1	1	3	3	16
N-DCC-Copula	3	3	2	1	1	2	3	3	18
DCC-GARCH	3	3	3	1	1	2	1	1	15
Static Dvine Copula	3	3	1	1	1	1	1	2	13

Table 6: This table presents the scorecard for the p-values of each back-testing procedure. The last column sums up the total score in each row. The asterisks indicate the wining models in each test (block of scores). The scorecard also provides a quick visual summary of the performance of each test in each quantile.

the first two concerns discussed above. That is if the p-values are very close to each other, models receive a similar score, and if a p-value is lower than 0.05, the result is excluded from the decision via a score of 0. Finally, in this study, the decision making is straightforward, since the SCAR-Dvine has received a higher total score in every single test. Nevertheless, one may envisage situations where the results are mixed, in the case of which a weighting scheme based on the reliability of the test might be helpful.

5. Conclusion

The electricity retail industry plays an integral role in the apparatus of modern energy markets. They are responsible for the efficient distribution of electricity through profiling the consumers consumption (i.e. load profile), aggregating the electricity demand and bidding for electricity in a competitive auction (e.g. NEM in Australia). On the other hand, they are constrained to provide energy at fixed rates to the end-users, while purchasing it at the spot price form the market. Therefore, effectively, they bare all of the market risk on the buy-side.

In managing the market risk, electricity retailers need to have a robust model for distributional forecasts that adequately capture extreme price movements. However, electricity price series exhibit highly stylised features, which are materially different to other financial time-series, namely, strong seasonality, extreme price spikes, non-linear price dependency between regional markets, and correlation asymmetry. Similar features might also be observed in other financial time-series, but not with the same magnitude and frequency; thus analysing electricity prices demands more sophisticated techniques.

This paper proposes the employment of SCAR-Dvine model for capturing complex dependence structures in daily Australian electricity prices. Our research design is based on the dynamic financial analysis approach, looking at an enterprise's risks holistically, as opposed to the traditional risk analysis which analyses risks individually. Modelling risk of the five markets as a complex interconnected system, as opposed to analysing markets individually, for the electricity retailers is more relevant than ever before. This is because, over the last two decades, many of the small market players have amalgamated into large enterprises operating in all five markets simultaneously. Seasonality is dealt with by using sophisticated models for long- and short-term seasonal components. The (conditional) marginal distributions are modelled using GARCH models with skewed t-distributed errors and goodness-of-fit tests suggest the appropriateness of these models. Furthermore, we provide empirical evidence that not only the SCAR-Dvine model fits the data very well in an in-sample period, but also preforms better than four competing models in predicting conditional quantiles in an out-of-sample back testing exercise.

Acknowledgement

Laleh Tafakori would like to thank the Australian Research Council for supporting this work through Laureate Fellowship FL130100039.

References

AAS, K., C. CZADO, A. FRIGESSI, AND H. BAKKEN (2009): "Pair-copula constructions of multiple dependence," *Insurance: Mathematics and economics*, 44(2), 182–198.

ALMEIDA, C., C. CZADO, AND H. MANNER (2016): "Modeling high dimensional time-varying dependence using D-vine models," *Applied Stochastic Models in Business and Industry*, 32, 621–638.

ANG, A., AND J. CHEN (2002): "Asymmetric correlations of equity portfolios," *Journal of financial Economics*, 63(3), 443–494.

AUSTRALIAN ENERGY REGULATOR (2012): "Retail Energy Market Update," *Australian Energy Regulator*, pp. 1–14.

BENNEDSEN, M. (2017): "A rough multi-factor model of electricity spot prices," *Energy Economics*, 63, 301–313.

BIERBRAUER, M., C. MENN, S. T. RACHEV, AND S. TRÜCK (2007): "Spot and derivative pricing in the EEX power market," *Journal of Banking & Finance*, 31(11), 3462–3485.

BRECHMANN, E. C., C. CZADO, AND K. AAS (2012): "Truncated regular vines in high dimensions with application to financial data," *Canadian Journal of Statistics*, 40(1), 68–85.

CARTEA, Á., AND P. VILLAPLANA (2008): "Spot price modeling and the valuation of electricity forward contracts: the role of demand and capacity," *Journal of Banking & Finance*, 32(12), 2502–2519.

CHANG, Y., C. S. KIM, J. I. MILLER, J. Y. PARK, AND S. PARK (2016): "A new approach to modeling the effects of temperature fluctuations on monthly electricity demand," *Energy Economics*, 60, 206–216.

CHERUBINI, U., E. LUCIANO, AND W. VECCHIATO (2004): *Copula methods in finance*. John Wiley & Sons.

CHOLLETE, L., A. HEINEN, AND A. VALDESOGO (2009): "Modeling international financial returns with a multivariate regime-switching copula," *Journal of financial econometrics*, p. nbp014.

CHRISTOFFERSEN, P., AND D. PELLETIER (2004): "Backtesting value-at-risk: A duration-based approach," *Journal of Financial Econometrics*, 2(1), 84–108.

CHRISTOFFERSEN, P. F. (1998): "Evaluating interval forecasts," *International economic review*, pp. 841–862.

COLLETAZ, G., C. HURLIN, AND C. PÉRIGNON (2013): "The Risk Map: A new tool for validating risk models," *Journal of Banking & Finance*, 37(10), 3843–3854.

CREAL, D., S. J. KOOPMAN, AND A. LUCAS (2013): "Generalized autoregressive score models with applications," *Journal of Applied Econometrics*, 28(5), 777–795.

CZADO, C. (2010): "Pair-copula constructions of multivariate copulas," in *Copula theory and its applications*, pp. 93–109. Springer.

DA SILVA FILHO, O. C., F. A. ZIEGELMANN, AND M. J. DUEKER (2012): "Modeling dependence dynamics through copulas with regime switching," *Insurance: Mathematics and Economics*, 50(3), 346–356.

DE JONG, C. (2006): "The nature of power spikes: A regime-switch approach," *Studies in Nonlinear Dynamics and Econometrics*, 10(3), 3.

DIETRICH, C. F. (1991): *Uncertainty, calibration and probability: the statistics of scientific and industrial measurement*. CRC Press.

ENGLE, R. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20(3), 339–350.

ENGLE, R. F., AND S. MANGANELLI (2004): "CAViaR: Conditional autoregressive value at risk by regression quantiles," *Journal of Business & Economic Statistics*, 22(4), 367–381.

GARCIA, R. C., J. CONTRERAS, M. VAN AKKEREN, AND J. B. C. GARCIA (2005): "A GARCH forecasting model to predict day-ahead electricity prices," *Power Systems, IEEE Transactions on*, 20(2), 867–874.

HAFNER, C. M., AND H. MANNER (2012): "Dynamic stochastic copula models: Estimation, inference and applications," *Journal of Applied Econometrics*, 27(2), 269–295.

HANSEN, E. B. (1994): "Autoregressive Density Estimation," *International Economic Review*, 35, 705–730.

HEINEN, A., A. VALDESOGO ROBLES, ET AL. (2009): "Asymmetric CAPM dependence for large dimensions: the canonical vine autoregressive model," Discussion paper, Universidad Carlos III de Madrid.

HIGGS, H. (2009): "Modelling price and volatility inter-relationships in the Australian wholesale spot electricity markets," *Energy Economics*, 31(5), 748–756.

HIGGS, H., AND A. WORTHINGTON (2008): "Stochastic price modeling of high volatility, mean-reverting, spike-prone commodities: The Australian wholesale spot electricity market," *Energy Economics*, 30(6), 3172–3185.

HONG, Y., J. TU, AND G. ZHOU (2007): "Asymmetries in stock returns: statistical tests and economic evaluation," *Review of Financial Studies*, 20(5), 1547–1581.

IGNATIEVA, K., AND S. TRÜCK (2016): "Modeling spot price dependence in Australian electricity markets with applications to risk management," *Computers & Operations Research*, 66, 415–433.

JANCZURA, J., S. TRÜCK, R. WERON, AND R. C. WOLFF (2013): "Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling," *Energy Economics*, 38, 96–110.

JOE, H. (2005): "Asymptotic efficiency of the two-stage estimation method for copula-based models," *J. Multivariate Anal.*, 94, 401–419.

JOE, H., H. LI, AND A. K. NIKOLOULOPOULOS (2010): "Tail dependence functions and vine copulas," *Journal of Multivariate Analysis*, 101(1), 252–270.

KELES, D., R. HARTEL, D. MÖST, AND W. FICHTNER (2012): "Compressed-air energy storage power plant investments under uncertain electricity prices: an evaluation of compressed-air energy storage plants in liberalized energy markets," *The Journal of Energy Markets*, 5(1), 53.

KIM, G., M. J. SILVAPULLE, AND P. SILVAPULLE (2007): "Comparison of semiparametric and parametric methods for estimating copulas," *Computational Statistics & Data Analysis*, 51, 2836 – 2850.

KOLE, E., K. KOEDIJK, AND M. VERBEEK (2007): "Selecting copulas for risk management," *Journal of Banking & Finance*, 31(8), 2405–2423.

KOLIAI, L. (2016): "Extreme Risk Modelling: An EVT–Pair-copulas Approach for Financial Stress Tests," *Journal of Banking & Finance*.

KUPIEC, P. H. (1995): "Techniques for verifying the accuracy of risk measurement models," *The J. of Derivatives*, 3(2).

LI, D. X. (1999): "On default correlation: A copula function approach," *Available at SSRN 187289*.

LIESENFELD, R., AND J. F. RICHARD (2003): "Univariate and multivariate stochastic volatility models: Estimation and diagnostics," *Journal of Empirical Finance*, 10, 505–531.

LINDSTRÖM, E., AND F. REGLAND (2012): "Modeling extreme dependence between European electricity markets," *Energy economics*, 34(4), 899–904.

LOW, R. K. Y., J. ALCOCK, R. FAFF, AND T. BRAILSFORD (2013): "Canonical vine copulas in the context of modern portfolio management: Are they worth it?," *Journal of Banking & Finance*, 37(8), 3085–3099.

MANNER, H., AND O. REZNIKOVA (2012): "A survey on time-varying copulas: Specification, simulations and application," *Econometric Reviews*, 31(6), 654–687.

MANNER, H., AND J. SEGERS (2011): "Tails of correlation mixtures of elliptical copulas," *Insurance: Mathematics and Economics*, 48(1), 153–160.

MASKIN, E., AND J. TIROLE (1988): "A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs," *Econometrica: Journal of the Econometric Society*, pp. 549–569.

MISIOREK, A., S. TRÜCK, AND R. WERON (2006): "Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models," *Studies in Nonlinear Dynamics & Econometrics*, 10(3).

MORAL-CARCEDO, J., AND J. VICENS-OTERO (2005): "Modelling the non-linear response of Spanish electricity demand to temperature variations," *Energy economics*, 27(3), 477–494.

NIKOLOULOPOULOS, A. K., H. JOE, AND H. LI (2012): "Vine copulas with asymmetric tail dependence and applications to financial return data," *Computational Statistics & Data Analysis*, 56(11), 3659–3673.

NOWOTARSKI, J., J. TOMCZYK, AND R. WERON (2013): "Robust estimation and forecasting of the long-term seasonal component of electricity spot prices," *Energy Economics*, 39, 13–27.

OH, D. H., AND A. J. PATTON (2017a): "Modelling dependence in high dimensions with factor copulas," *Journal of Business & Economic Statistics*, (forthcoming).

——— (2017b): "Time-Varying Systemic Risk: Evidence from a Dynamic Copula Model of CDS Spreads," *Journal of Business & Economic Statistics*, (forthcoming).

ON BANKING SUPERVISION, B. C. (2011): "Revisions to the Basel II Market Risk Framework," *Bank for International Settlements*.

PATTON, A. J. (2006): "Modelling asymmetric exchange rate dependence," *International economic review*, 47(2), 527–556.

PILIPOVIC, D. (2007): *Energy Risk: Valuing and Managing Energy Derivatives: Valuing and Managing Energy Derivatives*. McGraw Hill Professional.

POURKHANALI, A., J.-M. KIM, L. TAFAKORI, AND F. A. FARD (2016): "Measuring systemic risk using vine-copula," *Economic Modelling*, 53, 63–74.

RICHARD, J.-F., AND W. ZHANG (2007): "Efficient high-dimensional importance sampling," *Journal of Econometrics*, 141(2), 1385–1411.

SAVU, C., AND M. TREDE (2010): "Hierarchies of Archimedean copulas," *Quantitative Finance*, 10(3), 295–304.

SIBURG, K. F., P. STOIMENOV, AND G. N. WEISS (2015): "Forecasting portfolio-Value-at-Risk with nonparametric lower tail dependence estimates," *Journal of Banking & Finance*, 54, 129–140.

SKLAR, M. (1959): *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8.

SMITH, M. S., Q. GAN, AND R. J. KOHN (2012): "Modelling dependence using skew t copulas: Bayesian inference and applications," *Journal of Applied Econometrics*, 27(3), 500–522.

SO, M. K., AND R. K. CHAN (2014): "Bayesian analysis of tail asymmetry based on a threshold extreme value model," *Computational Statistics & Data Analysis*, 71, 568–587.

SUJIT K. SAHU, D. K. D., AND M. D. BRANCO (2003): "A new class of multivariate skew distributions with applications to Bayesian regression models," *The Canadian Journal of Statistics*, 31, 129–150.

WANG, S. (1998): "Aggregation of correlated risk portfolios: models and algorithms," in *Proceedings of the Casualty Actuarial society*, vol. 85, pp. 848–939.

WERON, R., AND A. MISIOREK (2008): "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models," *International Journal of Forecasting*, 24(4), 744–763.