Bayesian Predictive Distributions for Imbalance Prices with Time-varying Factor Impacts

Luana Marangon Lima, *Member, IEEE*, Paul Damien and Derek W. Bunn

***Abstract*—A dynamic Bayesian model is developed to estimate the time-varying nature of the drivers of the system imbalance prices in the British electricity market. We find that the key exogenous factors that significantly influence prices have impacts that evolve substantially over time. Thus, by modeling their evolution with time varying parameter estimation and making conditional forecasts on the latest estimates, more accurate forecasts are produced. Furthermore, using a Bayesian approach allows predictive distributions to be developed, as would be required for value-at-risk compliance purposes. These densities are also found to be more accurate at the extreme quantiles than a conventional GARCH model with static parameters. We validated the superior performance of this Bayesian time varying predictive density method with the same data as in a previously published benchmark model.**

***Index Terms*—Balancing market; Bayesian inference; Predictive Distributions; Pinball loss; Time-varying parameters.**

I. Introduction

Research on balancing markets has been growing rapidly as the uncertainties introduced by renewable resources and greater consumer engagement create significant real-time risks to market participants. If market participants (generators or retailers), produce or consume quantities different to their final nominations to the system operator for each trading period, they will be exposed to the imbalance prices on those volumes. In particular, if retailers fail to adequately hedge their demands, their subsequent exposure to imbalance prices can tip them into financial defaults. For example, in Britain in 2018 there were 21 retailers that went into default and, according to the settlement agency, this was mainly due to a lack of adequate hedging and their consequent exposure to imbalance prices . This has been followed by 30 more retailers that had to cease trading in 2021; see [1]. Imbalance risk is therefore a significant consideration among energy traders.

Research on the design of balancing markets has looked at how the markets should be administered ([2], [3], [4]), whether they should be designed with a single price or separate prices for positive and negative imbalances ([5], [6]) and how they may be susceptible to strategic behavior ([7], [8], [9]). Opportunistic trading strategies based upon predicting the direction of net volume imbalances at system level, whether positive or negative, and thereby creating physical positions in the opposite directions, have been demonstrated as profitable by [10], [11], [12]. Nevertheless, in the Nordic context, [13] argued that imbalances were not predictable at the day ahead stage. However, within the day, close to real-time, forecasting the sign and size of the net imbalance volume at system level has been undertaken successfully by several researchers using different methods, including [14] on the Nordic data. Forecasting the net imbalance volumes is of particular value to the system operators who have to balance the real-time demand and supply during the defined balancing market delivery periods.

Among the first to forecast net imbalance volumes were researchers on the British balancing market, including [15], who compared ARIMA, exponential smoothing and neural networks, as did [16]; the conclusions being that all worked well over short horizons but the simpler methods may be more robust. More recently in the British context, [17] compared machine learning with logistic regression for predicting imbalance volumes during 2017-2019 and concluded that the complexity of the machine learning was not obviously beneficial in terms of accuracy compared to the transparency value of the logistic regression. Outside the British context and going beyond point forecasts, [18] used quantile regressions for risk metrics on the German balancing market, as did [19] for the Austrian balancing volumes. [20] developed a full density forecasting for Austrian imbalance volumes using a skewed -t specification.

There is less research on forecasting imbalance prices, partly because the system operators are more concerned about their volume requirements. Nevertheless, the market participants (generators, retailers and speculators), who may be exposed to imbalance prices on their own trading positions, can benefit substantially from improved predictions. Thus, in the British imbalance prices, [21] recognized that there may be different price formation processes depending upon whether the system operator is seeking to increase or reduce generation in the system, and as a consequence, they found it beneficial to use a Markov-switching model to predict imbalance prices. In contrast, also on British balancing prices, [22] used machine learning techniques to explore the relevant of various real time variables with promising implications for forecasting. Although well-validated, with the focus being upon point forecasts, not density forecasts, they fail to deal with the full risk management needs of the participants. [23] sought to forecast the extreme balancing prices using statistical physics but did not specify the densities. However, from a risk management perspective, the chances of extreme price outcomes are crucial and so the full predictive distributions are often necessary in practice to provide the required precision in estimating the risks. A further consideration is that, even though switching between regimes is somewhat adaptive, it

Luana Marangon Lima is with Duke University, USA
Paul Damien is with University of Texas at Austin, USA
Derek W. Bunn is with London Business School, UK; dbunn@london.edu

assumes that the coefficients relating the various exogenous inputs to prices are constant over time. Yet we know that the market structure for power price formation is evolving and that the impact of some factors (eg coal prices) are declining, whilst others (eg wind speed) have been increasing, with the technological transition towards decarbonisation. Hence, we research an alternative approach to modeling the prices in balancing markets which can deal with both density prediction and time-varying parameters.

Our methodological approach to density forecasting with time varying parameters has its origins in the Bayesian dynamic linear model (DLM) of [24] and [25]. We adapt this to model the British electricity imbalance prices in the dataset used by [21], as previously published in these Transactions. This provides a basis for direct comparison with those previously-published regime switching results. Our new results show that the impacts of the exogenous driving factors for imbalance prices do vary over time and by taking this into account, more accurate forecasts can be produced. In our assessment of the density forecasts against conventional benchmark using the pinball loss functions, the DLM is not uniformly lower than the GARCH or AR-GARCH for all forecasting windows tested, but on average is lower. Importantly, for the forecasting window with the largest volatility, the DLM outperforms the GARCH and AR-GARCH. Also, using the root mean square error, for all of the forecasting windows, the DLM outperforms the GARCH and AR-GARCH. The paper is organized as follows. Section II describes the data and various exogenous variables used in the study. This is followed by the Bayesian model in Section III. Finally, the empirical analysis and results are detailed in Section III-E.

II. Data from the British Market

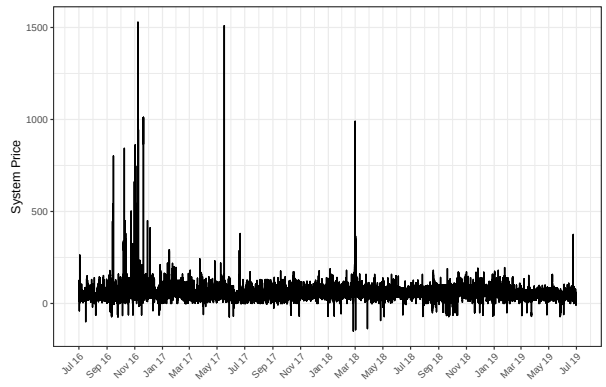
The British Balancing Mechanism operates on 30-minute intervals (”Settlement Periods”, SP) during which forward commitments are delivered and over which imbalance volumes are settled. An hour prior to each SP all participants must notify the System Operator (SO) of their expected physical positions for the SP, namely the amount of electricity to be generated or consumed during that half-hour. This ”gate-closure” point is critical for the balancing mechanism since at this point all flexible generators (and consumers) also inform the SO of their offers/bids to increase/reduce generation (or reduce/increase demand). Conditioned on these offers and bids, the SO produces an order book that is updated during the SP, predicated on the most economic bids and offers to balance the system. In practice many offers and bids are accepted on a minute-by-minute basis to control the balancing. There are 48 SPs in a day. Within each SP, the SO progressively accepts higher offers and lower bids as needed to balance the supply and demand for power and maintain frequency as efficiently as possible. This process repeats under each SP. This episodic nature of the SP process led [21] and others to consider the predictability of prices in the balancing market. Our focus is to consider how the marginal effects of the various exogenous factors that influence prices in each SP may evolve over time. Furthermore, we respond to the need for more information on the risks of extreme prices by seeking to estimate predictive densities from these evolving parameters. A Bayesian approach is natural in this framework.

The data—previously analyzed by [21]—were obtained from the Balancing Mechanism Reporting Service (BMRS) provided by ELEXON[1]. The sampling period is from 1st July 2016 to 30th June 2019, providing 3 years of historical data for developing and testing the statistical model. We use 2016 data to construct the learning phase of the Bayesian model. We set aside the last 2000 values in 2019 for out-of-sample validation. There are 48 data points per day, each representing an SP, resulting in a total of 52,560 data points within each time series.

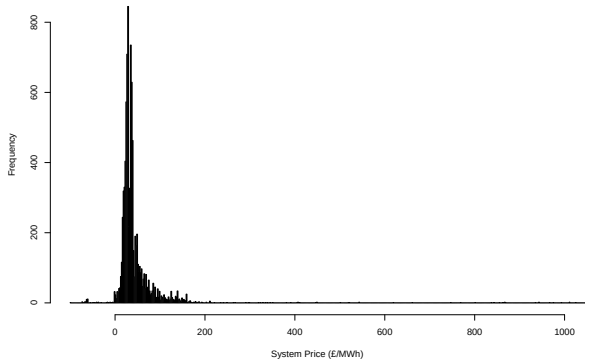
Consider Figure 1 that shows four plots. For the entire sample period, Panel A and B provide the time series and histogram plots of the price data. As already noted by [21] the series has some spikes and occasional negative prices and was therefore trimmed slightly in that analysis. There are pros and cons for such trimming, but in order to directly compare results, we followed the same data pre-processing before fitting the model: all observations larger than 140 were replaced with 140 and all negative prices were replaced with zeros. Panel C and D provide the ACF and PACF plots for the series in Panel A after the trimming process.

The exogenous variables selected for the model are classified as: (a) market state measures (System Price, Net Imbalance Volume (NIV)); (b) demand and supply forecast errors (Wind Error, Solar Error and Demand Error); (c) scarcity indicators in the supply and availability of power volumes (De-rated Margin, Non-Balancing Mechanism (NonBM) and Inter Delta). System Price is the imbalance price already described; NIV is the degree of imbalance per SP in terms of volume of power required for balancing; NonBM refers to extra reserve power sourced by the SO outside the balancing mechanism; Inter delta is the change of power inflows from interconnections from neighboring markets. The correlations of these exogenous factors, including their anticipated effects on system imbalance prices are in Table 1. Since this data is only revealed to the market about 10 minutes after each settlement period, and it would only then be useful for the subsequent period, all the exogenous variables were lagged two periods. All these time series are stationary; see [21]. Additionally, we use an AR(2) variable and a dummy variable. The latter is coded one if NIV is positive, zero otherwise. Along with the intercept term, the evolution of the regression parameter corresponding to the dummy variable (again lagged by two periods) will help distinguish the evolution of prices under the two NIV possibilities. The system imbalance price will be higher or lower than marginal cost dependent on whether system is short ($NIV < 0$) or long ($NIV > 0$). The relationship between system price and net imbalance is depicted in Figure 2.

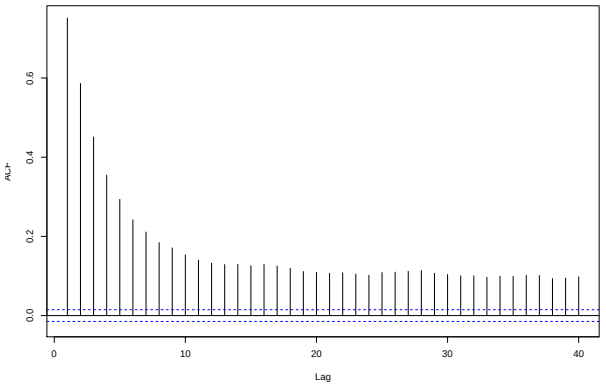
[1] www.elexon.co.uk



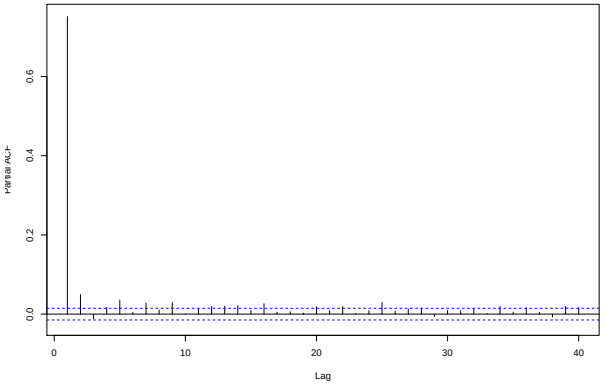
(a) Panel A



(b) Panel B



(c) Panel C



(d) Panel D

Fig. 1: System Price histogram, autocorrelation function (ACF) and partial autocorrelation function (PACF).

	Variable	Correlation Coefficient		
		Lag 0	Lag 1	Lag 2
X_1	Net Imbalance Volume (MWh)	0.70	0.59	0.47
X_2	DRM_1h (MW)	-0.21	-0.19	-0.17
X_3	NONBM (MWh)	0.37	0.33	0.26
X_4	Inter Delta (MWh)	-0.02	0.03	0.02
X_5	Wind Error (MW)	-0.02	-0.01	-0.002
X_6	Solar Error (MW)	-0.15	-0.16	-0.15
X_7	Demand Error (MW)	0.15	0.16	0.15

TABLE I: Correlations between System Price (£/MWh) and Exogenous Factors

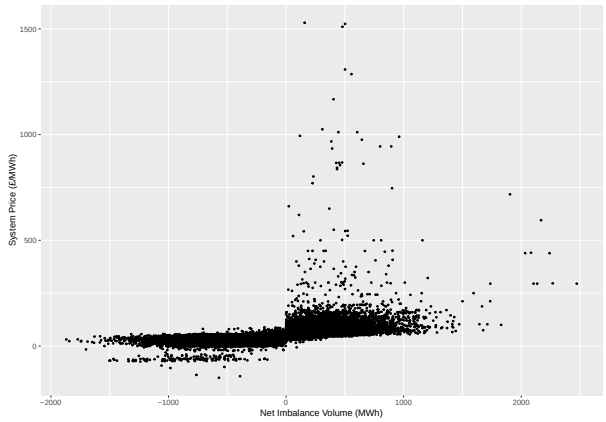


Fig. 2: System Price versus Net Imbalance Volume

III. A Bayesian Dynamic Linear model (DLM)

Following the use of state space models in econometric applications by [26], the subsequent advances to Bayesian inference were detailed in [24] and [25]. Throughout this section, we use bold typeface to denote vectors and matrices. Matrices are always uppercase, vectors can be either lowercase or uppercase.

A. Dynamic Linear Models

We start by formulating the general DLM modeling framework and then we include the context-specific meaning to the mathematical representation. Let $\mathbf{Y} = [Y_1, Y_2, ..., Y_N]$ be a vector comprising the imbalance prices. We aim to use the model for out-of-sample predictions; i.e., predict Y_t for $t = N + 1$. Note that Y_t is a scalar variable. Let $\mathbf{F}_t^{'}$ be a vector containing all the exogenous variables described in the previous section, and let $\boldsymbol{\theta}_t$ be a vector containing the corresponding regression coefficients for all the exogenous variables at time t. Following [24], the observation equation for the univariate normal DLM for time t is given by:

$$Y_t = \mathbf{F}_t^{'}\boldsymbol{\theta}_t + \nu_t, \nu_t \sim N[0, V_t], \tag{1}$$

where $N[0, V_t]$ is the normal distribution with mean 0 and variance V_t also known as the observational variance. For the univariate case, with n predictor variables, $\mathbf{F}_t^{'}$ is a $(1 \,\mathrm{x}\, n)$ vector and $\boldsymbol{\theta}_t$ is a $(n \,\mathrm{x}\, 1)$ vector. The above model is dynamic because the state vector $\boldsymbol{\theta}_t$ changes with t. It is useful to

note that this representation automatically also allows for non-stationarity, although in our dataset all the series are stationary. The dynamic aspect of the model is formally written as

$$\boldsymbol{\theta}_t = \mathbf{G}_t\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \boldsymbol{\omega}_t \sim N[\mathbf{0}, \mathbf{W}_t], \tag{2}$$

where $\mathbf{G}_t$ is a known, $(n \,\mathrm{x}\, n)$ state evolution matrix and $N[\mathbf{0}, \mathbf{W}_t]$ is a multivariate normal distribution with mean $\mathbf{0}$; $\mathbf{W}_t$ is an $(n \,\mathrm{x}\, n)$ evolution covariance matrix for $\boldsymbol{\theta}_t$. This equation captures the evolutionary changes in the regression parameters. The evolution and observation equations may also be expressed for each t as

$$(Y_t|\boldsymbol{\theta}_t) \sim N[\mathbf{F}_t^{'}\boldsymbol{\theta}_t, V_t], \tag{3}$$

$$(\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) \sim N[\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t]. \tag{4}$$

Let D_t be the set containing all information up to and including time t. At time $t = 0$, the DLM is specified as

$$(\boldsymbol{\theta}_0|D_0) \sim N[\mathbf{m}_0, \mathbf{C}_0], \tag{5}$$

for some prior values $\mathbf{m}_0$, $\mathbf{C}_0$. At any subsequent time period t, *before* the datum Y_t is observed, the state equation $(\boldsymbol{\theta}_t|D_{t-1})$ is given by

$$(\boldsymbol{\theta}_t|D_{t-1}) \sim N[a_t, R_t], \tag{6}$$

where

$$\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}, \tag{7}$$

$$\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^{'} + \mathbf{W}_t. \tag{8}$$

The one-step ahead forecast or predictive distribution is given by

$$(Y_t|D_{t-1}) \sim N[f_t, Q_t], \tag{9}$$

where

$$f_t = \mathbf{F}_t^{'}\mathbf{a}_t, \tag{10}$$

$$Q_t = \mathbf{F}_t^{'}\mathbf{R}_t\mathbf{F}_t + V_t. \tag{11}$$

After observing Y_t, we update the information set $D_t = \{D_{t-1}, y_t\}$, which will then be used to update the mean m_t and standard deviation C_t of the state vector. This update is essentially the posterior distribution of the state vector $(\boldsymbol{\theta}_t|D_t)$. These posterior quantities are computed via Bayes's theorem, and are given by

$$(\boldsymbol{\theta}_t|D_t) \sim N[m_t, C_t], \tag{12}$$

where

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t, \tag{13}$$

$$\mathbf{C}_t = \mathbf{R}_t - \mathbf{A}_t Q_t \mathbf{A}_t^{'}, \tag{14}$$

$$\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t Q_t^{-1}, \tag{15}$$

$$e_t = y_t - f_t. \tag{16}$$

The DLM is represented by the quadruple $\{\mathbf{F}_t^{'}, \mathbf{G}_t, V_t, \mathbf{W}_t\}$. Usually $\mathbf{G}_t$ and the relevant values of the sequence $\mathbf{F}_t^{'}$ are known. The evolution variance matrix $\mathbf{W}_t$ is also chosen by the user. Recall that $\mathbf{W}_t$ controls the stochastic variation of the model. If $\mathbf{W}_t = 0$ then there is no variation in the regression parameters, leading to a static regression model. Here we specify $\mathbf{W}_t$ using the discount factor approach described in [24], namely

$$\mathbf{W}_t = \frac{1-\delta}{\delta}\mathbf{G}\mathbf{C}_{t-1}\mathbf{G}^{'}. \tag{17}$$

The roles of $\mathbf{W}_t$, $\mathbf{F}_t^{'}$ and $\mathbf{G}_t$ are further discussed in section III-C. The remaining element of the quadruple, V_t, is often unknown and large relative to the system variance $\mathbf{W}_t$. [24] present a Bayesian learning procedure for unknown observational variance working in terms of the unknown precision parameter $\phi_t = 1/V_t$. A simple closed-form Bayesian analysis is still available if we impose a particular structure on the $\mathbf{W}_t$ sequence and on the initial prior for $\boldsymbol{\theta}_0$. This structure enables a conjugate sequential updating procedure for ϕ_t in addition to $\boldsymbol{\theta}_t$. The conjugate analysis is based on gamma distributions for ϕ_t. As [24] describe, conditional on V_t being known, the DLM will be defined by

$$\text{Obs. eqn.:}\quad (Y_t|\boldsymbol{\theta}_t) \sim N[\mathbf{F}_t^{'}\boldsymbol{\theta}_t, V_t], \tag{18}$$

$$\text{Sys. eqn.:}\quad (\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}, V_t) \sim N[\mathbf{G}_t\boldsymbol{\theta}_{t-1}, V_t\mathbf{W}_t^*], \tag{19}$$

Initial Information:

$$(\boldsymbol{\theta}_0|D_0, V_0) \sim N[\mathbf{m}_0, V_0\mathbf{C}_0^*]. \tag{20}$$

Note that all variances and covariances have V_t as a multiplier providing a scale-free model in terms of the starred variances $\mathbf{C}_0^*$ and $\mathbf{W}_t^*$. For V_t fixed, the model coincides with the original model given by equations (1), (2) and (5), with the scale factor V_t simply being absorbed into these matrices. If V_t is unknown, [24] show that the normal distribution is replaced by Student-t distributions in the DLM. Therefore, the new DLM structure is given by

$$\text{Obs. eqn.:}\quad (Y_t|\boldsymbol{\theta}_t) \sim N[\mathbf{F}_t^{'}\boldsymbol{\theta}_t, V_t], \tag{21}$$

$$\text{Sys. eqn.:}\quad (\boldsymbol{\theta}_t|\boldsymbol{\theta}_{t-1}) \sim T_{n_{t-1}}[\mathbf{G}_t\boldsymbol{\theta}_{t-1}, \mathbf{W}_t], \tag{22}$$

Initial Information:

$$(\boldsymbol{\theta}_0|D_0) \sim T_{n_{t-1}}[\mathbf{m}_0, \mathbf{C}_0], \tag{23}$$

$$(\phi_t|D_0) \sim G[n_0/2, n_0S_0/2], \tag{24}$$

with G denoting the Gamma distribution and $T_{n_{t-1}}[0, \mathbf{W}_t]$ is the Student-t distribution with n_{t-1} degrees of freedom. Note that in specifying the prior, we must choose the prior estimate S_0 and the associated degrees of freedom n_0 in addition to m_0 and C_0; S_0 is a prior point estimate of the observational variance. Now at any subsequent time period t, *before* the datum Y_t is observed, the new state equation $(\boldsymbol{\theta}_t|D_{t-1})$ is given by

$$(\boldsymbol{\theta}_t|D_{t-1}) \sim T_{n_{t-1}}[a_t, R_t], \tag{25}$$

where

$$\mathbf{a}_t = \mathbf{G}_t\mathbf{m}_{t-1}, \tag{26}$$

$$\mathbf{R}_t = \mathbf{G}_t\mathbf{C}_{t-1}\mathbf{G}_t^{'} + \mathbf{W}_t. \tag{27}$$

The new one-step ahead forecast or predictive distribution is given by

$$(Y_t|D_{t-1}) \sim T_{n_{t-1}}[f_t, Q_t], \tag{28}$$

where

$$f_t = \mathbf{F}_t^{'}\mathbf{a}_t, \tag{29}$$

$$Q_t = \mathbf{F}_t^{'}\mathbf{R}_t\mathbf{F}_t + S_t. \tag{30}$$

The new posterior distribution of the state vector $(\boldsymbol{\theta}_t|D_t)$ is given by

$$(\boldsymbol{\theta}_t|D_t) \sim T_{n_{t-1}}[m_t, C_t], \tag{31}$$

where

$$\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t, \tag{32}$$

$$\mathbf{C}_t = \frac{S_t}{S_{t-1}}\left(\mathbf{R}_t - \mathbf{A}_t\mathbf{A}_t^{'}Q_t\right), \tag{33}$$

$$\mathbf{A}_t = \mathbf{R}_t\mathbf{F}_t Q_t^{-1}, \tag{34}$$

$$e_t = Y_t - f_t, \tag{35}$$

and

$$(\phi_t|D_t) \sim G[n_t/2, n_t S_t/2], \tag{36}$$

where

$$n_t = n_{t-1} + 1, \tag{37}$$

$$S_t = S_{t-1} + \frac{S_{t-1}}{n_t}\left(\frac{e_t^2}{Q_t} - 1\right). \tag{38}$$

If interest lies in forecasting k steps ahead, the forecasting distributions are given by

$$(\boldsymbol{\theta}_{t+k}|D_t) \sim T_{n_t}[\mathbf{a}_t(k), \mathbf{R}_t(k)], \tag{39}$$

$$(Y_{t+k}|D_t) \sim T_{n_t}[\mathbf{f}_t(k), \mathbf{Q}_t(k)], \tag{40}$$

where

$$f_t(k) = \mathbf{F}_{t+k}^{'}\mathbf{a}_t(k), \tag{41}$$

$$Q_t(k) = \mathbf{F}_{t+k}^{'}\mathbf{R}_t(k)\mathbf{F}_{t+k} + S_t, \tag{42}$$

$$\mathbf{a}_t(k) = \mathbf{G}_{t+k}\mathbf{a}_t(k-1), \tag{43}$$

$$\mathbf{R}_t(k) = \mathbf{G}_{t+k}\mathbf{R}_t(k-1)\mathbf{G}_{t+k}^{'} + \mathbf{W}_{t+k}, \tag{44}$$

with starting values $\mathbf{a}_t(0) = \mathbf{m}_t$ and $\mathbf{R}_t(0) = \mathbf{C}_t$.

B. The three phases of the DLM process

The process of using the DLM consists of separating the data into three phases: first: prior construction via training data; second: estimation; third: prediction. For the dataset at hand, these phases are as following.

- Phase 1: Data from July to December 2016 correspond to the training phase of the model. These were used to construct the prior parameter values described in the previous section.
- Phase 2: Data from January 2017 to 21st May 2019 are used in the estimation phase of the Bayesian model detailed earlier. This is the in-sample component of the model.
- Phase 3: The out-of-sample data are from 22nd May 2019 to 30th June 2019. These were used to test the forecasting accuracy of the model.

We now describe each of these phases.

C. Phase 1: Prior specifications for the imbalance price model

The univariate DLM is classified by the quadruple $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$ and the initial information set D_0. In this section we go through the DLM specification by defining this quadruple and the prior distributions for $(\boldsymbol{\theta}_0|D_0)$ and $(\phi_0|D_0)$. Following recommendations in [24], we use historical data from June 2016 to formulate these priors, where these prior estimates are obtained via ordinary least squares.

As noted in the previous section, we also introduce a categorical variable to isolate the possibly differing price formation effects if the system operator is out of balance in a positive or negative direction (NIV positive or negative), via a dummy variable, X_t^d, which equals one if NIV is non-negative and zero otherwise.

Now we collate all our exogenous variables in one single vector to get $\mathbf{F}_t$. Explicitly,

$$\mathbf{F}_t^{'} = \left[\begin{array}{ccccccc} 1 & Y_{t-2} & X_{1,t-2} & X_{2,t-2} & ... & X_{t-2}^d \end{array}\right]. \tag{45}$$

The first component of $\mathbf{F}_t$ is 1 because it corresponds to the intercept term. Note that the intercept plays a special role in the formulation above. Ceteris paribus, the evolution of the intercept is equivalent to the evolution of price when NIV is less than zero. The $\boldsymbol{\theta}_t$ coefficients will be the corresponding regression coefficients.

The error term is given by $\nu_t \sim N[0, V_t]$. With respect to the observational variance V_t, we can see from Figure 1 that the data distribution is non-normal. Therefore we decided to work with a variance law for the observational variance given by

$$V_t = k(\mu_t)V_t, \tag{46}$$

where $\mu_t = \mathbf{F}_\mathbf{t}^{'}\boldsymbol{\theta}_\mathbf{t}$ is the level of the series at time t. Since we are working with the reciprocal of V_t, i.e., ϕ_t we have

$$V_t = k(\mu_t)\phi^{-1}. \tag{47}$$

Given the right-skewed histogram in Figure 1b, here we consider

$$k(\mu_t) = \mu_t^p, \tag{48}$$

with $p = 2$ since this would be equivalent to a log transformation of the data; see, [24]. Note that by adopting a variance law we actually allow V_t to change over time. The new variance law would change the updating equation (30) for the one-step ahead forecast and equation (42) for the k-step ahead forecast so that

$$Q_t = \mathbf{F}_t^{'}\mathbf{R}_t\mathbf{F}_t + \mu_t^p S_t, \tag{49}$$

$$Q_t(k) = \mathbf{F}_{t+k}^{'}\mathbf{R}_t(k)\mathbf{F}_{t+k} + \mu_t^p S_t, \tag{50}$$

where $\mu_t = f_t$, and f_t is given by equation (29) for the one-step ahead forecast or equation (41) for the k-step ahead forecast. For the system equations we assume the system matrix $\mathbf{G}_t$ is constant over time and equal to the identity matrix. Hence, the current state $\boldsymbol{\theta}_t$ is only dependent on the previous state $\boldsymbol{\theta}_{t-1}$. The error term vector given by $\boldsymbol{\omega}_t \sim N[\mathbf{0}, \mathbf{W}_t]$ represents purely random, unpredictable changes in level between time $t-1$ and t. For the evolution variance $\mathbf{W}_t$ we adopt the discount factor approach. A different δ can be

specified for each set of predictors. Suppose

$$\mathbf{P_t} = \mathbf{GC_{t-1}G}^{'}, \tag{51}$$

and let δ_T and δ_C be the discount factor associated with the autoregressive and exogenous variables, respectively. The evolution variance matrix is then defined as

$$\mathbf{W_t} = \text{block diag}\{\mathbf{P_{tT}}(\delta_{\mathbf{T}}^{-\mathbf{1}} - \mathbf{1}), \mathbf{P_{tC}}(\delta_{\mathbf{C}}^{-\mathbf{1}} - \mathbf{1})\}. \tag{52}$$

As described in [24], these $\delta^{'}$s usually take values between 0.8 and 1.0, with smaller values anticipating greater change in the model parameters at each stage and 1.0 being the static model. In our analysis, based on a little trial and error, we choose $\delta_T \approx 0.99$, and $\delta_C \approx 0.99$. Note that larger discount factors, which are bounded by one, imply that we are being diffuse in our beliefs.

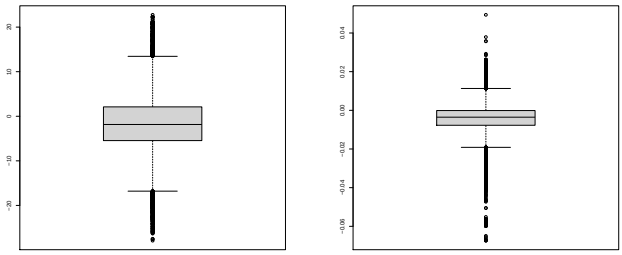
The use of 2016 data to obtain the prior moments $\mathbf{m_0}$ and $\mathbf{C_0}$ from (23) for the regression coefficients via a simple linear regression may be thought of as the training data. Because we are working with the reciprocal of V_t, ϕ_t, we also need to specify the moments n_0 and S_0 from (24). The degrees of freedom are usually given by $n_t = t - n$. Since we already used 2016's data for the prior, $n_0 = 24 - n$. We set S_0 to 0.1^2 to reflect vague prior knowledge.

D. Phase 2: Estimation

Since the estimation phase yields voluminous output, for brevity of presentation, without loss of generality, we only display the two variables, NIV and Solar Error. Addiitonal plots can be found on Appendix A. Consider Figure 3 comprising of two panels. Panel A (B) depicts the boxplots of the *mean* values of the regression parameters corresponding to NIV (Solar Error). Note that actually we have the entire posterior distributions corresponding to *each* value shown in these plots. It is evident that these posterior distributions are highly variable since even the *mean* values from these distributions range significantly. This confirms one of the primary hypotheses of this research, namely the time-varying nature of the factors that influence prices in the balancing market. Thus, from Panel A of Figure 3, the (mean) marginal effect of the dummy variable's coefficient corresponding to positive NIV values on system prices have an Interquartile Range of (-5.4572; 2.1070). Likewise the (mean) marginal effect of Solar Error has and IQR of (-0.007732; -0.0001156). It is important to model these fluctuating marginal impacts since they, in turn, will affect the conditional one-step-ahead predictive accuracy of the model, based upon market data lagged by two periods, in our application.

E. Phase 3: Prediction

Figure 4 depicts the 2,000 out-of-sample forecasts compared to the observed values. The DLM point forecasts follow the actual values quite closely. Note that with the Bayesian approach, instead of a point forecast, the model outcome is a distribution forecast at each time period and so we use the mean of the predictive distribution. Figure 5 depicts the one-step ahead (but with two period lagged market information)



(a) Boxplot of Posterior Means of NIV Coefficient (IQR: -5.4572; 2.1070)

(b) Boxplot of Posterior Means of Solar Error Coefficient (IQR: -0.007732; -0.0001156)

Fig. 3: Estimation Phase: Evolution of Posterior Means of $\boldsymbol{\theta}$

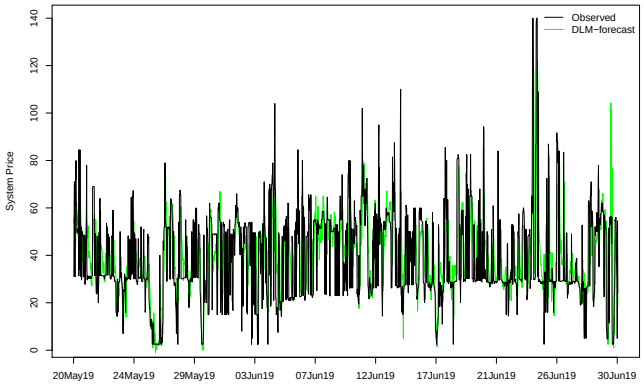


Fig. 4: Forecast versus observed system price for the last 2,000 observations

forecast distribution for prices for the last 100 observations. The green line represents the point forecasts $f_t(k)$, which are the means of the predictive distributions for the last 100 observations; the dotted lines provide the 95% prediction intervals for Y_{t+2} given D_t; and the black dots correspond to the observed prices. We note a further key point here. One convenient feature of the Bayesian DLM is that when a new observation arrives, the posterior distribution of the regression parameters are easily updated *without* having to go through the entire Phase 1 and Phase 2 steps. This is the "learning from experience" aspect implicit in Bayes's theorem. Also, once this update is made, it is straightforward to obtain any k-step ahead forecast, provided one has data available for the exogenous variables.

We next turn to assessing the forecasting accuracy of the model. For these data, [21] expressed concern about model performance for different forecasting windows; specifically, they were concerned with periods with higher variability than what was observed in 2019. For each forecasting window we considered the first 100 observations to establish the priors (Phase 1), we reserved the last 2000 observations to check performance (Phase 3) and all the observations in-between

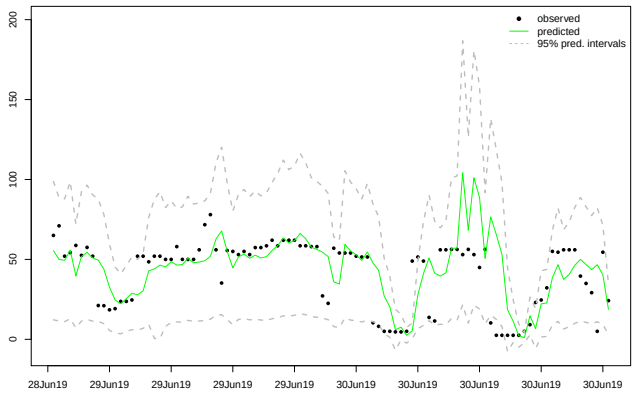


Fig. 5: Forecast distribution for the last 100 observations

for training (Phase 2). For example, from Figure 1, 2016 had a higher variability than subsequent years. Here we test the DLM for different forecasting windows and for comparison purposes we implement the GARCH(1,1) from [21] and a AR(2)-GARCH(1,1) with the same exogenous variables from Table I. Two measures are used to compare the forecasting accuracy. Firstly, the commonly used Root Mean Square Error (RMSE) for accuracy of the point forecasts; and secondly the Pinball loss function (see, for example, [27]) at the 1st, 5th, 95th and 99th quantiles of the predictive distributions.

The root mean square error (RMSE) given by

$$\text{RMSE} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t - \hat{y}_t}{y_t}\right)^2. \tag{53}$$

The Pinball loss function at time t, denoted P_t, for the q-th quantile is given by:

$$P_t = \begin{cases} (1-q)(\hat{y}_t^q - y_t), & y_t < \hat{y}_t^q \\ q(\hat{y}_t^q - y_t), & y_t \geq \hat{y}_t^q. \end{cases} \tag{54}$$

The out-of-sample RMSE for the Bayesian DLM and the other two models are given in Table III. And the Pinball loss values for all three models are given in Table II. For all the scenarios, we considered all previous observations to estimate the model parameters. First, consider the RMSE measure. The DLM outperforms GARCH(1,1) and the AR-GARCH for all the volatility windows. For the Nov-Dec 2016 window—the period with the largest volatility—the DLM RMSE is 20.57, which is better than the 22.18 from the AR-GARCH; importantly, it is also better than the 22.1 regime switching model proposed and reported in [21]. The latter authors did not produce density forecasts from their regime switching model. In our assessment of the density forecasts against conventional benchmark using the pinball loss functions, the DLM is not uniformly lower than the GARCH or AR-GARCH for all forecasting windows tested, but on average is lower. High prices are the most risky for any participant who is short. Hence, at the higher quantiles, the DLM outperforms GARCH in all the windows and outperforms AR-GARCH in four of the six windows. Overall, in conjunction with the RMSE results, the densities produced by the DLM are an improvement.

Forecasting Window	Pinball Loss 1%			Pinball Loss 5%		
	DLM	GARCH	AR-GARCH	DLM	GARCH	AR-GARCH
Nov-Dec 2016	6.85	8.00	7.35	6.82	8.03	7.35
May-Jun 2017	6.58	9.69	10.3	6.55	9.62	10.1
Nov-Dec 2017	6.53	4.92	4.25	6.52	5.08	4.45
May-Jun 2018	6.26	6.25	4.52	6.24	6.32	4.64
Nov-Dec 2018	7.58	8.21	4.41	7.53	8.24	4.6
May-Jun 2019	6.09	7.70	9.51	6.07	7.71	9.27

(a) 1st and 5th quantiles

Forecasting Window	Pinball Loss 95%			Pinball Loss 99%		
	DLM	GARCH	AR-GARCH	DLM	GARCH	AR-GARCH
Nov-Dec 2016	6.27	8.80	7.38	6.25	8.83	7.39
May-Jun 2017	5.87	8.09	5.22	5.84	8.02	5.01
Nov-Dec 2017	6.32	8.57	8.98	6.32	8.72	9.18
May-Jun 2018	5.76	7.90	7.37	5.73	7.97	7.49
Nov-Dec 2018	6.45	9.05	8.81	6.4	9.08	9.00
May-Jun 2019	5.55	7.83	4.01	5.52	7.84	3.78

(b) 95th and 99th quantiles

TABLE II: Pinball Loss over different forecasting windows

Forecasting Window	DLM	GARCH	AR-GARCH
Nov-Dec 2016	20.57	25.22	22.18
May-Jun 2017	16.85	22.14	18.50
Nov-Dec 2017	16.83	18.15	18.24
May-Jun 2018	15.38	17.83	16.11
Nov-Dec 2018	18.52	21.83	19.17
May-Jun 2019	15.92	18.88	17.07

TABLE III: RMSE out-of-sample over different forecasting windows

A similar plot to the one in Figure 5 is shown on Figure 6 for the Nov-Dec 2016 window. The plot also includes the forecasts obtained with the AR-GARCH model. Note that the Bayesian DLM tracks the price data quite well even during periods of higher variability in prices.

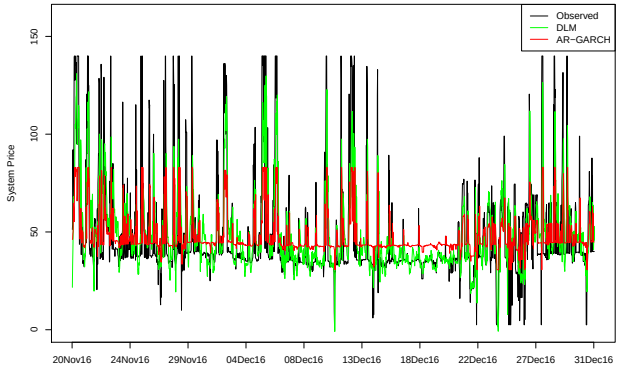


Fig. 6: Forecast versus observed system price for Nov-Dec 2016

IV. Conclusion

The research was motivated by two observations related to the forecasting of imbalance prices. Firstly, because imbalance prices are an increasing component of trading risk to market participants, it is important to understand the extreme price risks in the forecasts and so a predictive distribution is more useful than a point forecast. The second observation is that

the structural drivers of price formation are evolving quite rapidly in electricity markets. Thus, for example, we expect the impact of coal prices to become less and wind speed greater, as evident through the evolution of their respective coefficients in regression models of price formation. Thus, predictive models should therefore be conditional upon the latest estimates of coefficients, as estimated from time-varying parameter models. Putting these two requirements together, we implemented a Bayesian dynamic linear model.

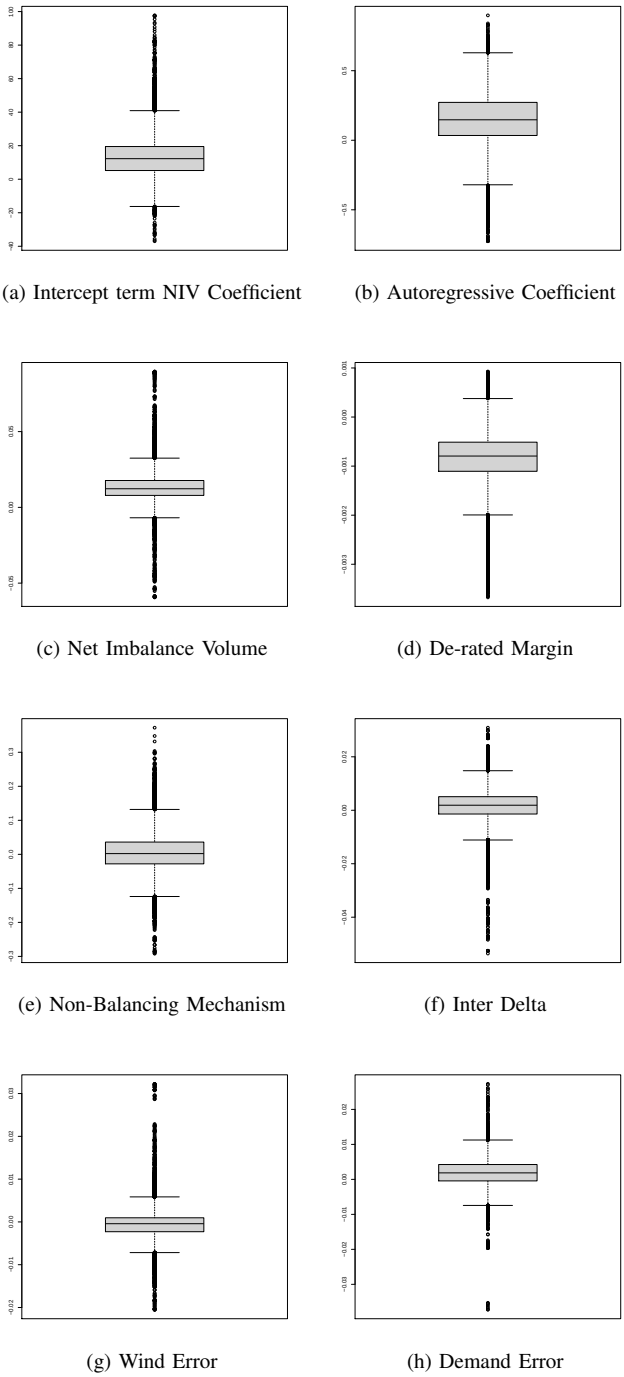
To assess the performance of this approach in the context of previous research, we used the same data as in a previous publication which advocated the use of regime switching for imbalance price forecasting. Our new Bayesian DLM was more accurate on the same data than the regime switching, vindicating the benefit of a model that captures gradual evolutionary parameter changes over one involving switching between distinct, constant parameter, specifications. Regarding the density forecasts, we bench-marked against the conventional GARCH and AR-GARCH models and again demonstrated superior performance in general.

We therefore consider that the predictive density model with Bayesian estimation of time varying parameters represents a substantial research contribution for imbalance forecasting. Evidently, balancing mechanisms vary considerably across markets, both in terms of price formation and in the timely availability of exogenous factors for market information. Nevertheless, we believe the methodology in this paper is generalizable, as the two fundamental driving observations—the need for density forecasts to model risk and the evolution of market structure—are unequivocally widespread.

Appendix A
Additional Plots

The boxplots of the *mean* values of the regression parameters corresponding to the other exogeous variables considered in the paper are shown in Figure 7. Once again these posterior distributions are highly variable.

The full forecasting distribution plot for all 2000 observations during Phase 3 of the analysis is shown in Figure 8. Note that the observed values are always withing the 95% forecasting distribution interval.



(a) Intercept term NIV Coefficient (b) Autoregressive Coefficient

(c) Net Imbalance Volume (d) De-rated Margin

(e) Non-Balancing Mechanism (f) Inter Delta

(g) Wind Error (h) Demand Error

Fig. 7: Estimation Phase: Evolution of Posterior Means of $\boldsymbol{\theta}$

References

[1] C. Cyrus, "Failed UK energy suppliers update," *Forbes Advisor*, January 24, 2022.

[2] L. Vandezande, "Design and integration of balancing markets in Europe," *Catholic University Leuven, PhD Thesis*, 2011.

[3] R. van der Veen, "The electricity balancing market: exploring the design challenge," *Utilities Policy*, vol. 43, pp. 186–194, 2016.

[4] K. Poplavskaya and L. de Vries, "Distributed energy sources and the organized balancing market: a symbiosis yet? case of three European balancing markets," *Energy Policy*, vol. 126, pp. 264–276, 2019.

[5] "OFGEM, Electricity Balancing Significant Code Review—Final Policy Decision," https://www.ofgem.gov.uk/sites/default/files/docs/2014/05/electricity_balancing_significant_code_review_-_final_policy_decision.pdf, 2014.

[6] T. Matsumoto, D. Bunn, and Y. Yamada, "Mitigation of the inefficiency in imbalance settlement designs using day-ahead prices," *IEEE Transactions on Power Systems*, vol. doi: 10.1109/TPWRS.2021.3135334, 2021.

[7] D. D. Boogert, "On the effectiveness of the anti-gaming policy between the day-ahead and real-time electricity markets in the Netherlands," *Energy Economics*, vol. 27(5), pp. 752–770, 2005.

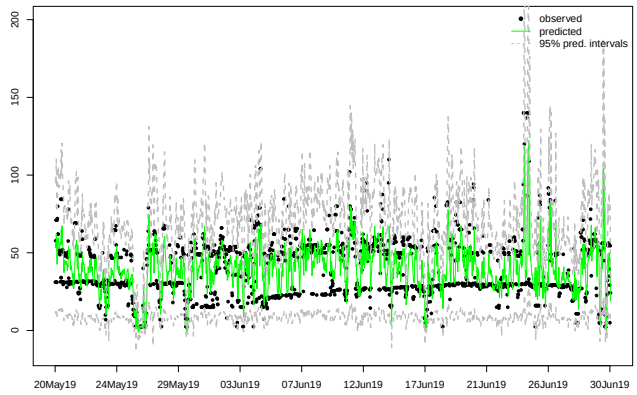


Fig. 8: Forecast distribution for the last 100 observations

[8] S. Just and C. Weber, "Strategic behavior in the German balancing energy mechanism: incentives, evidence, costs and solutions," *Journal of Regulatory Economics*, vol. 48:2, pp. 218–243, 2015.

[9] C. Möller, S. Rachev, and F. Fabozzi, "Balancing energy strategies in electricity portfolio management," *Energy Economics*, vol. 33:1, pp. 2–11, 2011.

[10] J. Browell, "Risk constrained trading strategies for stochastic generation with a single-price balancing market," *Energies*, vol. 11:6, 2018.

[11] C. Koch, "Intraday imbalance optimization: incentives and impact of strategic intraday bidding in Germany," *USAEE Working Paper*, vol. 19:421, 2019.

[12] F. Lisi and E. Edoli, "Analyzing and forecasting zonal imbalance signs in the Italian electricity market," *The Energy Journal*, vol. 39:5, pp. 1–20, 2018.

[13] G. Klaeboe, A. Erksrud, and S. Fleten, "Benchmarking time series based forecasting models for electricity balancing market prices," *Energy Systems*, vol. 6:1, pp. 43–61, 2015.

[14] T. Salem, K. Kathuria, H. Ramampiaro, and H. Langseth, "Forecasting intra-hour imbalances in electric power systems," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 30(01), pp. 9595–9600, 2019.

[15] M. Garcia and D. Kirschen, "Forecasting system imbalance volumes in competitive electricity markets," *IEEE Transactions on Power Systems*, vol. 21:1, pp. 240–248, 2006.

[16] J. Taylor, "Density forecasting for the efficient balancing of the generation and consumption of electricity," *International Journal of Forecasting*, vol. 22, pp. 707–724, 2006.

[17] P. Avis and J. Lee, "Evaluating system imbalance forecasting methods for the UK electricity market," *SOAS Working Paper: www.researchgate.net/publication/*, 2021.

[18] S. Goodarzi, H. Perera, and D. Bunn, "The impact of renewable energy forecast errors on imbalance volumes and electricity spot prices," *Energy Policy*, vol. 134, p. 110827, 2019.

[19] D. Bunn and S. Kermer, "Statistical aribitrage and information flow in an electricity balancing market," *The Energy Journal*, vol. 42(5), 2021.

[20] D. W. Bunn, A. Gianfreda, and S. Kermer, "A trading-based evaluation of density forecasts in a real-time electricity market," *Energies*, vol. 11, no. 10, 2018. [Online]. Available: https://www.mdpi.com/1996-1073/11/10/2658

[21] D. Bunn, J. Inekwe, and D. MacGeehan, "Analysis of the fundamental predictability of prices in the British balancing market," *IEEE Transactions on Power Systems*, vol. 36:2, pp. 1309–1316, 2021.

[22] A. Lucas, K. Pegios, E. Kotsakis, and D. Clarke, "Price forecasting for the balancing energy market using machine learning regression," *Energies*, vol. 13(5420), 2020.

[23] M. Mureddu and H. Meyer-Ortmanns, "Extreme prices in electricity balancing markets from an approach of statistical physics," *Physica A: Statistical Mechanics and its Application*, vol. 490, pp. 1324–1334, 2018.

[24] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York, 1997.

[25] G. Petris, S. Petrone, and P. Campagnoli, *Dynamic Linear Models in R*. Springer-Verlag, 2009.

[26] A. C. Harvey, *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, 1989.

[27] G. Tillman, "Quantiles as optimal point forecasts," *International Journal of Forecasting*, vol. 27:2, pp. 197–207, 2011.