Journal of Energy Markets **14**(3), 1–26
DOI: 10.21314/JEM.2021.004

Research Paper

One-week-ahead electricity price forecasting using weather forecasts, and its application to arbitrage in the forward market: an empirical study of the Japan Electric Power Exchange

Takuji Matsumoto and Misao Endo

Socio-economic Research Center, Central Research Institute of Electric Power Industry, 1-6-1 Otemachi, Chiyoda-ku, Tokyo 100-8126, Japan;
emails: matsumoto3832@criepi.denken.or.jp, endo@criepi.denken.or.jp

(Received March 23, 2021; revised June 9, 2021; accepted June 15, 2021)

ABSTRACT

Although forecasting one-week-ahead average electricity prices is necessary for decision-making such as evaluating forward contracts, its modeling has not been sufficiently studied. Therefore, to find a suitable forecasting approach, this study constructs and compares multiple parsimonious models using widely published weekly weather forecasts and then applies them to arbitrage trading in the forward market. In particular, we clarify the following empirical results using the data from Japan Electric Power Exchange. First, instead of using forecasted temperature directly as an explanatory variable, the two-step forecasting method using measured temperature as an intermediate variable is more likely to reduce forecast errors. Second, quantile regression has better density forecast accuracy than the generalized autoregressive conditional heteroscedasticity model. Third, the logarithmic conversion for prices tends to improve forecast accuracy. Fourth, one-week-ahead weather forecasts can significantly improve both the price forecast accuracy and the arbitrage profit. The proposed arbitrage strategy can be used by many participants because it can be flexibly changed according to the player's risk tolerance. In addition, our

Corresponding author: T. Matsumoto
Print ISSN 1756-3607 | Online ISSN 1756-3615

forecasting/trading method, based on published weather forecasts, has wide applicability in that it can be constructed even in markets where system information is not sufficiently disclosed.

Keywords: arbitrage trading; electricity spot price forecasting; forward market; quantile regression; weather forecast.

1 INTRODUCTION

Accurate forecasting of electricity market prices is key for electricity utilities to run their businesses economically. As shown systematically by Weron (2014), there have been many previous studies on electricity price forecasting, particularly for dealing with a short-term forecasting horizon from one hour ahead to a few days ahead. Electricity prices with a mid- or long-term horizon, on the other hand, are often dealt with in the context of risk management and derivative pricing, so the literature on their forecasting methods is relatively limited, but some distinctive studies have been conducted (Eydeland and Wolyniec 2003; Weron 2014). For long-term price forecasting, a typical method uses market equilibrium models that can take into account structural changes in the market (Bastian *et al* 1999; Ventosa *et al* 2005), which are often based on stylized fundamental assumptions of economic behavior (Green and Newbery 1992; Barquín and Vázquez 2008). For the mid-term horizon, from a few days to a few months ahead, a hybrid approach (Karakatsani and Bunn 2008; Marcos *et al* 2020) that can take into account both fundamental price formation and econometric recalibration has been reported to be effective (Bello *et al* 2017). However, for a one-week-ahead horizon, modeling approaches focusing on historical prices are often preferred (see, for example, Chen *et al* (2008), who used manifold learning, and Mandal *et al* (2006), who used an artificial neural network). Of these, however, there have been few related studies focusing especially on one-week-ahead average prices; to the best of our knowledge, such studies only include building univariate models from historical prices in the Singapore market (Loi and Le Ng 2018) and using historical prices plus lagged (measured) weather values in the Nord Pool market (Torró 2007). In practice, though, weekly average price forecasting is critical for electric utilities and is needed especially when evaluating prices of weekly electricity delivery contracts (including bilateral and forward contracts). The issue of how to effectively use week-ahead hedging instruments (futures and forward) is particularly important for (relatively small) retailers, where short-term risk management is a key management issue. For generators that need to submit weekly power generation plans to system operators (Suzuki 1991), week-ahead price forecasts can have a significant impact on their decisions. In particular, the operation of pumped storage is usually based on weekly scheduling (Popa *et al* 2010), and its operation scheduling has become more important with the recent expansion in renewable energy (Aihara and Yokoyama 2016). In reality, optimizing such operation scheduling requires granular forecasting, eg, hourly. However, to improve accuracy or reliability, it is important to practically verify forecasted week-ahead hourly prices separately by using the weekly average value.

Against this background, our paper focuses on the forecasting of one-week-ahead average prices, which has been relatively rarely studied. To do this effectively, we use widely published week-ahead weather forecasts, which have also received little attention in previous studies. Although weather forecasts with a long forecast horizon are known not to be very accurate (even the latest research says that the "practical predictability limit" is about 10 days (Zhang *et al* 2019)), it is nevertheless possible to improve the forecast accuracy of the one-week-ahead electricity price. Therefore, in this paper we construct several forecasting models for the one-week-ahead electricity spot price using published weekly weather forecasts, and explore a desirable modeling approach that can leverage the information value of weather forecasts. Further, in addition to forecasting the weekly average value, we also propose a density forecasting method using the quantile regression (QR) (Koenker 2005) and generalized autoregressive conditional heteroscedasticity (GARCH) models (Bollerslev 1986), and apply them to strategic arbitrage trading methods for weekly forward products.[1]

In particular, we tackle the following four problems that are thought to be useful for constructing accurate forecast models in practice.

(i) Which of the forecasted weather values or the measured (realized) weather values should be incorporated into the price regression formula?

(ii) Should we perform logarithmic conversion on the price series to be regressed?

(iii) Which of the QR or GARCH models should be used for forecasting the weekly price density?

(iv) What is the information value of the week-ahead weather forecast in terms of improving forecast accuracy and arbitrage profit?

These are all open questions regarding week-ahead average price forecasts. The more detailed motivations for addressing these issues are as follows. First, issue (i) arises from the hypothesis that constructing a price regression equation using measured weather may have higher explanatory power and greater forecast accuracy than using forecasted weather as explanatory variables, considering that the week-ahead

[1] This study uses the word "arbitrage" because transactions that profit from price differences between short-term electricity markets are generally called this in much of the literature, but it should be noted that the strategies proposed here carry certain risks (ie, they could also be called "speculative" trading).

weather forecast is inherently not very accurate. The viewpoint of (ii) corresponds to the problem of whether to use an additive model or a multiplicative model. The review by Weron (2014) positions them as the "two most important categories" among statistical forecasting methods but merely states that the additive model is "far more popular"; there seems to be no previous research focusing on the comparison between these classifications. Problem (iii) is based on the hypothesis that even if we use the weekly average electricity price, there may still be a distorted density shape, and the QR that can express an arbitrary density shape may be superior to GARCH, which assumes a (log)normal distribution in terms of forecast accuracy. Regarding (iv), previous studies have analyzed the effect of weather forecast information on forecasting the daily average (Huurman *et al* 2012) and hourly (Bigerna 2018) electricity prices, but there are no studies that have verified it using weekly average prices. Thus, these four problems have not been solved in previous studies; this study aims to fill those gaps. Considering the above, this study performs various empirical analyses using actual data from the Japanese market, for which the past weekly weather forecasts can be obtained in full, to observe useful implications for solving these open questions. In response to issues (i)–(iv), our results using out-of-sample-period data reveal that the adoption of a two-step model, a multiplicative model and a QR model, as well as the use of weather forecast information, has advantages in improving forecast accuracy.

In addition, in order to apply the constructed forecasting method to arbitrage strategies for the forward market, we build a trading strategy that can be flexibly changed according to the player's risk tolerance; the results of simulations demonstrate the effectiveness of the proposed trading strategy and the existence of considerable arbitrage opportunities in the Japan Electric Power Exchange (JEPX) forward market. The previous literature that dealt with trading strategies between multiple electricity markets based on price comparisons includes the following topics: arbitrage between the intraday market and the balancing market (Bunn and Kermer 2021; Just and Weber 2015), day-ahead and real-time prices (Boogert and Dupont 2005), incorporating storage batteries (Krishnamurthy *et al* 2017) and proposing a sales strategy for renewable electricity by comparing the prices of day-ahead and intraday markets (Maciejowska *et al* 2019). However, to the best of our knowledge, our study is the first attempt to perform an empirical analysis on the trading profit on one-week-ahead forward products. In addition, the previously proposed trading strategies that targeted shorter term markets did not consider weather forecasts, so this study also provides a new perspective in terms of measuring the monetary value of weather forecasts in the electricity trading market.

This paper is organized as follows. In Section 2, we construct a series of different types of forecast models for the weekly electricity price. Section 3 formulates the arbitrage strategies in the forward market. Section 4 validates the forecast accuracy of the models and demonstrates the effectiveness of the arbitrage strategies. Lastly, Section 5 presents our conclusions.

2 CONSTRUCTION OF FORECAST MODELS

In this section, in order to consider the key issues in constructing an effective forecast model for JEPX spot prices, we first review the spot price and weekly weather forecast data used in our forecast models, and then we build multiple models relating to the average value forecast and the density forecast.

2.1 Overview of data

2.1.1 Weekly average price distribution

Since several previous studies have shown that hourly or daily average prices have extremely distorted densities due to price spikes, the price density may still be distorted even if averaged per week. In order to grasp the rough density shape of the weekly average price, the histograms of the weekly average prices for three areas in the JEPX spot market (24-hour average base load and weekday 08:00 to 18:00 daytime loads) are shown in Figure 1, on which the fitted normal distribution and the lognormal distribution are superimposed. The weekly average price has a left-skewed distribution and is closer to the lognormal distribution than the normal distribution (the Akaike information criterion (AIC) given in the legend of each graph also indicates that the lognormal distribution is better fitted).

2.1.2 Weekly temperature forecast

Next, we review the forecast accuracy of the weekly weather forecast, which is the key factor in the model proposed in this study. Figure 2 shows the R-squared value when the measured value of temperature is regressed by the forecasted value of the seven different forecast horizons.[2] The next day's forecast has an explanatory power of about 77%, but as the forecast horizon becomes longer, the explanatory power gradually declines, reaching about 18% for seven days ahead.

2.2 Construction of weekly average price forecast model

In this section, we construct the forecast models for the weekly average price. As previously mentioned, there are several previous studies on statistical forecasting models for electricity prices using weather forecasts, but usually the weather forecast value is used directly as an explanatory variable. However, in the case of weekly

[2] We used historical data in Tokyo for each day in 2019, and the seasonality was removed by subtracting the climatological normal value. See Appendix A online for the original scatter plot.

FIGURE 1 Weekly average price histogram and fitted density.

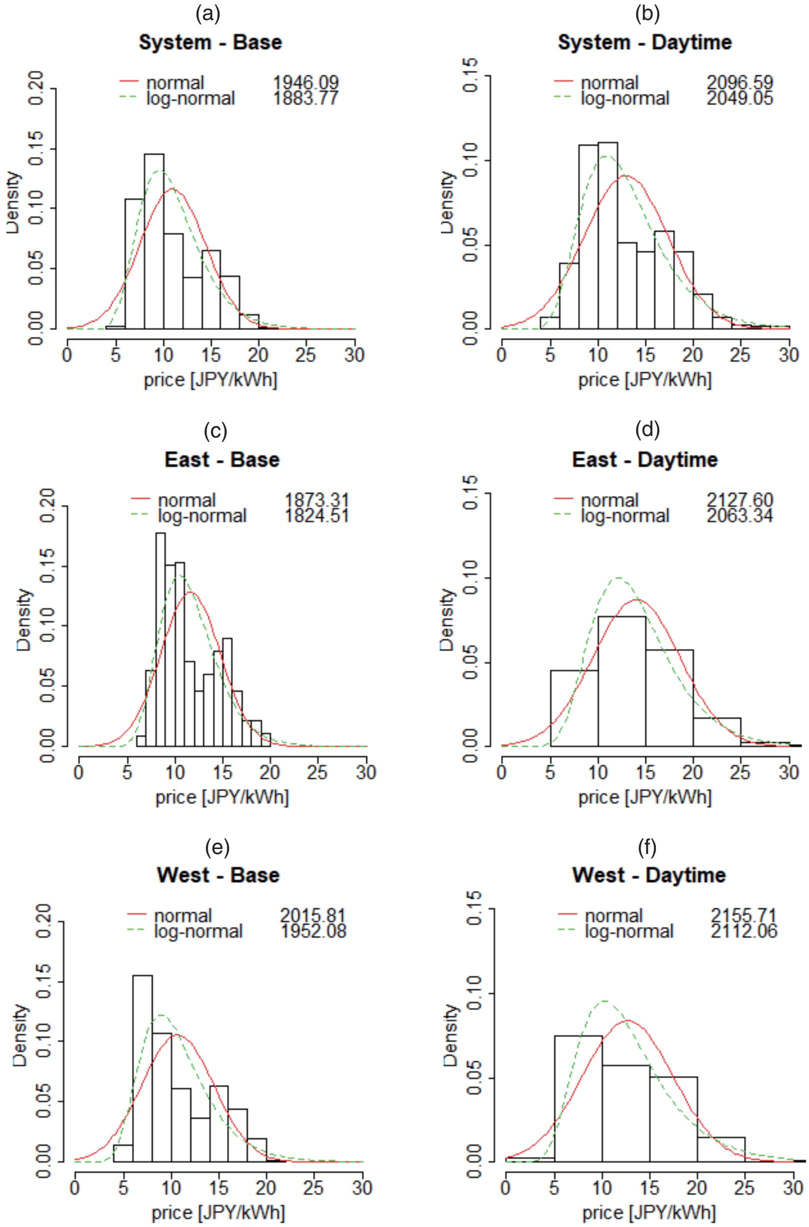


FIGURE 2 Contribution ratio of forecasted temperature to measured value.

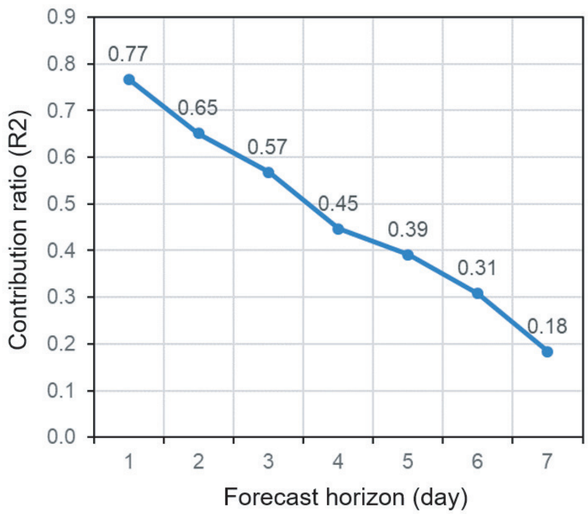
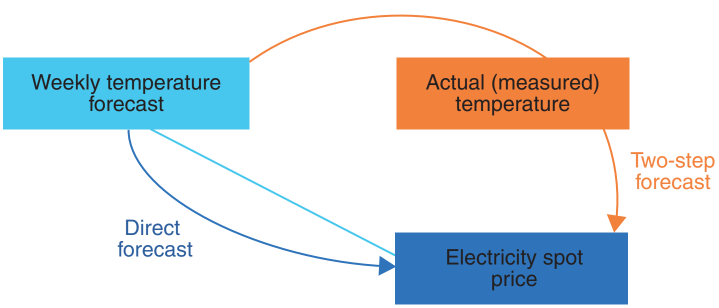


FIGURE 3 Conceptual diagram of comparison between direct forecast and two-step forecast.



weather forecasts, as seen in the previous section, the forecast accuracy deteriorates significantly when the forecast horizon is long. Therefore, if we adopt a model that directly uses the forecasted temperature as an explanatory variable (we call it a direct forecast model), it may be difficult to robustly model the intricate interactions that should inherently exist between the temperature and the electricity price (such as nonlinear seasonality and annual change of temperature sensitivity). Therefore, in this study, we construct a two-step forecasting model using both the forecasted temperature and the measured temperature. First, we estimate the price forecast regression using the measured temperature as an explanatory variable, and then we (separately) forecast the conditional probability density of the measured temperature in which the forecasted temperature is given. Lastly, we combine these two forecasts. In this manner, we compare the performance of these two approaches. A conceptual diagram of the comparison of these two models is shown in Figure 3.

Further, this study also compares two models, namely the price series model and the log-price series model, for each of the above two modeling approaches. The actual models are constructed in the following sections.

2.2.1 Price series model

Considering that several papers in the literature clarified that the JEPX spot price has a remarkable seasonal trend and strong correlations with fuel price and temperature (Matsumoto and Endo 2021; Matsumoto and Yamada 2021), in this study we will also build a model that can incorporate a seasonal trend and fuel price correlations. First, the following two types of ordinary least squares (OLSs) are formulated for the price series models (without logarithmic conversion), where model P1 uses the forecasted temperature and model P2 uses the measured temperature.

$$\text{Model P1}: \quad S_t = \alpha S_{t-1} + \beta(t)G_t + f(t) + g(t)\hat{\varepsilon}_t + \eta_t, \tag{2.1}$$

$$\text{Model P2}: \quad S_t = \alpha S_{t-1} + \beta(t)G_t + f(t) + g(t)\varepsilon_t + \eta_t, \tag{2.2}$$

where S_t is the weekly average spot price in week t, G_t is the spot price of liquid natural gas (Platts JKM LNG), which is observed as of week $t-1$ (the spot price for LNG cargo that will be delivered in the month following the month in which the start date of week t belongs), and ε_t ($\hat{\varepsilon}_t$) is the weekly average measured (forecasted) temperature residual (defined as the temperature minus its yearly cyclical trend, eg, $\varepsilon_t := \text{Temp}_t - \varphi(t)$, for measured temperature; see Appendix B online). α is a fixed coefficient, $\beta(t)$, $f(t)$ and $g(t)$ are variable coefficients (yearly cyclical trends) estimated based on the concept of a Fourier series expansion (see Appendix B online) and η_t is a residual term with an average of 0. Note that the forecasted temperature residual used in this study may be expressed as $\hat{\varepsilon}_{t|t-1}$ because it is observed in week $t-1$, but the notation $\hat{\varepsilon}_t$ is used for simplicity.

Here, as can be seen from (B4) in Appendix B online, the point of our model is that the intersection term – which consists of the periodic trend $\varphi_{4,1}(t)$, the elapsed week trend (Period_t) and the temperature residual ($\hat{\varepsilon}_t$ or ε_t) – is included in each price forecast model. This interaction term is included because a characteristic of the JEPX spot price is that the temperature sensitivity has changed over time due to the increase in the photovoltaic power generation capacity in Japan (Matsumoto and Endo 2021). The estimation result for this term will be detailed in Section 4.1.

The forecasted spot price (of week t forecasted in week $t-1$) of the price series model $\tilde{S}_{t|t-1}$ can be obtained by the following equation as the predictor part of both (2.1) and (2.2):

$$\tilde{S}_{t|t-1} = \alpha S_{t-1} + \beta(t)G_t + f(t) + g(t)\hat{\varepsilon}_t. \tag{2.3}$$

Note that in (2.3), the forecasted temperature residual is used even when the measured temperature is used (ie, for (2.2)). This is because, when the conditional expectation value is taken for both sides of (2.2), the measured temperature residual, which is not observed at the time of the forecast, changes into the forecasted (expected) temperature residual. Note also that the price series model is a linear model with respect to the temperature residual, so when obtaining the forecasted price, there is no need to consider the conditional variance of the measured temperature residual in which the forecasted temperature residual is given.

2.2.2 Log-price series model

Next, the following OLSs (models L1 and L2), which are isomorphic to (2.1) and (2.2), are constructed for the logarithmic series of spot prices.

$$\text{Model L1}: \quad \log S(t) = \gamma(t) + g(t)\hat{\varepsilon}_t + \eta_t, \tag{2.4}$$

$$\text{Model L2}: \quad \log S(t) = \gamma(t) + g(t)\varepsilon_t + \eta_t, \tag{2.5}$$

where $\gamma(t) := \alpha \log(S_{t-1}) + \beta(t)\log(G_t) + f(t)$.

When these log-price OLSs are used, it is not appropriate to use the predictor parts of (2.4) and (2.5) by just converting them to the original price ($\exp(\gamma(t) + g(t)\hat{\varepsilon}_t)$) as the desired forecast values. This is because this forecast method includes a downward bias, as the variances of the random variables are not considered (for convenience, we call this the forecast without bias correction). This fact can be understood by expanding (2.5) as follows:

$$\begin{aligned} \hat{S}_{t|t-1} &= E[S_t \mid \mathcal{F}_{t-1}] = E[\exp(\gamma(t) + g(t)\varepsilon_t + \eta_t) \mid \mathcal{F}_{t-1}] \\ &= \mathrm{e}^{\gamma(t)} E[\exp(g(t)\varepsilon_t) \mid \mathcal{F}_{t-1}] E[\mathrm{e}^{\eta_t} \mid \mathcal{F}_{t-1}] \\ &= \exp(\gamma(t) + g(t)\hat{\varepsilon}_t + \tfrac{1}{2}\{(g(t)\sigma_{\varepsilon,t})^2 + \sigma_{\eta,t}^2\}), \end{aligned} \tag{2.6}$$

where we assume that the random variables ε_t and η_t follow normal distributions independent of each other and that time series correlations do not exist for simplicity.[3] We assume a complete probability space $(\Omega, \mathcal{F}, P)$ and finite time horizon

[3] Even without making such an assumption of independence, this equation holds approximately; see Appendix C online.

$[0, T]$. $\{\mathcal{F}_t : t \in [0, T]\}$ represents the filtration generated by the price process, and we assume $\mathcal{F}_T = \mathcal{F}$. $E[\cdot \mid \mathcal{F}_{t-1}]$ denotes the conditional expected value in week $t - 1$ (at this time, the forecasted temperature has already been obtained), and $\sigma_{\varepsilon,t}$ and $\sigma_{\eta,t}$ denote the standard deviation of $\varepsilon_t - \hat{\varepsilon}_t$ and η_t, respectively.

In other words, when using a log-price series model such as (2.5), we need to forecast separately the standard deviation of the temperature forecast error and the standard deviation of the residual $\sigma_{\eta,t}$ of OLS (2.5), respectively, and by substituting them into (2.6), we can obtain the appropriate price forecast. The forecasted price obtained in this way is $\exp(\frac{1}{2}\{(g(t)\sigma_{\varepsilon,t})^2 + \sigma_{\eta,t}^2\})$ times larger than the forecast without bias correction. The same argument applies to the L1 model (2.4) using the forecasted temperature value, by setting the variance of the forecast error $\sigma_{\varepsilon,t}$ to 0. In this study, the forecasted values of $\sigma_{\varepsilon,t}$ and $\sigma_{\eta,t}$ are obtained by applying (separately) the GARCH(1,1) model to the series of $\varepsilon_t - \hat{\varepsilon}_t$ and η_t, respectively, up to week $t - 1$.

2.3 Construction of price density forecast model

In this section, we construct the density forecast model for the week-ahead spot price using QR, with a view to applying it to an arbitrage strategy in the forward market. For simplicity we do not consider the abovementioned two-step forecast; we deal only with the direct forecast method using the forecasted temperature as the explanatory variable. In order to evaluate the accuracy of the QR model, a model that combines OLS and GARCH (which we call the OLS + G model) is constructed for comparison.

For both QR and OLS + G, we construct the price series model as well as the log-price series model and verify the effect of logarithmic conversion on the forecast accuracy. In each case, we use regression formulas with exactly the same terms as (2.1) and (2.4), each of which is applied for QR or OLS + G. In the OLS + G model, after estimating the OLS, GARCH(1,1) is applied to the residual series to obtain the probability density forecast of one week ahead. Note that this study uses the R (version 3.6.3) package `fGarch` (prediction method by the garchFit($\cdot$) function) for GARCH, and the package `quantreg` (lasso penalized QR by the rq($\cdot$) function) for QR.

3 ARBITRAGE STRATEGY IN THE FORWARD MARKET

The forward product in JEPX is a contract that assumes that electricity will be delivered within a certain period at the forward price, contracted in advance, but the actual settlement is financially performed by the difference between the spot price realized in the future and the forward price (JEPX 2019). JEPX forward contracts are classified as a 24-hour type (for 24-hour delivery each day) or a daytime type (for delivery times from 08:00 to 18:00 on weekdays excluding weekends and holidays). Each type of contract has three different products with delivery periods of one week, one month and one year, respectively. Because the forward market is conducted in a continuous session, contracts are executed as soon as the conditions of the seller and the buyer are met while the market is open. As for the weekly forward products handled in this study, the contractable period is set from the 20th of the month before the month of the delivery start date to three days before the delivery start date. Note that because the delivery start date is every Saturday, every Wednesday is the last tradable day (JEPX 2020).[4]

The arbitrage-trading strategies dealt with in this study assume two forecasting cases, one using OLS-based average forecasting and the other using QR-based density forecasting. The decision variables for each case Y_{t-1}^{OLS} and $Y_{t-1}^{\mathrm{QR}(q)}$ (ie, variables of decision-making made in week $t-1$ regarding electricity delivered in week t) are defined as follows:

$$Y_{t-1}^{\mathrm{OLS}} = \begin{cases} 1 & \text{if } \hat{S}_{t|t-1} > F_{t|t-1}, \\ -1 & \text{if } \hat{S}_{t|t-1} < F_{t|t-1}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

$$Y_{t-1}^{\mathrm{QR}(q)} = \begin{cases} 1 & \text{if } \hat{S}_{t|t-1}^{(q)} > F_{t|t-1}, \\ -1 & \text{if } \hat{S}_{t|t-1}^{(1-q)} < F_{t|t-1}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

where $Y = 1$ means a long position in forward trading (buying the forward product that will be settled at the realized spot price after the spot market contract is executed), $Y = -1$ means a short position (the reverse of the above) and $Y = 0$ means no deal. $\hat{S}_{t|t-1}$ ($\hat{S}_{t|t-1}^{(q)}$) are the forecasted weekly average spot prices (the quantile forecasted prices at percentile $q \in \{0.05, 0.10, 0.25\}$), which are forecasted in week $t-1$. $F_{t|t-1}$ denotes the forward price for the electricity that will be delivered in week t observed in week $t-1$. At this time, the arbitrage profit π_t is calculated by

[4] The JEPX forward market also lists monthly products, but for these the final trading day is the 19th of the penultimate month before the delivery month, so the forecast horizon is extremely long (difficult to predict precisely) for arbitrage trading; hence, it is not dealt with in this study. Note also that, since JEPX is a market for the purpose of physical electricity trading, it is not permitted to carry out purely financial arbitrage virtually, as proposed here (JEPX 2020). However, even if we make a simple assumption that financial trading can be conducted in JEPX, the results obtained can have significant implications for traders' strategic forward market trading.

the following formula:

$$\pi_t = \begin{cases} S_t - F_{t|t-1} & \text{if } Y_{t-1} = 1, \\ F_{t|t-1} - S_t & \text{if } Y_{t-1} = -1, \\ 0 & \text{if } Y_{t-1} = 0. \end{cases} \tag{3.3}$$

Note that this paper ignores the impact of interest rates due to different trading points on futures and spots.

4 EMPIRICAL ANALYSIS

This section validates the forecast accuracy of each model and the effectiveness of the arbitrage strategy using the following empirical data:

(1) the spot price S_t and forward price F_t (¥/kWh) is the JEPX system price, Tokyo area price and Kansai area price;[5]

(2) the LNG spot price G_t (¥/MMBtu) is the Platts JKM spot price;[6]

(3) the measured temperature Temp_t (°C) is the maximum temperature measured by the Japan Meteorological Agency;[7] and

(4) the forecasted temperature $\widehat{\text{Temp}}_t$ (°C) is the maximum temperature forecasted by the Japan Meteorological Agency.[8]

All forecasted values are calculated as out-of-sample values, and the weekly average spot price and the series of explanatory variables observed in the past three years are used to forecast the weekly average price one week ahead, in a weekly rolling manner; we thereby obtain the price forecast for the four-year out-of-sample period from April 1, 2016 to March 31, 2020. The validation of the forecast accuracy performed in Section 4.2 uses all the data from the out-of-sample period, and the evaluation of the forward arbitrage profit in Section 4.3 will only use the sample when the forward price is observed in the out-of-sample period. That is, we assume the players can make the arbitrage trading decision, which requires a comparison between the forecasted spot price and the observed forward price, only when a forward contract by another player is realized. Note that, due to the extremely low liquidity of the JEPX forward market, there are many more weeks in which trading contracts are not executed than there are weeks in which they are executed.

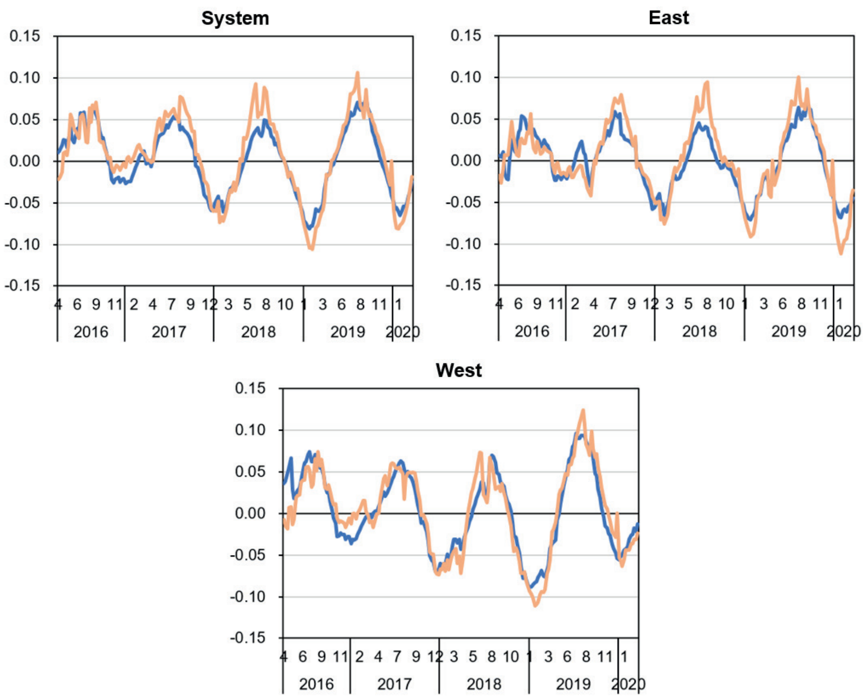
[5] URLs: http://jepx.org/market/index.html and http://jepx.org/market/forward.html.

[6] Platts, Benchmark Statement for LNG Japan/Korea Spot Crg DES.

[7] URL: http://data.jma.go.jp/gmd/risk/obsdl/index.php.

[8] URLs: http://weather-transition.gger.jp/ and http://pe-sawaki.com/WeatherForecast/.

FIGURE 4 Time series transition of estimated temperature sensitivity.



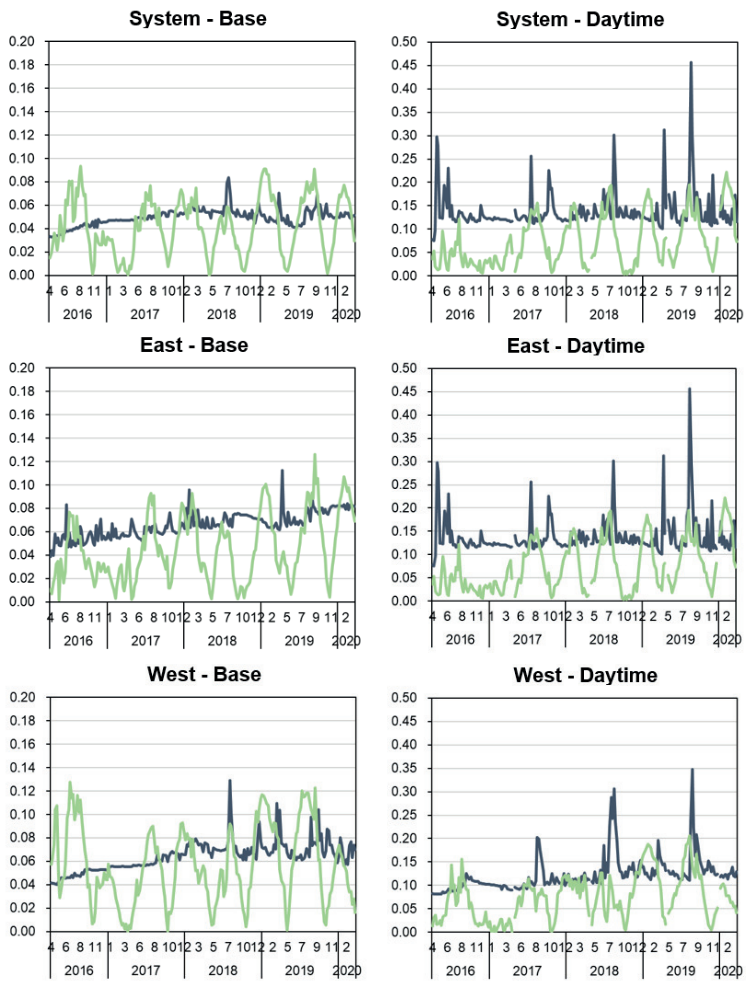
Blue line, base load. Orange line, daytime load.

Because the JEPX forward price has not been announced for each contract timing (ie, $F_{t|t-1}$ cannot be obtained), we use the average value of all the contracted forward prices in week t, $F_t := \text{Mean}[F_{t|\cdot}]$, which has been published. Regarding the temperatures needed for forecasting the Tokyo area price (east), the Kansai area price (west) and the system price, the maximum temperatures of Tokyo, Osaka and the average values of both cities are used. In addition, when forecasting the following week's average price, the weather forecast on Wednesday is used, which is the last day of the weekly forward trading.

4.1 Estimation result

This section summarizes the estimation results of the forecast model before verifying the forecast accuracy and analyzing the arbitrage profit. This is because, as we stated that using $\exp(\frac{1}{2}\{(g(t)\sigma_{\varepsilon,t})^2 + \sigma_{\eta,t}^2\})$ times the bias correction is necessary for the L2

FIGURE 5 Time series transition of estimated $(g(t)\sigma_{\varepsilon,t})^2$ (green line) and $(\sigma_{\eta,t})^2$ (blue line).



model, which is in the explanation of (2.6), we need to understand what time series data such as $(g(t)\sigma_{\varepsilon,t})^2$ and $(\sigma_{\eta,t})^2$ look like.

First, Figure 4 shows the time series transition of the estimated temperature sensitivity.[9] This is $g(t)$ (coefficient of temperature residual ε_t) estimated from (2.5),

[9] Color figures are available in the online version of this paper.

which was estimated as a one-period-ahead (one-week-ahead) value and plotted over all out-of-sample periods in a weekly rolling manner. It can be confirmed that the temperature sensitivity of the spot price is negative in winter and positive in summer, and that the absolute value of sensitivity is generally larger on the daytime load than on the base load. It is interesting to note that its amplitudes have increased significantly year by year. This is thought to be mainly due to the recent expansion of renewable energies (especially photovoltaic power generation in the case of Japan).

Second, Figure 5 shows the time series transition of the estimated $(g(t)\sigma_{\varepsilon,t})^2$ and $(\sigma_{\eta,t})^2$. As $\sigma_{\varepsilon,t}$ is the standard deviation of the temperature forecast error and $\sigma_{\eta,t}$ is the standard deviation of the residual of OLS (2.5), $(g(t)\sigma_{\varepsilon,t})^2$ and $(\sigma_{\eta,t})^2$ denote the variances resulting from the temperature prediction error and those from the others (residuals of the price model) regarding the forecasted densities of the logarithmic spot prices. As shown by the plots in Figure 5, both variances tend to expand over time for each of the six cases, which is probably due to increasing competitiveness ever since the full liberalization of the electricity retail business in 2016 as well as to the expansion in renewable energies. Regarding the daytime loads for times when price spikes are likely to occur, $(g(t)\sigma_{\varepsilon,t})^2$ is larger than $(\sigma_{\eta,t})^2$ overall. This reflects the fact that the price fluctuation risk resulting from spikes that occur probabilistically (not necessarily linked to the weather) is greater (ie, the predictability of spikes is lower) than that of the temperature prediction error.

4.2 Validation of forecast model

In this section, we compare the forecast performance of the different average value forecast and density forecast models defined in Section 2, and we verify the effect of the weather forecast on reducing the price forecast error. That is, we verify the forecast accuracy for each of the following comparisons.

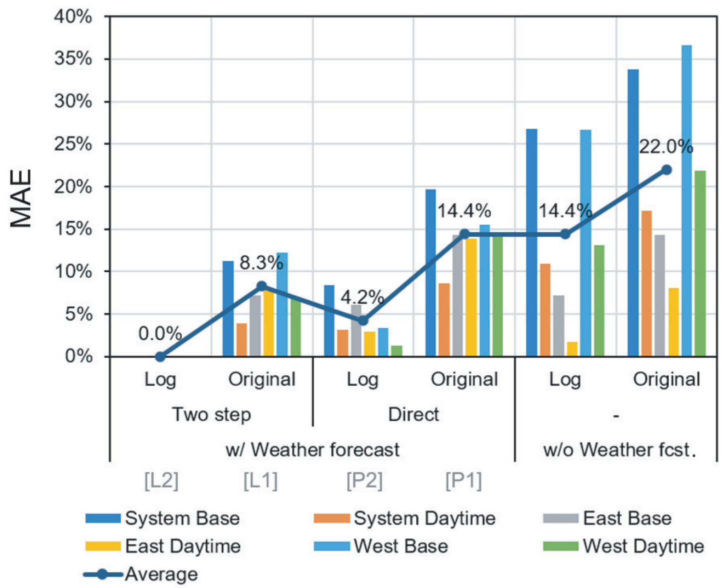
Average forecast models:

- direct forecast versus two-step forecast,
- price series versus log-price series,
- model with weather forecast versus model without weather forecast.

Density forecast models:

- price series versus log-price series,
- QR versus OLS + G,
- model with weather forecast versus model without weather forecast.

FIGURE 6 Comparison of MAE for the different modeling approaches.



4.2.1 Validation of weekly average forecast model

In this section, in order to verify the accuracy of the weekly average forecast model constructed in Section 2.2, the forecast error is measured by the mean absolute error (MAE). Figure 6 shows the MAE for the different modeling approach (the two-step/direct forecast and the price/log-price series). The case without weather information (Temp_t or $\widehat{\mathrm{Temp}_t}$) is also considered. In addition, in order to make the comparison between models easier to understand, the difference in the value of $\mathrm{MAE}_M/\mathrm{MAE}_{\mathrm{L2}} - 1$, ie, the ratio of the MAE of each model M (MAE_M) to the MAE of the log-price two-step forecast model ($\mathrm{MAE}_{\mathrm{L2}}$), is calculated for each of the six cases (3 areas $\times$ 2 loads) and displayed as a bar graph; the line graph shows the simple average values of the six cases.

First, the log-price series models have smaller MAEs overall than the price series models (this result corresponds to issue (ii) in Section 1). In addition, the MAEs of the two-step forecast are smaller than those of the direct forecast as a whole (this result corresponds to issue (i) in Section 1) and the MAE of the two-step forecast is 4.2% smaller on average even when comparing highly accurate log-price series models. We probably get this result because the regression formula using the weekly weather forecast with low forecast accuracy as the explanatory variable does not sufficiently express the nonlinearity and the interaction between the variables, and thus the explanatory power is low. On the other hand, in the two-step forecast approach, the relationship between the measured weather value and the spot price is expressed more robustly, so the model could be constructed more flexibly.

In terms of the effect of using the weather forecast information, the log-price series model confirmed a remarkable MAE improvement of 14.4%, which suggests that the weekly weather forecast is an important factor for improving the price forecast accuracy. In addition, the modeling approach of the popular additive model (the price series model) with the direct forecast approach (ie, the P1 model) has a larger MAE (14.4%) than the L2 model. In other words, the effect of devising the modeling approach is comparable with using the information of weekly weather forecasts, in terms of improving the forecast accuracy.

In addition, for the log-price series model of average value forecast, the bias correction as described in Section 2.2.2 was performed; as a result, it eliminated all the significant downward bias confirmed in some of the six cases (see Appendix D online).

4.2.2 Validation of density forecast model

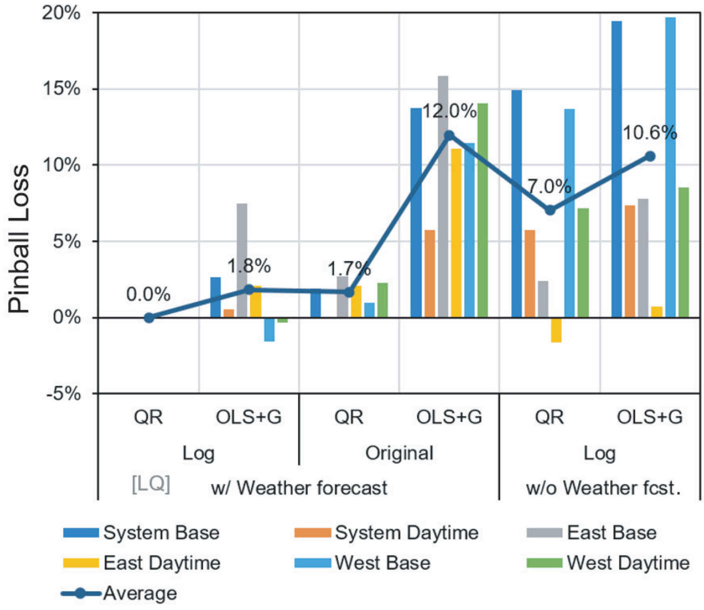
4.2.2.1 Pinball loss comparison between models. The goodness of fit of the density forecast model was evaluated using the pinball loss (PL) score (see, for example, Nowotarski and Weron 2018); we calculated the average PL of 19 percentiles ($q = 0.05, 0.01, \dots, 0.95$). Figure 7 shows the difference in the ratio of the PL of each model M to the PL of the model based on the log-price series QR model (the "LQ model"): $\mathrm{PL}_M/\mathrm{PL}_{\mathrm{LQ}} - 1$.

First, regarding the comparison between QR and OLS + G, the PL was smaller overall for QR, which can simulate an arbitrary density shape, regardless of whether logarithmic conversion was performed (this result corresponds to issue (iii) in Section 1).

Second, the log-price models generally have higher forecast accuracies than the price series models for both QR and OLS + G (this result corresponds to issue (ii) in Section 1). This result may be natural for the OLS + G model in light of the fact that the weekly price distributions are close to the lognormal distributions, as seen in Figure 1. On the other hand, even for the QR model that can forecast arbitrary density shapes, the result where logarithmic transformation improved the forecast accuracy suggests the superiority of the multiplicative model over the additive model.

4.2.2.2 Reduction of PL by weather forecast. Here, in order to help solve issue (iv) in Section 1, we consider the extent to which incorporating the weekly weather forecast can contribute to reducing the forecast error (PL, in this case) for each season

FIGURE 7 Comparison of PL in modeling approaches.



(month). Our problem awareness is similar to that in Huurman *et al* (2012), who verified this effect using the daily average price. As shown in Figure 8, when the weather forecast is incorporated (pink line), the forecast error is reduced, especially in summer (and slightly in winter), compared with when it is not incorporated (blue line). The reason why the forecast accuracy was hardly improved by the weather information in spring and autumn may be because the price (demand) has almost no relation to the temperature.

4.3 Profit by arbitrage

This section verifies the difference in arbitrage profit and loss risk between the various trading strategies by using the average-forecast-based OLS and quantile-forecast-based QR methods, each constructed in Section 3. Table 1 summarizes the profit and loss risk of arbitrage trading for each of the six cases (where OLS uses the L2 model and QR uses the LQ model, both incorporating the weather forecast). In Table 1, N is the number of arbitrage opportunities (ie, the number of observed weekly forward prices during the out-of-sample period), N_{arb} is the number of times the arbitrage was executed (ie, when Y is nonzero), N_{loss} is the number of times a loss

FIGURE 8 PL reduction effect from weather forecast (monthly average).

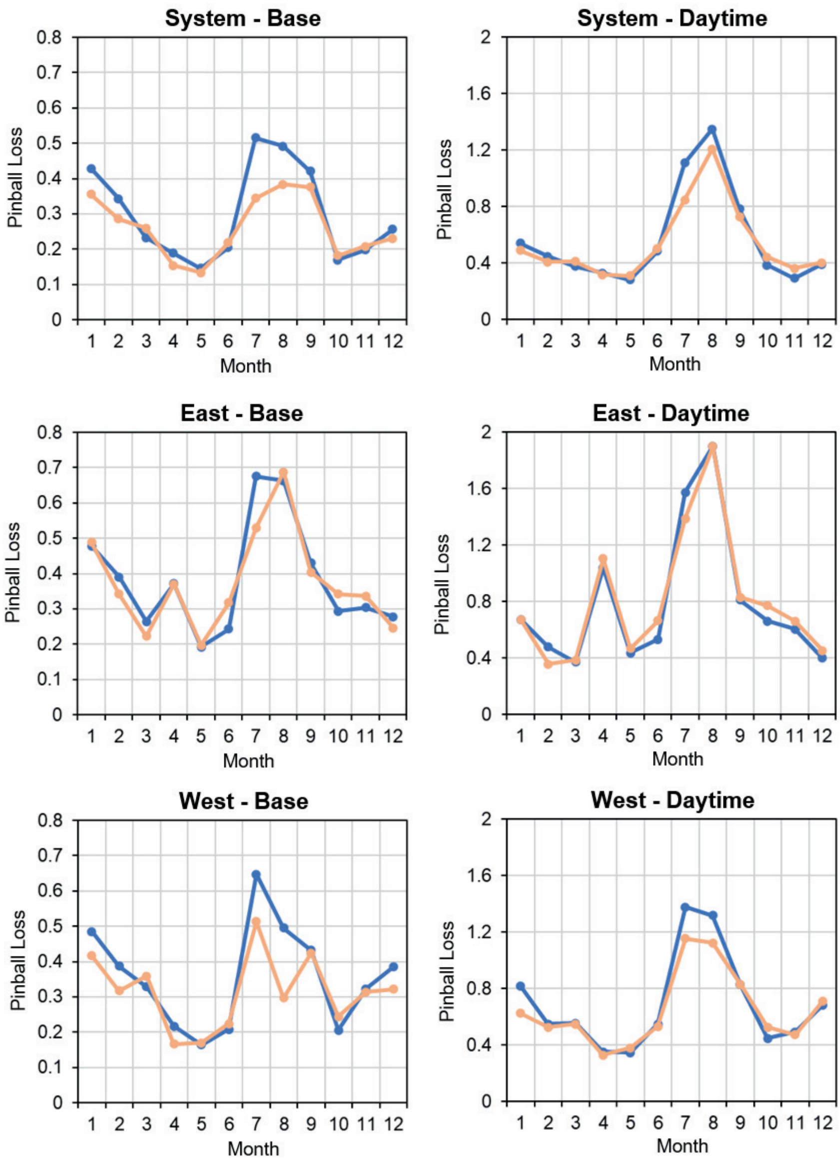


TABLE 1 Profit and loss risk brought about by arbitrage. [Table continues on the next page.]

(a) Base load

Area	N	Model	Percentile (%)	N_{arb}	N_{loss}	Loss (%)	VaR 95%	CVaR 95%	Mean profit	Mean profit (%)	p-value
System	35	QR	5	19	2	6	−0.18	−0.28	0.59	(6.6)	0.0013**
			10	23	2	6	−0.18	−0.28	0.82	(9.1)	0.0002**
			25	28	2	6	−0.18	−0.28	0.96	(10.7)	0.0000**
		OLS	—	35	6	17	−2.35	−2.45	0.81	(9.0)	0.0012**
East	17	QR	5	3	0	0	0.00	0.00	0.23	(2.3)	0.1227
			10	4	0	0	0.00	0.00	0.24	(2.4)	0.1005
			25	10	2	12	−1.70	−1.70	0.73	(7.4)	0.0525*
		OLS	—	17	5	29	−1.70	−1.70	1.04	(10.6)	0.0437**
West	7	QR	5	2	0	0	N/A	N/A	0.32	(3.7)	0.2949
			10	3	0	0	N/A	N/A	0.57	(6.7)	0.1418
			25	6	0	0	N/A	N/A	0.89	(10.4)	0.0192**
		OLS	—	7	0	0	N/A	N/A	1.07	(12.5)	0.0043**

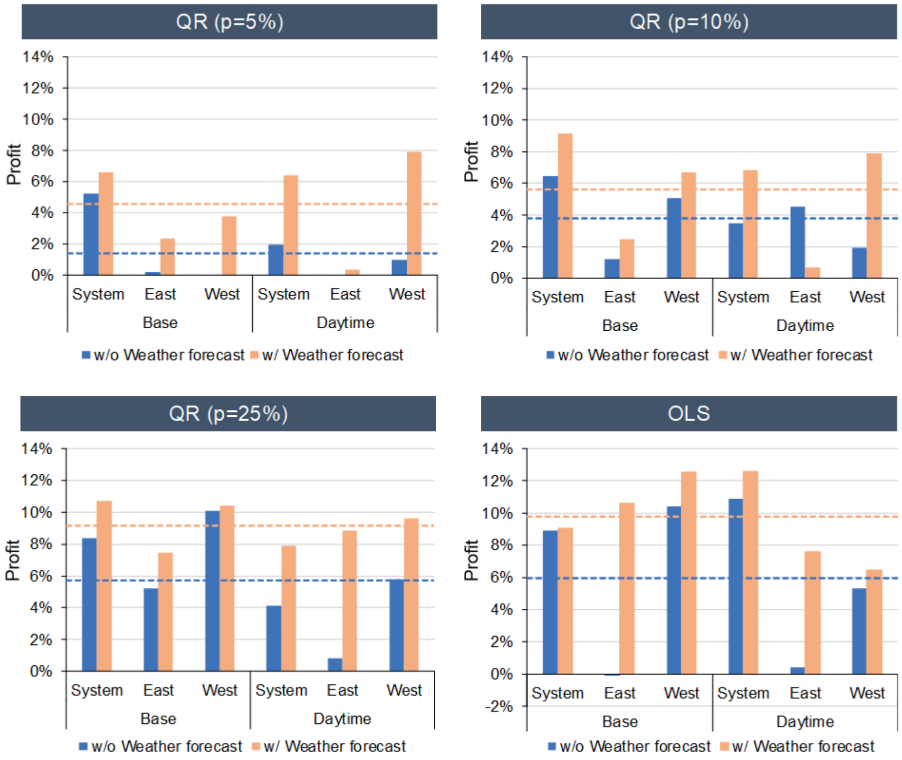
TABLE 1 Continued.

(b) Daytime load

Area	N	Model	Percentile (%)	N_{arb}	N_{loss}	Loss (%)	VaR 95%	CVaR 95%	Mean profit	Mean profit (%)	p-value
System	78	QR	5	19	2	3	0.00	−0.84	0.68	(6.4)	0.0022**
			10	25	5	6	−0.42	−1.42	0.73	(6.9)	0.0020**
			25	42	13	17	−1.89	−2.18	0.84	(7.9)	0.0021**
		OLS	—	78	25	32	−1.75	−2.09	1.34	(12.6)	0.0000**
East	19	QR	5	2	0	0	0.00	0.00	0.04	(0.3)	0.1901
			10	3	0	0	0.00	0.00	0.08	(0.7)	0.1261
			25	10	2	11	−0.88	−0.88	1.05	(8.8)	0.0776*
		OLS	—	19	6	32	−12.81	−12.81	0.9	(7.6)	0.3926
West	22	QR	5	6	0	0	0.00	0.00	0.82	(7.9)	0.0215**
			10	6	0	0	0.00	0.00	0.82	(7.9)	0.0215**
			25	13	3	14	−1.42	−1.42	1.00	(9.6)	0.0125**
		OLS	—	22	7	32	−6.43	−6.43	0.67	(6.5)	0.2238

The units of VaR95% and CVaR95% are ¥/kWh. The unit of mean profit is ¥/kWh. The percentage values in parentheses are the ones obtained by dividing the mean profits by the average spot prices during the out-of-sample period. The p-value is for the mean profit. ** and * denote statistical significance at the 5% and 10% confidence levels, respectively.

FIGURE 9 Difference in average arbitrage profit with and without weather forecast.



was incurred, and Loss is the ratio of N_{loss} to N. VaR95% means the 95th percentile value of the distribution of arbitrage profit (loss if minus) and CVaR95% means the average of the loss that was below the 95th percentile.

As a whole, when the quantile forecast using QR is performed, the risk of loss due to arbitrage (N_{loss}, VaR95% and CVaR95%) can be significantly suppressed compared with the OLS case. On the other hand, in terms of the average profit, OLS generally has a larger profit; thus, there is a trade-off between risk and return. In arbitrage using OLS, the average arbitrage profit should be maximized because the decision is made to always take the position where the expected return is positive regardless of the loss risk; hence, the calculation results obtained here are consistent. In addition, the fact that there are significant arbitrage opportunities of around 1 ¥/kWh on average (which corresponds to around 10% when viewed as a proportion of the average spot prices of each load and area during the out-of-sample period) may reflect the situation such that the JEPX forward market is quite illiquid and that few market participants will find extreme prices (far from the fair prices) and carry out arbitrage trading.

Finally, in order to evaluate the effect of using the weekly weather forecast from the viewpoint of improving the arbitrage profit, we also separately calculated the arbitrage profit obtained from the forecast model without the weather forecast, and we summarize the average profits of all the cases in Figure 9. The pink bar graph matches the value in parentheses for mean profit in Table 1 and the dotted line is a simple average profit with or without the weather forecast. In all the four cases of using OLS and QR, the arbitrage profit was significantly improved when the weekly weather forecast was used. We suggest that improving the accuracy of electricity price forecasts by incorporating the temperature forecast information could greatly contribute to more economic market trading.

As described above, it was empirically shown that the proposed arbitrage method works effectively in the JEPX market and can adjust the balance between risk and earned profit according to the player's risk tolerance. In particular, the empirical analysis revealing that weekly weather forecasts greatly contribute to higher transaction profit clearly indicates the information value of weather forecasts. The increment in earned profit due to weather forecasts (1.5–3.3 times for each strategy) is a significant amount compared with the reduction in forecasting error, which may mean that even a marginal improvement in forecast accuracy can significantly increase profitability in market trading.

5 CONCLUSION

In this study, we constructed a model using weekly weather forecasts for forecasting week-ahead average electricity prices, which has hardly been studied in the past, and we applied it to an arbitrage strategy in the forward market, demonstrating its effectiveness by using six different cases of market data from JEPX. The empirical analyses of this research clarified the following new perspectives on unresolved questions, yielding useful suggestions and ideas for the development and practical use of the model.

(1) When forecasting the weekly average price, instead of using the weather forecast directly as an explanatory variable, the two-step forecast approach, where the measured temperature values mediate, is more advantageous in reducing the forecast error.

(2) When forecasting the price density, QR, which can directly forecast the percentile values of any density shape, has better forecast accuracy than GARCH, which assumes a lognormal distribution.

(3) Common to both average forecast and density forecast, the multiplicative model using the logarithmic series tends to have greater forecast accuracy overall than the additive model using the original price series.

(4) The weekly weather forecast improves the forecast accuracy of the weekly average electricity price for the following week and plays an important role in earning profits in forward market trading.

Of these, regarding the case of using a log series in the two-step forecast (point 1 above), we pointed out that an estimation bias may occur if the probability density of the log price due to meteorological prediction errors and price residuals is not considered, and we provided a method to correct it, demonstrating its effectiveness.

Weekly weather forecasts are generally thought not to be particularly accurate due to the long forecast horizon, as also demonstrated in this paper, and so they may be overlooked by practitioners who forecast electricity prices, as evidenced by the lack of previous studies. However, the value of such information can be enhanced by devising modeling methods such as the logarithmic conversion and the two-step forecast approach proposed in this study. In particular, we showed that using weather forecasts can increase the arbitrage profit by 1.5–3.3 times.

The proposed arbitrage trading method that can consider the loss risk is likely to be used by many businesses as the strategy can be flexibly changed to suit the level of acceptable risk. In addition, our forecasting method using publicly available weather forecasts has wide applicability in many markets. It can be constructed even for the (relatively recently liberalized) markets where forecast information such as demand and renewable energy generation is not available. In future work, we will further explore whether our price forecasting and trading methods based on weather forecasts can also be applied to different time granularity or horizons.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES

Aihara, R., and Yokoyama, A. (2016). Optimal weekly operation scheduling of pumped storage hydropower plant considering optimal hot reserve capacity in a power system with a large penetration of photovoltaic generation. *Electrical Engineering in Japan* **195**(2), 35–48 (https://doi.org/10.1002/eej.22820).

Barquín, J., and Vázquez, M. (2008). Cournot equilibrium calculation in power networks: an optimization approach with price response computation. *IEEE Transactions on Power Systems* **23**(2), 317–326 (https://doi.org/10.1109/TPWRS.2008.919198).

Bastian, J., Zhu, J., Banunarayanan, V., and Mukerji, R. (1999). Forecasting energy prices in a competitive market. *IEEE Computer Applications in Power* **12**(3), 40–45 (https://doi.org/10.1109/67.773811).

Bello, A., Bunn, D. W., Reneses, J., and Muñoz, A. (2017). Medium-term probabilistic forecasting of electricity prices: a hybrid approach. *IEEE Transactions on Power Systems* **32**(1), 334–343 (https://doi.org/10.1109/TPWRS.2016.2552983).

Bigerna, S. (2018). Estimating temperature effects on the Italian electricity market. *Energy Policy* **118**, 257–269 (https://doi.org/10.1016/j.enpol.2018.03.068).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**(3), 307–327 (https://doi.org/10.1016/0304-4076(86)90063-1).

Boogert, A., and Dupont, D. (2005). On the effectiveness of the anti-gaming policy between the day-ahead and real-time electricity markets in the Netherlands. *Energy Economics* **27**(5), 752–770 (https://doi.org/10.1016/j.eneco.2005.03.006).

Bunn, D. W., and Kermer, S. O. (2021). Statistical arbitrage and information flow in an electricity balancing market. Working Paper, January 22, Social Science Research Network (https://doi.org/10.2139/ssrn.3101716).

Chen, J., Deng, S. J., and Huo, X. (2008). Electricity price curve modeling and forecasting by manifold learning. *IEEE Transactions on Power Systems* **23**(3), 877–888 (https://doi.org/10.1109/TPWRS.2008.926091).

Eydeland, A., and Wolyniec, K. (2003). *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley.

Green, R. J., and Newbery, D. M. (1992). Competition in the British electricity spot market. *Journal of Political Economy* **100**(5), 929–953 (https://doi.org/10.1086/261846).

Huurman, C., Ravazzolo, F., and Zhou, C. (2012). The power of weather. *Computational Statistics and Data Analysis* **56**(11), 3793–3807 (https://doi.org/10.1016/j.csda.2010.06.021).

JEPX (2019). Japan electric power exchange guide. (In Japanese.) URL: https://www.jepx.org/outline/pdf/Guide_2.00.pdf.

JEPX (2020). Japan electric power exchange trading regulations. (In Japanese.) URL: https://www.jepx.org/outline/pdf/tr_rules.pdf.

Just, S., and Weber, C. (2015). Strategic behavior in the German balancing energy mechanism: incentives, evidence, costs and solutions. *Journal of Regulatory Economics* **48**(2), 218–243 (https://doi.org/10.1007/s11149-015-9270-6).

Karakatsani, N. V., and Bunn, D. W. (2008). Forecasting electricity prices: the impact of fundamentals and time-varying coefficients. *International Journal of Forecasting* **24**(4), 764–785 (https://doi.org/10.1016/j.ijforecast.2008.09.008).

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press (https://doi.org/10.1017/CBO9780511754098).

Krishnamurthy, D., Uckun, C., Zhou, Z., Thimmapuram, P. R., and Botterud, A. (2017). Energy storage arbitrage under day-ahead and real-time price uncertainty. *IEEE Transactions on Power Systems* **33**(1), 84–93 (https://doi.org/10.1109/TPWRS.2017.2685347).

Loi, T. S. A., and Le Ng, J. (2018). Anticipating electricity prices for future needs: implications for liberalised retail markets. *Applied Energy* **212**, 244–264 (https://doi.org/10.1016/j.apenergy.2017.11.092).

Maciejowska, K., Nitka, W., and Weron, T. (2019). Day-ahead vs. intraday: forecasting the price spread to maximize economic benefits. *Energies* **12**(4), 631 (https://doi.org/10.3390/en12040631).

Mandal, P., Senjyu, T., Urasaki, N., Funabashi, T., and Srivastava, A. K. (2006). Short-term price forecasting for competitive electricity market. In *2006 38th North American Power Symposium*, pp. 137–141. IEEE Press, Piscataway, NJ (https://doi.org/10.1109/NAPS.2006.360135).

Marcos, R. A. D., Bunn, D. W., and Bello, A., Reneses, J. (2020). Short-term electricity price forecasting with recurrent regimes and structural breaks. *Energies* **13**(20), 5452 (https://doi.org/10.3390/en13205452).

Matsumoto, T., and Endo, M. (2021). Empirical analysis of risk premium in the Japanese electricity futures market. Report Y20004, Central Research Institute of Electric Power Industry. (In Japanese.)

Matsumoto, T., and Yamada, Y. (2021). Simultaneous hedging strategy for price and volume risks in electricity businesses using energy and weather derivatives. *Energy Economics* **95**, Paper 105101 (https://doi.org/10.1016/j.eneco.2021.105101).

Nowotarski, J., and Weron, R. (2018). Recent advances in electricity price forecasting: a review of probabilistic forecasting. *Renewable and Sustainable Energy Reviews* **81**, 1548–1568.

Popa, R., Popa, F., Popa, B., and Zachia-Zlatea, D. (2010). Optimization of the weekly operation of a multipurpose hydroelectric development, including a pumped storage plant. *IOP Conference Series: Earth and Environmental Science* **12**(1), Paper 012118 (https://doi.org/10.1088/1755-1315/12/1/012118).

Suzuki, M. (1991). Power supply and demand operation plan. *Measurement and Control* **30**(7), 557–564.

Torró, H. (2007). Forecasting weekly electricity prices at Nord Pool. Working Paper 88, June 7, Social Science Research Network (https://doi.org/10.2139/ssrn.991532).

Ventosa, M., Ballo, A., Ramos, A., and Rivier, M. (2005). Electricity market modeling trends. *Energy Policy* **33**(7), 897–913 (https://doi.org/10.1016/j.enpol.2003.10.013).

Weron, R. (2014). Electricity price forecasting: a review of the state-of-the-art with a look into the future. *International Journal of Forecasting* **30**(4), 1030–1081 (https://doi.org/10.1016/j.ijforecast.2014.08.008).

Zhang, F., Sun, Y. Q., Magnusson, L., Buizza, R., Lin, S. J., Chen, J. H., and Emanuel, K. (2019). What is the predictability limit of midlatitude weather? *Journal of the Atmospheric Sciences* **76**(4), 1077–1091 (https://doi.org/10.1175/JAS-D-18-0269.1).