Modelling Interregional Links in Electricity Price Spikes[☆]

A.E. Clements[a,*], R. Herrera[b], A.S. Hurn[a]

[a]*Queensland University of Technology, 2 George Street, Brisbane, Australia, 4001.*

[b]*Facultad de Economia y Negocios, Universidad de Talca, 2 Notre 685, Talca, Chile.*

Abstract

Abnormally high price spikes in spot electricity markets represent a significant risk to market participants. As such, a literature has developed that focuses on forecasting the probability of such spike events, moving beyond simply forecasting the level of price. Many univariate time series models have been proposed to deal with spikes within an individual market region. This paper is the first to develop a multivariate self exciting point process model for dealing with price spikes across connected regions in the Australian National Electricity Market. The importance of the physical infrastructure connecting the regions on the transmission of spikes is examined. It is found that spikes are transmitted between the regions, and the size of spikes is influenced by the available transmission capacity. It is also found improved risk estimates are obtained when inter-regional linkages are taken into account.

Keywords: Electricity prices, price spikes, point process, hawkes process, peaks over threshold, peaks over threshold, transmission capacity

JEL Classification C3, Q4

[☆]Clements and Hurn acknowledge the support of the Australian Research Council Discovery Project (DP120100837). Herrera thanks the Chilean Agency CONICYT for its financial support (FONDECYT 1150349).

[*]Corresponding author. Tel: +61-7-31382525. Fax: +61-7-31381500

Email addresses: a.clements@qut.edu.au (A.E. Clements), rodriherrera@utalca.cl (R. Herrera), s.hurn@qut.edu.au (A.S. Hurn)

1. Introduction

The National Electricity Market (NEM) in Australia, introduced in December 1998, operates one of the worlds largest interconnected power systems which comprises five regions, namely New South Wales, Victoria, Queensland, South Australia and Tasmania. Wholesale trading in this market is conducted as a spot market where supply and demand are instantaneously matched through a centrally-coordinated dispatch process. Retailers buy electricity from the wholesale grid at a market price, known as the spot price, and sell electricity to consumers at a heavily regulated price. An important feature both of this particular electricity market and of deregulated electricity markets worldwide is the periodic occurrence of abnormally high prices or price spikes in the spot electricity market (Barlow, 2002; de Jong and Huisman, 2003; Escribano, *et al.*, 2002; Lucia and Schwartz, 2002; Burger *et al.*, 2003; Byström, 2005; Cartea and Figueroa, 2005). Both the size of these irregular price events and their duration are particularly harmful to electricity retailers who cannot pass on price risk to customers, Anderson et al. (2006). Consequently, improving the understanding of factors contributing to the occurrence of extreme price events is important for risk management in the energy sector.

Early attempts to deal with price spikes utilised a range of traditional time series approaches. Autoregressive time-series models handle spikes through the use of thresholds (Misiorek et al. (2006)), Bernoulli and Poisson jump processes (Crespo Cuaresma et al., 2004; Knittel and Roberts, 2005) and a variety of heavy tailed error processes (Contreras et al., 2003; Byström, 2005; Garcia et al., 2005; Swider and Weber, 2007). Markov-switching models incorporate spikes by proposing different regimes, at least one of which is consistent with a state of system stress in which a spike is more likely to occur (de Jong and Huisman, 2003; Huisman and Mahieu, 2003; Weron et al., 2004; de Jong, 2006; Kosater and Mosler, 2006; Bierbrauer et al., 2007; Becker et al., 2007; Higgs and Worthington, 2008). Diffusion models of the spot price introduce spikes through the addition of a Poisson jump component with either a constant intensity parameter, Weron et al. (2004) or a time-varying intensity parameter, Knittel and Roberts (2005) in which the intensity of the jump process is typically a linear combination of deterministic seasonal and/or diurnal factors.

More recent approaches to modelling electricity prices treat spikes as events and therefore shift the focus away from modelling the entire price trajectory to consider the forecasting of price spikes only. This approach draws on the econometrics of point processes, which have become popular in the financial econometrics literature (see Bauwens and Hautsch (2009) for a relatively recent survey). An early adaptation of duration models (Engle and Russell, 1997) dealing with both the occurrence of the event and the size of the event (marks), is the Autoregressive Conditional Hazard (ACH) model developed by Hamilton and Jordà (2002), who considered predicting changes in the United States Federal funds target rate. The ACH model was implemented using electricity prices by Christensen, Hurn and Lindsay (2012). Other approaches in this tradition are those of Eichler et al. (2013), who employ the dynamic logit framework of Kauppi and Saikkonen (2008) for forecasting spike events which incorporates the history of spike events. Herrera and Gonzalez (2014) also use a duration based ACD-Peaks-Over-Threshold approach to model electricity price spikes.

As pointed out by Eichler et al. (2013), an important issue to consider is the multivariate behaviour of prices across regions. While studies such as Higgs (2009), Worthington et al. (2005) and Ignatieva and Trueck (2014) examine the volatility of, correlation and dependence between prices across regions, little attention has been paid to how extreme price spike events propagate across regions. Therefore, the focus of this paper shifts to modelling spikes in multiple regions simultaneously. Of interest is how spikes are transmitted across regions, and the importance of the physical infrastructure connecting the regions. The existing models of price spikes, while capable of dealing with the occurrence of the event and the size of the event simultaneously, suffer from the fundamental flaw of not being generalisable to a multivariate setting and therefore cannot provide a framework in which to examine the transmission of spikes across regions.

To solve this problem, this paper treats price spikes as a multivariate Hawkes process. This class of self-exciting point processes has become popular in high frequency financial applications given its simplicity, flexibility, and the ease with which the parameters can be interpreted in terms of self- and cross-excitation (clustering). Since the seminal work of Bowsher (2007), multivariate Hawkes processes

have been applied to various high-frequency financial problems including models of order book dynamics, high-frequency volatility and models of contagion. Building on earlier work by Herrera and Gonzalez (2014) and Korniichuk (2012), the model developed in this paper is a multivariate self-exciting marked point process (SEMPP) in which both the occurrence of spikes and their size in adjacent interconnected regions of the NEM are modelled.

A crucial element of the inter-regional transmission of price spikes is the availability of spare capacity on the interconnectors between the regions. In the case of the NEM, the regional markets of NSW and VIC are connected by a single interconnector, while those of QLD and NSW are linked via two interconnectors. Data on spare northerly (NSW to QLD, VIC to NSW) and southerly (QLD to NSW, NSW to VIC) interconnector capacity will be used to determine how physical infrastructure constraints influence the transmission of price spikes between the regions. The maintained hypothesis is that if spare import capacity is (not) available, future spikes should be smaller (larger) in size as generation capacity from the nearby region can (cannot) be transmitted to meet the local demand. Estimation results for a number of multivariate models will highlight the nature of the links between the regions and the role played by physical transmission constraints. The models will be compared to restricted univariate versions for each region. A forecasting exercise will also highlight the importance of inter-regional links in terms of forecasting the probability of spikes. All the analysis indicates that inter-regional links are important in that price spikes spill over between regions and the size of these spillovers are significantly related to interconnector capacity.

2. Institutional Background

The NEM operates as a pooled market in which all available supply to a region is aggregated and generators are dispatched so as to satisfy demand as cost effectively as possible through a centrally-coordinated dispatch process. A summary of the process for offers of generation, dispatch and calculation of the spot price is as

follows[1]. Prior to 12:30 pm on the day before production, generators provide offers of generation capacity to the Australian Energy Market Operator (AEMO). The offers consist of at most ten price-quantity pairs for each half-hour of the following day for prices between the market floor price and the market cap price.[2] Generators are free to change their offered quantities (commonly known as re-bidding) up to approximately five minutes before dispatch. Upon receipt of the offers from all generators, the supply curves are aggregated and generators are dispatched in line with the offered capacity so that demand is satisfied as inexpensively as possible. The dispatch price for each five minute interval is the offer price of the marginal generator dispatched into production. The spot price for each half-hour trading interval is then calculated as the arithmetic mean of the six five-minute interval dispatch prices observed within the half-hour, and all transactions occurring within the half-hour are settled at the spot price. If, in any given region, local demand exceeds local supply or electricity in a neighbouring region is sufficiently inexpensive to warrant transmission, then electricity is imported and exported between regions subject to the physical constraints of the interconnectors. Ability to import or export electricity is sometimes limited by the physical transfer capacity of the interconnector. When the technical limit of an interconnector is reached, the interconnector is said to be constrained.

Irregular price events (or price spikes) occur when the spot price of electricity exceeds a given price threshold. Whilst the actual threshold used is market-specific, the argument for using a threshold to define extreme events is generic (Mount et al., 2006; Kanamura and Ōhashi, 2007). For the purposes of this paper, the threshold is set at \$100 per megawatt hour (MWh) so that a price spike is defined as $P_t \geq \$100$ MWh, where P_t is the spot price for the 30 minute interval t. This threshold value has been commonly used in the Australian context (Becker et al., 2007; Christensen et al., 2009, 2012; Clements et al., 2013) and also used in the Pennsylvania – New Jersey – Maryland market (Mount et al. (2006). A price of \$100/MWh lies above

[1]The term bids is often used to denote offers of generation capacity from generators to the market operator.

[2]Currently the market floor price is $-\$1,000$ and market cap price is $\$13,500$ per megawatt hour (MWh), although for the sample period used in empirical work the cap price was $\$12,500$/MHh.

the 90th percentile in spot price for each half hour of the day in each of the regions and is also slightly larger than the marginal cost of gas-turbine generation capacity bought online during periods of market stress.

The choice of a fixed threshold, rather than a variable or endogenously determined one, also needs to be addressed. Extreme price events almost exclusively occur at times of day designated as "peak" load time, so the global definition of the threshold is not particularly problematic from the standpoint of differential load. Furthermore, the fact that the model estimated in this paper deals explicitly with the size (or mark) of the extreme event as part of the modelling process makes the choice of threshold value less critical. In applications where only the event itself is being modelled, the choice is arguably more important. Of course, making the threshold an endogenously determined parameter of the problem is an important technical problem. It is doubtful that the likelihood function is continuous in the threshold (because a different value of the threshold will imply fundamentally different point processes on which to estimate the model) but there may be other criteria that would provide a valid target function for an optimisation. This problem is a serious research question in its own right and beyond the immediate scope of this paper.

The standard explanation for the occurrence of abnormal price events is a simple micro-theoretic one. In the simplest possible world, supply can be regarded as horizontal until generation capacity is reached and thereafter becomes vertical. If demand rises to the point of system capacity (due perhaps to extreme weather conditions) or if a significant portion of supply suddenly goes offline due to generation failure then an abnormal price event results. In other words, price spikes are simply a manifestation of scarcity and are not necessarily due strategic behaviour on the part of market participants. Increased competition in electricity markets, however, may have had the undesirable side-effect of raising strategic awareness of both generators and retailers and consequently changing the nature of extreme price events. For example, strategic withholding (deliberately taking available capacity offline) or strategic bidding behaviour by generators aimed at pushing price up may now be as important in explaining price spikes as micro-theoretic ones.

Regions	Year	2005	2006	2007	2008	2009	2010	2011	2012	2013
QLD	Mean	720.80	522.56	214.80	1107.886	429.96	1178.89	623.04	124.30	224.68
	Std dev	1637.19	1420.09	818.61	2316.41	1155.36	2123.89	1609.33	307.65	531.61
	Min	0.09	0.31	0.07	0.19	0.33	0.15	0.82	0.15	0.03
	Max	7767.33	9057.27	8239.16	9820.99	8288.30	9107.97	8943.67	2792.65	6198.63
	N° Events	104	141	1553	195	261	56	184	257	747
NSW	Mean	694.36	435.09	293.4482	341.73	617.67	924.96	160.68	106.53	298.39
	Std dev	1644.69	1180.53	1115.14	955.63	1438.29	2126.49	410.34	138.14	519.56
	Min	1.18	0.10	0.02	0.14	0.13	0.45	0.05	0.09	0.16
	Max	9066.67	9638.95	9836.37	9900.00	9183.95	8245.79	12036.17	217.97	70.63
	N° Events	197	153	1529	150	303	93	201	120	84
VIC	Mean	291.24	569.05	163.61	211.01	718.24	1746.56	660.04	312.07	149.36
	Std dev	682.49	1412.74	701.69	960.77	2057.76	2946.32	1730.33	1276.96	504.91
	Min	0.11	0.13	0.03	0.14	0.14	0.11	0.79	0.14	0.04
	Max	3759.73	9034.14	9900.00	8666.23	9900.00	9898.59	9496.53	9874.42	4181.90
	N° Events	117	198	1453	233	197	101	55	171	169

Table 1: Descriptive statistics for the electricity price spikes for QLD, NSW and VIC defined as prices above AUS$100/MWh for the period 2005-2013. The year 2013 only considers observations until 31 July 2013. Both samples are spaced every 30 minutes representing 48 trading intervals in each 24-hour period.

This study considers 30 minute observations from the Queensland (QLD), New South Wales (NSW) and Victorian (VIC) markets representing 48 trading intervals in each 24-hour period. Data for the NSW, QLD pair is available from 1 January 2005 to 31 December 2013, while for the NSW, VIC pair (due to the availability of historical interconnector data), 1 July 2008 to 31 December 2013. Analysis of a three region model is restricted to the sample period starting on 1 July 2008. For the subsequent empirical analysis, data to 31 December 2012 is employed as an estimation period while 1 January 2013 to 31 July 2013 is used for forecast evaluation. Table 1 presents descriptive statistics of the marked point process obtained from the spikes for the three markets based on the spike definition of $P_t \geq \$100$ MWh. To complement Figure 1 we divide the full sample into calendar years to highlight the variability in the intensity of the spikes and the size of the marks. In relation to the spike frequency over time it can be seen that the year 2007 was exceptional with more than 50% of the observed events occurring during this time across all three regions. This period coincides with the height of extreme drought conditions in the Australian Eastern states which had significant impacts on physical infrastructure. In relation to the size of the marks, there is not a large degree of variation across time. The mean size of the marks is much larger than the threshold selected for both markets, with the standard deviations being large and varying a

great deal over time. With the standard deviation close to three times the observed mean over-dispersion is clearly present.

Figure 1 plots the realised spike events for the NSW, VIC and QLD markets for the full period 1 January 2005 to 31 December 2013 with a number of patterns emerging. As noted by Christensen et al. (2009), the memory, or clustering in the spikes within each of the regions is clearly evident. When comparing the regions, many of the spikes seem to occur at similar times. This may be taken as prima facie evidence of a degree of association between the regions, the central issue of this paper. While relatively large spikes have continued to occur in VIC and QLD during the latter part of the sample, spikes in NSW during this time have been very small in comparison.

Spare interconnector capacity will be used to help explain the transmission of price spikes between the regions, defined as import capacity - metered interconnector flow. Excess capacity from NSW to QLD, $EC_{NSW\rightarrow QLD,t}$ and QLD to NSW, $EC_{QLD\rightarrow NSW,t}$ are computed by comparing capacity and flow on the northerly and southerly directions on the interconnector. The same logic is applied to construct excess capacity from NSW to VIC, $EC_{NSW\rightarrow VIC,t}$ and VIC to NSW, $EC_{VIC\rightarrow NSW,t}$. Figure 2 shows the excess interconnector capacities between the two pairs of regions. The top two panels show $EC_{QLD\rightarrow NSW,t}$ (top panel) and $EC_{NSW\rightarrow QLD,t}$ (second panel) and reveal that excess capacity is often low or zero, more so in the southerly direction into NSW. The last two panels show $EC_{VIC\rightarrow NSW,t}$ (third panel) and $EC_{NSW\rightarrow VIC,t}$ (bottom panel) reveal that capacity is lower in the northerly direction into NSW, and is often constrained. The plot of $EC_{NSW\rightarrow VIC,t}$ in the bottom panel show that flow in the southerly direction is not frequently constrained. The greater the excess capacity, a priori the greater the amount of power that can be imported to combat price spikes due to stresses in a region. When the excess capacity reaches zero, there is no opportunity to import power and the interconnector is said to be constrained and essentially the region is decoupled from the neighbouring region. Variables $EC_{QLD\rightarrow NSW,t}$ and $EC_{NSW\rightarrow QLD,t}$ ($EC_{VIC\rightarrow NSW,t}$ and $EC_{NSW\rightarrow VIC,t}$) will be used to explain the size of spikes in QLD and NSW (VIC and NSW) as imported power into a region can be used to offset the impact of factors such as loss in generation capacity and mitigate the size of spikes.

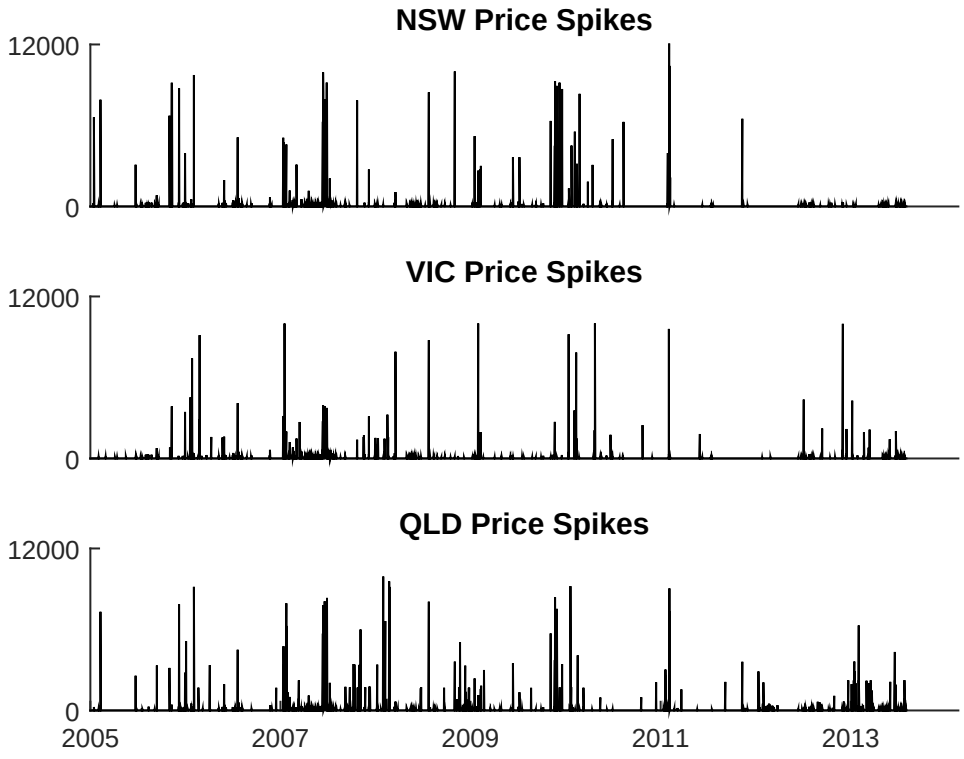

Figure 1: Realised spike events, $P_t \geq \$100/MWh$ for the NSW (top panel) and VIC (middle panel) and QLD (lower panel) markets.

3. Methodology

In this section we introduce a multivariate self-exciting marked point process (SEMPP) in order to characterise the electricity price spikes from the point of view of their frequency, impact and size. In particular, we concentrate on a Hawkes point process (Hawkes, 1971; Embrechts, Liniger and Lin, 2011), which is a special class of SEMPP that can be used to build probabilistic models to capture the instantaneous behaviour of random events based on the history of the process by means of its conditional intensity.

Let $N(t)$ be a $d-$variate marked point process $N = \{N_1, \ldots, N_d\}$ enumerating the

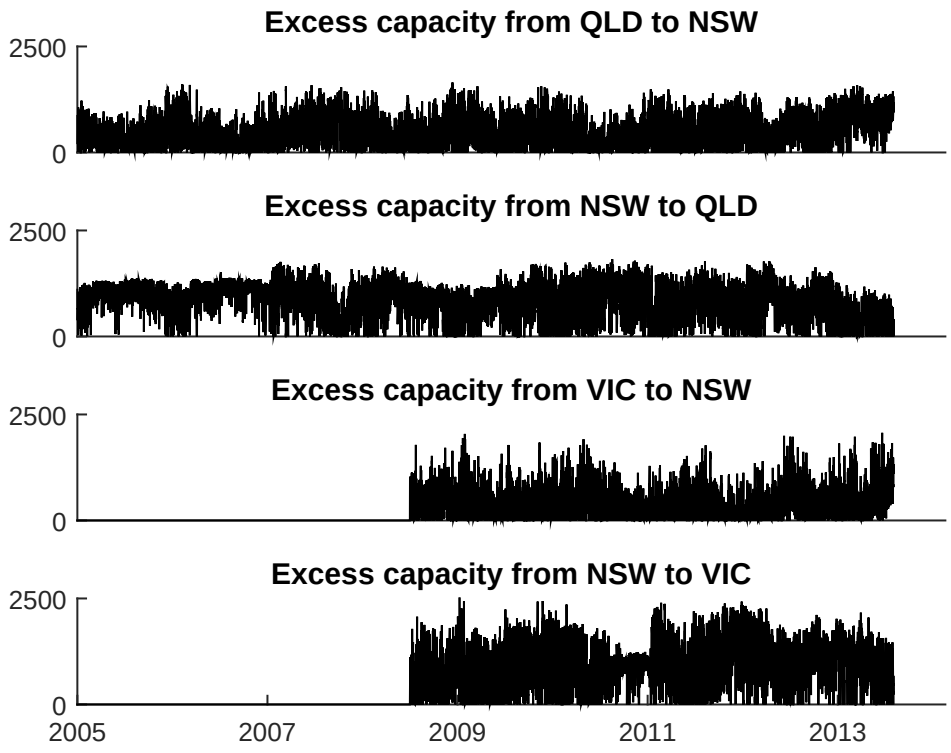

Figure 2: Plots of excess interconnector capacity from QLD to NSW $EC_{QLD\to NSW,t}$ (top panel), NSW to QLD $EC_{NSW\to QLD,t}$ (second panel), VIC to NSW $EC_{VIC\to NSW,t}$ (third panel) and NSW to VIC $EC_{NSW\to VIC,t}$ (bottom panel.)

occurrence of all the events $\{(t_i, Y_i)\}$ of a stochastic process up to time t, where $t_i \in \mathbb{R}$ and Y_i are marks with probability density function f. In this paper, $N(t)$ is either a two-dimensional object comprising price spikes in QLD and NSW, or NSW and VIC in two bivariate models, or a three-dimensional object representing spikes in QLD, NSW and VIC. In all cases, each event has an associated mark capturing the size of the spike.

The defining characteristic of the process $N(t)$ is the intensity with which events

(price spikes) occur. The intensity is given by

$$\lambda\left(t,y \mid \mathscr{H}_t\right) = \lambda_g\left(t \mid \mathscr{H}_t\right) f\left(y \mid t, \mathscr{H}_t\right)$$

in which $\mathscr{H}_t$ may be interpreted as the available information set up to but not including time t. The first term on the right hand side is the ground conditional intensity process, $\lambda_g\left(t \mid \mathscr{H}_t\right)$, which describes the intensity with which events (price spikes) occur. The second term on the right hand side is the probability density function of the marks, Y (size of the price spike). The appearance of this term captures the idea that the ground conditional intensity of the point process depends on the size of the mark, because the size of the price spike may be indicative of the degree of stress in the market.

The ground conditional intensity is assumed to take the following form

$$\lambda_g^j(t,y \mid \mathscr{H}_t) = \left(\mu_j + \sum_{k=1}^{d} \eta_{jk} \int_{(-\infty,t)\times\mathbb{R}} \phi_k\left(y\right) h_j\left(t-s\right) N_k\left(ds \times dy\right)\right) \quad (1)$$

in which the following assumptions, which ensure the existence and uniqueness of a multivariate stationary Hawkes process (Embrechts et al., 2011), are satisfied

$$\int_0^\infty h_j\left(t\right) dt = 1, \qquad \int_0^\infty t h_j\left(t\right) dt < \infty, \qquad \int_{-\infty}^{\infty} \phi_k\left(y\right) f_k(y \mid t, \mathscr{H}_t) dy = 1 \quad (2)$$

for all $j,k \in \{1,\ldots,d\}$. A common specification for h_j is a kernel exhibiting exponential decay which implies a Markov property for the model

$$h_j\left(t-s\right) = \alpha_j e^{-\alpha_j(t-s)}, \qquad s < t, \text{ with } \alpha_j > 0 \text{ for } j \in \{1,\ldots,d\}.$$

In practical applications, the integral in (1) with respect to the counting process N can replaced by its discrete version, thus allowing the ground conditional intensity function to take the simple form

$$\lambda_g^j(t,y \mid \mathscr{H}_t) = \mu_j + \sum_{k=1}^{d} \eta_{jk} \sum_{i:t_i^k<t} \phi_k\left(y\right) \alpha_j e^{-\alpha_j\left(t-t_i^k\right)}. \quad (3)$$

This parameterisation of the ground conditional intensity in equation (3) allows for two important channels of influence on the intensity. The specification adopted for

the form of the kernel function h_j, means that a new event causes the conditional intensity to jump up, and then decay back at speed α_j towards the baseline intensity μ_j. This parameterisation clearly captures the stylised fact that price spikes tend to cluster or self-excite. The first summation in the second term of (3) is over the regions, $k = \{1, \cdots, d\}$. This means that events in other regions affect the intensity in the current region, j, and cross-excitation is therefore allowed. This feature of the model is a significant departure from the existing literature on spot electricity markets.

Another important channel of influence on the intensity concerns the link between the marks, Y, and the intensity of the point process. This effect is captured by means of the impact function, ϕ_k, which takes the following normalised polynomial form

$$\phi_k(y) = \frac{a_k + b_k y + c_k y^2}{a_k + b_k E[Y] + c_k E[Y^2]}$$

with a_k, b_k, c_k unrestricted for either $k \in \{1,2\}$ or $k \in \{1,2,3\}$. This form ensures that the third condition in (2) is satisfied. In order to specify this polynomial function fully in terms of the moments of Y, the probably distribution for the marks of the point process must be specified. Since price spikes, by definition, are extreme events that occur in a very small fraction of the full sample, the generalised Pareto distribution (GPD) is a natural candidate to describe their distribution. The asymptotic tail distribution results given by the Pickands-Balkema-deHaan theorem[3], suggest that the GPD function is the best choice of limiting function for modelling extremely large price movements. The probability density function of the GPD is defined by

$$f_k(y \mid t, \mathscr{H}_t) = \begin{cases} (1/\sigma_k(t))\,(1 + \xi_k y/\sigma_k(t))^{-1-1/\xi_k}; y > 0, & \xi_k \neq 0 \\ (1/\sigma_k(t))\, e^{-y/\sigma_k(t)} & \xi_k = 0, \end{cases} \tag{4}$$

where $\sigma_k(t) > 0$ and ξ_k are scale and shape parameters respectively. Note that $0 \leq y < \infty$ if $\xi_k \geq 0$, and $0 \leq y < -\sigma(t)_k/\xi_k$ if $\xi_k < 0$ for $k \in \{1,2\}$ or $k \in \{1,2,3\}$. The $n-$central moments for the GPD exist if $\xi_k < 1/n$, and are given by $E[Y^n] =$

[3] See Theorem 3.4.13 property (b) in Embrechts et al. (1997).

$\left(n!\sigma(t)_k^n\right)/\prod_{i=1}^{n}(1-i\xi_k)$. This now allows the impact function to be specified in terms of the scale parameter of the GPD

$$\phi_k(y)=\frac{\left(1+b_k y+c_k y^2\right)(1-2\xi_k)(1-\xi_k)}{(1-\xi_k+b_k\sigma(t)_k)(1-2\xi_k)+2c_k\sigma(t)_k^2} \tag{5}$$

with b_k, c_k unrestricted for either $k\in\{1,2\}$ or $k\in\{1,2,3\}$.

The shape parameters, ξ_k, are simply estimated as constant parameters of the problem. However, the scale parameters $\sigma_k(t)$ are modelled in such a way as to be dependent on conditions prevalent in region k at the time of the price spike. A simple linear functional form will be adopted in which scale parameters in the GPD distribution for the marks for each region are determined by

$$\log\sigma_k(t)=\beta_{\mathbf{k}}'\mathbf{X}_{\mathbf{k},\mathbf{t}} \tag{6}$$

where $\beta_{\mathbf{k}}$ is a vector of coefficients and $\mathbf{X}_{\mathbf{k},\mathbf{t}}$ is a vector of covariates observed at time t, $k\in\{1,\ldots,d\}$. This linear specification is fairly flexible. In particular, beyond unexpected load (to be discussed in detail below) it allows the state of the available interconnector capacity between regions at the time of the price spike to enter as explanatory variables in determining $\sigma_k(t)$. This is an additional mechanism by which inter-regional influences are enabled within the structure of the model. As such $\mathbf{X}_{\mathbf{k},\mathbf{t}}$ is specified as $[EC_{k,t};UL_{k,t}]'$ where $UL_{k,t}$ is the unexpected load for region k.

Given the definitions of the ground intensity process, $\lambda_g^j(t,y\mid\mathscr{H}_t)$ in equation (1) and the density for the marks, $f(y\mid t,\mathscr{H}_t)$ in equation (4), the log-likelihood function is

$$\begin{aligned}\log L &= \sum_{j=1}^{d}\int_{\mathscr{T}\times\mathbb{R}}\log\lambda_g^j(s\mid\mathscr{H}_s)N_j(ds\times dl)\\ &+ \sum_{j=1}^{d}\int_{\mathscr{T}\times\mathbb{R}}\log f_j(l\mid s,\mathscr{H}_s)N_j(ds\times dy)-\sum_{j=1}^{d}\Lambda_j(T),\end{aligned} \tag{7}$$

where $\Lambda_j(T)=\int_0^T\lambda_g^j(s\mid\mathscr{H}_s)\,ds$ defines the compensator for all $j\in\{1,\ldots,d\}$.

4. Modelling Load

In order to obtain a series for unexpected load it is necessary to build a model for forecasting load. Building a complete forecasting model for load is well beyond the scope of this paper and is indeed a major undertaking in its own right. What is required here is a simple benchmark model that does an acceptable job of capturing the major features of the load profile and is easy to forecast. Since load exhibits strong seasonal and diurnal patterns a simple forecasting model that takes account of these features is sufficient for the purposes of the current study. Despite the simplicity of the model it has been shown to produce very accurate load forecasts, at least for the Queensland region of the NEM. The model is fully explained in Clements, Hurn and Li, (2015).

A model structure that captures half-hourly variability in load while respecting the well-known features of the load profile is one in which each half hour is modelled separately. Let the logarithm of the load at half hour h and day d be given by L_{hd}, then, the ARMA structure of the prototype model for a given half hour period is

$$L_{hd} = \theta_{h0} + \theta_{h1}L_{hd-1} + \theta_{h2}L_{hd-7} + \phi_{h1}\varepsilon_{hd-1} + \phi_{h2}\varepsilon_{hd-7} + \varepsilon_{hd}\,,$$

in which $h = 1, \cdots, 48$ and ε_{hd} is the disturbance term. So for each half-hour, h, the parameters are estimated based on a subset of the data which only contains the observations at that interval. In this way, the partial correlation between load and lagged load is allowed to differ in a daily pattern by the different parameter values across equations. A minimal lag structure requires L_{hd} to be explained by load in the same half hour on the previous day, L_{hd-1} and the load in the same half hour of the same day in the previous week, L_{hd-7}. For the same reasoning, the unexpected changes in load in the same half hour on the previous day, ε_{hd-1} and the previous week, ε_{hd-7}, are included.

There are three important improvements that can be made to this prototype model.

(i) In order to allow for the coefficients on one-day lagged load to differentiate between days, the one-day lagged load, L_{hd-1}, is interacted with day-of-the-week dummy variables, $\mathbb{W}_{dp}$, $p = 1, \cdots, 7$. Attempts to reduce the number of dummy variables in the specification, for example by using one for week-

days and one for weekends, or defining the dummy variables in terms of whether the day before and after is a weekday or in weekend, produced inferior results.

(ii) It is reasonable to posit that the load in consecutive half hours will be correlated so that in addition to observed load in last half-hour period of the day prior to the making a forecast, L_{48d-1}, each equation should also contain the lagged load from the immediately preceding half hour, L_{h-1d}. Additional lags of consecutive half-hour periods were tried but the improvement in forecast performance was minimal.

(iii) An annual pattern in the load in all the regions of the NEM is allowed for by specifying Fourier polynomials with annual cycles interacted with the one-week lagged load, L_{hd-7}. The degree of the Fourier polynomials in the series expansion is four. While this choice is not tested formally, experimentation showed that little is to be gained by increasing the degree of the polynomials.

Consequently, the preferred multiple equation time series model used for forecasting load is

$$\begin{aligned} L_{hd} =& \theta_{h0} + \theta_{hd1} L_{hd-1} + \theta_{hd2} L_{hd-7} + \theta_{h4} L_{48d-1} + \theta_{h5} I_{h>1} L_{h-1d} \\ &+ \phi_{h1} \varepsilon_{hd-1} + \phi_{h2} \varepsilon_{hd-7} + \varepsilon_{hd}, \end{aligned} \tag{8}$$

in which

$$\theta_{hd1} = \sum_{p=1}^{7} \eta_{hp} \mathbb{W}_{dp},$$

$$\theta_{hd2} = \tau_{h1} + \sum_{q=1}^{4} \left[\tau_{h2q} \sin\left(2q\pi\left(\frac{hd}{17472}\right)\right) + \tau_{h3q} \cos\left(2q\pi\left(\frac{hd}{17472}\right)\right) \right],$$

and $I_{h>1}$ denotes an indicator function which is equal to 1 when $h > 1$ and 0 otherwise. This modification turns the 48 equations for the half hours of a day into a recursive system. Once again, repeated application of ordinary least squares can be used to estimate the system, it provides a parsimonious way of capturing the intra-day load correlation without increasing computational complexity significantly. Experimentation indicates that the more efficient estimation method with taking into account of intra-day error correlation does not generally improve fore-

cast accuracy.

There is some evidence in the literature to suggest that the response of load to temperature is nonlinear in nature and the challenge is to model this nonlinear response but at the same time maintain a model specification that is linear in parameters. Clements et al. (2015) use a flexible spline method to provide a piecewise linear specification in temperature. The advantages of including the temperature variables are marginal even when actual rather than forecast temperature values are used, at least for the Queensland region. The reason for this is that actual load varies quite widely for any given temperature due mainly to the fact that the temperature record is a non-representative one because it is taken at a specific location and then used as a proxy for the temperature in the entire region. This is a questionable assumption given the size of the regions of the NEM.

In this section, the forecast performance of the preferred model in (8) is compared against the industry standard reported by the market operator AEMO. AEMO as the operator of the NEM, provides short-term load forecasts in pre-dispatch IS reports for the next trading day.[4] Among the horizons of the load forecast, 12-hour ahead forecasts provide important information for dispatch planning for the next day. To monitor 12-hour ahead load forecast accuracy, the monthly averaged MAPE of the 12-hour ahead forecasts is reported by AEMO as a benchmark for assessing forecast performance.[5] Although the details of the specification of the AEMO forecasting procedure are not available, it is the main forecasting model chosen by the market operator and may therefore be taken to be representative of the state of art in terms of load forecasting models. Given the limited historical data publicly available from AEMO with respect to their published load forecasts, the period from July 2012 to November 2013 is used for subsequent comparison. A comparison of the performance of the preferred multiple equation model (with and without temperature variables) shown in Table 2

[4] See, `http://www.nemweb.com.au/REPORTS/CURRENT/PreDispatchIS_Reports/`.

[5] See,`http://www.aemo.com.au/Electricity/Data/PreDispatch-Demand-Forecasting-Performance`

Table 2: Summary comparison of 12-hour ahead forecast by equation (8) and the AEMO forecasts for the period from July 2012 to November 2013. The † symbol indicates that the forecast was generated using actual temperature data and the flexible spline procedure described in Clements et al. (2015).

	Eq. (8)†	Eq. (8)	AEMO
Overall MAPE	1.21%	1.37%	1.88%
Max. APE	20.21%	20.26%	-
No. APE $\geq 5\%$	384	585	-
No. APE $\geq 10\%$	38	44	-
No. APE $\geq 15\%$	7	7	-
No. APE $\geq 25\%$	0	0	-
Max. monthly MAPE	1.84%	2.02%	3.2%
Obs.	24864	24864	-

Although this period is only 17 months, the advantage of the proposed model is shown clearly in Table 2, with the monthly MAPEs well below the AEMO forecasts and an improvement of around 0.67% in the overall MAPE over the AMEO forecasts. Since AEMO forecasts are based on temperature forecasts instead of real temperature, the results from the proposed model obtained by omitting the variables for current temperature are also reported. While there is a fall in accuracy relative to the situation when actual temperature is used, this effect is very small and the model is still more accurate than the AEMO forecast under all criteria (0.51% lower in the overall MAPE).

5. Estimation Results

Table 3 reports the parameter estimates and standard errors for the three models, namely, the bivariate QLD-NSW model, the bivariate NSW-VIC model and the QLD-NSW-VIC model. The interpretation of the parameters of the models reported in Table 3 is made easier by recognising that the parameters are indexed by up to three indices, $m = 1,2,3$ for three models, $j = 1,2,3$ for the regions QLD, NSW and VIC, respectively, and k is either a dummy index for the repetition of the j index or represents the number of explanatory variables in equation. It is useful to bear in mind the interpretations of the parameters of the model.

(i) The parameters μ_{mj} are the constant terms in the intensity equation (3).

(ii) The parameters η_{mjk} summarise the self- and cross-excitation properties of the model in the intensity equation (3). When $j = k$ the parameter is a self-excitation parameter and when $j \neq k$ the term captures a cross-excitation effect.

(iii) The parameters α_{mj} are the estimated rates of exponential decay of the intensity equation (3).

(iv) The parameters b_{mj} and c_{mj} are coefficients of the impact function in equation (5) which control the shock to intensity conditional on the size of a spike.

(iv) The parameters ξ_{mj} are the shape parameters of the of GPD distribution.

(vi) The parameters β_{mjk} with $k = 0,1,2,3$ are parameters relating to the scale of the GPD function in equation (6). The parameters β_{mk0} are the constant terms. If $k > 0$ and $j = k$ then β_{mjk} are the coefficients on unexpected load. If $k > 0$ and $j \neq k$ then β_{mjk} are on excess capacity.

As a further aid to the interpretation of the results in Table 3, the coefficients governing the self-excitation of the intensity of the SEMPP are coloured grey, as are the coefficients governing the effect of unexpected load on the scale of the GPD distribution.

In all regions and for all models, the coefficients η_{mjk} with $j = k$ (coloured grey) relating to the self-excitation of the SEMPP are statistically significant. This result confirms that there is clustering in the occurrence of extreme price events in all regions of the NEM. Furthermore, in all but one case (η_{121} in the QLD-NSW model) the coefficient estimates on the cross excitation terms are positive and significant. This indicates that the occurrence of a spike in one region will increase the chance of a spike in the neighbouring regions. While the results are not reported here, restricted models with no cross excitation are also estimated. For all three models, the value of the log-likelihood function is significantly reduced.[6]

[6]The restricted and unrestricted values of the log-likelihood functions, respectively, are -17831.18 to -17927.50 for QLD-NSW, -5116.16 to -5135.15 for NSW-VIC and -8460.076 to -8568.90 for the three region model.

		Unrestricted Hawkes QLD-NSW				Unrestricted Hawkes VIC-NSW					Unrestricted Hawkes QLD-NSW-VIC					
		QLD		NSW		VIC		NSW			QLD		NSW		VIC	
j		m=1,j=1		m=1, j=2		m=2,j=1		m=2,j=2			m=3,j=1		m=3,j=2		m=3,j=3	
$\lambda_g^j(t \mid \mathscr{H}_t)$	μ_{mj}	0.0031	(0.0002)	0.0030	(0.0002)	0.0016	(0.0001)	0.0017	(0.0002)	μ_{mj}	0.0024	(0.0002)	0.0017	(0.0001)	0.0016	(0.0001)
	η_{mj1}	0.6672	(0.0235)	0.0095	(0.0060)	0.6558	(0.0475)	0.0402	(0.0128)	η_{mj1}	0.5732	(0.0633)	0.0109	(0.0072)		
	η_{mj2}	0.1565	(0.0170)	0.8308	(0.0187)	0.0307	(0.0091)	0.7545	(0.0368)	η_{mj2}	0.1685	(0.0272)	0.7916	(0.0335)	0.0344	(0.0099)
	η_{mj3}									η_{mj3}			0.0377	(0.0121)	0.6503	(0.0477)
	α_{mj}	0.3758	(0.0145)	0.4049	(0.0129)	0.4402	(0.0310)	0.4897	(0.0262)	α_{mj}	0.4267	(0.0255)	0.4895	(0.0242)	0.4438	(0.0311)
	b_{mj}	0.7205	(0.1827)	0.8075	(0.1493)	0.4248	(0.2269)	0.4975	(0.1930)	b_{mj}	0.7599	(0.3721)	0.1347	(0.0668)	0.4229	(0.2264)
	c_{mj}	-0.1614	(0.0439)	-0.1579	(0.0351)	-0.0804	(0.0503)	-0.0924	(0.0418)	c_{mj}	-0.1760	(0.0931)	-0.0008	(0.0037)	-0.0796	(0.0502)
$f_j(y \mid t, \mathscr{H}_t)$	ξ_{mj}	0.2207	(0.0246)	0.0919	(0.0186)	0.3509	(0.0477)	0.1804	(0.0404)	ξ_{mj}	0.2681	(0.0656)	0.1554	(0.0392)	0.3523	(0.0479)
	β_{j0}	-0.5183	(0.0406)	-0.4431	(0.0379)	-0.4165	(0.1089)	-0.2933	(0.0908)	β_{mj0}	-0.4733	(0.0968)	0.0158	(0.1024)	-0.4172	(0.1093)
	β_{mj1}	0.8966	(0.1493)	-1.9322	(0.1288)	0.4949	(0.1766)	-2.1335	(0.1876)	β_{j1}	0.2410	(0.2594)	-2.2008	(0.3099)		
	β_{mj2}	-1.0020	(0.1260)	1.7309	(0.1199)	-2.3420	(0.2683)	1.2445	(0.2363)	β_{mj2}	-0.2246	(0.3559)	0.8048	(0.2558)	-2.3162	(0.2684)
	β_{mj3}									β_{mj3}			-2.0290	(0.1943)	0.4902	(0.1782)

Table 3: Estimated maximum likelihood parameters for the unrestricted Hawkes models in the period 1 January 2005 to 31 December 2012, for the pair QLD-NSW and from 1 July 2008 to 31 December 2012 for the pair VIC-NSW. Figures in parentheses are standard errors. The coefficients governing the self-excitation of the intensity of the SEMPP are coloured grey, as are the coefficients governing the effect of unexpected load on the scale of the GPD distribution. The log-likelihood value for the QLD-NSW model is -17831.18 and for the VIC-NSW model is -5116.16. In the three region model, the sample is from 1 July 2008 to 31 December 2012 and the log-likelihood value is -8460.076. The estimation was conducted on log-prices which simplifies the estimation process.

The results in Table 3 indicate that the size of the spikes are significantly related to both excess capacity and unexpected load. The positive estimates of β_{mjk}, which are coloured in grey, indicate that the unexpected load shocks have a significant positive impact on the size of price spikes, through their effect on the scale coefficient in equation (6). The single exception is the estimate for QLD in the three region model, β_{311}, which is positive but not significant. Significant negative estimates for β_{mjk} when $j \neq k$ are as expected: the greater the excess capacity on an interconnector into a region, the smaller the expected size of the spike.

The robust conclusion to emerge from Table 3 is that inter-regional influences matter. Not only is cross-excitation in the intensity important, but the flow of electricity across regional boundaries via the interconnectors is also a significant factor in determining the expected size of price spikes. It is clear that a multivariate framework should be preferred to a univariate one when modelling spikes in electricity prices.

Figure 3 plots the impact functions, $\phi_k(y)$ for the bivariate QLD-NSW model. These surfaces reveal the impact on intensity of the size of the event (y) and the observed covariates through the scale parameters $\sigma_k(t)$ in equation (6), *given that a price spike has occurred*. It appears that as the scale coefficient increases (due to either higher unexpected load and/or lower excess capacity) there is a general increase in the value of the impact function. The influence of the size of a price spike (marks), on the other hand appears to be quadratic. One possible explanation to this nonlinear effect is the phenomenon of rebidding by base load generators. For larger spikes there is a strategic incentive for generators to rebid generation capacity at the market floor price in order to ensure their bids are dispatched (see, for example, Hurn, Silvennoinen and Teräsvirta, 2015). This extra capacity can have a calming impact on the intensity of further spikes becomes smaller as the size of the price spike grows. Results for the impact functions for the bivariate NSW-VIC model are identical in nature and not reported here.

Figure 4 shows the impact functions from the three-region model. Interestingly, the estimates of $\phi_k(y)$ for QLD and VIC are the same as those obtained from the bivariate models, but the pattern for NSW is different. The effect of the scale coefficient is slightly negative and very small, a result which is probably due to the interaction of two excess capacities from the other two regions. The effect of

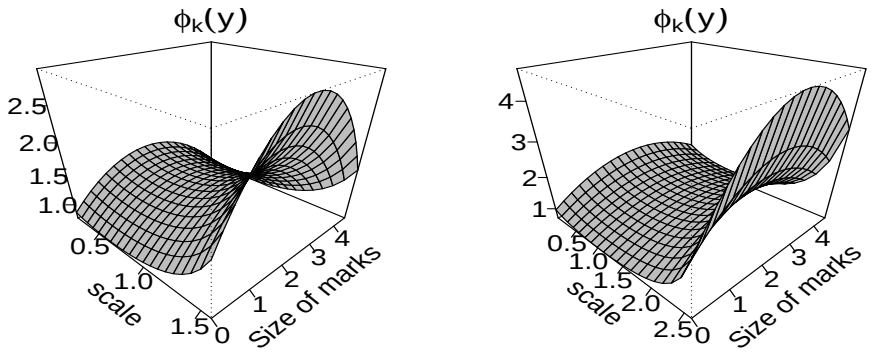

Figure 3: Impact function $\phi_k(y)$ as a function of both the size of the marks, y, and scale parameter $\sigma_k(t)$ for the bivariate QLD-NSW model. The QLD function is in the left panel and the NSW function is in the right panel.

the size of the spike on intensity is monotonically increasing, though the effect is smaller than in either of the bivariate models.

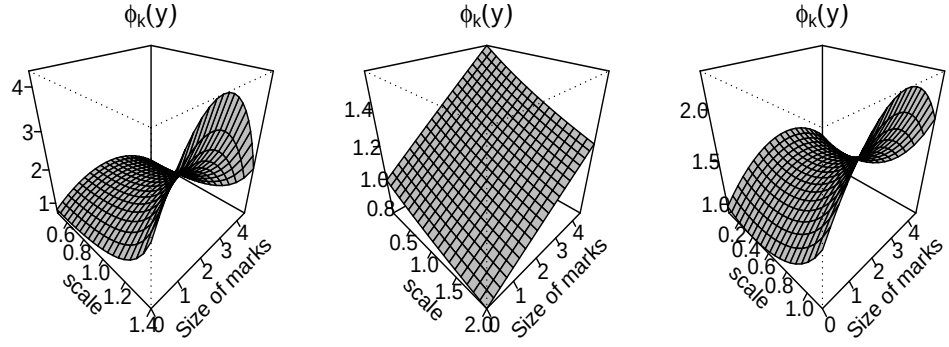

Figure 4: Impact function $\phi_k(y)$ as a function of both the size of the marks, y, and scale parameter $\sigma_k(t)$ for the QLD-NSW-VIC model. The impact functions for QLD, NSW and VIC are shown in the left, centre and right panels, respectively.

A common approach to gauge the goodness of fit of a Hawkes model is by means of the residual analysis proposed by Ogata (1988). The idea is to obtain the residual process of the model through the compensator

$$\tau_{i+1}^j = \Lambda_j\left(t_i^j, t_{i+1}^j\right) = \int_{t_i^j}^{t_{i+1}^j} \lambda_g^j\left(s \mid \mathscr{H}_s\right) ds$$

for each dimension $j \in \{1,\ldots,d\}$. According to the time change property for a point process, the residual $\{\tau_i^j\}$ should closely resemble a realisation of a unit rate Poisson process if the model is well defined. If this property is satisfied, the exceedence times of the process will be exponentially distributed with unit rate.

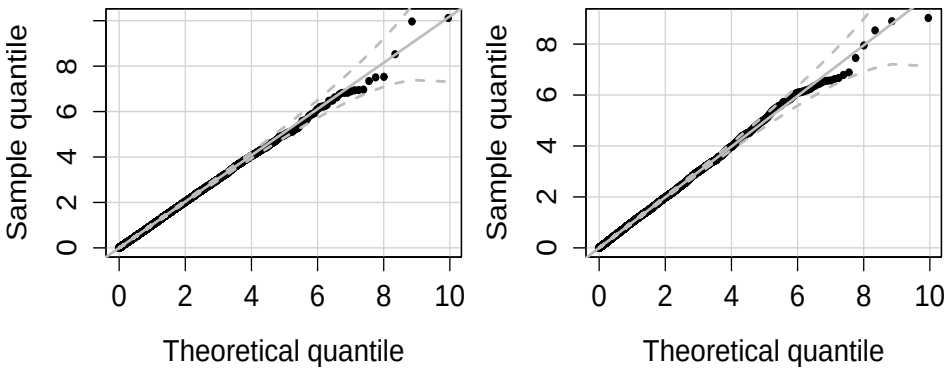

Figure 5: Goodness of fit in-sample: QQ-plots of the residual process for the bivariate QLD-NSW Hawkes model. Results for QLD in the left panel and NSW in the right panel.

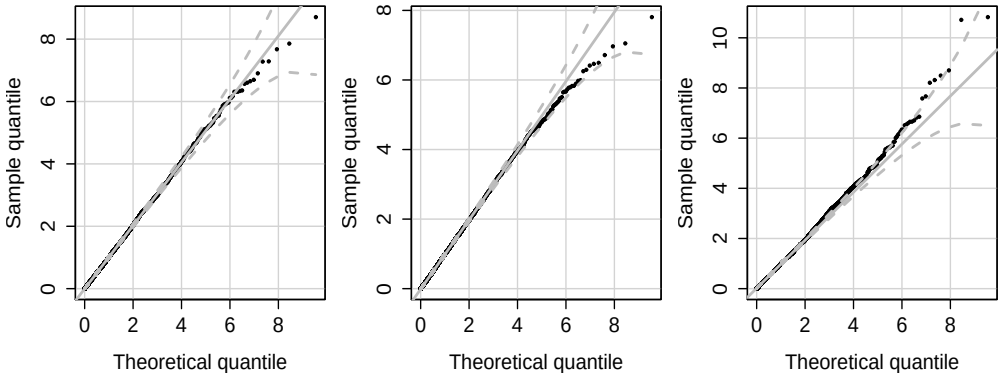

Figure 6: Goodness of fit in-sample: QQ-plots of the residual process for the three region Hawkes model. Results for QLD in the left panel, NSW in the centre panel and VIC in the right panel.

Figure 5 displays the QQ-plots realisations for the residual process for both regions in the QLD-NSW bivariate model against an exponential distribution. Clearly the model captures the clustering in spikes well for both regions because the residuals are consistent with the exponential distribution. This result is also true for the three-region model shown in Figure 6 and for the NSW-VIC bivariate model, which is

not reported.

6. Forecast Accuracy

The results in Section 5 show that inter-regional linkages lead to a superior in-sample fit for the SEMPP model. The task of this section is to determine whether or not they lead to superior forecasts. Providing a standard benchmark model against which to measure performance is not entirely straightforward, simply because there is no existing method that deals with both the intensity and the size of the spike, while also being generalisable to many regions. An early adaptation of the Autoregressive Conditional Hazard (ACH) model developed by Hamilton and Jordà (2002) to electricity prices (Christensen et al., 2012), which models both the occurrence of the event and the size of the event (marks), cannot be generalised to the multivariate setting. The same is true for the duration based ACD-Peaks Over Threshold approach of Herrera and Gonzalez (2014). The dynamic logit framework used by Eichler et al. (2013) is potentially generalisable to the multivariate setting, but it cannot deal with marked point processes. Given the lack of a direct competitor, the importance of the inter-regional links in predicting the intensity of future spikes is assessed by comparing the Hawkes models to their restricted counterparts in which all cross-excitation terms are set to zero, $\eta_{mjk} = 0$ for $j \neq k$. Where appropriate the results will also be related to those reported by Christensen et al. (2012) and Clements et al. (2013).

Forecasts of the intensity from the restricted and unrestricted models are based on the history of spike times, but both use covariates from the next half hour forecast period to determine the expected size of a spike so that no stance is taken here on a forecasting model for the covariates. Forecasting performance is assessed by using the parameters of the models estimated in-sample. These estimates are then used to provide half-hour ahead forecasts for the period 1 January 00:00 to 31 July 23:30 2013. Although new observations are available every half-hour, the model is not re-estimated because this proved to be too computationally expensive. The forecast analysis is based on the one-step-ahead probability of an event occurring at time $t+1$ instead the conditional intensity. This probability can be obtained for

each region j through the intensity measure as follows:

$$P\left\{N_j(t+1) - N_j(t) = 1 \mid \mathscr{H}_t\right\} \quad = \quad 1 - \exp\left(-\Lambda_j\left([t,t+1)\times(y,\infty)\right)\right)$$

where

$$\Lambda_j\left([t,t+1)\times(y,\infty)\right) = \int_t^{t+1}\int_y^{\infty} \lambda^j(s,l \mid \mathscr{H}_s)$$

defines the intensity measure. Solving this integral we have that

$$\Lambda_j\left([t,t+1)\times(y,\infty)\right) = \lambda_g^j\left(t+1 \mid \mathscr{H}_{t+1}\right) F_j(y \mid t+1, \mathscr{H}_{t+1}),$$

where F_j is the cumulative distribution function for the GPD defined in equation 4. Figure 7 displays the probability forecasts for price spikes from the unrestricted three region Hawkes model and shows, informally, that the model does a good job of forecasting spikes and adapts relatively well to clusters of spike events.

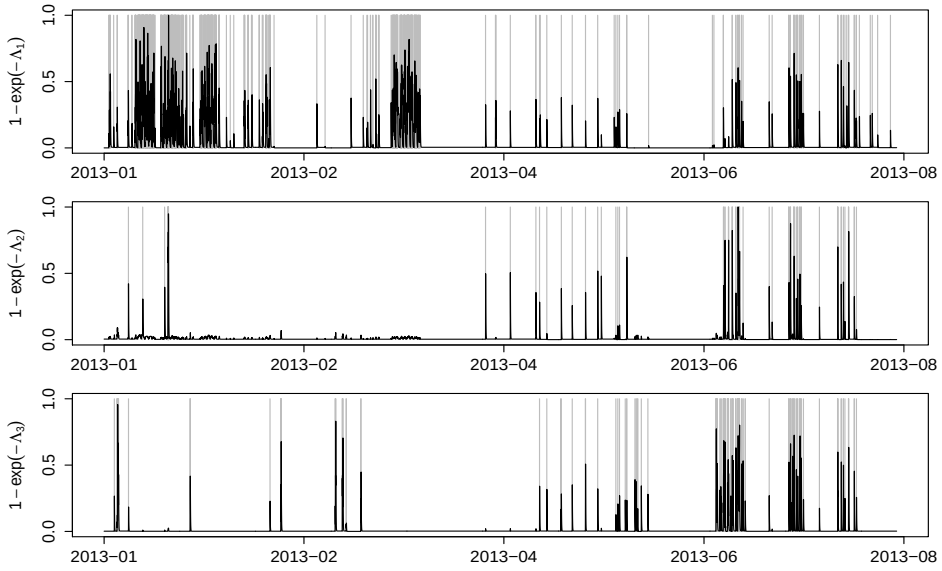

Figure 7: Half-hourly step-ahead probability forecasts from the unrestricted three region Hawkes model for the period 1 January 00:00 hrs to 31 July 23:30 hrs 2013. Price spikes P_t higher than AUS$ 100/MWh are displayed with a gray vertical bar. The top panel shows results for QLD, middle panel NSW and bottom panel VIC. $\Lambda_j := \Lambda_j\left([t,t+1)\times(y,\infty)\right)$ defines the intensity measure for the region j.

A formal evaluation of forecast accuracy follows the procedure of Christensen et al.

(2012) and assumes that extreme events are correctly forecast if a spike occurs at time $t+1$ and the probability of this event is higher than 0.5. On the contrary, a false alarm occurs if an event does not occur at time $t+1$ but the probability forecast is higher than 0.5. The choice of 0.5 as the probability cutoff is essentially arbitrary. A lower threshold improves the percentage of correct forecasts but also increases the probability of false alarms.

Table 4 reports the probability forecasts results from both the restricted and unrestricted bivariate models for the QLD-NSW model. Price spikes as classified as mild if $\$100 \geq P_t \leq \$300$, and severe if $\$300 \geq P_t \leq \$12500$, where all prices are per MWh. For both regions, the general unrestricted bivariate Hawkes model gives the best forecasting results. In QLD, a large number of price spikes occurred during the forecast period. During this time, the unrestricted model correctly forecast 60% of overall spikes (66% for mild and 34% of severe spikes) with a false detection rate of 32%. In comparison, the restricted model correctly forecast spikes at a rate of only 42% with false detection at 57%. Many fewer spikes (only much smaller spikes) occur in NSW during the forecast period. The restricted model in fact performs slightly better in this region. Of co-spikes (simultaneous events in both regions), the unrestricted model accurately predicts 29% of their occurrence and produces 10% false alarms. Given the threshold of 0.5, 26% of first spikes in QLD can be correctly forecast, 13% of first spikes in VIC and only 6% in NSW. This result is consistent in that spikes in NSW are fewer and smaller during the forecast period and hence are more difficult to predict.

Table 5 reports the equivalent results for the bivariate NSW-VIC model. Results for NSW from the unrestricted model are virtually identical to those from the QLD-NSW model, with the restricted model being less accurate in this case. For the VIC region, the unrestricted model correctly forecasts 45% of overall spikes (45% for mild and only 1 of 13 severe spikes) with a false detection rate of only 13%. Results for the three region model are not included as they are very similar to those from the bivariate models with only slight reductions in accuracy. This result is of little surprise given the shorter sample period along with an increase in the number of parameters to be estimated increases leading to greater estimation uncertainty. Overall is appears that the inclusion of inter-regional linkages in the model is the

most important single factor that improves forecast performance.

Region	Model	Event Classification	Types of Events		
			Mild	Severe	All
QLD	Hawkes	True N°	594	153	747
		Correct	394	52	446
		False alarms			237
	Restricted Hawkes	Correct	278	32	310
		False alarms			427
NSW	Hawkes	True N°	84	0	84
		Correct	34	0	34
		False alarms			39
	Restricted Hawkes	Correct	14	0	14
		False alarms			26
QLD - NSW	Hawkes	True N° co-spikes	72	0	72
		Correct			21
		False alarms			7

Table 4: Results of the half-hourly step-ahead probability forecasts for the regions QLD NSW, during the period 1 January 00:00 hrs to 31 July 23:30 hrs 2013. The price spikes P_t are classified in mild if $\$100 \geq P_t \leq \$300$, and severe if $\$300 \leq P_t \geq \$12500$ where all prices are per MWh. The estimation sample is 1 January 2005 to 31 December 2013.

In terms of the accuracy of probability forecasts, the results from the SEMPP models are superior to those of Christensen et al. (2012, Table 5). Not only is the percentage of correct forecasts at least as good as the ACH model forecasts, the SEMPP model is vastly superior in terms of the prediction of the size of the marks. In fact, the ACH model proved incapable of accurately picking any severe events at all in 3 of the 4 regions of the NEM considered. The semi-parametric approach of Clements et al. (2013) provides a solid forecasting performance for spikes, which gives superior accuracy (around 70-75%) to the results reported here. Note, however, that these results are not directly comparable to those of the SEMPP because they refer to different sample periods and the semi-parametric model is only capable of predicting the occurrence of an event and not its mark.

A final point about the forecasts concerns the ability of the model to pick the first spike in a cluster correctly. The results show that 26% of first spikes in QLD can be correctly forecast, 13% of first spikes in VIC but only 6% in NSW if a threshold probability of 0.5 is adopted. NSW is clearly a problem area here but, as we em-

Region	Model	Event Classification	Types of Events		
			Mild	Severe	All
VIC	Hawkes	True N°	155	13	169
		Correct	75	1	76
		False alarms			22
	Restricted Hawkes	Correct	53	1	54
		False alarms			46
NSW	Hawkes	True N°	84	0	84
		Correct	35	0	35
		False alarms			39
	Restricted Hawkes	Correct	18	0	18
		False alarms			38
NSW - VIC	Hawkes	True N° co-spikes	65	0	65
		Correct			29
		False alarms			11

Table 5: Results of the half-hourly step-ahead probability forecast for the regions VIC and NSW, during the period 1 January 00:00 hrs to 31 July 23:30 hrs 2013. The price spikes P_t are classified in mild if AUS$ 100/MWh$< P_t <$AUS$ 300/MWh, and severe if AUS$ 300/MWh$\leq P_t \leq$AUS$ 12500/MWh. Model estimation is based on the sample, July 1, 2008 to December 31, 2012.

phasise in the paper, the sample period we deal with shows very few spikes in NSW during the forecast period making them particularly difficult to model. While these numbers may appear on the low side, they are actually quite respectable in terms of the published literature. Given the asymmetric nature of the loss involved when failing to predict a spike, Clements et al. (2013) suggest a simple rule for varying the threshold to obtain better predictions, specifically, they suggest that using a threshold of 0.1 for predicting the first spike, followed by raising the threshold to 0.75 for subsequent spikes gives the best forecast performance. The percentage of first spikes correctly forecast using this rule reported by Clements et al. (2013) are 32.7%, 11.4%, and 20.8% of the first spikes for NSW, QLD and VIC respectively. This avenue of inquiry has not been pursued here.

A different approach to measuring the forecast accuracy of the models is by means of estimating Value at Risk (VaR) for different levels of confidence. In a traditional finance setting, Value-at-Risk (VaR) represents a dollar loss amount that we are 95% confident will not be exceeded in a particular period (where, for the sake of transparency, a conventional confidence level is assumed). A related measure of

risk is Expected Shortfall (ES) which represents the size of the expected loss given that the VaR has been exceeded. In the context in which it is used in this paper, VaR measures the level of price that we are 95% confident will not be exceeded in a particular period, specifically the next half hour. ES then measures the expected level of price in a half hour period in which the VaR level of price has been exceeded. These risk estimates allow market participants to gain an understanding of their cashflow risks (either costs or revenues) as a result of abnormally high wholesale prices. Chavez-Demoulin and McGill (2012) shows that for the j-th dimension of a Hawkes POT model, the conditional VaR for a confidence level α is

$$VaR_{j,\alpha}^{t+1} = u_j + \frac{\sigma_j(t)}{\xi_j}\left(\left(\frac{1-\alpha}{\lambda_g^j(t \mid \mathscr{H}_t)}\right)^{\xi_j} - 1\right),$$

while the associated conditional Expected Shortfall (ES) is calculated as

$$ES_{j,\alpha}^{t+1} = \frac{VaR_{j,\alpha}^{t+1} + \sigma_j(t) - \xi_j u_j}{1-\xi_j}.$$

In order to determine the accuracy of VaR estimates, a range of statistical tests are applied. A detailed explanation of each test is outside the scope of this paper, however, they are common in the literature relating to VaR estimation, and a description of each of test can be found in Herrera (2013). An important concept used in the construction of these tests is that of an *exception*, which is defined as a price spike whose value exceeds the estimated VaR,

$$I_t(\alpha) = \begin{cases} 1 & \text{if } P_t > VaR_\alpha^{t+1} \\ 0 & \text{otherwise.} \end{cases}$$

Four tests to measure the accuracy of the VaR estimates are employed. The first three tests correspond to Likelihood-Ratio (LR) tests introduced by Christoffersen (1998). The first is an unconditional coverage test (LR_{uc}) to measure whether the fraction of exceptions obtained for the VaR is indeed its expected value. The second corresponds to a test of independence (LR_{ind}) to test independence among the exceptions. The third is the conditional coverage test (LR_{cc}), which is a combination of the last two tests to determine independence and correct coverage. Finally,

Regions	QLD			NSW			VIC			NSW		
α	0.995	0.999	0.9995	0.995	0.999	0.9995	0.995	0.999	0.9995	0.995	0.999	0.9995
Hawkes Model												
Exc.(%)	0.56	0.18	0.09	0.34	0.05	0.02	0.52	0.10	0.03	0.4	0.16	0.04
LR_{uc}	0.33	0.02	0.06	0.02	0.06	0.11	0.71	0.87	0.29	0.14	0.06	0.57
LR_{ind}	0.42	0.79	0.89	0.12	0.94	0.98	0.29	0.88	0.97	0.01	0.81	0.96
LR_{cc}	0.45	0.05	0.17	0.02	0.18	0.28	0.54	0.98	0.57	0.01	0.18	0.85
DQ	0.42	0.79	0.89	0.12	0.94	0.98	0.29	0.88	0.97	0.01	0.81	0.96
Restricted Hawkes Model												
Exc.(%)	0.47	0.18	0.09	0.43	0.21	0.1	0.57	0.08	0.03	0.4	0.1	0.09
LR_{uc}	0.75	0.02	0.14	0.30	0.00	0.03	0.30	0.43	0.29	0.14	0.87	0.14
LR_{ind}	0.49	0.79	0.90	0.02	0.76	0.88	0.00	0.91	0.97	0.56	0.88	0.90
LR_{cc}	0.75	0.06	0.33	0.03	0.01	0.09	0.01	0.72	0.57	0.28	0.98	0.33
DQ	0.49	0.79	0.90	0.02	0.76	0.88	0.00	0.91	0.97	0.56	0.88	0.90

Table 6: Predictive performance for forecasting of the bivariate models based on the accuracy of the VaR estimates. Entries in the rows are the significance levels (p-values) of the respective tests, with exception of the confidence level for the VaR (α), and the percentage of exceptions (Exc. (%)). For the pair QLD - NSW we use the long sample, while for the pair VIC - NSW the short sample.

the Dynamic Quantile test (*DQ*) proposed by Engle and Manganelli (2004) is also employed. The idea of the *DQ* test, is to capture the possible dependence among the exceptions.

The tests are conducted at different confidence levels, with the lowest being 0.995. This is because the threshold selected to define a spike, \$100/MWh, is equivalent to a quantile of 0.992 for events in NSW, 0.929 in QLD and 0.985 in VIC. Results are shown in terms of $p-$values, only for confidence levels over the higher of the two levels across the regions, to make the results comparable over time. Table 6 reports the results of these tests for both the restricted and unrestricted bivariate models for QLD-NSW and NSW-VIC. Overall, across all regions, all tests and levels of significance α, the forecasts from the unrestricted models produce slightly fewer rejections, six against ten from the restricted models. The extra rejections are evenly spread across the different tests. For both models, most of the rejections occur from forecasts in NSW, which is unsurprising given the lack of large spikes during the forecast period. In terms of QLD forecasts, there is little distinguish between the restricted and unrestricted models. In VIC, the unrestricted model fails to adequately forecast the VaR at the lowest level of significance $\alpha = 0.995$. The same patterns are evident in the VaR forecasts from the three region model. While there are not great differences here in terms of VaR forecast accuracy between the restricted and unrestricted models, when considering the superiority of the unrestricted model in terms of forecasting spike probabilities, overall the importance of

the unrestricted model and the interregional links is evident.

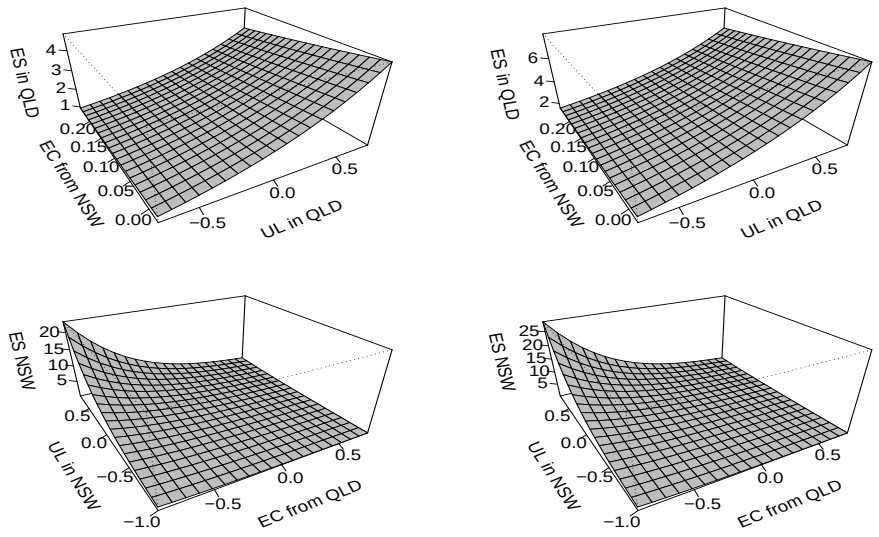

Figure 8: Graphical description of the ES ($\alpha = 0.995$ and $\alpha = 0.9995$ in the left and right columns) as a function of the covariates for the regions of QLD (top panel) and NSW (bottom panel). Observe that the Unexpected load has a positive impact, while the Excess capacity of other region has a negative impact in the estimated ES value.

It is also possible to represent graphically the impact of the covariates on ES at a given confidence level. This occurs through impact of the covariates, namely unexpected load (UC) and excess interconnector capacity (EC), on the scale parameter $\sigma_k(t)$ in the density for the marks. Figure 8 shows a graphical description of the risk in terms of ES as a function of covariates for the regions of QLD (top panel) and NSW (bottom panel) at $\alpha = 0.995$ and $\alpha = 0.9995$ in the left and right columns respectively. Consistent with the parameter estimates, unexpected load has a positive impact, while the excess capacity into each region has a negative impact in the scale specification, and therefore, on the final ES value. The levels of ES become larger moving to the higher levels of confidence. Plots for NSW-VIC are identical in nature and are omitted here.

7. Conclusion

Price spikes in wholesale electricity prices pose a great risk to market participants and accurate forecasting of these events is needed in order to aid the task of managing price risk. The existing literature has modelled the occurrence of price spikes in a univariate framework which confines attention to individual market regions. This paper uses a multivariate point process to model the occurrence and size of extreme price events and in so doing highlights the effect of physical infrastructure on the transmission of price spikes in interconnected regions of the Australian electricity market. This multivariate approach leads to improved in-sample fit, forecasts and estimates of risk measures such as value-at-risk and expected shortfall. Although the results generated by the multivariate self-exciting point process model are promising, a limitation of this study is that there is no simple competitor against which to benchmark the results. An important area of future research will therefore be to place these results in a broader context by developing competing models that allow multivariate transmission of price spikes.

E. J. Anderson, X. Hu, and D. Winchester. Forward contracts in electricity markets: The australian experience. *Energy Policy*, 35:3089–3103, 2006.

M. T. Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12: 287–298, 2002.

L. Bauwens and N. Hautsch. Modelling financial high frequency data using point processes. In T.G. Andersen, R.A. Davis, J-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series*. Springer, 2009.

R. Becker, A. S. Hurn, and V. Pavlov. Modelling spikes in electricity prices. *Economic Record*, 83:371–382, 2007.

M. Bierbrauer, C. Menn, Rachev S. T., and Trück. Spot and derivative pricing in the eex power market. *Journal of Banking and Finance*, 31:3462–3485, 2007.

C.G. Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141:876–912, 2007.

M. Burger, B. Klar, A. Mueller, and G. Schindlmayr. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4:109–122, 2003.

H. N. E. Byström. Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, 14:41–55, 2005.

Á. Cartea and M. G. Figueroa. Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12:313–335, 2005.

V. Chavez-Demoulin and JA McGill. High-frequency financial data modeling using hawkes processes. *Journal of Banking and Finance*, 55, 2012.

T. Christensen, S. Hurn, and K.A. Lindsay. It never rains but it pours: Modeling the persistence of spikes in electricity prices. *The Energy Journal*, 30(1):25–48, 2009.

T. Christensen, S. Hurn, and K.A. Lindsay. Forecasting spikes in electricity prices. *International Journal of Forecasting*, 28:400–411, 2012.

P.F. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39:841–862, 1998.

A. Clements, J. Fuller, and S. Hurn. Semi-parametric forecasting of spikes in electricity prices. *The Economic Record*, 89:508–521, 2013.

A.E. Clements, A.S. Hurn, and Z. Li. Forecasting day-ahead electricity load using a multiple equation time series approach. NCER Working Paper #103, 2015.

J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo. Arima models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18(18):1014–1020, 2003.

J. Crespo Cuaresma, J. Hlouskova, C. Kossmeier, and M. Obersteiner. Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, 77:87–106, 2004.

C. de Jong. The nature of power spikes: A regime-switching approach. *Studies in Nonlinear Dynamics and Econometrics*, 10(3), 2006.

C. de Jong and R. Huisman. Option pricing for power prices with spikes. *Energy Power Risk Management*, 7:12–16, 2003.

M. Eichler, O. Grothe, H. Manner, and D. Türk. Models for short-term forecasting of spike occurrences in the Australian electricity market: a comparitive study. Technical report, 2013.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events*. Springer,Berlin, 1997.

Paul Embrechts, Thomas Liniger, and Lu Lin. Multivariate hawkes processes: an application to financial data. *Journal of Applied Probability*, 48:367–378, 2011.

R. F. Engle and J. R. Russell. Forecasting the frequency of changes in quoted foreign exchange prices with the autoregressive conditional duration model. *Journal of Empirical Finance*, 4:187–212, 1997.

R.F. Engle and S. Manganelli. Caviar. *Journal of Business and Economic Statistics*, 22:367–381, 2004.

Á. Escribano, J. I. Peña, and P. Villaplana. Modelling electricity prices: International evidence, working paper 02-27, universidad carlos iii de madrid. 2002.

R. C. Garcia, J. Contreras, M. van Akkeren, and J. B. Garcia. A garch forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*, 20:867–874, 2005.

J.D. Hamilton and Ò. Jordà. A model of the federal funds rate target. *Journal of Political Economy*, 110:1135–1167, 2002.

A.G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90, 1971.

R. Herrera. Energy risk management through self-exciting marked point process. *Energy Economics*, 38:64–76, 2013.

R. Herrera and N. Gonzalez. The modeling and forecasting of extreme events in electricity spot markets. *Forthcoming in International Journal of Forecasting*, 2014.

H. Higgs. Modelling price and volatility inter-relationships in the aus- tralian wholesale spot electricity markets. *Energy Economics*, 31:748–756, 2009.

H. Higgs and A. Worthington. Stochastic price modelling of high volatility, mean-reverting, spike-prone commodities: The australian wholesale spot electricity market. *Energy Economics*, 30:3172–3185, 2008.

R. Huisman and R. Mahieu. Regime jumps in electricity prices. *Energy Economics*, 25:425–434, 2003.

A.S. Hurn, A. Silvennoinen, and T. Teräsvirta. A smooth transition logit model of the effects of deregulation in the electricity market. *forhtcoming in Journal of Applied Econometrics*, 2015.

K. Ignatieva and S. Trueck. Modeling spot price dependence in australian electricity markets with applications to risk management. Technical report, Working paper, 2014.

T. Kanamura and K. Ōhashi. On transition probabilities of regime switching in electricity prices. *Energy Economics*, 30:1158–1172, 2007.

H. Kauppi and P. Saikkonen. Predicting u.s. recessions with dynamic binary response models. *The Review of Economics and Statistics*, 90:777–791, 2008.

C. R. Knittel and M. R. Roberts. An empirical examination of restructured electricity prices. *Energy Economics*, 27:791–817, 2005.

V. Korniichuk. Forecasting extreme electricity spot prices. Technical report, Cologne Graduate School Working Paper Series 03-14, Cologne Graduate School in Management, Economics and Social Sciences, 2012.

P. Kosater and K. Mosler. Can markov regime-switching models improve power-price forecasts? evidence from german daily power prices. *Applied Energy*, 83, 2006.

J. J. Lucia and E. S. Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5:5–50, 2002.

A. Misiorek, S. Trück, and R. Weron. Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics and Econometrics*, 10, 2006.

T. Mount, Y. Ning, and X. Cai. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*, 28:62–80, 2006.

Y. Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, pages 9–27, 1988.

D. J. Swider and C. Weber. Extended arma models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research*, 77: 583–593, 2007.

R. Weron, M. Bierbrauer, and S. Trück. Modelling electricity prices: jump diffusion and regime switching. *Physica A*, 336:39–48, 2004.

A. Worthington, A. Kay-Spratley, and H. Higgs. Transmission of prices and price volatility in australian electricity spot markets: A multivariate garch analysis. *Energy Economics*, 27:337–350, 2005.