Doctoral Dissertation

A Hawkes Model Approach to Modeling Price Spikes in the Japanese Electricity Market

Bikeri Adline Kerubo

March 17, 2023

Graduate School of Information Science
Nara Institute of Science and Technology

A Doctoral Dissertation
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of ENGINEERING

Bikeri Adline Kerubo

Thesis Committee:

Professor Kazushi Ikeda	(Supervisor)
Professor Shoji Kasahara	(Co-supervisor)
Professor Yoshinobu Kawahara	(Osaka University)
Assistant Professor Hieida Chie	(Co-supervisor)
Assistant Professor Fukushima Makoto	(Co-supervisor)

Dedication

This thesis is dedicated to the memory of my loving Dad. I miss him every day. I'm glad that he saw the beginning of this process offering his wisdom and support and would have loved to see the end...

"The art of living... is neither careless drifting on the one hand nor fearful clinging to the past on the other. It consists in being sensitive to each moment, in regarding it as utterly new and unique, in having the mind open and wholly receptive."

— Alan Wilson Watts

Acknowledgements

I would like to thank my PhD advisor, Professor Kazushi Ikeda, for welcoming me into this laboratory and for his dedicated guidance, support, and invaluable mentoring throughout my research. Without him this work would not have been possible. In addition, I would like to thank all the professors in my lab for their collective support and valuable feedback during my research progress seminar sessions that helped me succeed. I also thank all members of the Mathematical Informatics Laboratory for their friendship, generosity, and encouragement especially my friends Mario, Rae and Igor.

Additionally, I give my sincere gratitude to the administration of NAIST for the financial support that they extended to me throughout my PhD study.

A special word of gratitude goes to my friends Midori and Donald Gagner who have shaped my life during the time that I have spent living and studying in Japan. Thank you for being super nice and supportive.

Special thanks to my extended family in Kenya, my siblings, and my lovely parents Lilian and the late Alan. Thank you, Mum and Dad- words cannot convey how much I owe you for the guidance that you have shown me. Thanks for influencing my life with your love, wisdom, intelligence, integrity, values and humility.

Finally, I would like to say that I owe a great debt of gratitude to my family- my husband and children, who have an unwavering support for me. I love you all so much. Most of all, I thank my husband for his amazing heart, love and invaluable support in all things.

A Hawkes Model Approach to Modeling Price Spikes in the Japanese Electricity Market*

Bikeri Adline Kerubo

Abstract

The Japanese Electric Power eXchange (JEPX) provides a platform for trading of electric energy in a manner similar to more traditional financial markets. As the number of market agents increase, there is an increasing need for effective price forecasting models. Electricity price data is observed to exhibit periods of relatively stable i.e., low-magnitude, low-variance prices interspersed by periods of higher prices accompanied by larger uncertainty. The price data time series therefore exhibits a temporal non-stationarity characteristic that is difficult to capture with typical time-series modeling frameworks. This thesis focuses on models for the occurrence of price spikes where we define spikes as observing prices above a pre-defined threshold. For the purpose of modelling and analysis, the price spikes threshold in the JEPX is set at 25 Yen/kWh. The price spikes time series is observed to be a set of rare events that occur in clusters.

This work proposes to model the data as a Hawkes process whereby the occurrence of a spike event increases the probability of observing more spikes in the period immediately following a price spike event. Apart from the classical Hawke's model formulation, this work proposes two variations for modelling the price spikes time series in the JEPX. The first variation models the change in the magnitude of the underlying intensity as a function of the magnitude of the price spike while the second variation models the change in the decay rate of the underlying intensity as a function of the magnitude of the price spike. An analysis on the forecasting performance of the original Hawkes model, the proposed variations compared to a baseline persistence model shows that the variable magnitude variation of the Hawkes model best captures the underlying characteristics of the process generating the price spike events. The model also performs best in forecasting the occurrence of price spike events.

---

*Doctoral Dissertation, Graduate School of Information Science,
Nara Institute of Science and Technology, March 17, 2023.

Keywords:

Electricity markets; Electricity Price Spikes; Japanese Electric Power Exchange; Hawkes Process

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

Over the past three decades, the electric energy sub-sector in many countries has undergone several significant changes that have considerably changed the way in which power is delivered to consumers [1, 2]. Of these changes, market liberalization in many countries has probably been the most significant. Liberalization of electricity markets refers to the splitting up of vertically integrated electricity utilities into smaller units, each responsible for a section of the electricity system [3, 4]. Liberalization has completely changed the dynamics of the entire system with a big impact on electricity prices consequently having a major impact on all market players from electricity generators to consumers. Electricity prices in liberalized markets are largely dictated by the natural forces of supply and demand meaning that they have become much more complex.

Fundamentally, electricity supply has characteristics that are quite different from most traded commodities. For example, there is a need for a constant real-time balance between production and consumption to keep steady supply [5]. This operational detail becomes more complex when there are several generators and retailers in the market. While the function of guarantee of grid stability is still left to a single market operator, they have to work together with other market players whose objective is profit maximization, unlike traditional power utilities whose objective was cost minimization while guaranteeing supply reliability [6]. Operational difficulty is increased by the dependence of electricity consumption on factors such as time of day or day of week. Furthermore, not only demand, but generation too are influenced by factors such as weather conditions, neighboring markets etc. [7].

The factors stated above means that electricity prices are highly volatile with sudden and unexpected price spikes [8]. Market players are continuously exposed to a highly uncertain environment where profits can be easily wiped out in a few days of extreme prices. In many countries, the uncertain nature of electricity prices is intensified by increased penetration of variable renewable energy sources such as solar and wind power in the grid [9]. While these sources provide much needed clean energy, the original energy source i.e. solar radiation or blowing wind is highly variable resulting in a significant effect on the balance of supply and demand and consequently on the final hourly or half-hourly prices [10].

In Japan, the Japanese Electric Power eXchange (JEPX) offers a day-ahead market for power trading [11]. The JEPX was founded in 2003 and begun trading in wholesale electric power in 2005 with the aim of processing electricity transactions on the exchange. A day-ahead electricity market such as JPEX provides electricity suppliers and retailers with a mechanism to trade electricity in an environment where the price is driven by market forces of supply and demand. The primary goal for JEPX was for market participants to operate efficiently and economically on the electricity market, ensure fair competition and revitalize the business of transmission and distribution of electrical energy across the country.

Given the background of uncertainty and accompanying risks, it is important for the market participants to continuously define the market characteristics through comprehensive analysis and accurate analytics. The need for mathematical models describing the characteristics of the system and accompanying tools for risk assessment and risk management are greatly increased. There is an increased need for development of mathematical models that can define, describe and detect extreme market conditions. In addition, there is increased need for models that can be used to generate good price forecasts capturing the highly uncertain nature of the markets.

This thesis works specifically on modelling extreme prices in electricity markets focusing specifically on the JEPX. Electricity price data in the JEPX is observed to exhibit periods of relatively stable prices interspersed by periods of higher prices that last for a period of time. These “high price periods” are defined as spikes and have a significant effect on the bottom lines of market participants. The occurrence of a price spike is defined as the observation of a price above a certain pre-defined threshold where the threshold is a risk value that would mean significant losses for the market player. Price spike events are modelled using the Hawkes model [12] which is typically used to model non-stationary point processes [13]. The Hawke's model has been used extensively to model high frequency financial time series data with time varying intensities and self-exciting features for intra-day trading [14]. Apart from an analysis of the classical Hawke's model approach, two modifications of the model are proposed and the effectiveness of the models are tested and detailed. The approaches proposed in this thesis can form a basis for market participants to deal with the high price uncertainty which has a big impact on their financial success in the markets.

1.2 Problem Statement

While a lot of attention has been put on the modeling of prices in electricity markets, there are very few studies focusing specifically on modelling price spikes. Most market crises in the industry are caused by periods of extremely high prices even when those periods last for just a short period of time. Price spikes can cause losses to the tunes of millions of dollars even for relatively small market players and it is therefore important to add to the body of knowledge in this area.

In addition to focusing on spot price modelling, most literature model market indices such as the average daily price or number of extreme price events in a day. While such indices are important in defining the overall underlying state of the market, they run the risk of oversimplifying the problem. Electricity price characteristics exhibit more complex characteristics. For example, the probability distributions of prices are very dependent on the time of day and disproportionately affect periods of high demand. There is therefore need for more detailed modeling approaches.

A review of existing literature shows that very few studies have been carried out on the Japanese electricity market. The Japanese market is particularly unique given existence of different frequency zones in the grid and a lack of an external connection. There are also 9 areas operated by the spot market with differing inherent characteristics as well.

1.3 Justification

Deregulated electricity markets are a relatively new concept when compared to traditional vertically integrated markets that have existed for over a century. Consequently, there remains several gaps to be filled by research and development. The Japanese electricity market is especially relatively young having only achieved full liberalization in 2016.

The work presented in this study targets filling the gap on the understanding and modelling of price spike events in the JEPX which is very important for participants in the market. It adds to the body of knowledge by proposing different variants of classical Hawkes models for modeling price spike occurrence probabilities and testing the effectiveness of the models in short-term forecasting of the occurrence of electricity price spike events in the Japanese market.

1.4 Research Objectives

The following were the main objectives of the research:

1. Developing models for forecasting electricity price spikes with higher accuracy than existing approaches. Such models would be very useful for market agents especially in developing bidding strategies and making hedging decisions.

2. Modelling the underlying dynamics of price spike events. This would be useful for detecting periods of market shocks and consequently in developing plans for protecting end users from extreme prices.

1.5 Thesis Contributions

This thesis focuses on modeling processes that captures underlying dynamics of deregulated electricity markets. Such models could be used by market agents for forecasting purposes and hence developing bidding strategies or may be used by market operators and regulators to detect extreme circumstances. While there is lots of literature on the modelling and forecasting of electricity prices time series, research dealing specifically on the extreme price events are significantly fewer. However, recent events have shown that these extreme prices have the biggest impacts on market players and there have been several cases of electricity retailers going bankrupt after a short run of price spikes. Unlike most literature on the modelling of prices in electricity markets, this study focuses specifically on modelling price spikes by dis-integrating spikes from "normal" periods rather than modelling the electricity market prices in their entirety. Available studies take an aggregate approach when modelling price spikes – modelling the number of events in a day – thereby losing information on the time period in which the spike would occur. Given that extreme prices seem to coincide with high demand periods, it is important to provide models that specifically isolates the event occurrence time periods. In addition, while existing literature typically generate one-step ahead forecasts, a method for generating forecasts for a few-days ahead is provided in this study.

The price spikes time series are modeled using the Hawkes model which is typically used to model non-stationary point processes. Results are presented that demonstrate the effectiveness of a modified form of the Hawkes model in the short-term forecasting of the occurrence of price spike events. Modelling is done at half-hourly time resolutions as opposed to average day-ahead prices since prices vary throughout the day depending on the time of day. No assumptions or simplification are made on electricity prices either, so weekend prices are not ignored. Modifications on the classical Hawke's model are presented showing the effect of including spike magnitudes information on the spike event occurrence forecasting performance. Finally, simulations on the Japanese electric power exchange, on which there are very few studies, are presented.

1.6 Thesis Organization

This thesis is organized as follows:

Chapter 1

This chapter presents an introduction to the thesis, stating the problem statement, justification and research objectives.

Chapter 2

This chapter gives a background of the research with reviews of various literature and discussions related to the research concepts including:

- A review of price forecasting in electricity spot markets.
- Components of the electric power system, deregulated electricity markets and the Japanese electricity market.

Chapter 3

In this chapter, a fundamental analysis of price data and price spikes data in the Japan electric power exchange (JEPX) is presented highlighting the JEPX Data structure and modelling of the price spikes time series.

Chapter 4

The models implemented in this research are explained in this chapter. These are:

- The classical Hawke's model
- Hawke's model with a variable intensity jump
- Hawke's model with a variable effect decay speed

The application of the models on the JEPX price spikes data is illustrated showing the effectiveness of the proposed approach.

References

[1] G. Cervigni and D. Perekhodtsev, *Wholesale electricity markets*. Edward Elgar Publishing, 2013.

[2] P. Farré, C. García, and H. Hervitz, "Deregulation and liberalization in the european electricity sector," *International Business & Economics Research Journal (IBER)*, vol. 9, 12 2010. doi: https://doi.org/10.19030/iber.v9i2.529.

[3] L. A. Reisch and H.-W. Micklitz, "Consumers and deregulation of the electricity market in germany," *Journal of Consumer Policy*, vol. 29, no. 4, pp. 399–415, 2006. doi: https://doi.org/10.1007/s10603-006-9016-z.

[4] N. Fatras, Z. Ma, H. Duan, and B. N. Jørgensen, "A systematic review of electricity market liberalisation and its alignment with industrial consumer participation: A comparison between the nordics and china," *Renewable and Sustainable Energy Reviews*, vol. 167, p. 112793, 2022. doi: https://doi.org/10.1016/j.rser.2022.112793.

[5] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, 2017. doi: https://doi.org/10.1109/MPE.2016.2637122.

[6] R. Ghorani, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "Optimal bidding strategy of transactive agents in local energy markets," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5152–5162, 2019. doi: https://doi.org/10.1109/TSG.2018.2878024.

[7] P. A. Steenhof and W. Fulton, "Factors affecting electricity generation in china: Current situation and prospects," *Technological Forecasting and Social Change*, vol. 74, no. 5, pp. 663–681, 2007. doi: https://doi.org/10.1016/j.techfore.2006.09.005.

[8] C. J. Ullrich, "Realized volatility and price spikes in electricity markets: The importance of observation frequency," *Energy Economics*, vol. 34, no. 6, pp. 1809–1818, 2012. doi: https://doi.org/10.1016/j.eneco.2012.07.003.

[9] F. Paraschiv, D. Erni, and R. Pietsch, "The impact of renewable energies on eex day-ahead electricity prices," *Energy Policy*, vol. 73, pp. 196–210, 2014. doi: https://doi.org/10.1016/j.enpol.2014.05.004.

[10] S. R. Sinsel, R. L. Riemke, and V. H. Hoffmann, "Challenges and solution technologies for the integration of variable renewable energy sources—a review," *Renewable Energy*, vol. 145, pp. 2271–2285, 2020. doi: https://doi.org/10.1016/j.renene.2019.06.147.

[11] T. Nakajima, "Inefficient and opaque price formation in the japan electric power exchange," *Energy Policy*, vol. 55, pp. 329–334, 2013. doi: https://doi.org/10.1016/j.enpol.2012.12.020.

[12] P. J. Laub, T. Taimre, and P. K. Pollett, "Hawkes processes," 2015. doi: https://doi.org/10.48550/ARXIV.1507.02822.

[13] M.-A. Rizoiu, Y. Lee, S. Mishra, and L. Xie, "A tutorial on hawkes processes for events in social media," 2017. doi: https://doi.org/10.48550/ARXIV.1708.06401.

[14] A. G. Hawkes, "Hawkes processes and their applications to finance: a review," *Quantitative Finance*, vol. 18, no. 2, pp. 193–198, 2018. doi: https://doi.org/10.1080/14697688.2017.1403131.

2 Background

2.1 A Review of Price Forecasting in Electricity Spot Markets

Electricity Price Forecasting (EPF) is becoming an increasingly important topic in the modern deregulated electricity market. This is due to the high volatility characteristics that is observed in most markets. Over the past decade, a variety of methods for electricity price forecasting have been proposed based on different modelling approaches [1–3]. This sections gives an overview of recent work on EPF including the few proposals targeted specifically on modelling price spikes.

2.1.1 Electricity Prices Modelling

Proposed methodologies for electricity price forecasting (EPF) range from those based on classical statistical models to newer machine learning models. A variety of model hybrids combining various models have also been proposed. Some of the complexities of available solutions, their strengths and weaknesses are detailed in this section.

Classical Models

Regression models, Auto-Regressive models (AR), Auto-Regression with Exogenous inputs (ARX) and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) are considered under the category of classical statistical models

Regression analysis is a set of statistical methods used to estimate relationships between a dependent variable and one or more independent variables [4]. It is used to assess the strength of the relationship between variables and to model relationships between them. Reference [5] focused on modelling the impact of various explanatory variables on the electricity price through a multiple linear regression analysis for long-term electricity price forecasting in the Iberian electricity market. The quality of the estimated models obtained validated the use of statistical or causal methods, such as the Multiple Linear Regression (MLR) Model, as a plausible strategy to achieve causal forecasts of electricity prices in medium and long-term electricity price forecasting. From the evaluation of the electricity price forecasting for Portugal and Spain, in the year of 2017, the mean absolute percentage errors (MAPE) were 9.02% and 12.02%, respectively. In 2018, the MAPE, evaluated for 9 months, for Portugal and Spain equaled 7.12% and 6.45%, respectively.

Autoregressive (AR) models predict future behavior of a time series based on past behavior. They are used for forecasting when there is some correlation between values in a time series and the values that precede and succeed them [6]. In [7] they provided a method to predict next-day electricity prices based on the Auto-Regressive Integrated Moving Average (ARIMA) methodology. A detailed explanation of the ARIMA models and results from mainland Spain and Californian markets are presented. Average errors in the Spanish market of around 10% with and without explanatory variables, and around 5% in the stable period of the Californian market are reported. In Spain, explanatory variables were only needed in months with high correlation between available hydro production and price. In any other months, the effects cancelled out. For both markets, these were considered reasonable errors, taking into account the complex nature of price time series.

Auto-Regressive Integrated Moving Average with exogenous inputs (ARIMAX) models extend ARIMA models through the inclusion of exogenous variables. An ARIMAX (p, d, q) model is defined for some time series data y_t and exogenous data X_t, where p is the number of autoregressive lags, d is the degree of differencing and q is the number of moving average lags [8]. Authors in [9] argue that since electricity prices have seasonal variation and vary depending on multiple external factors, then the Seasonal Auto-Regressive Integrated Moving Average model with eXogenous variables (SARIMAX) is a plausible approach. Electricity prices follow a seasonal pattern controlled by various external factors thus SARIMAX models would be preferable for short-term forecasting including forecasting of day-ahead prices. By applying algorithm rules for differencing to remove continuing trends, the data becomes stationary and parameters, 14 external factors, were chosen to predict day ahead electricity prices. The presented experimental results show reasonable low Root Mean Square Error (RMSE) values for predicted day-ahead electricity prices.

Generalized Auto-Regressive Conditional Heteroscedasticity (GARCH) models aim to model the conditional volatility of a time series [10]. In [11] electricity price behavior in the Nordic electric power market is forecasted with both the Markov-Switching Generalized Auto-Regressive Conditional Heteroscedasticity (MS-GARCH) model and a set of different volatility models. The MS-GARCH model is estimated with two regimes, representing periods of low and high volatility. This study demonstrated that electricity price volatility is not only highly volatile but also strongly regime-dependent. The empirical results show that the MS-GARCH model enabled more accurate forecasting than the standard GARCH models, according to tail loss and reality check tests for one- and multi-step ahead forecasts. The results suggested better price forecasts using the proposed MS-GARCH model would achieve benefits for both electricity generation companies and consumers.

Machine Learning Models

Artificial Intelligence (AI) techniques such as expert systems, neural networks, and fuzzy logic and other Machine Learning (ML) techniques have also been proposed for solving various technical challenges in the new electricity markets [12–14]. Several authors have shown that machine learning models can achieve better performance than traditional classical statistical models which are mostly employed in forecasting of electricity prices. ML approaches such as Deep Neural Networks (DNNs) may be capable of adequately representing nonlinear relationships that are clearly inherent in electricity price time series and there is an increasing number of research papers have addressed the use of these models for electricity price forecasting.

Reference [15] employed an Artificial Neural Networks (ANN) model on electricity price data. The forecast model depends on appropriate input parameter sets and the research focus was set on the selection and preparation of fundamental data that had noticeable impact on electricity prices. The aim of the paper was to develop a model based on artificial neural networks (ANN) to forecast hourly electricity prices at the European Power Exchange (EPEX) day-ahead market. The major contributions of this work included the appropriate selection and preparation of input data applying clustering algorithms and, finally, the determination of the best fitting ANN configuration including the appropriate activation function, training algorithm, the ANN learning rate and momentum. The proposed model is shown to achieve better results than baseline models such as seasonal ARIMA models and other ANN-type models.

A Recurrent Neural Network (RNN) based model is proposed in [16] where the authors use multi-layer Gated Recurrent Units (GRUs) as a new technique for electricity price forecasting. A variety of algorithms are trained with a three-year rolling window and simulation results are compared to classical RNNs. In their experiments, three-layered GRUs outperformed other neural network structures and state-of-the-art statistical techniques in a statistically significant manner using data from the Turkish day-ahead market.

A modeling framework for forecasting electricity prices using four different deep learning models – DNN, LSTM, GRU, CNN to forecast electricity prices is proposed in [17]. The authors compared the forecasting accuracy of the deep learning models to 27 common electricity price forecasting approaches. Their benchmark results showed that the proposed deep learning models outperform the state-of-the-art methods and obtained statistically significant results. Conclusions from their studies included: machine learning methods generally provide better accuracy than statistical models and moving average terms do not improve prediction accuracy. Finally, interestingly, they observed that hybrid models do not outperform their simpler counterparts.

A dynamic trees model for predicting electricity prices using data from the Iberian market is proposed in [18]. Dynamic trees are a tree-based method based on Bayesian inference, where the trees can remain unchanged, be pruned or grow according to the new values arriving in the online process. The leaves that the data is associated with can provide predictions based on the two types of rules, either in data mean or in a linear model. Optionally, data can be retired from the tree, either the oldest data or data discarded through active learning techniques, and a choice can also be made to rejuvenate the tree after data retirement. The authors compared the results to another tree-based technique, random forest - a widely used method that has proven its good results in many areas. Simulations included several versions of the dynamic trees approach for both very short-term EPF approach (one hour ahead) and short-term approach (one day ahead) and the results showed that dynamic trees can improve the performance of traditional random forest models for both very short-term EPF and short-term EPF.

According to [19], electricity prices depend heavily on the seasonality of different time scales; therefore, any forecast of electricity prices must take this into account. Neural networks have proven successful in short-term price forecasting, but complicated architectures like LSTM can be used to integrate seasonal behavior. Their paper showed that simple neural network architectures such as DNNs with an embedding layer for seasonal information can produce competitive forecasting. The embedding-based processing of calendar information also opens up new applications for neural networks in electricity trading, such as generating price-forward curves. In addition to the theoretical foundation, they also provided an empirical multi-year study on the German electricity market for both applications and derived business insights from the embedding layer. The study showed that in the short-term price prediction, the mean absolute error of the proposed neural networks with an embedding layer is better than the LSTM and time-series benchmark models, and even slightly better than their best benchmark model with a sophisticated hyper parameter optimization. The results were supported by statistical analysis using Friedman and Holms tests.

Authors in [20] have explored a feed forward neural network model known as the Extreme Learning Machine (ELM). ELM is a single hidden layer feed-forward neural network (SLFN) whose input weights and biases randomly generated and its output weights analytically calculated. The critical idea behind ELM is to transform difficult issues arising from nonlinear optimization, like the optimal determination for input weights, hidden layer biases, output weights, to a simple least square problem of deciding the optimal output weights. The proposed approach showed improved price intervals forecast accuracy by incorporating bootstrapping method for uncertainty estimations. Case studies based on chaos time series and Australian National Electricity Market (NEM) price series showed that the proposed method can effectively capture the non-linearity from the highly volatile price data series with less computation time compared to other methods. The results highlight the potential of this approach for online spot market price forecasting.

Hybrid Models

Hybrid models in time series forecasting combine statistical models and machine learning models with the idea that the combination compensates for the limitations of one approach with the strengths of the other [21]. Reference [22] proposes a hybrid model that exploits the features and strength of the Auto-Regressive Fractionally Integrated Moving Average (ARFIMA) model as well as the feed-forward neural networks model and examine the effectiveness of the proposed model using data from the Nordpool electricity market. Similarly, in [23] a long- short-term memory (LSTM) deep neural network combined with feature selection algorithms for electricity price prediction under the consideration of market coupling is proposed. LSTM models improve the model's performance in handling nonlinear and complex problems and processing time series data. The effectiveness of the proposed approach is illustrated using data from the Nordic market. The authors specifically highlight that feature selection is essential to achieving accurate predictions.

In [24] an outlier-robust hybrid model for electricity price forecasting that combined a basic forecasting engine called the outlier-robust extreme learning machine model and three new algorithms is developed. In particular, a new optimizer called chaotic sine-cosine algorithm was developed to obtain the ideal parameters for phase space reconstruction, and then a novel feature selection method was implemented to construct the optimal features in electricity price modeling. In addition, an effective data pre-processing method for effective forecasting by capturing electricity price characteristics has been proposed. Subsequently, experiments based on electricity price data from the electricity markets of Australia and Singapore showed that the proposed model is superior to other benchmark models. In addition, the model can be a reliable forecasting method not only in electricity market management, but also when modeling time series with complex non-linear properties and outliers.

A comprehensive empirical study on the optimal way of implementing the least absolute shrinkage and selection operator (LASSO) for electricity price forecasting models is carried out in [25]. The authors addressed the three issues: i) optimal structure of the baseline model; ii) the choice of the LASSO tuning (or regularization) parameter; and iii) the use of variance stabilizing transformations (VSTs). On variable (or feature) selection, i.e., the optimal structure of the baseline model, they identified the most important variables which provided guidelines to structure better performing expert models. In particular, they found that large sets of potential regressors are not a problem for the LASSO procedure. Although the LASSO typically uses only a small fraction of the initial set of explanatory variables, providing additional information in the underlying model significantly improves the accuracy of the obtained forecasts. Secondly, regarding the choice of the LASSO tuning parameter, they found one λ for all days and hours in the test period to be an acceptable option but this is recommended only if the computational time needs to be significantly reduced. To increase the forecast accuracy, it was found that it is better to reselect the tuning parameter on a daily basis. Lastly, the concept of VSTs, was confirmed to not only increase the forecasting accuracy of LASSO-estimated models but also the gains from using an appropriate VST increase with the complexity of the model.

A price forecasting algorithm based on the wavelet transform combined with ARIMA and GARCH models is proposed in [26]. Through the wavelet transform, the historical price time series is decomposed and reconstructed into an approximation series and some detail series. Each sub-series is then separately predicted by a suitable time series model and the final forecast is obtained by composing the forecasted results of each subseries. The method was examined for market clearing price (MCP) prediction in the Spanish market and locational marginal price (LMP) prediction in the PJM market and compared with existing price forecast techniques.

2.1.2 Electricity Price Spikes Modelling

An economic analysis of price spikes is presented in [27] where the authors investigate the factors and mechanisms determining spikes in the Italian electricity market. Based on the market data, they performed a specific analysis of the auctions mechanisms and of the hourly bid and offer of electricity characteristics to determine how and why price spikes occur. Their results showed that rigidity, which characterizes both the demand curve and part of the supply curve, is the fundamental determinant of prices pikes. More precisely, when the demand values are high, both curves are characterized by a high rigidity and even small fluctuations in the demand quantities or generation have a large impact on the resulting clearing price. The fluctuations in renewable energy sources (RES) production also proved to be fundamental. Even if the amount of production from RES does not seem to have any effect at first glance, since they are usually offered at zero price, they shift the entire supply curve and the resulting effects are very powerful.

In [28], a stochastic regime-switching model with time-varying parameters is shown to capture the type of volatile price behavior observed in many deregulated spot markets for electricity. The mean prices in two price regimes and the transition probabilities were specified as functions of the offered reserve margin and the system load. The high-price regime corresponded to the observed price spikes that typically occur during the summer months. In addition, the structure of the model was consistent with the actual hockey stick shape of the offers submitted by suppliers into the PJM market. Most capacity is offered at relatively low prices, and a few units are offered at much higher prices up to the price cap ($1000/MWh in PJM). Specifying Markov-switching in the models allow the high-price regime to be more persistent than is the case with a simple binomial jump process. Using on-peak daily data for PJM, the analysis showed that the model replicated the observed price volatility. Consequently, this type of model is potentially useful for evaluating forward contracts and investment decisions in electricity markets since standard financial models of prices do not allow for the unusual asymmetric type of volatility (i.e., infrequent price spikes) found in deregulated electricity markets.

Dramatic rises in electricity prices can be observed during periods of market stress as highlighted in [29]. The authors treat abnormal episodes or price spikes as count events and propose to build a model of the spiking process. In contrast to prior literature which either ignored temporal dependence in the spiking process or attempted to model the dependence solely in terms of deterministic variables, like seasonal and day of the week effects, the authors argue that persistence in the spiking process is an important factor in building an effective model. They adapt a Poisson autoregressive framework for integer-valued time series that accounts for the number of simultaneous stresses remaining latent and provided a model that could be estimated by maximum likelihood. The arrival and survival rates of price spikes were found to be dependent upon extreme temperature events and peak load. However, the model's ability to capture the intrinsic persistence in price spikes was cited as more significant. The proposed model's effectiveness in generating simulated price spikes time series that were characteristically similar to those observed in practice was illustrated. In addition, the model produced forecasts of the probability of price spike events with higher accuracy than simpler benchmark models.

The forecasting of extreme price events, the occurrence of which is treated as a realization of a discrete time point process, is the focus of [30]. An Autoregressive Conditional Hazard (ACH) framework was used to analyze the drivers of the process and to forecast the probability of extreme price events occurring in real time. Abnormal loads were found to have a significant impact on the probability of a price spike and on the severity of the spike. Importantly, stochastic factors capturing the history of the process were found to be significant in explaining the occurrence of extreme price events. Specifically, the durations between price spikes were found to depend nonlinearly on previous expected and observed durations. The ACH model was shown to provide rolling half-hour ahead forecasts of price spikes that are superior to the forecasts made by a memoryless model using the same set of exogenous information. In addition, the returns generated from a simple synthetic futures trading scheme based on the one-step-ahead forecast probabilities of the ACH model provide further evidence of the strength of the model in forecasting electricity price spikes.

An argument that there is increasing empirical evidence of increased price volatility and spikes in electricity markets as a result of fluctuating renewable energy production, extreme weather events and other factors is presented in [31]. While price spikes are necessary to cover the fixed costs of power plants, they can also indicate market imperfections and anti-competitive behavior. Regulators have set market price caps to protect consumers and prevent abusive behavior by vendors. Additionally, some regulators have imposed temporary price caps during or after major events. In weak institutional frameworks, however, these ceilings may be driven by political motives rather than economic logic. This paper assessed the welfare effect of the temporary price cap introduced in 2017 on the Turkish electricity market. Using matching and panel data methods, they show that the temporary price cap reduced overall welfare but did not impact the market clearing price and projected supply. Their analysis further showed that this decision was driven by non-economic motives and identified a number of fundamental issues in the Turkish market that limit the effective functioning of the market.

2.2 Components of the Electric Power System

Electricity is a form of energy that has a broad range of applications. It is easy to control, non-polluting at the location of its usage and convenient, used in the applications of heat, light and power. As a secondary energy source, it is generated from the conversion of other energy sources, like coal, natural gas, oil, nuclear power, hydro-power and other renewable sources [32].

The electric power grid, is an interconnected network of components used to deliver electricity from generators to consumers usually over vast distances, it is a highly complex and intricate system to balance as electricity is consumed at the same time it is generated and generation must meet demand at all times. The main function of the electric grid is to deliver electricity to the end users in the most economical and reliable manner. The grid infrastructure consists of three basic components: generation, transmission, and distribution as shown in Figure 2.1 [33].

2.2.1 Generation

Electricity generation process converts mechanical energy into electrical energy by the use of a generator (with the exception of solar power, which uses photovoltaic cells). Electricity is typically generated from three main categories of energy sources i.e. fossil fuels, nuclear and renewable sources [34]

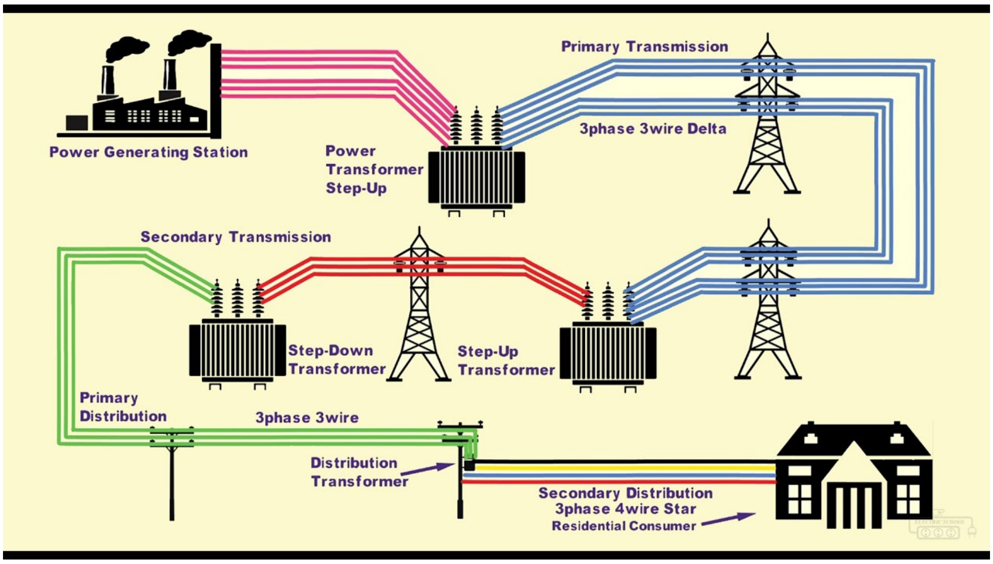

Figure 2.1: Structure of the power grid. source: Adopted Jr Electric.

2.2.2 Transmission

After generation, the voltage is stepped up and transported over long distances through the transmission lines at ultra high voltages. Stepping up voltage during transmission reduces currents and consequently minimizes system losses. The transmission network serves as a link between generation and distribution components [35].

2.2.3 Distribution

Electricity from the transmission grid is fed to the distribution network via step-down transformers. Here, the aggregated energy also includes power output of other generators embedded to the distribution system. Distribution network components include power transformers that step down voltages, service lines to deliver electricity to consumers, and energy meters to measure amount of electricity being used. Distribution lines are categorized by voltage levels - high voltage lines at 22kV, medium voltage at 6.6kV, and low-voltage network at 240V to 100V. Electricity from the distribution substations is at medium voltage and is stepped down to low voltages of 100V or 200V for domestic use [36].

2.3 Electricity Markets

2.3.1 Evolution of Electricity Markets

Several transformations in the electric energy industry over the recent past have encouraged a re-look into the operations of this very important sub-sector. For decades, electric power systems were operated under the vertically integrated models where all functions of the system (generation, transmission and distribution) were under a single utility i.e. it was a monopolistic market. However, over the last two-to-three decades, all over the world, the electric energy sub-sector has been undergoing significant changes that have necessitated a re-look into the various operation procedures [37]. The most significant change has been deregulation of the industry in many countries.

Electric power system deregulation refers to the process of changing rules and regulations that control the industry to allow for competition in supply of the resource which allows customers to choose their electricity suppliers [38]. These suppliers could be retailers, traders or offshoots of the original monopolies. Vertically integrated utilities own generating plants as well as a transmission and distribution network. In a traditional regulated environment, such a company has a monopoly for the supply of electricity over a given geographical area. Following the liberalization of the electricity market, its generation and network activities are likely to be separated.

In this new competitive framework, the vertically integrated systems are unbundled into generation companies (GENCOs), transmission companies (TRANSCOs) and distribution companies (DISCOs) as shown in Figure 2.2. Furthermore, as illustrated in Figure 2.2, new market agents enter the market including consumer companies, electricity resellers, the power exchange (PX) and the independent system operator (ISO). The main objectives of deregulation are improved economic efficiency of the production and use of electricity as well as increased system reliability and security as agents compete for a share of the market. In addition, it is aimed to provide better incentives for capital formation, better incentives for consumers to reduce their electricity consumption when costs exceed their benefits, and better incentives for research and development [39].

Deregulation of the electricity market has changed the dynamics of electric-

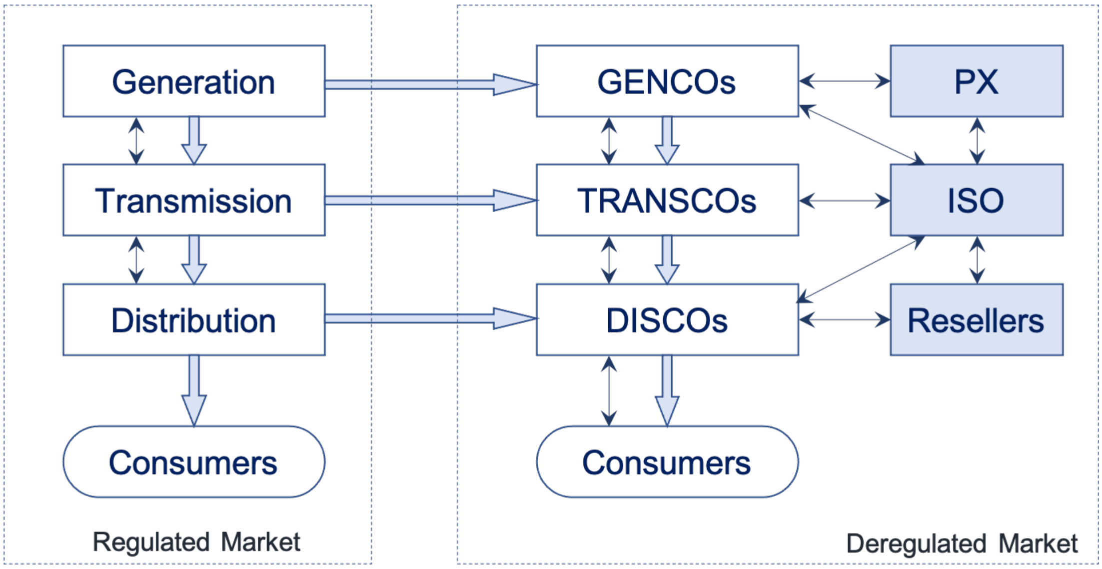

Figure 2.2: Deregulated electricity market structure.

ity prices making it a complex phenomenon as prices fluctuate over short time duration. Additionally, electricity trade has characteristics that are quite uncommon: e.g. constant balance between production and consumption, dependence of the consumption on the time, e.g. hour of the day and load and generation that are influenced by weather conditions and neighboring markets [40]. These characteristics mean that electricity prices are highly volatile with sudden and unexpected price peaks. The increased penetration of variable renewable energy sources (RES) in the system intensifies this behavior [41]. Therefore, for the market players, improved accuracy in price forecasting is important for defining bids for the spot market, setting up contract policies, and formulating expansion plans.

2.3.2 Electricity Market Structure

Market Players

The restructuring of the electricity market required the disintegration of all three components of the power systems. The market is thus organized in such a way that each market sector represents a component of the power system leading to different market players as illustrated in Figure 2.2. Generating companies (GENCOs) participate on the generation level, transmission companies (TRANSCOs) and distribution companies (DISCOs) on transmission and distribution levels respectively. Other market agents include: the market operator(MO), independent system operator (ISO), market regulator and retailers who resell electricity from the wholesale market. The market structure generally describes the way in which market participants interact with each other to generate electricity and deliver to the consumer. These agents are briefly described below:

1. **Generating Companies (GENCOs)** - They produce and sell electrical energy. They can also sell services such as regulation, voltage control and reserve that the system operator needs to maintain the quality and security of the electricity supply. A GENCO may own a single plant or a portfolio of generators of different sources of generation. There are GENCOs that coexist with vertically integrated utilities and are referred to as independent power producers (IPP) [39].

2. **Transmission Companies (TRANSCOs)** – They own transmission assets such as lines, cables, transformers and reactive compensation devices. They operate this equipment according to the instructions of the independent system operator. TRANSCOs transmits electricity using a high-voltage, bulk transport system from GENCOs to DISCOs for delivery to customers. It is composed of an integrated network that is shared by all participants and radial connections that join generating units and large customers to the network. TRANSCOs are sometimes subsidiaries of companies that also own generating plants. An independent transmission company (ITC) is a transmission company that does not own generating plants and also acts as an independent system operator [39].

3. **Distribution Companies (DISCOs)** – They own and operate distribution networks. In a traditional environment, they have a monopoly for the sale of electrical energy to all consumers connected to their network. The management of the distribution network is quite different to that of the transmission network. The infrastructure is itself far more distributed and

real time management of the systems are more 'passive' than for transmission. The increase of embedded generation, particularly of generation that is intermittent, or causing the requirement of reactive power, can substantially change the flow and the requirement to manage distribution networks 'actively' in real time [39]. Distribution network operators offer services such as connections and have the highest interactions with customers particularly through metering activities. In a fully deregulated environment, the sale of energy to consumers is decoupled from the operation, maintenance and development of the distribution network. Retailers then compete to perform this energy sale activity. One of these retailers may also be a subsidiary of the local distribution company [42].

4. **Retailers** - They buy electrical energy on the wholesale market and resell it to consumers who do not wish, or are not allowed, to participate in this wholesale market. Retailers do not have to own any power generation, transmission or distribution assets. Some retailers are subsidiaries of generation or distribution companies. All the customers of a retailer do not have to be connected to the network of the same distribution company [39].

5. **The Independent System Operator (ISO)** - Has the primary responsibility of maintaining the security of the power system. It has the authority to commit and dispatch some or all system resources and to curtail loads for maintaining the system security (i.e., remove transmission violations, balance supply and demand, and maintain the acceptable system frequency). Also, the ISO ensures that proper economic signals are sent to all market participants, which in turn, should encourage efficient use and motivate investment in resources capable of alleviating constraints. It is called independent because in a competitive environment, the system must be operated in a manner that does not favor or penalize one market participant over another. The ISO must have powerful computing tools for market surveillance, ancillary auctions, and congestion management to fulfil its responsibility [39]

6. **Market Operator (MO)** - The function of the market operator is separate to that of the system operator, although it can be performed by the same entity [39]. The MO typically runs an algorithm that matches the bids and offers that buyers and sellers of electrical energy have submitted. It also takes care of the settlement of the accepted bids and offers. This means that it forwards payments from buyers to sellers following delivery of the energy. The independent system operator (ISO) is usually responsible for running the market of last resort, that is, the market in which load and generation are balanced in real time. Markets that close some time ahead of real time are typically run by independent for-profit market operators [39].

7. **Market Regulator** - It is a government-independent entity whose function is to oversee the market and to ensure its competitive and adequate functioning. Additionally, the regulator promotes and enforces orders and regulations [39].

The overall structure of the deregulated electricity market is as shown in Figure 2.3. The system is an interaction of various participants with the exchange of energy, ancillary services, money and information. After purchasing primary fuel

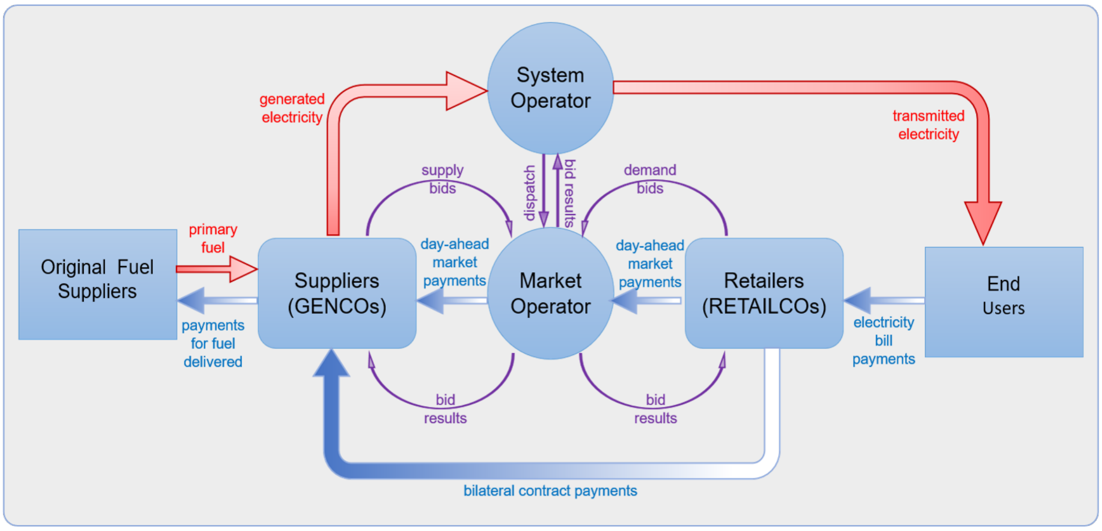

Figure 2.3: Structure of the deregulated electricity market illustrating the flow of energy, money and bidding information.

from fuel suppliers, the generation company will generate electricity and deliver it to the end-user through the physical power system. The system, which is the physical network of power transmission and distribution lines, is operated by a system operator. Even though, technically, energy is delivered to the end user directly by the system operator, most end-users will pay electricity bills to a retailer. The retail company (RETAILCO) is an intermediary between the power suppliers and the end user whose job is to handle the financial transactions. They indirectly purchase power from the GENCOs and sell to the end user without handling the technical aspect of power flow.

Deregulation also set up a market similar to traditional commodity markets where electricity is traded between suppliers and consumers (usually resellers) [40]. This market, referred to as a power exchange (PX), is operated by a market operator. The retailer therefore has the option of purchasing power from the PX instead of setting up a bilateral agreement directly with the GENCOs. The advantage of the bilateral contract is predictability of prices which makes financial planning easier. On the other hand, being a competitive market place, the PX could provide lower prices and thus economic benefits. Just as the retailer has the option to purchase power from the PX, the GENCO also has the option to sell power at the PX. The Market Operator receives bids for energy supplies from the suppliers and offers of energy purchases from the consumers. They then carry out a market clearing operation that matches the bids and offers. Market clearing is done through an exchange of information with the system operator who confirms the feasibility of the physical system being able to supply the energy in a secure manner. Once the market clearing process is concluded, the bidding results are announced to the agents. In addition to delivering energy to the end user, the system operator also carries out the traditional system operation tasks that ensure system security and reliability such as supply of reserve and ensuring the N-1 criterion is met. The market operator also carries out the service of operating the market on behalf of the suppliers and consumers. The cost of these services are finally passed to the end-user in their electricity bills.

Electricity Markets Categories

Electricity markets can be broken down into the following categories [40]:

1. **Forwards and futures market:** A futures market is an auction market in which participants buy and sell physical or financial products for delivery on a specified future date. These products are called derivatives or derivative products. The most salient feature of futures markets is that they allow trading physical or financial products in the future at today prices. Thus, futures markets are useful if the price of electricity is highly uncertain in the spot market.

2. **Day-ahead market:** In the day-ahead market products which are traded today are delivered on the next day. Day-ahead products are common spot products and can be traded either on a power exchange or as bilateral agreement.

3. **Intra-day market:** The intra-day market is for products with a delivery on the same day. This market allows the producers a short-term load-dependent optimisation of their generation and is typically not a market for pure trading purposes. Intra-day products are traded either on a power exchange or bilaterally [40].

4. **Balancing and reserve market**: There are different definitions of the terms “balancing market” and “reserve market”, because these markets depend on the regulator and are market specific. In this context, the reserve market is the market allowing the ISO to purchase the products needed for compensating imbalances between supply and demand in the electricity system at short notice. The balancing market (also referred to as the real-time market) denotes the market where a merchant purchases or sells the additional energy for balancing his accounting grid. Since the balancing service is provided by the ISO, the ISO usually charges or reimburses the merchant for additional energy and only in some national markets does the merchant have the possibility to buy or sell this balancing energy from or to someone else. Therefore, the balancing market can be regarded as a market only in a broad sense [40]

2.3.3 The Day Ahead Market

The system operator receives energy supply bids from power generation companies detailing the amount of energy they are willing to supply at a a specified trading period and at what price. These bids are usually delivered before a preset deadline on the day before the actual delivery of energy (hence the term day-ahead market). For example, in Japan the deadline is 10 AM on the day prior to the delivery of energy. In a two-sided auction, such as the Japanese market, consumers also submit demand offers detailing the amount of energy they would like to purchase and at what price for a given trading period. The demand offers are also submitted before the bidding deadline [43].

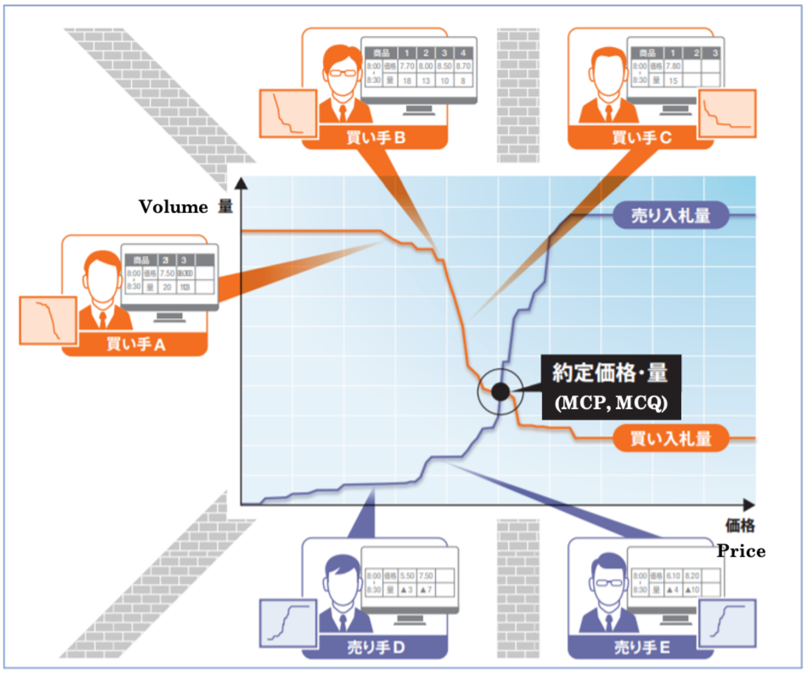

Figure 2.4: Illustration of market clearing at the Japan electric power exchange (JEPX) [44]

The market operator then does a market clearing operation that matches the supply bids and demand offers. A supply curve is constructed by arranging the bids from the lowest to the highest in terms of price. Similarly, a demand curve is constructed by arranging the offers from the highest to the lowest in terms of price. The point where the supply and demand curves intersect is termed the market clearing point and the corresponding demand and price are the market clearing quantity (MCQ) and market clearing price (MCP) respectively as illustrated in Figure 2.4. In systems where power flow between various areas are limited by the physical capacities of the connecting transmission lines, market clearing has to be carried out considering the constraints of how much power can be transferred between areas leading to the possibility of different MCPs for different areas.

Suppliers are allocated portions of the MCQ dependent on whether their bids were lower or higher than the MCP. For those suppliers who receive allocations, they are paid the MCP irrespective of their original bid price. On the other side, consumers whose offers were higher than the MCP receive energy from the market operator while paying the MCP. The results of the market clearing process are also communicated to the system operator who retains the traditional function of scheduling the available generation while ensuring security and reliability of the physical system.

2.3.4 Illustration of the Market Clearing Procedure

The potential for volatility of prices is illustrated in this subsection with a simple simulation of clearing of a PX. Consider a market trading period with ten supply bids and ten demand offers as shown in Table 2.1. The market operator would order the bids based on the bid/offer price as shown in Table 2.2. Also shown in Table 2.2 is a cumulative sum column for the bid quantities.

The supply curves and the demand curves are then constructed from the ordered bids and plotted as shown in Figure 2.5. Based on the obtained curve, the MCP and MCQ would be 6.3 yen/kWh and 370 MWh as shown. Now consider the same bids above but with the price of supply bid S7 changed from 4.00 yen/kWh to 7.00 yen/kWh. The effect of this change in bid price (i.e. the effect of the alteration of S7's bidding strategy) on the MCP and MCQ will be as illustrated in Figure2.6. In this case, the MCP is raised to 6.5 yen/kWh and the MCQ is reduced to 280 MWh. The power allocations will be different and the revenues received and costs incurred by the various market agents will be different. Market agents are constantly engaged in forming strategies to maximize their benefits which translate to shifts in the price and hence uncertainty in the economic performance of the participants.

Table 2.1: Illustrative supply bids and demand offers.

Supply Bid no.	**Bid price [yen/kWh]**	**Bid quantity [MWh]**	**Demand offer no.**	**Offer price [yen/kWh]**	**Offer quantity [MWh]**
S1	5.10	100	D1	8.00	150
S2	6.10	30	D2	6.30	30
S3	7.10	20	D3	5.20	30
S4	4.50	50	D4	6.50	50
S5	5.00	40	D5	6.10	30
S6	5.50	40	D6	7.40	40
S7	4.00	90	D7	6.30	60
S8	8.00	20	D8	7.50	20
S9	4.50	20	D9	4.50	40
S10	7.50	10	D10	7.00	50

Table 2.2: Illustrative aggregated supply bids and demand offers.

Supply Bid no.	**Bid Price [Yen/kWh]**	**Bid Quantity [MWh]**	**Cuml. Quantity [MWh]**	**Demand offer no.**	**Offer Price [Yen/kWh]**	**Offer Quantity [MWh]**	**Cuml. Quantity [MWh]**
S4	4.50	50	50	D1	8.00	150	650
S9	4.50	20	70	D8	7.50	20	670
S5	5.00	40	110	D6	7.40	40	710
S1	5.10	100	210	D10	7.00	50	760
S6	5.50	40	250	D4	6.50	50	810
S2	6.10	30	280	D2	6.30	30	840
S7	7.00	90	370	D7	6.30	60	900
S3	7.10	20	390	D5	6.10	30	930
S10	7.50	10	400	D3	5.20	30	960
S8	8.00	20	420	D9	4.50	40	1000

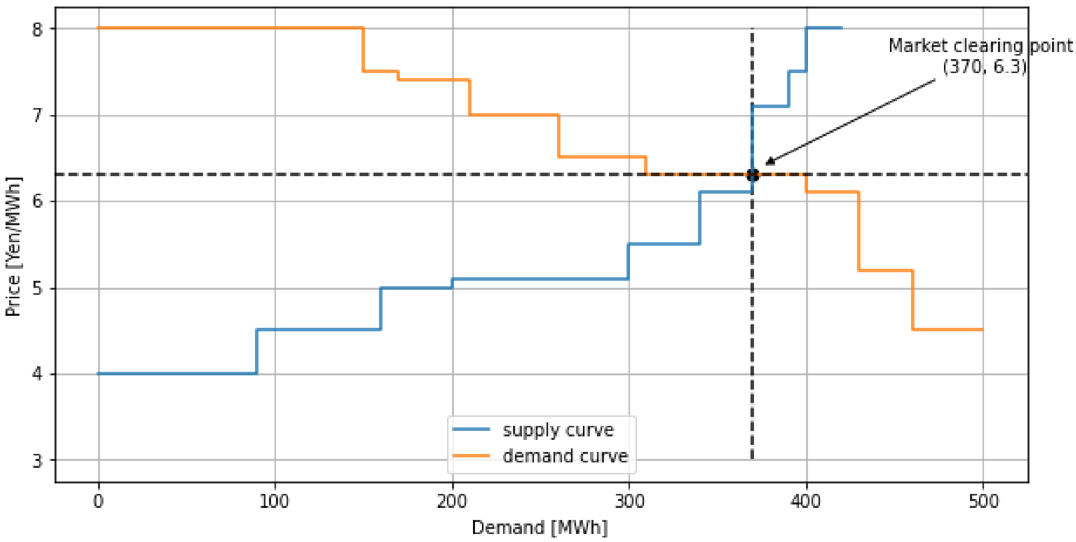

Figure 2.5: Illustration of electricity market clearing

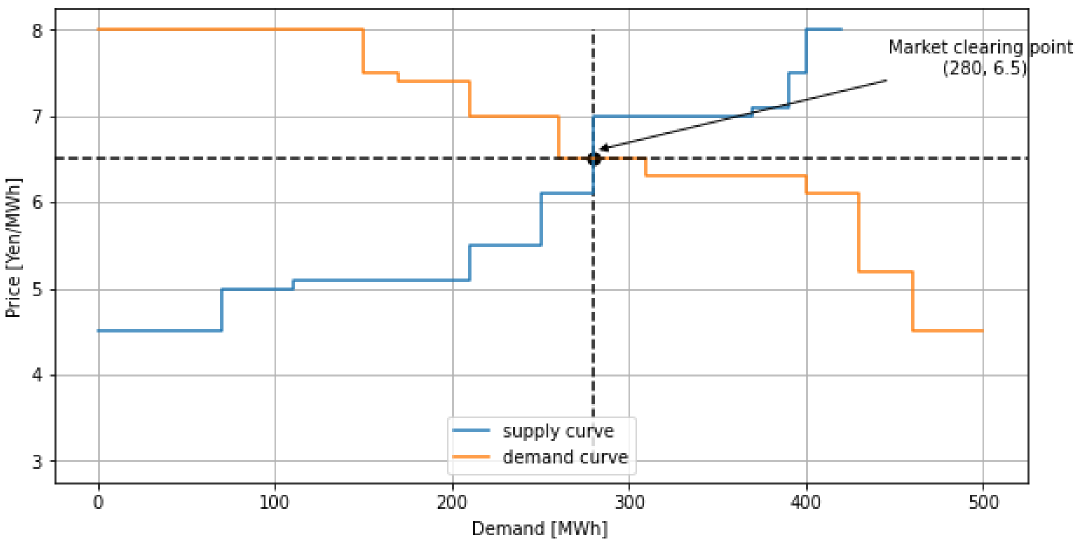

Figure 2.6: Illustration of the effect of altering bids on the market clearing results

2.4 Japanese Electricity Market

2.4.1 Background of the Japanese Electricity Industry

Japan's electric power industry came into being in 1886 with the commencement of operations by the Tokyo Electric Light Company which was formed using private equity. Numerous electric utilities were subsequently established around Japan to serve the growing demand for electricity driven by increasing industrialization [43]. By the early 1930s there were more than 800 utilities. However, as Japan headed into World War II, the electric utilities came under the control of the government. In 1939, the government established the Japan Electric Generation and Transmission Company and electricity generation and transmission facilities came under centralized control. The government also consolidated the electricity distribution business into nine separate regional blocks [43].

Following World War II, the Japan Electric Generation and Transmission Company was dissolved in May 1951, the company's facilities and functions were transferred to nine privately owned electricity distribution utilities. As a result, a regime of regional monopolies was established based on integrated systems of electricity generation and transmission in nine regions. The number of utilities then increased to 10 with the establishment of Okinawa Electric Power Company following the reversion of Okinawa to Japanese control in 1972. These electric utilities made focused investments in power supply facilities to meet a growing demand for electricity driven by Japan's rapid economic growth [43]. Hence, they sufficiently contributed to the Japanese rapid economic growth.

2.4.2 Liberalization of the Japanese Electricity Market

Following the trend toward deregulation in the electric power industry in Western countries, the liberalization of entry into the electricity generation sector started in 1995 in Japan, followed by retail supply liberalization for customers receiving extra-high voltage (20kV or above) in 2000. The scope of deregulation was expanded further in stages thereafter: to high voltage (6kV) customers with contracted demand of 500 kW or above, in principle, in April 2004, and to all customers in the high voltage category (those with a contracted demand of 50kW or above) in April 2005 [45].

However, power shortages and other issues caused by the 2011 Great East Japan Earthquake prompted discussion of the ideal configuration of the nation's electric power system and its reform. Based on this discussion, full liberalization the electricity market was pursued since 2015. The liberalization of the electric power retailing and generation sectors was completed in April 2016. The legal separation of transmission and distribution from vertically integrated businesses was implemented in April 2020, resulting in the spin-off of new transmission and distribution companies from the former general electricity utilities [43].

While these changes have several obvious and important benefits, they make the system more complex than before. There is significantly more volatility and uncertainty in the system especially in the balancing of supply and demand and consequently in the clearing of the electricity market. This means that electricity market prices are especially volatile even at the intra-hour timescale which presents financial risks especially to small-scale electricity resellers as they try to compete with the traditional large utilities. Today, Japan's electrical power industry comprises three major sectors: electricity generation, transmission and distribution, and retailing. The number of operators in these sectors is 986, 49, and 730 respectively [43].

2.4.3 The Japan Electric Power Exchange (JEPX)

The Japan Electric Power Exchange (JEPX) was established in November 2003 as a day ahead electricity market and commenced trading in April 2005. The purpose of JEPX is to provide a mechanism for power suppliers and customers to trade electricity in an environment where the price is dictated by the market forces of supply and demand. This mode of operation has several advantages including competitive prices for consumers and improved operational efficiency. The primary goal for market participants is to operate efficiently and economically in the electricity market.

The JEPX day-ahead market is a two-sided auction where power generation companies (GENCOs) place supply bids while electricity retailers (RETAILCOs) place demand offers two days before the trading day [43]. The market operator (MO) then clears the market based on ranking of supply bids and demand offers taking into consideration constraints on power that can be transmitted between areas. The market is run daily and the results which include the supply volume, demand volume, system price and area specific prices are posted on the publicly viewable JEPX website [44]

The Japanese power grid is divided into two frequency systems: a 60 Hz system in eastern Japan and a 50 Hz system in western Japan with no international connection [45]. There are ten service areas as shown in Figure 2.7. Also shown in Figure 2.7 are the transmission capacity limits in the connections between areas. These capacity limits, result in transmission congestion hence differences in MCPs between areas. The principal market participants are the electricity generation utilities and electricity retailers involved in wholesale power transactions.

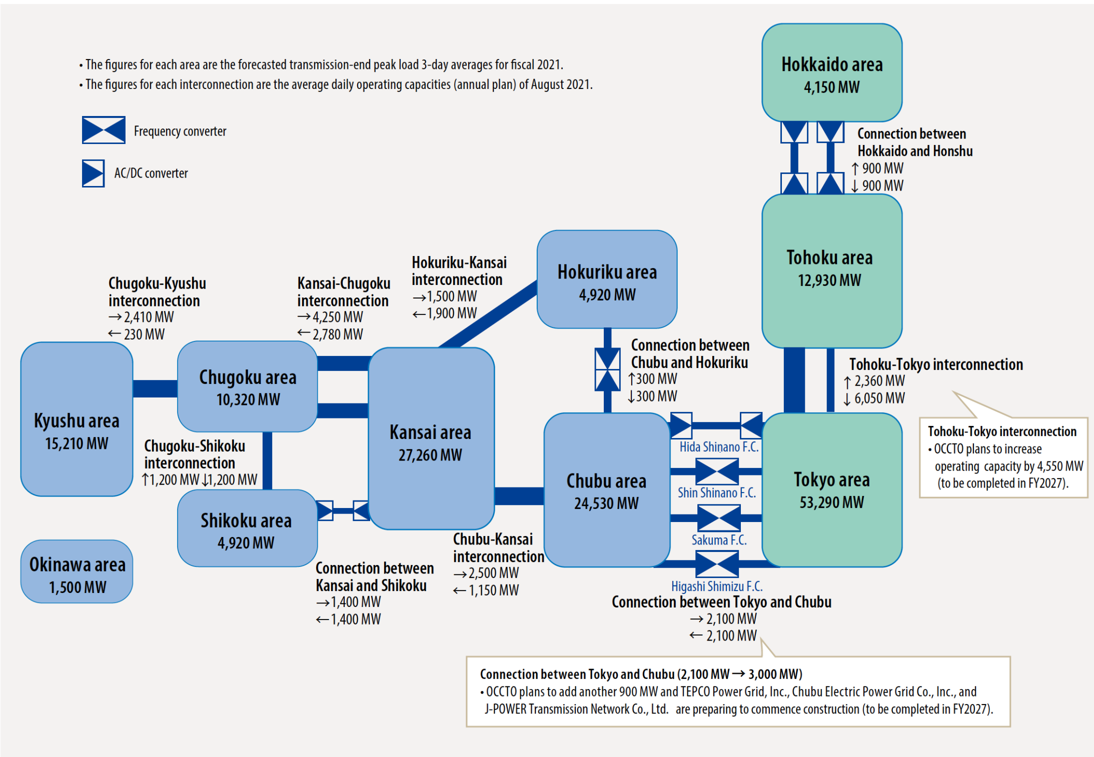

Figure 2.7: Illustration of the 10 market areas in the Japanese electricity market. source: Adopted from the Japan Electric Power Information Center (JEPIC) [43]

The JEPX currently provides a marketplace for the three main electricity transactions:

1. **Spot market:** Trading in 30-minute increments of electricity for next-day delivery. Accounts for approximately 99% of JEPX trading.

2. **Forward market:** Trading in electricity for delivery over the course of a specified future period. Products are created by packaging together specific periods and times, such as monthly 24-hour products or weekly daytime products.

3. **Intra-day market:** A market for correcting unexpected misalignment between supply and demand occurring between a spot market transaction and delivery (a minimum of one hour later).

2.5 Summary

This chapter presents the background of some of the main concepts related to the research carried out in this thesis. A brief introduction to the structure of electricity markets, their development, functioning and characteristics has been given, along with an insight into the Japanese electricity market. The chapter also reviews recent literature on approaches for modelling and forecasting electricity market prices and price spikes with a realization that considerable effort has gone into finding suitable models for forecasting electricity prices. However, there is still need for models dealing specifically with the electricity spikes time series data.

===

REFERENCES

References

[1] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014. doi: https://doi.org/10.1016/j.ijforecast.2014.08.008.

[2] L. Jiang and G. Hu, "A review on short-term electricity price forecasting techniques for energy markets," in *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 937–944, 2018. doi: https://doi.org/10.1109/ICARCV.2018.8581312.

[3] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1548–1568, 2018. doi: https://doi.org/10.1016/j.rser.2017.05.234.

[4] L. Fahrmeir, T. Kneib, S. Lang, and B. D. Marx, *Regression Models*, pp. 23–84. Berlin, Heidelberg: Springer Berlin Heidelberg, 2021.

[5] J. Ramos, A. Ferreira, and P. Fernandes, "A linear regression pattern for electricity price forecasting in the iberian electricity market," *Revista Facultad de Ingeniería Universidad de Antioquia*, 05 2019. doi: https://doi.org/10.17533/udea.redin.20190522.

[6] H. Akaike, *Autoregressive Model Fitting for Control.* New York, NY: Springer New York, 1998.

[7] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, "Arima models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003. doi: https://doi.org/10.1109/TPWRS.2002.804943.

[8] S. M. Ulyah, M. F. F. Mardianto, and Sediono, "Comparing the performance of seasonal arimax model and nonparametric regression model in predicting claim reserve of education insurance," *Journal of Physics: Conference Series*, vol. 1397, p. 012074, dec 2019.

[9] C. McHugh, S. Coleman, D. Kerr, and D. McGlynn, "Forecasting day-ahead electricity prices with a sarimax model," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1523–1529, 2019. doi: https://doi.org/10.1109/SSCI44817.2019.9002930.

[10] M. Fakhfekh and A. Jeribi, "Volatility dynamics of crypto-currencies' returns: Evidence from asymmetric and long memory garch models," *Research in International Business and Finance*, vol. 51, p. 101075, 2020.

[11] A. Cifter, "Forecasting electricity price volatility with the markov-switching garch model: Evidence from the nordic electric power market," *Electric Power Systems Research*, vol. 102, pp. 61–67, 2013. doi: https://doi.org/10.1016/j.epsr.2013.04.007.

[12] E. Cruz May, A. Bassam, L. J. Ricalde, M. Escalante Soberanis, O. Oubram, O. May Tzuc, A. Y. Alanis, and A. Livas-García, "Global sensitivity analysis for a real-time electricity market forecast by a machine learning approach: A case study of mexico," *International Journal of Electrical Power & Energy Systems*, vol. 135, p. 107505, 2022. doi: https://doi.org/10.1016/j.ijepes.2021.107505.

[13] E. Foruzan, S. D. Scott, and J. Lin, "A comparative study of different machine learning methods for electricity prices forecasting of an electricity market," in *2015 North American Power Symposium (NAPS)*, pp. 1–6, 2015. doi: https://doi.org/10.1109/NAPS.2015.7335095.

[14] M. Castelli, A. Groznik, and A. Popovič, "Forecasting electricity prices: A machine learning approach," *Algorithms*, vol. 13, no. 5, 2020. doi: https://doi.org/10.3390/a13050119.

[15] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218–230, 2016. doi: https://doi.org/10.1016/j.apenergy.2015.09.087.

[16] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, 2018. doi: https://doi.org/10.3390/en11051255.

[17] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018. doi: https://doi.org/10.1016/j.apenergy.2018.02.069.

[18] J. Pórtoles, C. González, and J. M. Moguerza, "Electricity price forecasting with dynamic trees: A benchmark against the random forest approach," *Energies*, vol. 11, no. 6, 2018. doi: https://doi.org/10.3390/en11061588.

[19] A. Wagner, E. Ramentol, F. Schirra, and H. Michaeli, "Short- and long-term forecasting of electricity prices using embedding of calendar information in neural networks," *Journal of Commodity Markets*, vol. 28, p. 100246, 2022. doi: https://doi.org/10.1016/j.jcomm.2022.100246.

[20] X. Chen, Z. Y. Dong, K. Meng, Y. Xu, K. P. Wong, and H. W. Ngan, "Electricity price forecasting with extreme learning machine and bootstrapping," *IEEE Transactions on Power Systems*, vol. 27, no. 4, pp. 2055–2062, 2012. doi: https://doi.org/10.1109/TPWRS.2012.2190627.

[21] Z. Hajirahimi and M. Khashei, "Hybrid structures in time series modeling and forecasting: A review," *Engineering Applications of Artificial Intelligence*, vol. 86, pp. 83–106, 2019.

[22] N. Chaâbane, "A novel auto-regressive fractionally integrated moving average–least-squares support vector machine model for electricity spot prices prediction," *Journal of Applied Statistics*, vol. 41, no. 3, pp. 635–651, 2014. doi: 10.1080/02664763.2013.847068.

[23] W. Li and D. M. Becker, "Day-ahead electricity price prediction applying hybrid models of lstm-based deep learning methods and feature selection algorithms under consideration of market coupling," *Energy*, vol. 237, p. 121543, 2021. doi: https://doi.org/10.1016/j.energy.2021.121543.

[24] J. Wang, W. Yang, P. Du, and T. Niu, "Outlier-robust hybrid electricity price forecasting model for electricity market management," *Journal of Cleaner Production*, vol. 249, p. 119318, 2020. doi: https://doi.org/10.1016/j.jclepro.2019.119318.

[25] B. Uniejewski and R. Weron, "Efficient forecasting of electricity spot prices with expert and lasso models," *Energies*, vol. 11, no. 8, 2018. doi: https://doi.org/10.3390/en11082039.

[26] Z. Tan, J. Zhang, J. Wang, and J. Xu, "Day-ahead electricity price forecasting using wavelet transform combined with arima and garch models," *Applied Energy*, vol. 87, no. 11, pp. 3606–3610, 2010. doi: https://doi.org/10.1016/j.apenergy.2010.05.012.

[27] S. Zedda and G. Masala, "Price spikes in the electricity markets: how and why," 07 2019.

[28] T. D. Mount, Y. Ning, and X. Cai, "Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters," *Energy Economics*, vol. 28, no. 1, pp. 62–80, 2006. doi: https://doi.org/10.1016/j.eneco.2005.09.008.

[29] T. Christensen, A. Hurn, and K. Lindsay, "It never rains but it pours: Modeling the persistence of spikes in electricity prices," *Energy Journal -Cambridge Ma then Cleveland Oh-*, vol. 30, 07 2008. doi: https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No1-2.

[30] T. Christensen, A. Hurn, and K. Lindsay, "Forecasting spikes in electricity prices," *International Journal of Forecasting*, vol. 28, no. 2, pp. 400–411, 2012. doi: https://doi.org/10.1016/j.ijforecast.2011.02.019.

[31] S. M. Sirin and I. Erten, "Price spikes, temporary price caps, and welfare effects of regulatory interventions on wholesale electricity markets," *Energy Policy*, vol. 163, p. 112816, 2022. doi: https://doi.org/10.1016/j.enpol.2022.112816.

[32] R. Tidball, J. Bluestein, N. Rodriguez, and S. Knoke, "Cost and performance assumptions for modeling electricity generation technologies," doi: https://doi.org/10.2172/993653.

[33] J. D. Glover, M. S. Sarma, and T. Overbye, *Power system analysis & design, SI version*. Cengage Learning, 2012.

[34] A. Von Meier, *Electric power systems: a conceptual introduction*. John Wiley & Sons, 2006.

[35] J. J. Grainger, *Power system analysis*. McGraw-Hill, 1999.

[36] E. Lakervi and E. J. Holmes, *Electricity distribution network design*. No. 212, IET, 1995.

[37] K. Mayer and S. Trück, "Electricity markets around the world," *Journal of Commodity Markets*, vol. 9, pp. 77–100, 2018. doi: https://doi.org/10.1016/j.jcomm.2018.02.001.

[38] E. R. Larsen and D. W. Bunn, "Deregulation in electricity: understanding strategic and regulatory risk," *Journal of the Operational Research Society*, vol. 50, no. 4, pp. 337–344, 1999. doi: https://doi.org/10.1057/palgrave.jors.2600695.

[39] D. Kirschen and G. Strbac, *Fundamentals of Power System Economics*. Wiley, 2019.

[40] M. Shahidehpour, H. Yamin, and Z. Li, *Market operations in Electric Power Systems: Forecasting, scheduling, and Risk Management*. Institute of Electrical and Electronics Engineers, Wiley-Interscience, 2002.

[41] J. Bushnell and K. Novan, "Setting with the sun: The impacts of renewable energy on wholesale power markets," Working Paper 24980, National Bureau of Economic Research, August 2018.

[42] L. Barroso and A. Conejo, "Decision making under uncertainty in electricity markets," in *2006 IEEE Power Engineering Society General Meeting*, pp. 3 pp.–, 2006. doi: https://doi.org/10.1109/PES.2006.1709323.

[43] "The Electric Power Industry in Japan 2020." https://www.jepic.or.jp/pub/pdf/epijJepic2020.pdf. Accessed: 2022-08-08.

[44] "The Japan Electric Power Exchange Website." http://www.jepx.org/english/index.html. Accessed: 2022-08-08.

[45] K. Ofuji and S. Kanemoto, "Price forecasting of japan electric power exchange using time-varying ar model," in *2007 International Conference on Intelligent Systems Applications to Power Systems*, pp. 1–6, 2007. doi: https://doi.org/10.1109/ISAP.2007.4441658.

3 Fundamental Analysis of Price Spikes in the Japanese Electric Power Exchange

3.1 Introduction

3.1.1 Background

Following a worldwide trend of electricity markets liberalization, the Japanese Electric Power Exchange (JEPX) provides a platform for trading of electric energy in Japan [1]. However, such a market structure introduces new operational dynamics that did not exist in the era of a centralized system. One major effect is increased uncertainty in the price of electricity even though in the long run, the average price of may be lowered. A related impact is a risk of reduced system reliability as electric power generators are more concerned with maximizing profits in contrast to the objectives of the original power utilities whose main aim was to ensure availability of the resource.

As with any other commodity markets, electricity markets can experience market shocks – a situation in which prices are driven much higher that normal due to extreme market conditions. Such a scenario was observed in the JEPX around late December 2020 / early January 2021 where prices hit a high of 220Yen/kWh which is almost 30 times the average price of around 7 Yen/kWh.

Given this backdrop, there is need for mathematical models of the system that may be used to understand the underlying dynamics that lead to such extreme events. In this chapter we present a preliminary analysis of prices and extreme prices (price spikes) in the JEPX. We also present result fitting a classical Hawkes model [2] to the price spikes time series data. Such models could be used by market agents for forecasting purposes and hence developing bidding strategies or may be used by market operators and regulators to detect extreme circumstances.

3.1.2 Literature Review

Several papers have studied the nature of prices in various deregulated electricity markets. These studies have variously pointed out the "spiky" nature of the electricity prices time series. The day ahead and real time market data for the Turkish power market for the period 2012–2015 is analyzed in [3] to classify price spikes and their causes. They further investigated the levels of deviation between the day ahead market values and the real time market values. They defined price deviation and load deviation ratios to measure the level of deviation both in price and demand. The analysis for the load is based on load shedding and cycling values. They analyzed the mean and standard deviation in market prices and determined the price spike as a two-sigma deviation from the mean value. It was shown that 60% of the price deviation ratios were in the range of (±20%), while 44% were in the range of (±10%) and 35% are in the range of (±5%). They also showed that 56.9% of the spikes are due to problems in the generation of natural gas-based power plants which affect the day ahead and real time prices. A total of 29.2% of the spikes are due to power plant and system failures that affect only real time prices. They drew the conclusion that extreme differences between the day ahead planning and real time market could be an indicator of system management problems. They also found that the most drastic price differences are due to natural gas shortage problems.

In [4] a study of the historical bidding behavior is carried out to see how power suppliers and demand service providers were actually bidding in the California day-ahead energy market. Based on their observations, they formulate a Prisoner's dilemma matrix game and introduce the notion of "opportunistic tacit collision" to explain strategic bidding behaviors in which suppliers withhold generation capacity from the market to drive up prices. This explanation is applicable with or without market power, transmission constraints, and insufficient supply, and is only enhanced by these factors. Their analysis of historical bid curves from the California day-ahead energy market also suggested that withholding during the study period was not a routine practice. Withholding, however, was observed during every price spike, and illustrated in the two examples studied in the paper.

A data mining-based approach for predicting the occurrence of the electricity market price spikes together with the ability of predicting normal range prices is presented in [5]. After applying feature selection techniques and statistical analysis of relevant factors, the authors proposed method showed a promising result in price spike occurrence prediction. The case studies also showed that among many existing classification algorithms, SVM can give a reliable spike occurrence prediction. Moreover, the result of the probability classifier can be combined with SVM to improve the prediction accuracy and provide more information. In their case studies, they combined the spike forecast with the expected price forecast to give a complete forecast of market prices.

A multi-feature based approach with the incorporation of variable thresholds is developed in [6] to detect electricity price spikes in the national electricity market of Australia. The variable thresholds, which were determined using a weighted sliding window average and an adjusted standard deviation to help to segregate spikes from normal price variations. Also, significant features were extracted from the market after analyzing the underlying causes resulting into the price spikes. These features are employed as inputs to a support vector machine to classify electricity prices as spikes or non-spikes. A case study was conducted using a dataset acquired from the state of New South Wales, Australia.

In [7], the authors explain that price spike forecasting has two main aspects: prediction of price spike occurrence and value. In their paper, a novel technique for price spike occurrence prediction was presented composed of a new hybrid data model, a novel feature selection technique and an efficient forecast engine. The hybrid data model included both wavelet and time domain variables as well as calendar indicators, comprising a large candidate input set. The set was refined by the proposed feature selection technique evaluating both relevancy and redundancy of the candidate inputs. The forecast engine was a probabilistic neural network, which were fed by the selected candidate inputs of the feature selection technique and predict price spike occurrence. The efficiency of the whole proposed method for price spike occurrence forecasting was evaluated by means of real data from the Queensland and PJM electricity markets.

3.1.3 Chapter Organization

In this chapter a fundamental analysis of prices and price spikes in the JEPX is carried out. In section 3.2, the prices time series is introduced including the dependence on the supply-demand balance in the market. The spatial and temporal dependence of prices is also illustrated. Finally, the definition of price spikes and its justification is given. In section 3.3, the modelling of price spikes time series using the classical Hawkes model is presented. Extracted model parameters are shown and the evolution of the intensity function (spike occurrence probability) is given. Finally, chapter conclusions are given in section 3.4.

3.2 JEPX Data

Electricity price and demand data for the JEPX is publicly available on the exchange's website [8]. The available data includes traded energy volumes and area prices in thirty minute resolutions giving 48 commodities per area per day. Figure 3.1 shows a snapshot of the raw data downloaded from the exchange's website. The trading volume in the spot market has been on the rise since 2016, reaching approximately 312.8 TWh in fiscal 2020. This means that more than 30% of all electricity sold in Japan is sold through JEPX. New electricity retailers procure more than 80% of their electricity from the spot market [1].

3.2.1 Evolution of System Prices and Clearing Volumes

The average system price has ranged between the 7–9 Yen/kWh mark since the 2016 fiscal year. Prices fell in April 2020 as demand dropped from the impact of measures against COVID-19 [9]. In that month, trading occurred at the lowest possible system price of 0.01 yen/kWh. Later, prices have skyrocketed from the effects of a summer heat wave and cold winter months which have been further exacerbated by the fuel supply crunch.

Figure 3.2 shows the average daily market clearing prices (in blue) and the 30-day moving average (in orange) between April 2016 and March 2022. The plot exhibits trends of generally increasing and decreasing prices over the period. There is also seasonality aspect where prices rise during the summer and winter

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Date	Timecode	Supply Volume (kWh)	Purchase Volume (kWh)	Clearing Volume (kWh)	System Price (円/kWh)	Area Price Hokkaido (円/kWh)	Area Price Tohoku (円/kWh)	Area Price Tolyo (円/kWh)	Area Price Chubu (円/kWh)	Area Price Hokuriku (円/kWh)	Area Price Kansai (円/kWh)	Area Price Chugoku (円/kWh)	Area Price Shikoku (円/kWh)	Area Price Kyushu (円/kWh)
2	4/1/2020	1	16592450	18033600	15772950	6.43	6.84	6.76	6.76	4	4	4	4	4	4
3	4/1/2020	2	16986950	18180600	15766550	5.77	6.62	6.62	6.62	3.73	3.73	3.73	3.73	3.73	3.73
4	4/1/2020	3	17263950	18234450	15990350	5.18	6.61	6.61	6.61	4.09	4.09	4.09	4.09	4.09	4.09
5	4/1/2020	4	17583250	18565050	16078100	4.87	6.51	6.51	6.51	4.04	4.04	4.04	4.04	4.04	4.04
6	4/1/2020	5	17943550	18738950	16338200	4.87	6.51	6.51	6.51	4.09	4.09	4.09	4.09	4.09	4.09
7	4/1/2020	6	18184000	18802350	16447650	4.87	8.3	6.5	6.5	4.43	4.43	4.43	4.43	4.43	4.43
8	4/1/2020	7	18271250	18893350	16503500	4.5	6.5	6.5	6.5	4.43	4.43	4.43	4.43	4.43	4.43
9	4/1/2020	8	18421100	18892350	16592600	4.43	6.5	6.5	6.5	4.04	4.04	4.04	4.04	4.04	4.04
10	4/1/2020	9	18372800	18952050	16682800	4.43	6.5	6.5	6.5	4	4	4	4	4	4
11	4/1/2020	10	18385550	18940300	16588650	4.34	6.51	6.51	6.51	3.73	3.73	3.73	3.73	3.73	3.73
12	4/1/2020	11	18261400	18989700	16462450	4.45	6.61	6.61	6.61	3.73	3.73	3.73	3.73	3.73	3.73
13	4/1/2020	12	18129800	19082750	16372850	4.87	6.62	6.62	6.62	4.04	4.04	4.04	4.04	4.04	4.04
14	4/1/2020	13	18253600	19397500	16435600	4.87	8.41	8.41	8.41	4.43	4.43	4.43	4.43	4.43	4.43
15	4/1/2020	14	18380850	19864600	16437150	5.18	10.13	8.41	8.41	4.44	4.44	4.44	4.44	4.44	4.44
16	4/1/2020	15	18500000	20550950	16652850	6.57	10.22	8.65	8.65	4.87	4.87	4.87	4.87	4.87	4.87
17	4/1/2020	16	18568000	21207900	16901250	6.85	10.04	8.65	8.65	4.87	4.87	4.87	4.87	4.87	4.87
18	4/1/2020	17	19168500	22235000	17645150	7.74	10.14	9	9	5.83	5.83	5.83	5.83	5.83	5.83
19	4/1/2020	18	19302900	23296150	17805900	8.71	11.67	9.17	9.17	7.69	7.69	7.69	7.69	7.69	6.22
20	4/1/2020	19	19112100	23826750	18023000	9.43	11.89	10.76	10.76	10.35	10.35	10.35	10.35	10.35	6.22
21	4/1/2020	20	19309250	23588350	18067050	9.12	11.98	8.6	10.35	10.35	10.35	10.35	10.35	10.35	6.22
22	4/1/2020	21	19573650	23262650	18008500	8.75	11.16	7.72	9.24	9.24	9.24	9.24	9.24	9.24	6.22
23	4/1/2020	22	19741550	23461700	17991350	8.68	7.72	7.72	9.24	9.24	9.24	9.24	9.24	9.24	6.22
24	4/1/2020	23	19725750	23351000	17924600	8.63	7.58	7.58	9.09	7.69	7.69	7.69	7.69	7.69	6.22
25	4/1/2020	24	19596150	23260050	17791750	8.63	7.56	7.56	9.09	7.06	7.06	7.06	7.06	7.06	6.22
26	4/1/2020	25	19560950	22728950	17580050	7.38	6.43	6.43	8.84	4.87	4.87	4.87	4.87	4.87	4.87
27	4/1/2020	26	19722750	22616100	17615250	7.03	6.43	6.43	8.84	4.87	4.87	4.87	4.87	4.87	4.87
28	4/1/2020	27	19684450	23068550	17767700	8.19	8.96	8.96	8.96	5.83	5.83	5.83	5.83	5.83	5.83
29	4/1/2020	28	19596550	23158900	17703100	8.52	11.95	8.96	8.96	6.15	6.15	6.15	6.15	6.15	6.15
30	4/1/2020	29	19731400	22657250	17836200	8.43	15.95	9.02	9.02	5.18	5.18	5.18	5.18	5.18	5.18
31	4/1/2020	30	19529650	22573800	17881100	8.43	9.02	9.02	9.02	4.87	4.87	4.87	4.87	4.87	4.87
32	4/1/2020	31	19267600	22781450	17827450	8.42	14.59	9.09	9.09	4.87	4.87	4.87	4.87	4.87	4.87
33	4/1/2020	32	19042100	22808300	17798000	8.7	25	11.52	11.52	5.83	5.83	5.83	5.83	5.83	5.83
34	4/1/2020	33	18909250	22766000	17663450	8.78	11.95	11.64	11.64	7.4	7.4	7.4	7.4	7.4	7.4
35	4/1/2020	34	18730200	22416000	17450050	8.86	18.94	11.52	11.52	7.59	7.59	7.59	7.59	7.59	7.4
36	4/1/2020	35	19253200	21808250	17742900	8.75	16	9.29	9.29	5.18	5.18	5.18	5.18	5.18	5.18
37	4/1/2020	36	19253750	21542650	17575750	8.75	11.54	9.29	9.29	5.83	5.83	5.83	5.83	5.83	5.83
38	4/1/2020	37	19603950	21637100	17632800	8.83	15.5	9.24	9.24	6.22	6.22	6.22	6.22	6.22	6.22
39	4/1/2020	38	19581700	21668550	17674350	8.83	15.4	9.57	9.57	5.83	5.83	5.83	5.83	5.83	5.83

Figure 3.1: Snapshot of JEPX data showing the 30-minutes resolution day ahead market clearing results [8].

months. The corresponding average daily market clearing quantities and the 28-day moving averages are shown in Figure 3.3. This plot shows a clear rising trend in traded volumes over the period as more participants enter the market. Similarly, seasonality is exhibited in the traded volumes with increases during the summer and winter months.

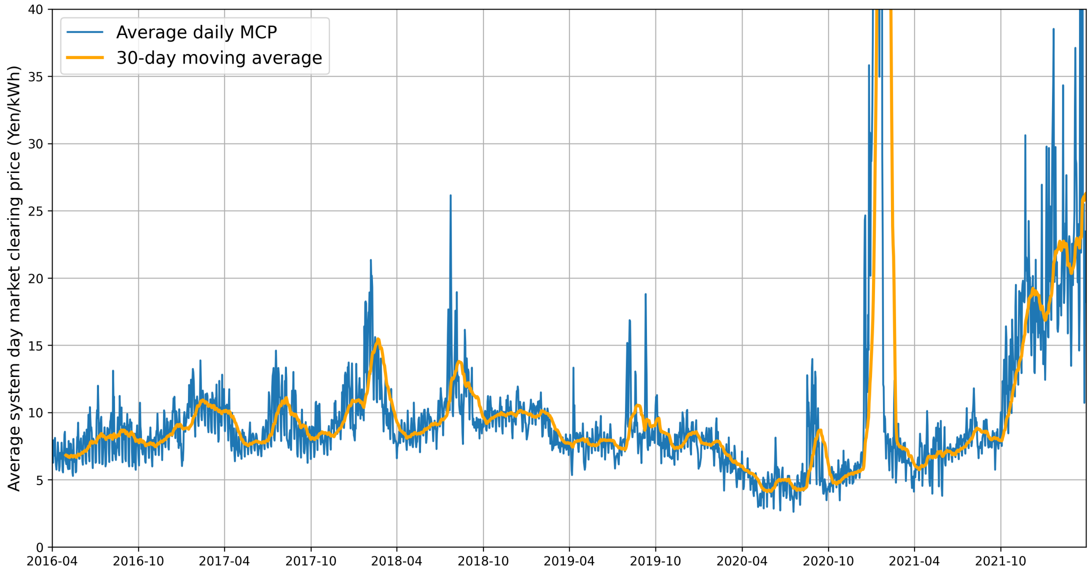

Figure 3.2: Average daily system market clearing price (MCP) and the corresponding 30-day moving average.

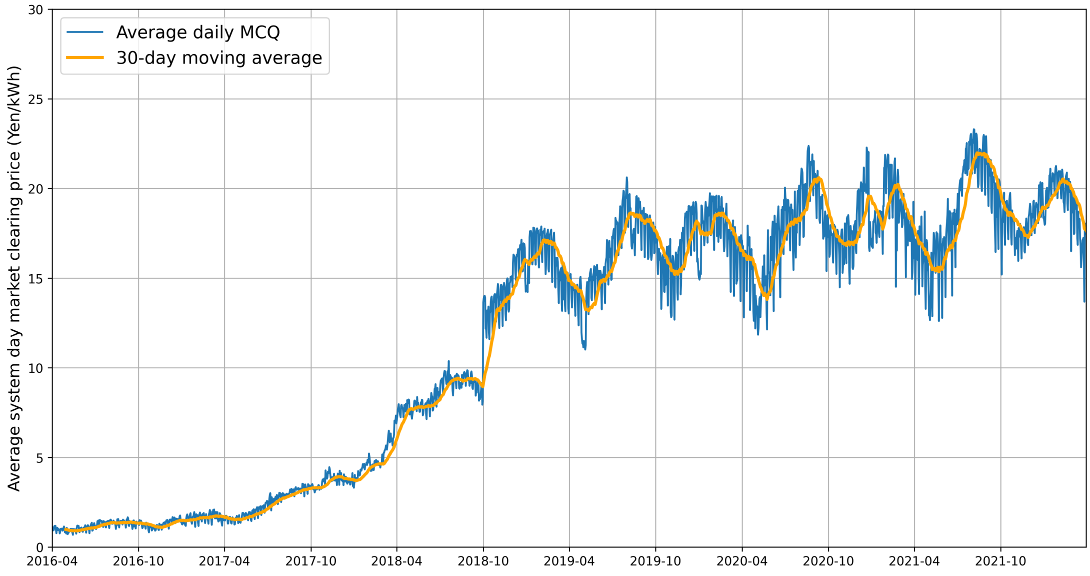

Figure 3.3: Average daily system market clearing quantity (MCP) and the corresponding 30-day moving average.

3.2.2 Spatial and Temporal Dependence of Price Data

Price characteristics are dependent on the trading area. An analysis of the raw data shows price differences between trading areas where differences are caused by constraints on electricity transmission which may limit the transferability of cheaper energy between areas. Figure 3.4 shows a heat map of the median prices for the 48 time-codes in each of the 9 areas. The plot shows the differences in price by area with Hokkaido being generally more expensive followed by the two eastern regions of Tohoku and Tokyo. The median prices in the Western regions are generally similar but the prices in the furthest west region of Kyushu are slightly lower than the other regions due to high penetration of solar energy.

Figure 3.5 on the other hand shows a heat map of the median system prices for the 48 time-codes for each type of day. The plot shows the differences in price by day of week with weekdays being generally more expensive than weekends. The median prices on holiday days are very similar in characteristics to price on Saturdays.

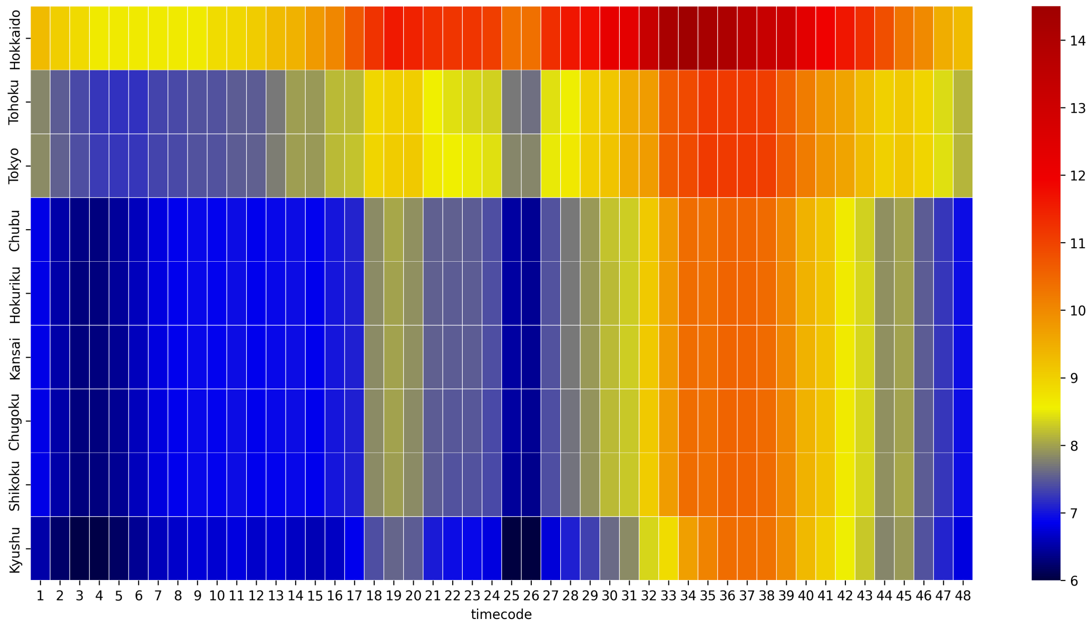

Figure 3.4: Heat map of median area prices grouped by trading time-slot. Demonstrates price dependence on area.

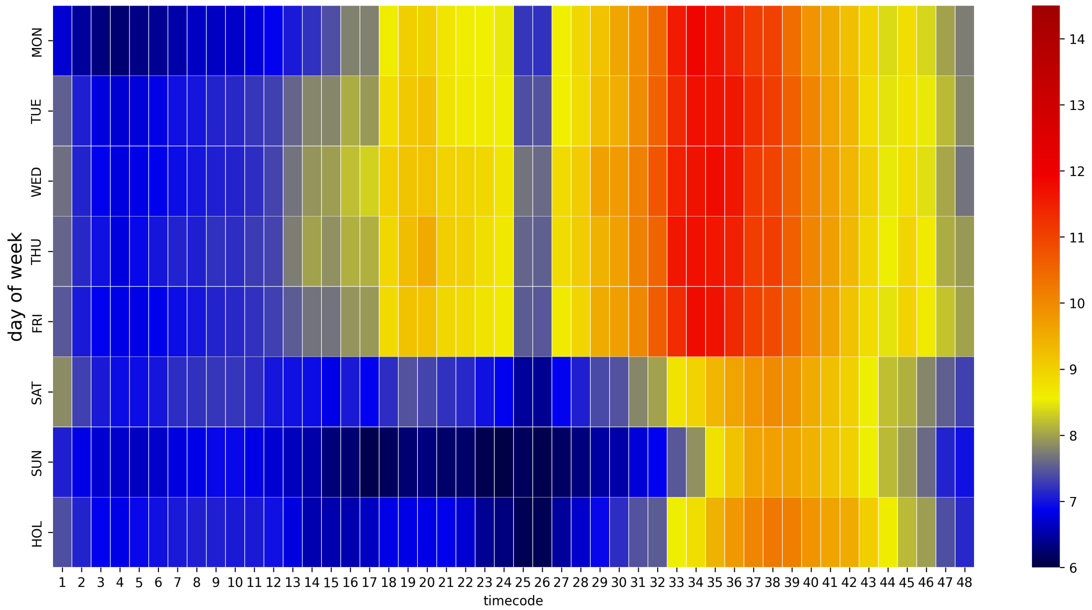

Figure 3.5: Heat map of median system price grouped by trading time-slot. Demonstrates price dependence on day of week.

3.2.3 Definition of Price Spike Events

The plots of electricity price data shown in Figure 3.2 show a trend of relatively small fluctuations interspersed by periods of much larger fluctuations. These characteristics are explained by the nature of the underlying processes generating the data. MCP is largely the result of the balance between supply and demand in the market. Typically, there will be small daily variations as power suppliers and retailers tweak their supply and demand bids respectively in an effort to maximize the gains from the market. These actions, in conjunction with day-to-day variations in power demand and/or supply will result in the observed relatively small day-to-day fluctuations.

However, from time to time, there will be significant imbalances between supply and demand. The sudden loss of a large generating unit for example will have the effect of a sudden significant reduction in cheaper supply since larger units tends to be cheaper. The net effect is a sudden increase in the MCP which will be recorded as a price spike. During the period immediately following such an event, there will be lots of movements among both the supply and demand agents all trying to maximize their benefits while the situation lasts. Once the underlying event passes and the market balance returns to normal, the prices will tend to return to typical levels. Note that events could also affect demand, a good example being the unexpected extreme weather conditions that was observed during the winter of December 2020 – January 2021.

The dependence of price on trading area and time-codes is further evidenced by the plots of Figure 3.6 which show the range of prices (10^{th} percentile, 50^{th} percentile, and 90^{th} percentile) for each of the 9 regions. These plots emphasize the need for models that vary by trading area and trading time-code. From Figure 3.6, it can be observed that the peak of the 90th percentile curve in most areas (other than Hokkaido) is about 25 Yen/kWh. We therefore pick this value as the threshold above which we define spikes as having spiked.

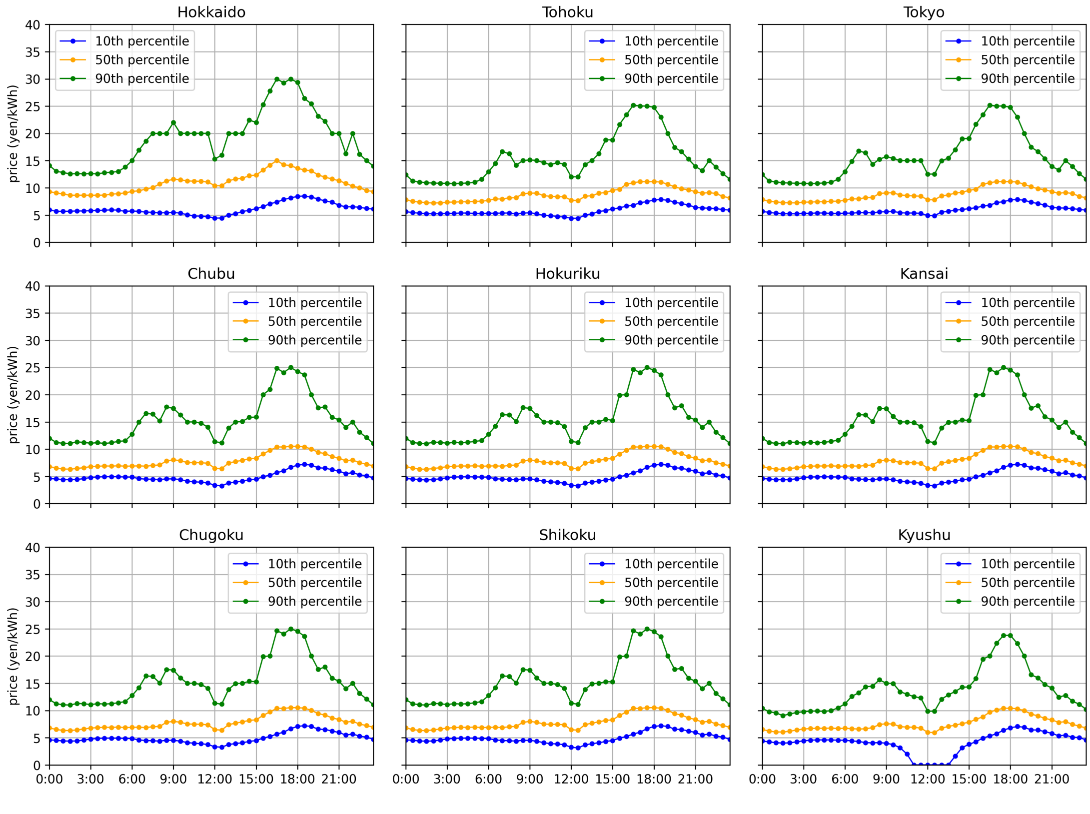

Figure 3.6: Range of area prices (10^{th} percentile, 50^{th} percentile, and 90^{th} percentile)

3.3 Modelling of Price Spikes Time Series

3.3.1 Price Spikes Time Series Data

The preceding explanation motivates the choice of the mathematical model. We characterize the data to comprise two sources of noise: 1) low variance IID noise that explains small inter-day variations in the data and, 2) relatively larger magnitude disturbances that explain periodic spikes in prices. We define a price spike as a price above a specified threshold. This threshold will be representative of a risk value above which the retailer would typically hedge their stakes.

The series of data points showing prices above the pre-defined threshold value

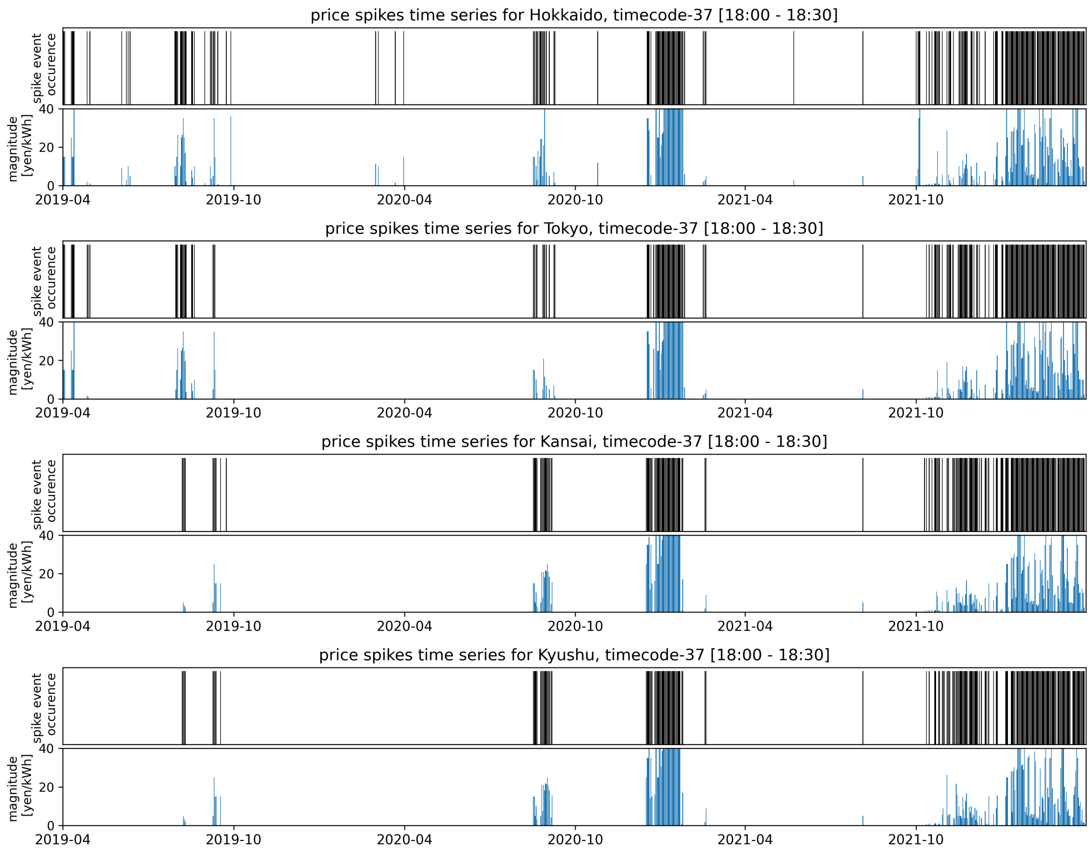

Figure 3.7: Bar plots of price spikes for different areas for the time-slot 16:30 17:00. Plots show the occurrence or non-occurrence of events (spikes) and corresponding magnitudes

is the price spikes time series. This time series will have a presence / absence component and a magnitude component. The temporal evolution of price spikes for different trading areas is illustrated in the bar plots of Figure 3.7 The barcode plot (top plot) illustrates the presence or absence of spikes while the bottom plot shows the magnitudes of the spikes.The plots of Figure 3.7 highlight the following characteristics of the price spikes time series data:

1. There is a tendency for the spikes to appear in clusters
2. Spike occurrence is not dependent on climatic season
3. Spike occurrence is dependent on both area and time-code

The price spikes are modelled as inhomogeneous Poisson processes with varying intensities. More specifically, the occurrence of a price spike increases the probability of more spikes occurring in the period immediately following the spike.

3.3.2 Price Spikes Modelling - Hawkes Process

Certain natural events e.g., earthquakes display the characteristics of clustering in time i.e., the occurrence of an event increases the chance of recurrence in the period immediately following the event resulting in a sequence of data with temporally varying intensities. The Hawkes process [2] is suitable for modelling such events and has been used in literature to model financial markets [10], disease occurrences [11] and social media interactions [12].

The Hawkes model is typically used to model "self-exciting" point processes i.e. a process in which an arrival increases the rate of future arrivals for some period of time. The original Hawkes process describes a self-excitation process using a conditional intensity function $\lambda_d^{a,t}$ given by:

$$\lambda_d^{a,t} = \mu^{a,t} + \sum_{d_i=0}^{d} \phi^{a,t}(d - d_i) \tag{3.1}$$

where $\mu^{a,t} \geq 0$ is the base intensity for area a and time-slot t while the second term on the right-hand-side of (4.4) is the self-excitation component modeling the influence of past events on the current value of the excitation function and therefore governing the clustering of the point process. As is typical in literature, an exponential decay of the form:

$$\phi^{a,t}(d-d_i) = u_{d_i}^{a,t} \cdot \gamma^{a,t} \cdot e^{-(d-d_i)/\tau^{a,t}} \tag{3.2}$$

is used as the excitation function where $\gamma^{a,t}$ and $\tau^{a,t}$ are model parameters. $\gamma^{a,t}$ governs the magnitude of the increase in the intensity function whenever there is a spike i.e. $u_d^{a,t} = 1$, while $\tau^{a,t}$ is the time constant for the decay of the effect of the spike thereby giving information on the influence of the price spike on future event occurrences. The value of the intensity function $\lambda_d^{a,t}$ defines the probability of occurrence of a spike in area a at time period t on day d. The price spikes time series can therefore be thought of as stemming from a parameter varying Poisson process, where $u_d^{a,t}$ is a Bernoulli's random variable occurring with the probability $\lambda_d^{a,t} \leq 1$ (conversely, not occurring with the probability $1 - \lambda_d^{a,t}$).

3.3.3 Evaluation of Price Spike Occurrence Intensity

Extracted Parameters

Using a Bayesian approach, the Hawkes model parameters were extracted for various areas and timecode combinations. Tables 3.1 to 3.4 shows some of the values of the extracted parameters. The value $\mu^{a,t}$ represents the probability of a spontaneous spike and it can be seen to be very low. The values can get as low as 0.002 for Kyushu area at timecode 1. Even at the highest value, the value of $\mu^{a,t}$ is only 0.0196 for Hokkaido at timecode 37. These results show that the probability of a spike is very low and spike events are very rare. The parameter $\tau^{a,t}$ is the time constant of the persistence of the effect of spike events in days. These values range from 1.75 to 3.56 days showing that the spike events, even when they occur, die out very quickly. Finally, the parameter $\gamma^{a,t}$ represents the jump in spike intensity function after the occurrence of a spike event. These values are found to range from 0.22 to 0.37. This means that a two consecutive spike events are required for the spike occurrence probability to increase above 0.5. In general, these parameters show that price spike events are very rare in occurrence, occurrence of an event increases the value of the intensity function only modestly, and the effect of the spike event occurrence dies out very quickly.

Table 3.1: Extracted Hawkes model parameters for Hokkaido area.

timecode	period	$\mu^{a,t}$	$\tau^{a,t}$	$\gamma^{a,t}$
1	00:00 - 00:30	0.0060	2.86	0.256
9	04:00 - 04:30	0.0052	1.99	0.332
15	07:00 - 07:30	0.0060	2.86	0.266
25	12:00 - 12:30	0.0060	2.98	0.228
37	18:00 - 18:30	0.0196	3.4	0.231
41	20:00 - 20:30	0.0124	3.56	0.221

Table 3.2: Extracted Hawkes model parameters for Tokyo area.

timecode	period	$\mu^{a,t}$	$\tau^{a,t}$	$\gamma^{a,t}$
1	00:00 - 00:30	0.0044	2.36	0.301
9	04:00 - 04:30	0.0052	1.99	0.327
15	07:00 - 07:30	0.0028	3.25	0.254
25	12:00 - 12:30	0.0084	2.86	0.224
37	18:00 - 18:30	0.0100	2.86	0.283
41	20:00 - 20:30	0.0044	3.11	0.247

Table 3.3: Extracted Hawkes model parameters for Kansai area.

timecode	period	$\mu^{a,t}$	$\tau^{a,t}$	$\gamma^{a,t}$
1	00:00 - 00:30	0.0028	2.54	0.279
9	04:00 - 04:30	0.0044	3.25	0.232
15	07:00 - 07:30	0.0020	3.56	0.223
25	12:00 - 12:30	0.0044	2.54	0.225
37	18:00 - 18:30	0.0076	2.54	0.299
41	20:00 - 20:30	0.0036	3.11	0.251

Table 3.4: Extracted Hawkes model parameters for Kyushu area.

timecode	period	$\mu^{a,t}$	$\tau^{a,t}$	$\gamma^{a,t}$
1	00:00 - 00:30	0.0020	1.75	0.371
9	04:00 - 04:30	0.0036	1.87	0.325
15	07:00 - 07:30	0.0022	4.36	0.188
25	12:00 - 12:30	0.0044	2.64	0.225
37	18:00 - 18:30	0.0084	2.36	0.317
41	20:00 - 20:30	0.0028	2.45	0.304

Evolution of intensity function

Using the parameters extracted as shown in Tables 1 to 4, the spike occurrence intensity function was evaluated for various area and timecode combinations. The results are shown in Figures 1 to 4. The figures show that for all cases, the intensity function is nearly zero most of the time as spikes occur rarely. For the morning peak timeslot of 7:00 - 7:30 am, spike occurrence is rarer than the evening peak timeslot of 18:00 - 18:30. In all cases, the high intensity function that lasted for about a month around January 2021 stands out. During this period, prices consistently hit the 200 Yen/kWh price cap as a combination of severe weather and fuel shortages put a strain on the market. It is also interesting to note that the intensity function since October 2021 stayed consistently high. This coincided with the increased demand for crude oil as most economies begun opening up after ending COVID-19 related restrictions. The situation was exacerbated by the start of the Russia – Ukraine war in February 2022. The plots of Figures 1 to 4 show that the Hawkes model is able to capture the underlying dynamics generating the price spikes time series data.

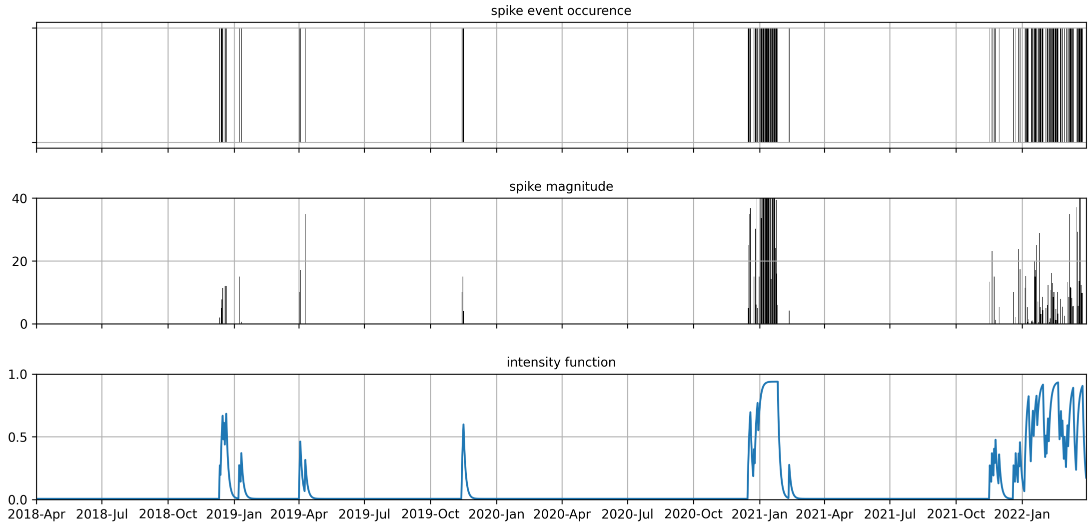

Figure 3.8: Evolution of price spikes time series and extracted intensity function for Hokkaido area and timeslot 07:00-07:30

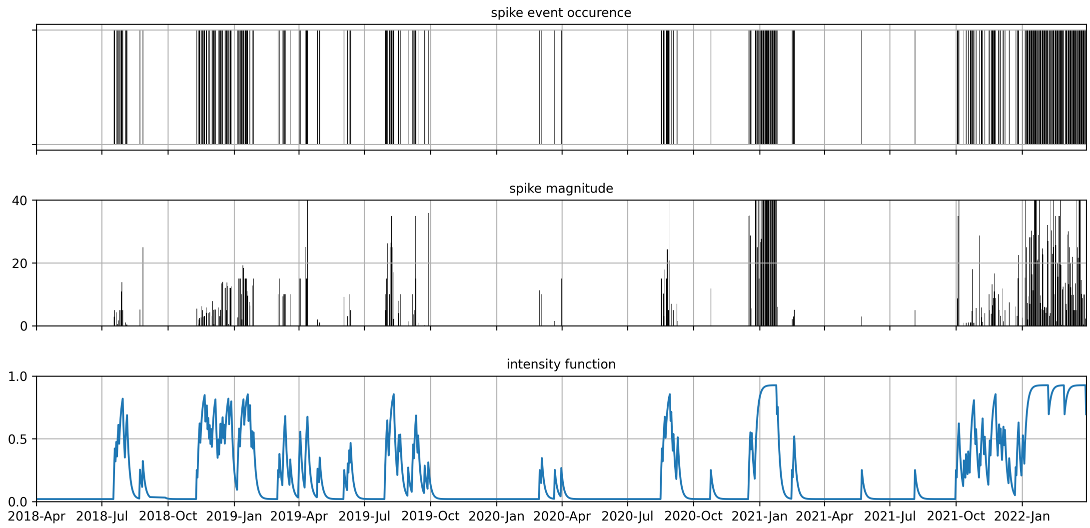

Figure 3.9: Evolution of price spikes time series and extracted intensity function for Hokkaido area and timeslot 18:00-18:30

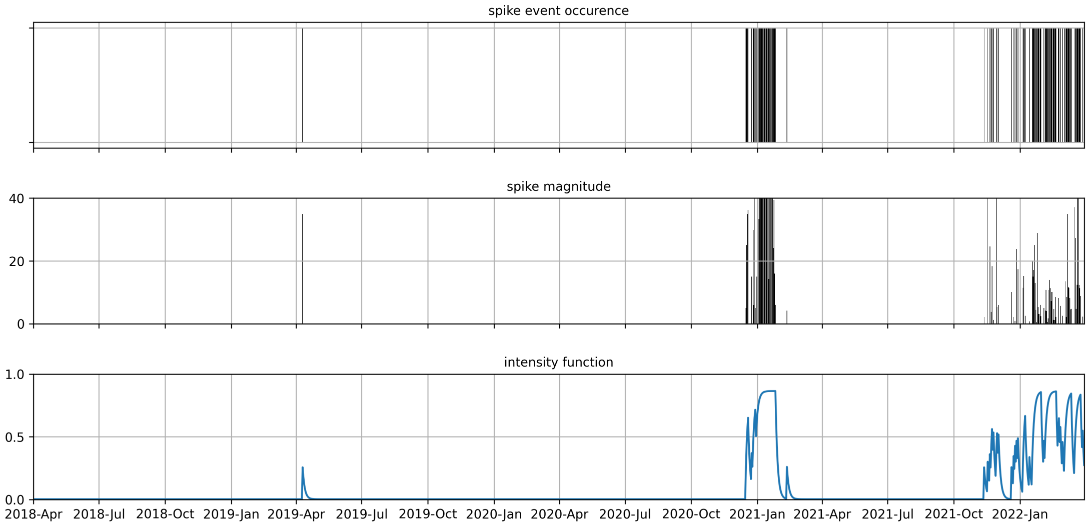

Figure 3.10: Evolution of price spikes time series and extracted intensity function for Tokyo area and timeslot 07:00-07:30

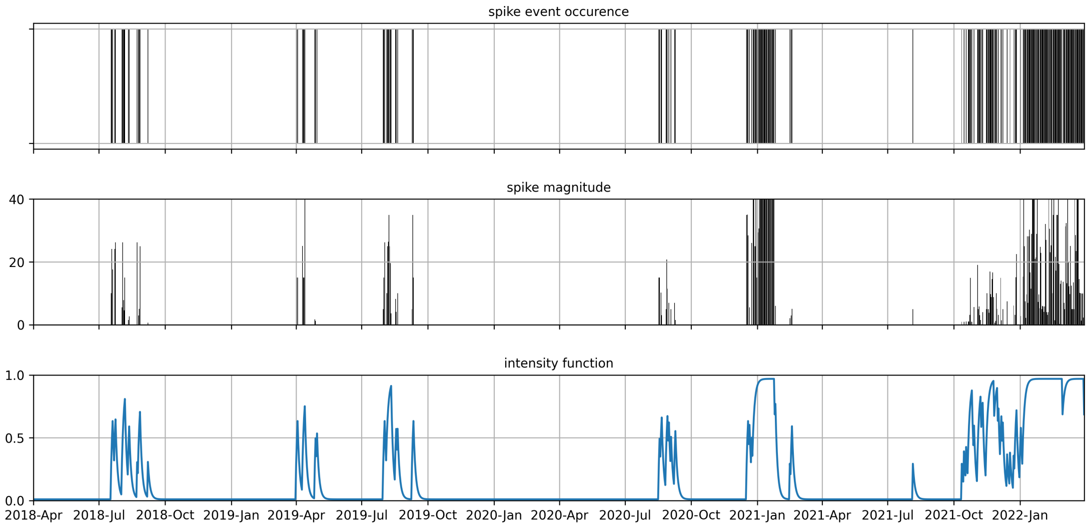

Figure 3.11: Evolution of price spikes time series and extracted intensity function for Tokyo area and timeslot 18:00-18:30

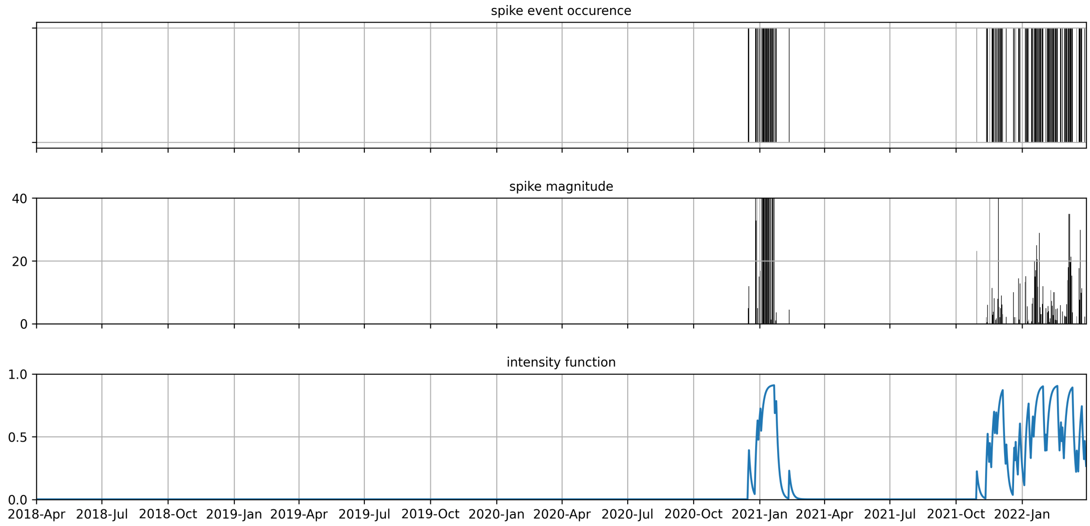

Figure 3.12: Evolution of price spikes time series and extracted intensity function for Kansai area and timeslot 07:00-07:30

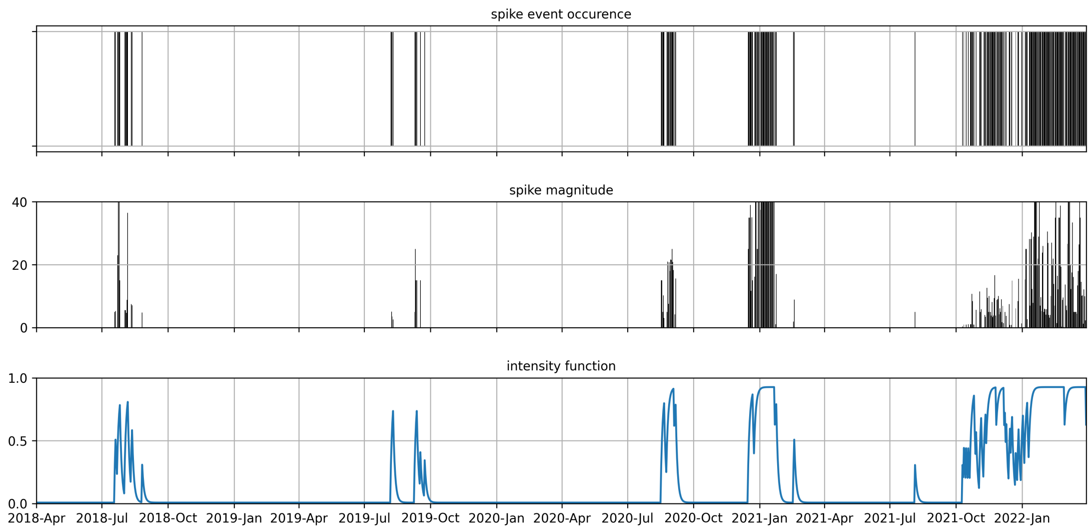

Figure 3.13: Evolution of price spikes time series and extracted intensity function for Kansai area and timeslot 18:00-18:30

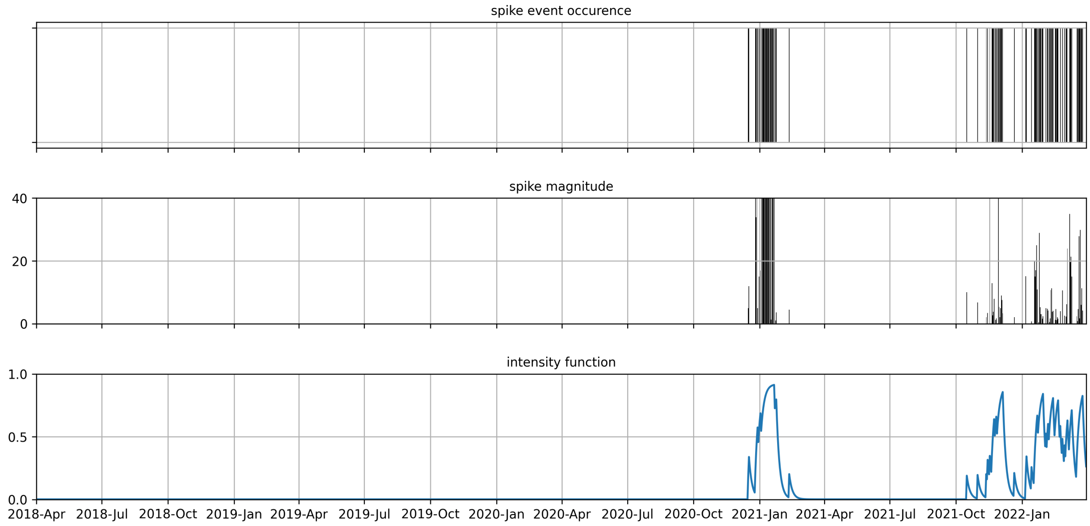

Figure 3.14: Evolution of price spikes time series and extracted intensity function for Kyushu area and timeslot 07:00-07:30

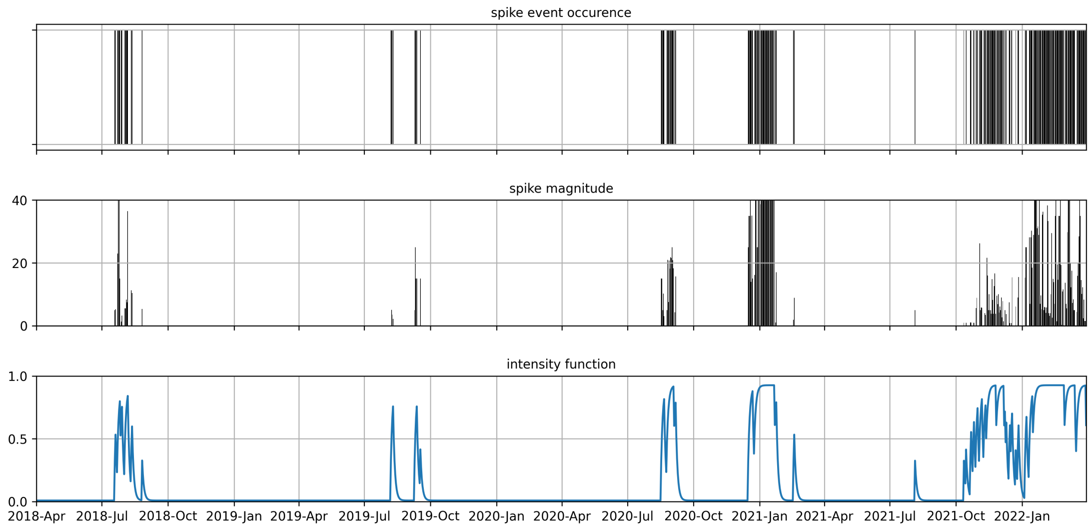

Figure 3.15: Evolution of price spikes time series and extracted intensity function for Kyushu area and timeslot 18:00-18:30

References

[1] "The electric power industry in japan 2022." https://www.jepic.or.jp/pub/pdf/epijJepic2022.pdf. Accessed: 2022-08-08.

[2] A. G. Hawkes and D. Oakes, "A cluster process representation of a self-exciting process," *Journal of Applied Probability*, vol. 11, no. 3, p. 493–503, 1974. doi: https://doi.org/10.2307/3212693.

[3] G. Gayretli, A. Yucekaya, and A. H. Bilge, "An analysis of price spikes and deviations in the deregulated turkish power market," *Energy Strategy Reviews*, vol. 26, p. 100376, 2019. doi: https://doi.org/10.1016/j.esr.2019.100376.

[4] X. Guan, Y.-C. Ho, and D. Pepyne, "Gaming and price spikes in electric power markets," *IEEE Transactions on Power Systems*, vol. 16, no. 3, pp. 402–408, 2001. doi: https://doi.org/10.1109/59.932275.

[5] J. H. Zhao, Z. Y. Dong, X. Li, and K. P. Wong, "A framework for electricity price spike analysis with advanced data mining methods," *IEEE Transactions on Power Systems*, vol. 22, no. 1, pp. 376–385, 2007. doi: https://doi.org/10.1109/TPWRS.2006.889139.

[6] D. H. Vu, K. M. Muttaqi, A. P. Agalgaonkar, and A. Bouzerdoum, "A multi-feature based approach incorporating variable thresholds for detecting price spikes in the national electricity market of australia," *IEEE Access*, vol. 9, pp. 13960–13969, 2021. doi: https://doi.org/10.1109/ACCESS.2021.3051313.

[7] N. Amjady and F. Keynia, "Electricity market price spike analysis by a hybrid data model and feature selection technique," *Electric Power Systems Research*, vol. 80, no. 3, pp. 318–327, 2010. doi: https://doi.org/10.1016/j.epsr.2009.09.015.

[8] "The japan electric power exchange website." http://www.jepx.org/english/index.html. Accessed: 2022-08-08.

[9] "The electric power industry in japan 2020." https://www.jepic.or.jp/pub/pdf/epijJepic2020.pdf. Accessed: 2022-08-08.

[10] P. J. Laub, T. Taimre, and P. K. Pollett, "Hawkes processes," 2015. doi: https://doi.org/10.48550/ARXIV.1507.02822.

[11] A. G. Hawkes, "Hawkes processes and their applications to finance: a review," *Quantitative Finance*, vol. 18, no. 2, pp. 193–198, 2018. doi: https://doi.org/10.1080/14697688.2017.1403131.

[12] M.-A. Rizoiu, Y. Lee, S. Mishra, and L. Xie, "A tutorial on hawkes processes for events in social media," 2017. doi: https://doi.org/10.48550/ARXIV.1708.06401.

4 A Hawkes Model Approach to Modeling Price Spikes in the Japanese Electricity Market

4.1 Introduction

4.1.1 Background

Liberalization of the electric energy sub-sector in many countries has allowed for electricity to be traded in financial markets in a manner similar to other commodity markets such as company stocks [1–3]. The day-ahead electricity market such as the Japanese Electric Power Exchange (JEPX) [4] provides a mechanism for power suppliers and customers to trade electricity in an environment where the price is dictated by the market forces of supply and demand. These markets were mainly introduced with the goal of improving service quality and curtailing monopolistic tendencies of the original regional or national utilities [4]. Trading occurs under the supervision of an independent market operator (MO) who works closely with the actual system operator (SO) who keeps the original role of ensuring high levels of system reliability [5]. The primary goal for market participants is to operate efficiently and economically in the electricity market which requires the design of proper strategies and tools based on power system requirements [6].

Prices in deregulated electricity markets are highly variable due to the dependence on several factors both on the energy supply and demand side [7]. Several factors including weather conditions, fuel costs, power plant operating costs, and regulations contribute to the increasing price uncertainty in the market [8]. On the supply side, the cost of fuel to run thermal generators is typically dependent on oil prices which is quite variable in itself. Furthermore, increase in power supplied by solar and wind based renewable energy generators escalates uncertainty in the availability of power since these sources depend on relatively uncertain primary sources of energy [9]. The management of these non-dispatchable energy resources introduces new significant challenges in the context of a competitive market environment [10]. However, the largest source of uncertainty is probably the strategies adopted by various companies during the bidding process as they jostle for a slice of the market share. On the consumer side, well documented and relatively predictable variations in power demand has an effect on prices [11]. Furthermore, in a two-sided market such as the JEPX, large power consumers or resellers participate in the bidding process with the objective of driving down prices to reduce their electricity bills. Certain aspects of the physical system such as limits in power that can be sent though certain power lines also has an effect on prices. These uncertainties mean that electricity prices are characterized by large spikes and dips which has an effect on the financial benefits that suppliers and consumers can gain from the market. The first step in mitigating against the volatility in the market is to set up an accurate price forecasting system. The obtained forecasts form an important basis for decision making by investors in the electricity market [12].

As with other commodity markets, electricity markets also experience periods of market shocks – a situation in which prices are driven much higher that normal due to extreme market conditions [13]. Price spikes can be the results of various issues in the market including power plant and system failures or extreme weather conditions that drive up demand [14]. Such a scenario was observed in the JEPX around late December 2020 / early January 2021 where prices hit a high of 220 Yen/kWh which is 25-to-30 times the average price of around 7-to-8 Yen/kWh. While the end consumers of may not have felt the effects of these spikes since most are in fixed tariff plans, a similar occurrence in the Texas' market ERCOT in the United States left households whose typical bills are 100 to 200 dollars with bills to the tune of tens of thousand of dollars. This further reinforces the need for mathematical models of the system that may be used to understand the underlying dynamics that lead to such extreme events.

4.1.2 Literature Review

Modeling of Electricity Prices

An overview of existing literature shows that numerous approaches have been proposed for the modelling and forecasting of short-term prices in electricity markets. References [15] and [16] give very good reviews on the topic. Similarly, a review specifically on probabilistic forecasting can be found in [17].

In [18], the authors provide a method for predicting next-day electricity prices using the autoregressive integrated moving average (ARIMA) methodology. A detailed analysis of the electricity prices time-series using the ARIMA models precedes model fitting and analysis based on the mainland Spain and Californian markets. Classical autoregressive integrated moving average (ARIMA) models with various AR and MA orders are also tested in [19]. Simulation results using data from the UK electricity market selects the ARIMA (4,1,2) model as the best and is found to achieve better results than persistence or a typical artificial neural network (ANN) model. A seasonal auto-regressive integrated moving average model with exogenous variables (SARIMAX) for modelling electricity prices is proposed in [20]. The model is chosen so as to capture the seasonal variations of electricity prices. In [21] electricity price behavior in the Nordic electric power market is forecasted using both the Markov-switching generalized autoregressive conditional heteroscedasticity (MS-GARCH) model and a set of different volatility models. The GARCH models aim to model the conditional volatility of the price time series. The study demonstrated that electricity price volatility is not only highly volatile but also strongly regime-dependent.

Given the highly non-linear nature of the electricity price time series data, several authors have presented approaches based on newer machine learning models. An exhaustive analysis of existing forecasting models is carried out in [22] where 27 common approaches are analyzed with a conclusion that generally, deep learning models outperform statistical models. They also conclude that hybrid models do not outperform simpler counterparts. A Recurrent Neural Network (RNN) based Model is proposed in [23] where multi-layer Gated Recurrent Units are proposed for electricity price forecasting. The use of Dynamic Trees for both very short and short term electricity price forecasting and the improvement in forecasting performance when compared to a typical Random Forest approach is illustrated in [24] with tests on the Iberian market. Reference [25] employed an Artificial Neural Networks (ANN) model with a focus on the selection and preparation of fundamental data that had noticeable impact on electricity prices.

In [26] a hybrid modelling approach that combines the features and strengths of the auto-regressive fractionally integrated moving average (ARFIMA) model and the least-squares support vector machine (SVM) model is proposed. Similarly, a long short-term memory (LSTM) deep neural networks combined with feature selection algorithms for electricity price prediction under the consideration of market coupling is proposed in [27]. An outlier-robust hybrid model for forecasting electricity prices that combines a simple outlier-robust extreme learning machine model and several other algorithms is developed in [28]. Their experiments based on electricity price data from the Australia and Singapore markets demonstrate the effectiveness of the model especially in dealing with the complex nonlinear characteristics and numerous outliers present in the price data.

Electricity Price Spikes Modelling

While there has been a lot of focus on the modeling and forecasting of electricity prices, there is less literature on modelling price spikes i.e. the occurrence of extreme prices which has a significant effect on the operation of market agents. An economic analysis of price spikes is presented in [13] where the authors investigate the factors and mechanisms determining spikes in the Italian electricity market. Based on the market data, they performed a specific analysis of the auctions mechanisms and of the hourly bid and offer of electricity characteristics to determine how and why price spikes occur. Their results showed that rigidity, which characterizes both the demand curve and part of the supply curve, is the fundamental determinant of prices pikes. Fluctuations in renewable energy sources (RES) production also proved to be fundamental.

In [29], a stochastic regime-switching model with time-varying parameters is shown to capture the type of volatile price behavior observed in many deregulated spot markets for electricity. The mean prices in two price regimes and the transition probabilities were specified as functions of the offered reserve margin and the system load. The high-price regime corresponded to the observed price spikes that typically occur during the summer months. In addition, the structure of the model was consistent with the actual hockey stick shape of the offers submitted by suppliers into the PJM market. Most capacity is offered at relatively low prices, and a few units are offered at much higher prices up to the price cap.

Dramatic rises in electricity prices can be observed during periods of market stress as highlighted in [30]. The authors treat abnormal episodes or price spikes as count events and propose to build a model of the spiking process. The importance of persistence in the spiking process and its significance in building an effective model is highlighted. They adapt a Poisson autoregressive framework for integer-valued time series that accounts for the number of simultaneous stresses remaining latent and provided a model that could be estimated by maximum likelihood. The arrival and survival rates of price spikes were found to be dependent upon extreme temperature events and peak load. However, the model's ability to capture the intrinsic persistence in price spikes was cited as more significant.

The forecasting of extreme price events, the occurrence of which is treated as a realization of a discrete time point process, is the focus of [31]. An Autoregressive Conditional Hazard (ACH) framework was used to analyze the drivers of the process and to forecast the probability of extreme price events occurring in real time. Abnormal loads were found to have a significant impact on the probability of a price spike and on the severity of the spike. Importantly, stochastic factors capturing the history of the process were found to be significant in explaining the occurrence of extreme price events.

An argument that there is increasing empirical evidence of increased price volatility and spikes in electricity markets as a result of fluctuating renewable energy production, extreme weather events and other factors is presented in [32]. While price spikes are necessary to cover the fixed costs of power plants, they can also indicate market imperfections and anti-competitive behavior. Regulators have set market price caps to protect consumers and prevent abusive behavior by vendors. Additionally, some regulators have imposed temporary price caps during or after major events. In weak institutional frameworks, however, these ceilings may be driven by political motives rather than economic logic.

4.1.3 Chapter Organization

The remainder of this chapter is organized as follows: Section 4.2 gives an introduction to the Japanese Electricity market including the characteristics of the electricity grid, a look at the prices datasets and the definition of price spikes. Section 4.3 presents the modeling approach introducing the proposed models, the parameter extraction approach and the generation of short term price spikes forecasts. Simulation results based on data from the JEPX are presented and discussed in section 4.4 and the study conclusions are drawn in section 4.5.

4.2 The Japanese Electricity Market

4.2.1 Introduction to the Spot Market and the Power Grid

Following the trend toward deregulation in the electric power industry in Western countries, the liberalization of the electricity generation sector started in 1995 in Japan, followed by retail supply liberalization for customers receiving extra high-voltage (20 kV or above) in 2000 [33]. The scope of deregulation was expanded in different stages afterwards. However, power shortages and other issues caused by the 2011 Great East Japan Earthquake prompted discussion about the ideal structure of the country's electric power system and its reform. Based on this, full liberalization of the Japanese electricity market was achieved in 2016. The Japan Electric Power Exchange (JEPX) was established in November 2003 and begun trading in April 2005. The purpose of JEPX is to handle electricity transactions on the exchange. This research focuses on forecasting prices spikes in the JEPX day ahead market.

The general structure of the Japanese electricity market follows closely those of more established markets such as the PJM (Pennsylvania, Jersey, and Maryland) Power Pool interconnection in the US [34] and the European Nord Pool covering the Northern European countries, such as Norway, Sweden, Denmark, and Germany [35]. While most of the energy is traded in the day-ahead market, there is an intra-day market for settlement of hour-ahead forecasted load demand. However, unlike the markets in the United States, a real-time market [36] is yet to be implemented. Market clearing is carried out on a 30-minute time resolution unlike the Italian or Spanish markets where clearing is done on a 1-hour time resolution. The Japanese market also uses a zonal marginal pricing policy with nine trading areas in a structure similar to most European markets but unlike the PJM which uses a Locational Marginal Pricing (LMP) policy [37]. Similar to the Swedish, Spanish, Italian and most European markets, the market operator (MO) in the Japanese market is separate from the system operator (SO) but unlike the Australian, PJM and UK markets where market and system operation functions are carried out by the same entity. Apart from the market structure, the Japanese system is unique in having two system frequencies (50Hz in the east and 60Hz in the west) [4] within the same market, sometimes leading to significant differences in prices within the market. In addition, there is no international connection unlike the highly connected European markets which means that localized mismatches in supply and demand cannot be offset by imports or exports from or to nearby grids.

There are 10 operational areas (9 in the main island of Honshu and the Okinawa area that covers the southern islands) as shown in Figure 4.1 [33]. These areas correspond to the regions originally operated by the main power utilities before system deregulation are currently operated by separate system operators. The JEPX handles transactions for the 9 main areas on a thirty-minute time resolution resulting in 48-trading periods per area per day on the day-ahead electricity market.

The transmission capacity limits in the connections between areas. These limitations result in transmission congestion hence differences in prices between areas. Particularly the HVDC interconnection linking Tokyo and Chubu areas result in differences between prices in the Eastern grid operated at 50Hz and Western grid operated at 60 Hz. The overall structure of the electricity market is as shown in Figure 4.2 and the main market participants are the electricity generation companies and electricity retailers involved in wholesale power transactions. The number of participating generators in the market are 986 and a total of 730 retailers as of September 2022 [38].

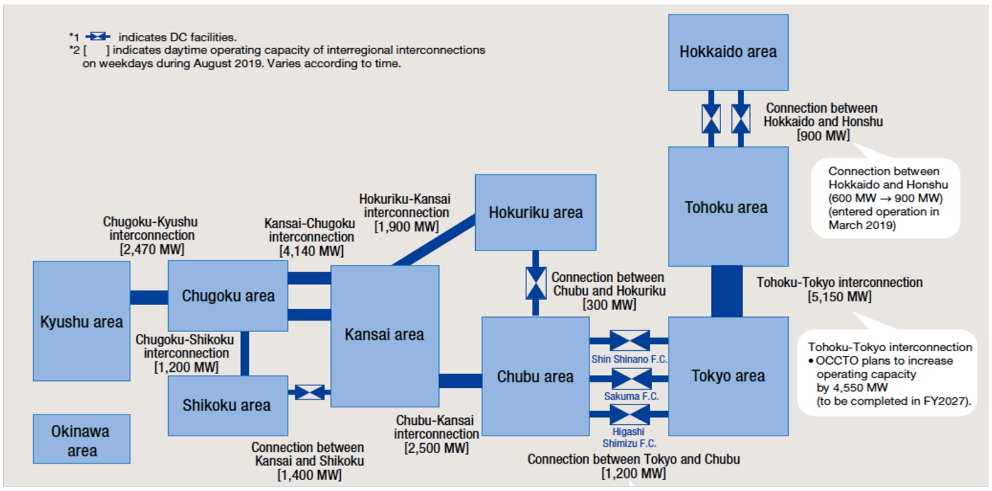

Figure 4.1: Illustration of the trading areas in the Japanese electricity market and the physical interconnections. Source: The Japan Electric Power Information Center (JEPIC) [33].

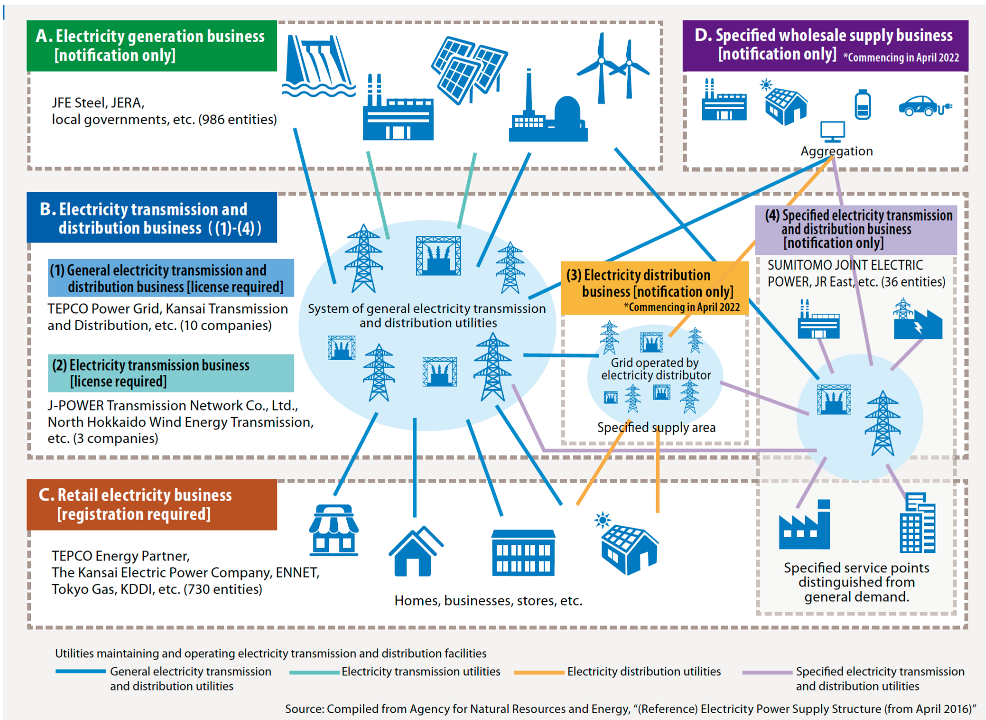

Figure 4.2: General organization of the Japanese electricity market. Source: The Japan Electric Power Information Center (JEPIC) [38].

4.2.2 JEPX Dataset

The JEPX electricity price and traded energy volumes data are publicly available on the exchange's website [39]. The available data includes traded volumes and corresponding system and area market clearing prices (MCPs) in 30-minute resolutions giving 48 commodities per area, per day. Table 4.1 shows the summary statistics of the MCP data for each of the nine areas in the JEPX for the period spanning April 1, 2016 to March 31, 2022. The values clearly indicate differences in the grid with prices in the eastern areas of Hokkaido, Tohoku, and Tokyo being generally higher than prices in the Western grid. In addition, prices in the Northern island of Hokkaido are generally highest while prices in the south western region of Kyushu being the lowest. Prices in the western regions of Chubu, Hokuriku, Kansai, Chugoku, and Shikoku are very similar indicating adequate transmission capacity between these areas leading to almost always a single MCP between them. Table 4.1 also shows the 95th percentile value of the prices time series. These values give a sense of what would be considered as extreme prices in the market.

Table 4.1: Summary statistics of area prices (Yen/kWh) in the JEPX for the period April 1, 2016 to March 31, 2022.

Area	**mean**	**median**	**std. dev**	**skewness**	**95th percentile**
Hokkaido	12.72	10.54	12.01	8.90	25.67
Tohoku	10.82	8.67	11.75	9.83	23.00
Tokyo	10.93	8.71	11.85	9.67	23.37
Chubu	9.91	7.66	10.96	9.48	21.87
Hokuriku	9.89	7.66	10.94	9.53	21.54
Kansai	9.88	7.66	10.93	9.57	21.24
Chugoku	9.87	7.65	10.93	9.57	21.24
Shikoku	9.86	7.65	10.95	9.61	21.22
Kyushu	9.16	7.36	10.69	10.21	19.45

A fundamental analysis of the JEPX spot-market time series data shows that prices have a clear spatial and temporal dependence. This dependence is illustrated in Figure 4.3 which shows the average area MCPs (calculated on a thirty-minute resolution) for the period spanning April 1, 2016 to March 31, 2022 covering the first six years since the start of full market liberalization that allowed for competitive retail of electricity to individual consumers. The data is grouped to show the average prices for both workdays and non-workdays (weekends and holidays).

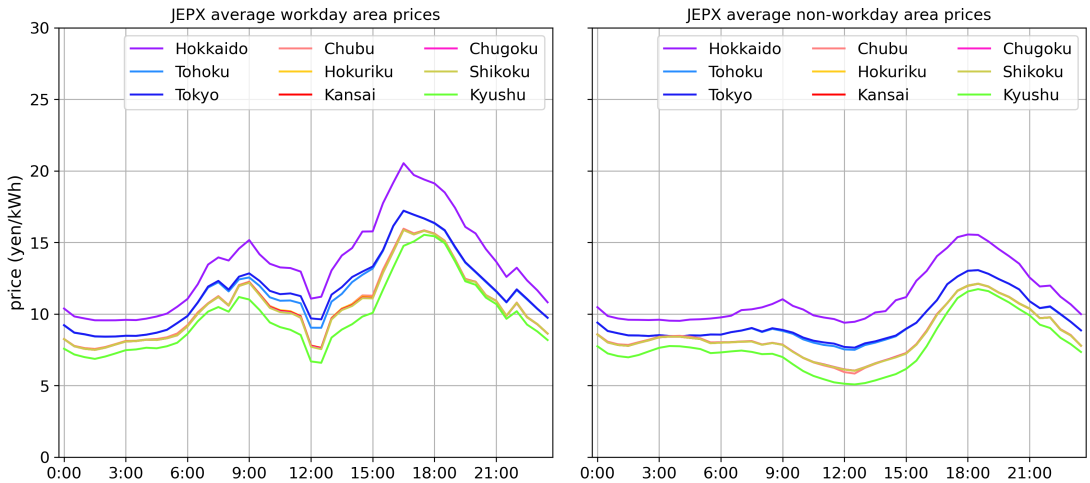

Figure 4.3: Average area prices for the JEPX for the period spanning April 1, 2016 to March 31, 2022.

On the spatial axis, prices are seen to be different for different areas. Generally, prices are observed to be higher in the eastern regions of Hokkaido, Tohoku and Tokyo and lower on the western side. The south-western region of Kyushu exhibits especially relatively lower prices. This dependence is explained by constraints on the amount of power transferable through the interconnections between regions. Capacity limits of the HVDC interconnection linking the Tokyo and Chubu areas especially results in significant differences between prices in the eastern grid operated at a frequency of 50-Hz and the western grid operated at 60-Hz.

Temporally, MCPs are dependent on the type of day i.e. workday or non-work day; and the time-of-day i.e. there are clear peak and off-peak periods within a day. This temporal dependence is explained by the socio-economic behavior of electricity users with morning and evening peaks. Reduced energy intensive activities on weekends and holidays results in generally reduced demand and consequently, lower prices during these days as compared to typical work days.

Considering the above characteristics of JEPX price data, the use of individual models for each area and each trading time period is proposed to handle the spatial dependence and time-of use dependence respectively. In addition, a transformation approach is proposed to handle the type of day dependence. The difference in prices due to type of day is handled by transforming the weekend/holiday prices $y_d^{a,t\prime}$ to equivalent weekday prices $y_d^{a,t}$ using scaling factors $k_d^{a,t}$ as:

$$y_d^{a,t} = k_d^{a,t} \times y_d^{a,t\prime} \tag{4.1}$$

where $k_d^{a,t}$ is the ratio of the average workday price to the average non-workday price up to day d for area a and trading period t.

Figure 4.4 shows the effect of the price transformation technique on the empirical cumulative distribution curves (cdfs) of the MCPs for Tokyo area. In the original form of Figure 4.4(a), the MCPs for workdays and non-workdays can be thought of as belonging to different probability distributions which would require

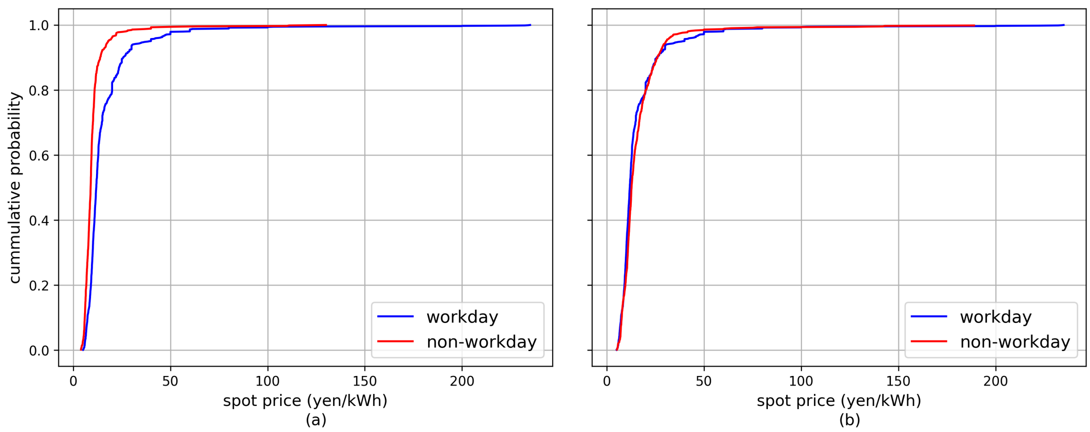

Figure 4.4: Empirical cumulative distribution curves (cdfs) for the Tokyo area MCPs (trading period 33: 16:00-16:30) (a) without the application of transformation of equation (1) and (b) with the application of the transformation of equation (4.1).

a regime switching type model to properly handle this characteristic. However, as seen from Figure 4.4(b), the transformation of equation (4.1) “shifts” the distribution of the non-workday prices to be similar to that of the workday prices allowing for the fitting of a single model and avoiding the breaking of the time series to workday and non-workday portions. The models are then fit on the transformed time series and carry out the inverse of the above transformation when determining the magnitude of the forecasted price spikes.

4.2.3 Definition of Price Spikes

Extreme prices in electricity markets can have devastating impacts on market participants even if they last for just a short period of time. Cases of electricity resellers that have been driven to bankruptcy due to these "price spikes" have been observed in the last few years even in more established markets in the US, UK and Europe. In an environment of tight margins, even with a conservative hedging strategy, one can be left exposed to extreme price risks especially since spikes tend to occur during peak demand periods where retailers will be typically under-hedged.

A price spike on day d and in area a and trading period t, $u_d^{a,t}$, is defined as an observed price value $y_d^{a,t}$ above a pre-defined threshold y^δ. Mathematically, this is represented as:

$$u_d^{a,t} = \begin{cases} 0, & \text{if } y_d^{a,t} \le y^\delta \\ 1, & \text{otherwise} \end{cases} . \tag{4.2}$$

For the electricity retailer procuring energy from the spot market, this price threshold would define a risk value above which the potential of loss becomes significant.

Figure 4.5 shows the price ranges (10th to 90th percentile) of area spot prices in the JEPX for the data spanning April 1, 2016 to March 31, 2022. Here, only four representative areas are shown since the price characteristics in several areas are quite similar as seen from the data in Table 4.1 and Figure 4.3. Apart from Hokkaido area in which prices were a little higher, the peak of the 90th percentile curve is about 25 Yen/kWh and prices above this value can be considered extreme. In fact, in the entire dataset, only 3.7% of the area prices are greater than this

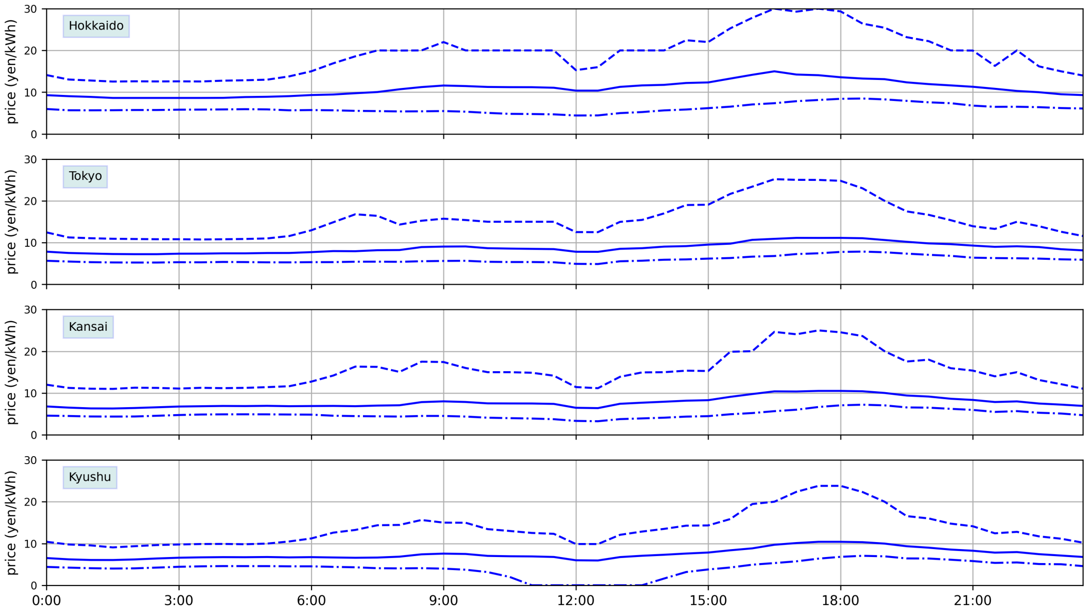

Figure 4.5: The 10th percentile (dot-dashed line), median (full line) and the 90th percentile (dashed line) for the JEPX area prices time series.

threshold. A value of 25 Yen/kWh is therefore adopted as a consistent price spike threshold throughout this study, though the analysis and results would be similar with any other reasonable threshold value.

The series of data points showing prices above the threshold value is the price spikes time series. This time series has two components: (1) the occurrence or non-occurrence of a spike and (2) the magnitude of a spike should it occur. The temporal evolution of price spikes for different trading areas and selected trading periods is illustrated in the plots of Figure 4.6. Each price spikes time series is presented as two plots: the barcode plot (top plot) illustrates the occurrence or non-occurrence of spikes while the bottom plot shows the corresponding spike magnitudes. From the plots of Figure 4.6, it can be observed that there is a tendency for spikes to occur in clusters i.e. there are specific periods within the time series where the probability of spike occurrence is clearly higher than others. This characteristic can be explained by the underlying process driving market prices i.e. the balance between supply and demand. Periods of high prices i.e. continuously occurring spikes, are usually the result of a short term imbalance

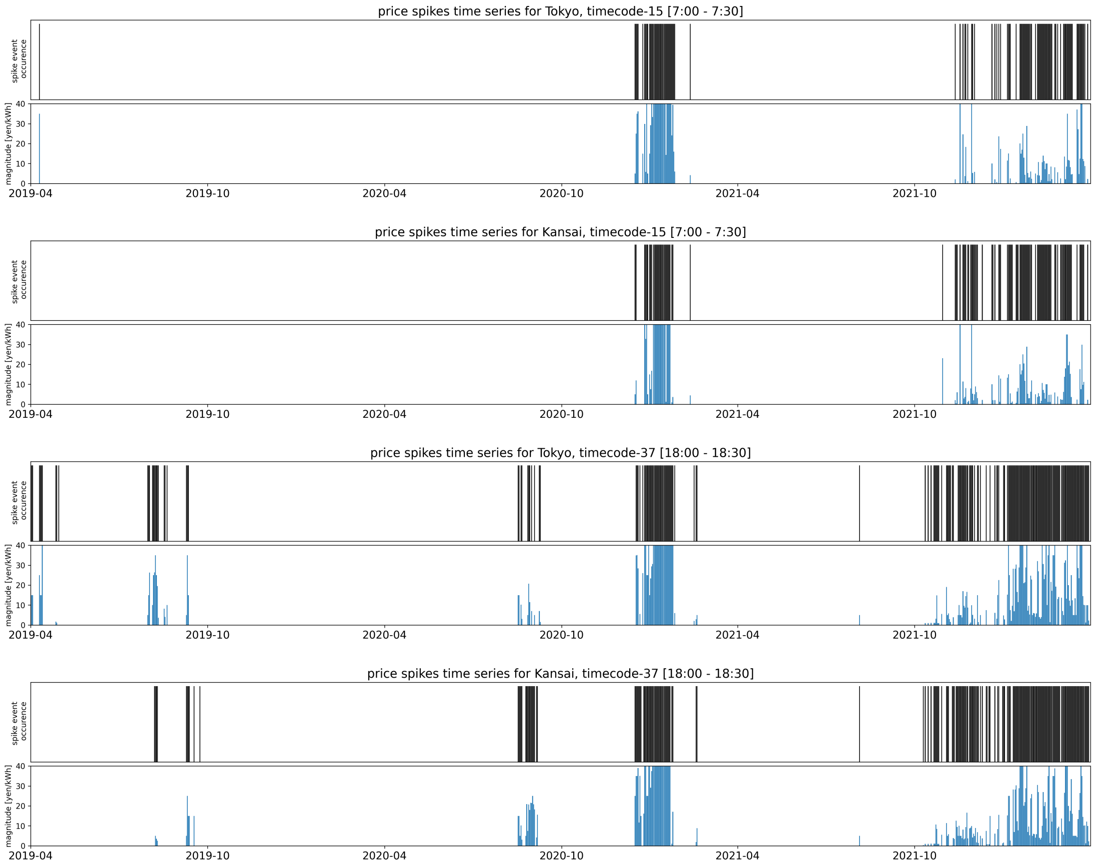

Figure 4.6: Temporal evolution of price spikes time series for selected areas and trading time periods. The vertical axis of the magnitude plots is truncated at 40 Yen/kWh though much higher spike magnitudes have also been observed. The axis truncation achieves better clarity of the plots.

between supply and demand e.g. due to fuel shortages or unforeseen shutdowns of large generators.

4.3 Methodology

4.3.1 Notation

Indexes a, t, and d are used to denote the area, trading time period and day respectively so that $y_d^{a,t}$ denotes the day-ahead MCP for area a time period t and day d. The price spike threshold is defined as y^δ so that a spike occurs in area a, trading period t and day d if $y_d^{a,t} > y^\delta$ and vice-versa. The occurrence (presence or absence) of a spike is then given by a binary variable $u_d^{a,t}$ as shown in equation (4.2). Similarly, the magnitude of a spike $x_d^{a,t}$ is given by:

$$x_d^{a,t} = \begin{cases} 0, & \text{if } u_d^{a,t} = 0 \\ y_d^{a,t} - y^\delta, & \text{if } u_d^{a,t} = 1 \end{cases} \tag{4.3}$$

4.3.2 Hawkes Model

The Hawkes model is typically used to model "self exciting" point processes i.e. a process in which an arrival of an event increases the rate of future arrivals for some period of time [40, 41]. In the case of the electricity price data, sudden "jumps" (spikes) in prices are modeled as excitation signals that increase (or reduce) the price for some period of time after occurrence.

Fundamentally, the use of equations proposed by Hawkes is proposed to model the price spikes time series data [42]. The Hawkes model defines the occurrence of an event in terms of an intensity function given by:

$$\lambda_d^{a,t} = \mu^{a,t} + \sum_{d_i=0}^{d} \phi^{a,t}(d - d_i) \tag{4.4}$$

where $\mu^{a,t} \geq 0$ is the base intensity for area a and time-slot t while the second term on the right-hand-side of (4.4) is the self-excitation component modeling the influence of past events on the current value of the excitation function and therefore governing the clustering of the point process. As is typical in literature, an exponential decay of the form:

$$\phi^{a,t}(d - d_i) = u_{d_i}^{a,t} \cdot \gamma^{a,t} \cdot e^{-(d-d_i)/\tau^{a,t}} \tag{4.5}$$

is used as the excitation function where $\gamma^{a,t}$ and $\tau^{a,t}$ are model parameters. $\gamma^{a,t}$ governs the magnitude of the increase in the intensity function whenever there is a spike i.e. $u_d^{a,t} = 1$, while $\tau^{a,t}$ is the time constant for the decay of the effect of the spike thereby giving information on the influence of the price spike on future event occurrences. The value of the intensity function $\lambda_d^{a,t}$ defines the probability of occurrence of a spike in area a at time period t on day d. The price spikes time series can therefore be thought of as stemming from a parameter varying Poisson process, where $u_d^{a,t}$ is a Bernoulli's random variable occurring with the probability $\lambda_d^{a,t} \leq 1$ (conversely, not occurring with the probability $1 - \lambda_d^{a,t}$).

Replacing (4.5) in (4.4), the Hawkes model can be re-written as:

$$\lambda_d^{a,t} = \mu^{a,t} + \sum_{d_i=0}^{d} u_{d_i}^{a,t} \cdot \gamma^{a,t} \cdot e^{-(d-d_i)/\tau^{a,t}} \tag{4.6}$$

$$= \mu^{a,t} + \sum_{d_i=0}^{d-1} u_{d_i}^{a,t} \cdot \gamma^{a,t} \cdot e^{-(d-d_i)/\tau^{a,t}} + \gamma^{a,t} u_d^{a,t} \tag{4.7}$$

$$= \mu^{a,t} + e^{-1/\tau^{a,t}} \sum_{d_i=0}^{d-1} u_{d_i}^{a,t} \cdot \gamma^{a,t} \cdot e^{-(d-d_i-1)/\tau^{a,t}} + \gamma^{a,t} u_d^{a,t} \tag{4.8}$$

$$= \mu^{a,t} + \alpha^{a,t}\left(\lambda_{d-1}^{a,t} - \mu^{a,t}\right) + \gamma^{a,t} u_d^{a,t} \tag{4.9}$$

$$= \alpha^{a,t}\lambda_{d-1}^{a,t} + \beta^{a,t} + \gamma^{a,t} u_d^{a,t} \tag{4.10}$$

where $\alpha^{a,t} = e^{-1/\tau^{a,t}}$ and $\beta^{a,t} = (1 - \alpha^{a,t})\mu^{a,t}$.

The above form of the Hawkes model reveals a structure similar to a typical auto-regressive model with $\beta^{a,t}$ as the constant term, $\alpha^{a,t}$ as the auto-regressive coefficient and $\gamma^{a,t} u_d^{a,t}$ as the noise term. Re-writing the Hawkes model in this form reveals a more intuitive understanding of the model. In the absence of a spike, the intensity function drops back towards its natural value $\mu^{a,t}$ at a speed dictated by the coefficient $\alpha^{a,t}$ while when a spike event occurs, the intensity function experiences a jump governed by the value of $\gamma^{a,t}$

4.3.3 Modified Hawkes Model

In the original Hawke's equation (4.10) the parameter $\gamma^{a,t}$ that defines the increase in intensity after the occurrence of a spike is taken as a constant value implying that any observed price spike will lead to an increase in the spike's occurrence probability by the same magnitude. The price spikes time series however exhibits a correlation between the magnitude of the spikes and the number of subsequent spikes suggesting a dependence of the underlying intensity function on the the magnitude of observed spike events. Therefore, while testing the suitability of the original Hawke's model (4.10), denoted as *Hawkes-I*, on fitting the price spikes time series, the implementation of two variants of the basic model is also proposed.

The first of these, denoted as *Hawkes-II*, is a variable intensity jump model where the parameter $\gamma^{a,t}$ in (4.10) – which defines the increase in the magnitude of the intensity function given a spike – is dependent on the magnitude of the price spike i.e.:

$$\gamma_d^{a,t} = \gamma^{a,t,0}\left(1 - e^{-x_d^{a,t}/x_0^{a,t}}\right) \tag{4.11}$$

where $x_0^{a,t}$ is the expected value of $x^{a,t}$ given as the average magnitude of price spikes observed up to day d. With this formulation, a large magnitude spike – which typically indicates significant stress on the supply-demand balance in the market – will result in a relatively larger jump in the magnitude of the intensity function when compared to a spike of lower magnitude.

The second variation, denoted as *Hawkes-III*, is a variable effect decay speed model where the parameter $\tau^{a,t}$ - which defines the rate of decay of the intensity function given a spike - is dependent on the magnitude of the price spike i.e.:

$$\tau_d^{a,t} = \tau^{a,t,0}\left(1 - e^{-x_d^{a,t}/x_0^{a,t}}\right) \tag{4.12}$$

This formulation suggests that the effect of a large magnitude spike will last longer than that of a relatively smaller spike event.

4.3.4 Parameter Extraction

A Bayesian approach to estimating the Hawkes model parameters is taken in this study. The model parameters are treated as random variables for which posterior distributions are estimated based on the observations up to a given day. The joint posterior distribution of the model parameters for area a, and time slot t, given a set of observations up to day d, $p(\Theta_d^{a,t})$ is given by:

$$p(\Theta_d^{a,t}|X_d^{a,t}) = \frac{p(\Theta_{d-1}^{a,t}|X_{d-1}^{a,t}) \cdot p(x_d^{a,t}|\Theta_{d-1}^{a,t})}{\Delta} \tag{4.13}$$

where the prior $p(\Theta_{d-1}^{a,t}|X_{d-1}^{a,t})$ is the posterior distribution on day $d-1$ and $p(x_d^{a,t}|\Theta_{d-1}^{a,t})$ is the likelihood of the observation $x_d^{a,t}$ given the parameters on day $d-1$. Intuitively, the likelihood function is obtained from the Hawkes model as:

$$p(x_d^{a,t}|\Theta_{d-1}^{a,t}) = \begin{cases} 0, & x_d^{a,t} < 0 \\ 1-\lambda_d^{a,t}, & x_d^{a,t} = 0 \\ \lambda_d^{a,t}, & x_d^{a,t} > 0 \end{cases} \tag{4.14}$$

Starting from a uniform prior distribution on day 0, iterations through equations (4.13) and (4.14) are carried out to obtain the posterior distribution on day d.

4.3.5 Short-term Forecasting

Given the magnitude of the intensity function $\lambda_d^{a,t}$ on day d, and the model parameters $\alpha^{a,t}$, $\beta^{a,t}$ and $\gamma^{a,t}$, it's n-days ahead forecast $\hat{\lambda}_{d+n}^{a,t}$ is obtained by the iterating through the equation:

$$\hat{\lambda}_{d+k}^{a,t} = \alpha^{a,t}\hat{\lambda}_{d+k-1}^{a,t} + \beta^{a,t} + \gamma^{a,t}\hat{u}_{d+k}^{a,t} \tag{4.15}$$

for $k = 1, 2, 3, \ldots, n$ where $\hat{u}_{d+k}^{a,t} = \hat{\lambda}_{d+k-1}^{a,t}$ and $\hat{u}_{d+1}^{a,t} = \lambda_d^{a,t}$. Since the spike occurrence forecasting is essentially a binary classification problem, the binary forecast for the occurrence of a spike $\bar{u}_{d+n}^{a,t}$ is arrived at by comparing the forecast value $\hat{u}_{d+k}^{a,t}$ to a pre-defined threshold δ so that:

$$\bar{u}_{d+k}^{a,t} = \begin{cases} 0, & \text{if } \hat{u}_{d+k}^{a,t} \leq \delta \\ 1, & \text{otherwise} \end{cases} \tag{4.16}$$

The decision threshold δ adjusts the conservativeness of the forecasting model. As δ tends to zero, the model will forecast more 1's which will reduce false negative errors and vice versa.

4.4 Results

4.4.1 Data

The JEPX electricity market price data used in this study is publicly available on the exchange's website [39]. The value of 25 Yen/kWh is used as the price spike threshold though the results are similar for any other reasonable threshold values. The modelling and analysis is carried out using data for the 6-year period spanning the start of the fully deregulated market on April 1, 2016 to March 31, 2022 – the end of the 2021 Japanese financial year.

Figure 4.7 shows the spike occurrence probability for two selected study areas (Tokyo and Kansai) over the study period. The spike occurrence probability is calculated simply as the ratio of number of price spike events in the time series to the total number of events. Figure 4.7 shows that the spike occurrence probability is generally lower than 3% apart from during the evening peak periods where it rises to a maximum of around 6%. The probability is however clearly higher in the Tokyo region (in the 50-Hz eastern grid) than in Kansai (part of the 60-Hz western grid).

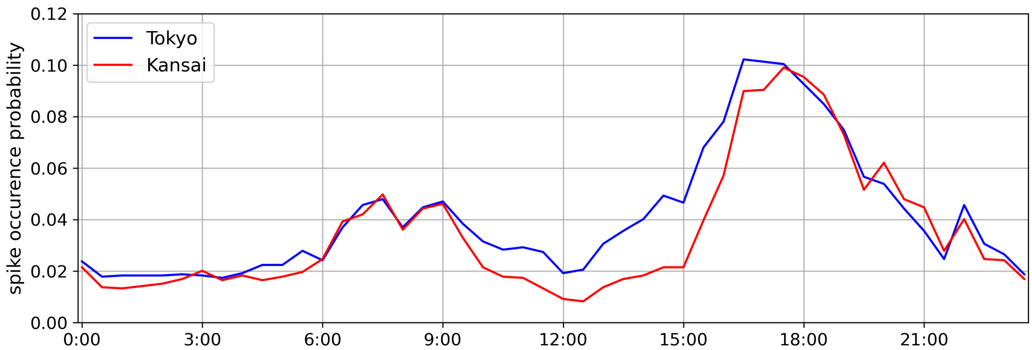

Figure 4.7: The variation of spike occurrence probability by trading time period for the two selected study areas (Tokyo and Kansai).

Generating individual models for each trading time period for each of the 9 areas results in 432 different models. Since it is impractical to display the results for all models here, four representative models are selected for detailed discussions. The selected models correspond to two representative areas – Tokyo for the Eastern grid and Kansai for the Western grid – and two trading time periods – 08:00-08:30 (timecode 15) representing the morning peak time period and – 18:00-18:30 (timecode 37) representing the evening peak time period. These selections give four area-timecode combinations referenced as Tokyo-15, Tokyo-37, Kansai-15 and Kansai-37 in the following discussions.

4.4.2 Baseline Model - Persistence Model

A persistence model is used to get baseline results for forecasting accuracy from which the performance of the proposed models are compared. The persistence model uses the prices on a given day d to forecast prices over the next N days under the assumption that the present day conditions persist over the forecasting horizon. For the price spikes time series, the algorithm states that if a spike occurs in area a, timecode t on day d, the spike is forecasted to persist over the next N days. Mathematically, this is written as:

$$\hat{x}_{d+i}^{a,t} = x_d^{a,t} \qquad \text{where } \ i = 1, 2, \ldots, N \tag{4.17}$$

The persistence model estimates the intensity function as the conditional probability of observing a price spike given the last observation. This is calculated as the ratio of the number of spike events that follow the last made observation $u_d^{a,t}$ for the spikes data observed up to day d. While quite simple in formulation, the persistence model can achieve very good performance for such data where correlations between consecutive observations are high and can set relatively high benchmarks for other more complex models under consideration.

4.4.3 Model Performance: Goodness of Fit

Following the Bayesian approach, parameters for the three versions of the Hawke's model were extracted. The model parameters are re-estimated daily for the entire dataset with the first two years of data used to generate the first set of parameter values. The daily updated model parameters are used to generate 14-days ahead forecasts for price spike events as described by (4.16) with a classification threshold of $\delta = 0.5$.

Using the extracted parameters, the intensity function evolution was evaluated for the dataset. Figure 4.8 shows the evolution of the Hawke's model intensity $\lambda_d^{a,t}$ for the four selected area time-code combinations. These plots show near zero-values during no-spikes periods highlighting rare nature of price spikes during "normal" market conditions. The plots also clearly highlight periods of increased stress in the system with high spike occurrence probabilities. An interesting observation is the prolonged high intensity period that lasts from October 2021 to March 2022.

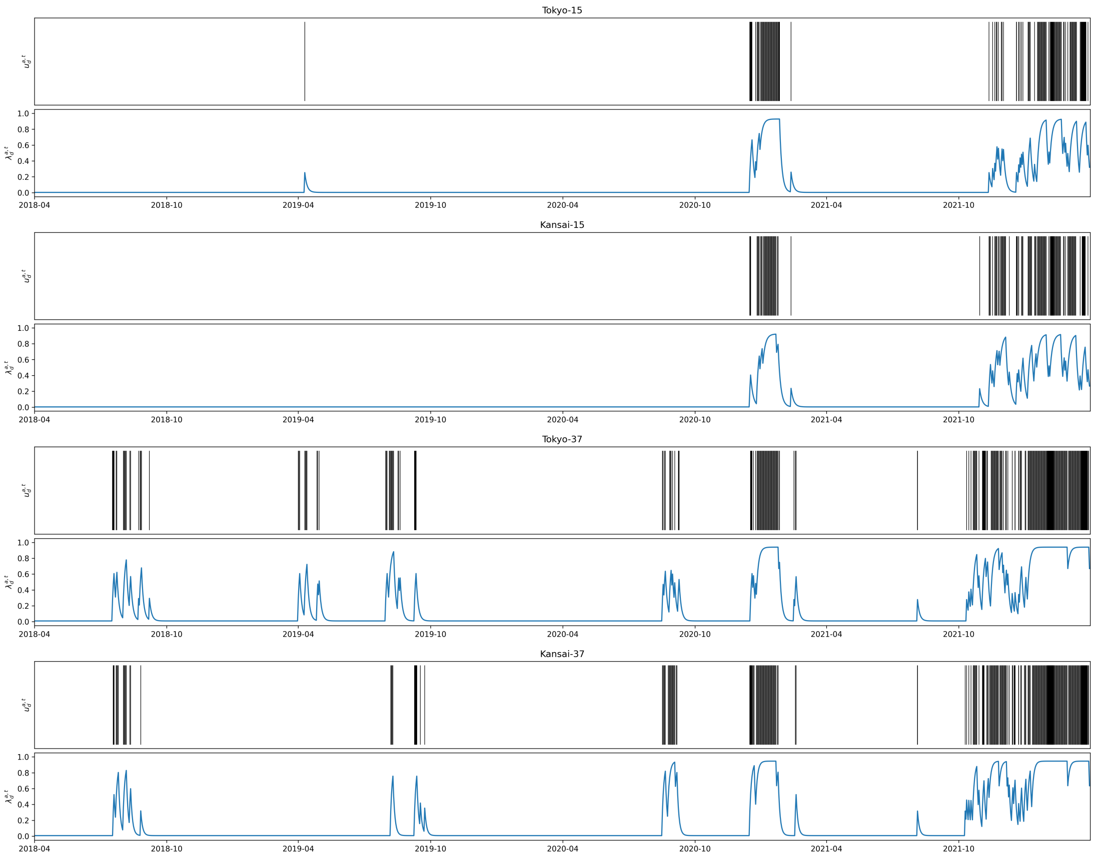

Figure 4.8: Evolution of the Hawke's model intensity $\lambda_d^{a,t}$ for the four selected area time-code combinations.

The goodness of fit of the models on the training data is assessed by evaluating the respective values of the log-likelihood functions. Given the values of the intensity function $\lambda_d^{a,t}$ and the observations $u_d^{a,t}$, the log likelihood function for the process that generates the observations is given by:

$$\log L = \sum_{d_i=0}^{d} u_{d_i}^{a,t} \log \lambda_{d_i}^{a,t} + (1 - u_{d_i}^{a,t}) \log(1 - \lambda_{d_i}^{a,t}) \tag{4.18}$$

The values of $\log L$ were evaluated for the persistence model and the three proposed Hawkes modelling approaches and the results are shown in Table 4.2. From

Table 4.2: Values of the log-likelihoods $\log L$ corresponding to the four investigated models.

$\log \mathbf{L}$	**Tokyo-15**	**Kansai-15**	**Tokyo-37**	**Kansai-37**
Persistence	-207.7	-204.5	-333.9	-271.3
Hawkes I	-141.9	-147.8	-256.4	-209.9
Hawkes II	-133.0	-135.4	-248.2	-199.6
Hawkes III	-200.9	-198.6	-317.1	-263.4

Table 4.2, it can be clearly deduced that the Hawkes modelling approaches outperform the persistence model in capturing the underlying characteristic of the spike occurrence intensity. Amongst the Hawkes models, the *Hawkes II* model consistently outperforms the other versions for all area-timecodes. This result suggests that the magnitude of price spikes has a stronger effect on the change in magnitude of the intensity function $\gamma^{a,t}$ as captured by the *Hawkes II* model than on the autocorrelation effect captured by the decay parameter $\tau^{a,t}$ in the *Hawkes III* model. In fact, the *Hawkes III* model performs worse than the original *Hawkes I* model and quite similar to the baseline persistence model.

4.4.4 Model Performance: Spike Event Occurrence Forecasting

While the log-likelihood values of Table 4.2 give an indication of the closeness of the generated intensity function to the day ahead price spike occurrence, they do not give an insight on the ability of the models to generate good short-term price spike forecasts. The forecasting performance of the models is assessed based on the ability to forecast the occurrence of spike events for a number of days ahead.

The first metric reported on the forecasting performance of the studied models is the Mean Absolute Error (MAE) of the intensity function (spike occurrence probability) time series. The MAE is given by:

$$\text{MAE} = \frac{\sum_{d=1}^{N_d} \sum_{k=1}^{N_k} |\hat{u}_{d+k}^{a,t} - u_{d+k}^{a,t}|}{N_k \cdot N_d} \tag{4.19}$$

where $\hat{u}_{d+k}^{a,t}$ is the forecasted value of the intensity function, $u_{d+k}^{a,t}$ is the spike occurrence variable (1 for a spike event and 0 for no spike event), N_d is the number of days in the dataset and N_k is the length of the forecasting horizon in days. The MAE values are tabulated in Table 4.3 for the four candidate models. This metric gives the average deviation of the magnitude of the intensity function from the actual observations. Similar to the results of the log-likelihood values shown in Table 4.2, the *Hawkes II* model is found to outperform the other models in generating spike occurrence probability values close to the observations.

Table 4.3: Mean Absolute Error values of the intensity functions for the four investigated models.

MAE	**Tokyo-15**	**Kansai-15**	**Tokyo-37**	**Kansai-37**
Persistence	0.0730	0.0738	0.1331	0.1112
Hawkes I	0.0733	0.0771	0.1343	0.1168
Hawkes II	0.0676	0.0702	0.1251	0.1045
Hawkes III	0.0741	0.0806	0.1301	0.1201

The k-days ahead spike occurrence forecasting results are categorized as true negatives (TN) i.e. $\bar{u}_{d+k}^{a,t} = 0$ and $u_{d+k}^{a,t} = 0$, true positives (TP) i.e. $\bar{u}_{d+k}^{a,t} = 1$ and $u_{d+k}^{a,t} = 1$, false negatives (FN) i.e. $\bar{u}_{d+k}^{a,t} = 0$ and $u_{d+k}^{a,t} = 1$, and false positives (FP) i.e. $\bar{u}_{d+k}^{a,t} = 1$ and $u_{d+k}^{a,t} = 0$. The typical performance index is the accuracy which measures the number of true forecasts in the dataset given by:

$$\text{ACC} = \frac{TN + TP}{TN + FN + FP + TP} = 1 - \frac{FN + FP}{TN + FN + FP + TP} \tag{4.20}$$

The accuracy index for a time series of binary variables is equivalent to 1 minus the mean absolute percentage error (MAPE) for a time series of continuous variables. It is however noted that false negatives results would typically have a bigger impact on the bottom lines of market participants than false positives. To capture this characteristic, a slightly different metric, the weighted accuracy (WACC), is adopted. The WACC is given by

$$\text{WACC} = 1 - \frac{a \times FN + (2-a) \times FP}{TN + FN + FP + TP} \tag{4.21}$$

where $1 \leq a \leq 2$ is the weight placed on the false negative observations. The larger the value of a, the greater the weight placed on false negative errors and vice versa. In the analysis, a value of $a = 1.6$ is used weighting the false negatives four times more than false positives in (4.21).

The WACC for the forecasting performance of the four models in generating 14-days ahead forecasts were calculated and the results are given in Table 4.4. The results show that the *Hawkes II* model outperforms the other models in all cases. However, the performance of the *Hawkes I* and *Hawkes III* models are very comparable to the persistence model. Its important to note the high values of WACC (> 0.89) due to the large number of true negatives in the forecasts. This however means that even the seemingly slight improvements shown in Table 4.4 correspond to significant reductions in the number of false negatives in the generated forecasts.

Table 4.4: Values of weighted accuracy corresponding to the four investigated models.

WACC	Tokyo-15	Kansai-15	Tokyo-37	Kansai-37
Persistence	0.9419	0.9384	0.8929	0.9120
Hawkes I	0.9441	0.9435	0.8923	0.9164
Hawkes II	0.9466	0.9468	0.8980	0.9211
Hawkes III	0.9422	0.9401	0.8947	0.9128

The results shown in Table 4.4 are calculated for the 14-day forecasting horizon. Figure 4.9 shows the variation in forecasting performance as the forecasting horizon increases. As expected, the weighted accuracy drops as the forecasting horizon increases. However, in all four cases, the *Hawkes-II* model is generally better than the other three models. It is also noticeable that the forecasting performance is almost the same for the 1-day ahead forecasts. The *Hawkes-II* model

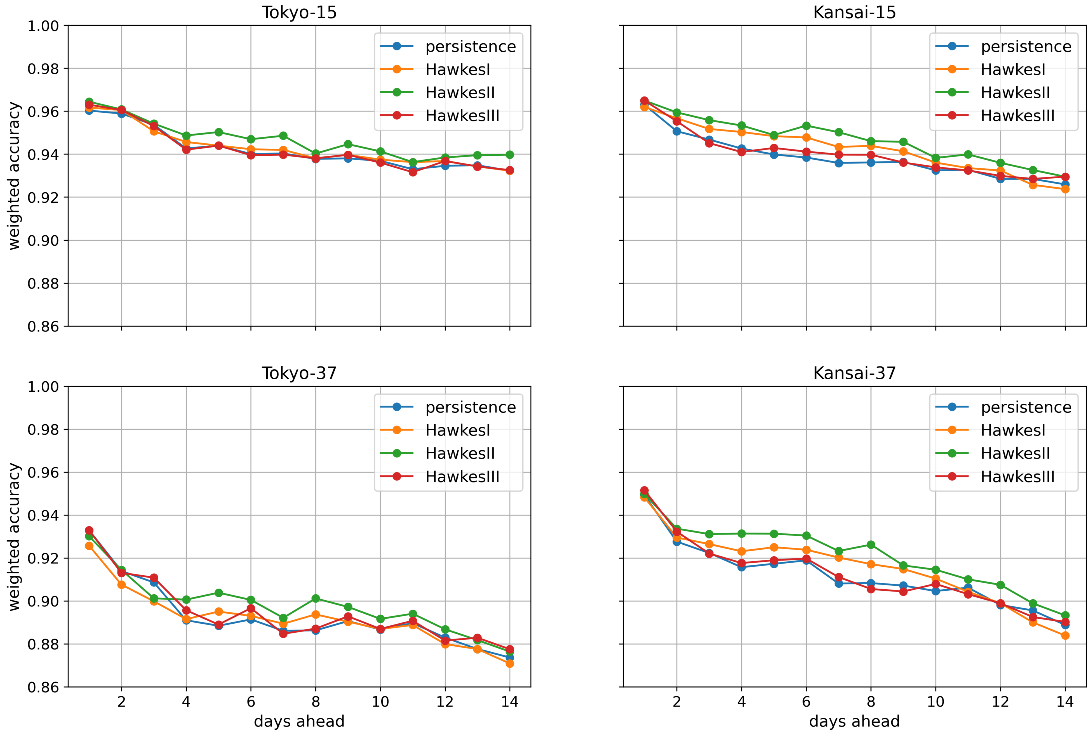

Figure 4.9: Variation of weighted accuracy WACC with forecasting horizon for the selected areas and trading time periods.

is generally better for longer forecasting horizons confirming that it performs better at capturing short-term variations in the intensity function characteristics.

The model's forecasting performance is also assessed using the Matthews correlation coefficient (MCC) which is typically used to measure the quality of binary classifications [43]. The MCC is similar to the typical Pearson correlation coefficient for continuous variables and is given by:

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4.22}$$

The MCC gives the correlation between observed and predicted variables and takes values between -1 and +1. As a performance metric, it is generally regarded as a balanced measure which can be used even with unbalanced classes as is the case with the price spikes dataset [44]. Table 4.5 shows the values of the MCC for the forecasts generated by the test models. As seen from the table, and confirmed by the previous performance metrics, the *Hawkes II* model generally outperforms the other models in this metric as well.

Table 4.5: Values of Matthews correlation coefficients corresponding to the four investigated models.

MCC	Tokyo-15	Kansai-15	Tokyo-37	Kansai-37
Persistence	0.5937	0.5845	0.5917	0.6576
Hawkes I	0.6043	0.6065	0.592	0.6676
Hawkes II	0.6317	0.6339	0.6167	0.6915
Hawkes III	0.5995	0.5903	0.6034	0.6569

Similar to Figure 4.9 showing the weighted accuracy metric against the length of

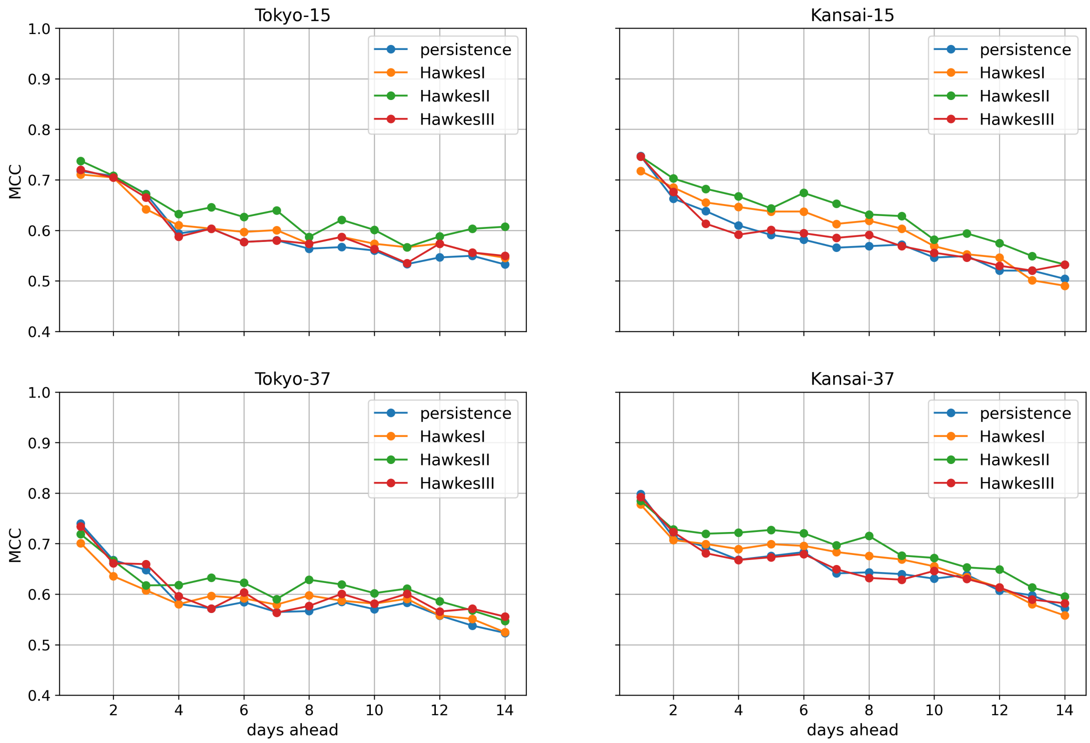

Figure 4.10: Variation of Matthews correlation coefficients with forecasting horizon for the selected areas and trading time periods.

the forecasting horizon, the MCC for the 14-days ahead forecasts are also plotted in Figure 4.10. These plots mirror the results shown in Figure 4.9 indicating that the *Hawkes II* model outperforms the other models especially for longer forecasting horizons. The MCC falls from between 0.7 to 0.8 for 1-day ahead forecasts to between 0.5 and 0.6 for 14-day ahead forecasts with the averages of between 0.6 and 0.7 shown in Table 4.5.

4.5 Conclusions

Two variations of the classical Hawkes model for modelling the price spikes time series in the Japanese electric power exchange (JEPX) are proposed in this study. The first variation models the change in the magnitude of the underlying intensity as a function of the magnitude of the price spike while the second variation models the change in the decay rate of the underlying intensity as a function of the magnitude of the price spike. An analysis on the goodness of fit to the training data of the original Hawkes model, the proposed variations compared to a baseline persistence model shows that the variable magnitude variation of the Hawkes model best captures the underlying characteristics of the process generating the price spike events. This is illustrated by achieving lower log-likelihood values compared to the other three models. The modified Hawkes model also performs best in generating short-term (a few days ahead) forecasts of the occurrence of price spike events. The improved performance is demonstrated using three metrics: (1) the MAE of the spike occurrence probability, (2) a modified accuracy index that weighs false negative forecasts more than false positives, and (3) the Mathews correlation coefficient (MCC) that tests the correlation between predictions and observations. The modified Hawkes model especially outperforms the other candidate models as the length of the forecasting horizon increases.

References

[1] K. Bhattacharya, M. H. Bollen, and J. E. Daalder, *Operation of restructured power systems*. Springer Science & Business Media, 2012.

[2] P. L. Joskow, "Lessons learned from electricity market liberalization," *The Energy Journal*, vol. 29, pp. 9–42, 2008.

[3] A. Moran and R. Sood, *Evolution of Global Electricity Markets.* Academic Press: Waltham, MA, USA, 2013.

[4] K. Hohki, "Outline of japan electric power exchange (jepx)," *IEEJ Transactions on Power and Energy*, vol. 125, no. 10, pp. 922–925, 2005.

[5] C. Harris, *Electricity markets: Pricing, structures and Economics.* Wiley, 2008.

[6] M. Shahidehpour, H. Yamin, and Z. Li, *Market operations in Electric Power Systems: Forecasting, scheduling, and Risk Management.* Institute of Electrical and Electronics Engineers, Wiley-Interscience, 2002.

[7] J. N. Tsitsiklis and Y. Xu, "Pricing of fluctuations in electricity markets," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 457–464, 2012. doi: https://doi.org/10.1109/CDC.2012.6426096.

[8] Z. Gligoric, S. S. Savic, A. Grujic, M. Negovanovic, and O. Music, "Short-term electricity price forecasting model using interval-valued autoregressive process," *Energies*, vol. 11, no. 7, 2018. doi: https://doi.org/10.3390/en11071911.

[9] K. Tselika, "The impact of variable renewables on the distribution of hourly electricity prices and their variability: A panel approach," *Energy Economics*, vol. 113, p. 106194, 2022. doi: https://doi.org/10.1016/j.eneco.2022.106194.

[10] C. Johnathon, A. P. Agalgaonkar, J. Kennedy, and C. Planiden, "Analyzing electricity markets with increasing penetration of large-scale renewable power generation," *Energies*, vol. 14, no. 22, 2021. doi: https://doi.org/10.3390/en14227618.

[11] F. E. Benth, J. S. Benth, and S. Koekebakker, *Stochastic modelling of electricity and related markets*, vol. 11. World Scientific, 2008.

[12] A. J. Conejo, M. Carrión, and J. M. Morales, *Decision making under uncertainty in electricity markets*. Springer Science+ Business Media, LLC, 2010.

[13] S. Zedda and G. Masala, "Price spikes in the electricity markets: how and why," 07 2019.

[14] G. Gayretli, A. Yucekaya, and A. H. Bilge, "An analysis of price spikes and deviations in the deregulated turkish power market," *Energy Strategy Reviews*, vol. 26, p. 100376, 2019. doi: https://doi.org/10.1016/j.esr.2019.100376.

[15] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014. doi: https://doi.org/10.1016/j.ijforecast.2014.08.008.

[16] L. Jiang and G. Hu, "A review on short-term electricity price forecasting techniques for energy markets," in *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 937–944, 2018. doi: https://doi.org/10.1109/ICARCV.2018.8581312.

[17] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1548–1568, 2018. doi: https://doi.org/10.1016/j.rser.2017.05.234.

[18] J. Contreras, R. Espinola, F. Nogales, and A. Conejo, "Arima models to predict next-day electricity prices," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014–1020, 2003. doi: https://doi.org/10.1109/TPWRS.2002.804943.

[19] G. Gao, K. Lo, and F. Fan, "Comparison of arima and ann models used in electricity price forecasting for power market," *Energy and Power Engineering*, vol. 9, no. 4B, pp. 120–126, 2017.

[20] C. McHugh, S. Coleman, D. Kerr, and D. McGlynn, "Forecasting day-ahead electricity prices with a sarimax model," in *2019 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1523–1529, 2019. doi: https://doi.org/10.1109/SSCI44817.2019.9002930.

[21] A. Cifter, "Forecasting electricity price volatility with the markov-switching garch model: Evidence from the nordic electric power market," *Electric Power Systems Research*, vol. 102, pp. 61–67, 2013. doi: https://doi.org/10.1016/j.epsr.2013.04.007.

[22] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018. doi: https://doi.org/10.1016/j.apenergy.2018.02.069.

[23] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, 2018. doi: https://doi.org/10.3390/en11051255.

[24] J. Pórtoles, C. González, and J. M. Moguerza, "Electricity price forecasting with dynamic trees: A benchmark against the random forest approach," *Energies*, vol. 11, no. 6, 2018. doi: https://doi.org/10.3390/en11061588.

[25] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied Energy*, vol. 162, pp. 218–230, 2016. doi: https://doi.org/10.1016/j.apenergy.2015.09.087.

[26] N. Chaâbane, "A novel auto-regressive fractionally integrated moving average–least-squares support vector machine model for electricity spot prices prediction," *Journal of Applied Statistics*, vol. 41, no. 3, pp. 635–651, 2014. doi: https://doi.org/10.1080/02664763.2013.847068.

[27] W. Li and D. M. Becker, "Day-ahead electricity price prediction applying hybrid models of lstm-based deep learning methods and feature selection algorithms under consideration of market coupling," *Energy*, vol. 237, p. 121543, 2021. doi: https://doi.org/10.1016/j.energy.2021.121543.

[28] J. Wang, W. Yang, P. Du, and T. Niu, "Outlier-robust hybrid electricity price forecasting model for electricity market management," *Journal of Cleaner Production*, vol. 249, p. 119318, 2020. doi: https://doi.org/10.1016/j.jclepro.2019.119318.

[29] T. D. Mount, Y. Ning, and X. Cai, "Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters," *Energy Economics*, vol. 28, no. 1, pp. 62–80, 2006. doi: https://doi.org/10.1016/j.eneco.2005.09.008.

[30] T. Christensen, A. Hurn, and K. Lindsay, "It never rains but it pours: Modeling the persistence of spikes in electricity prices," *Energy Journal*, vol. 30, 07 2008. doi: https://doi.org/10.5547/ISSN0195-6574-EJ-Vol30-No1-2.

[31] T. Christensen, A. Hurn, and K. Lindsay, "Forecasting spikes in electricity prices," *International Journal of Forecasting*, vol. 28, no. 2, pp. 400–411, 2012. doi: https://doi.org/10.1016/j.ijforecast.2011.02.019.

[32] S. M. Sirin and I. Erten, "Price spikes, temporary price caps, and welfare effects of regulatory interventions on wholesale electricity markets," *Energy Policy*, vol. 163, p. 112816, 2022. doi: https://doi.org/10.1016/j.enpol.2022.112816.

[33] "The electric power industry in japan 2020." https://www.jepic.or.jp/pub/pdf/epijJepic2020.pdf. Accessed: 2022-08-08.

[34] L. Monitoring Analytics, "State of the market report for pjm," *Monitoring Analytics: Norristown, PA, USA*, 2015.

[35] N. Flatabo, G. Doorman, O. Grande, H. Randen, and I. Wangensteen, "Experience with the nord pool design and implementation," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 541–547, 2003. doi: https://doi.org/10.1109/TPWRS.2003.810694.

[36] K. Bradbury, L. Pratson, and D. Patiño-Echeverri, "Economic viability of energy storage systems based on price arbitrage potential in real-time u.s. electricity markets," *Applied Energy*, vol. 114, pp. 512–519, 2014. doi: https://doi.org/10.1016/j.apenergy.2013.10.010.

[37] A. Ott, "Experience with pjm market operation, system design, and implementation," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 528–534, 2003. doi: https://doi.org/10.1109/TPWRS.2003.810698.

[38] "The electric power industry in japan 2022." https://www.jepic.or.jp/pub/pdf/epijJepic2022.pdf. Accessed: 2022-08-08.

[39] "The japan electric power exchange website." http://www.jepx.org/english/index.html. Accessed: 2022-08-08.

[40] A. G. Hawkes and D. Oakes, "A cluster process representation of a self-exciting process," *Journal of Applied Probability*, vol. 11, no. 3, p. 493–503, 1974. doi: https://doi.org/10.2307/3212693.

[41] P. Embrechts, T. Liniger, and L. Lin, "Multivariate hawkes processes: an application to financial data," *Journal of Applied Probability*, vol. 48, no. A, p. 367–378, 2011. doi: https://doi.org/10.1239/jap/1318940477.

[42] A. G. HAWKES, "Spectra of some self-exciting and mutually exciting point processes," *Biometrika*, vol. 58, pp. 83–90, 04 1971. doi: https://doi.org/10.1093/biomet/58.1.83.

[43] D. Chicco and G. Jurman, "The advantages of the matthews correlation coefficient (mcc) over f1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, 2020. doi: https://doi.org/10.1186/s12864-019-6413-7.

[44] S. Boughorbel, F. Jarray, and M. El-Anbari, "Optimal classifier for imbalanced data using matthews correlation coefficient metric," *PloS one*, vol. 12, no. 6, p. e0177678, 2017.

Publication List

[1] B. Adline and K. Ikeda, “A Hawkes Model Approach to Modeling Price Spikes in the Japanese Electricity Market,” Energies, vol. 16, no. 4, p. 1570, Feb. 2023, doi: 10.3390/en16041570.