Probabilistic forecasting of electricity prices using an augmented LMARX-model

BY Jonas Andersson and Samaneh Sheybanivaziri

DISCUSSION PAPER

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 11/2023

ISSN: 2387-3000
July 2023

Probabilistic forecasting of electricity prices using an augmented LMARX-model

Jonas Andersson* and Samaneh Sheybanivaziri†

Norwegian School of Economics, Department of Business and Management Science

July 10, 2023

Abstract

In this paper, we study the performance of prediction intervals in situations applicable to electricity markets. In order to do so we first introduce an extension of the logistic mixture autoregressive with exogenous variables (LMARX) model, see (Wong, Li, 2001), where we allow for multiplicative seasonality and lagged mixture probabilities. The reason for using this model is the prevalence of spikes in electricity prices. This feature creates a quickly varying, and sometimes bimodal, forecast distribution. The model is fitted to the price data from the electricity market forecasting competition GEFCom2014. Additionally, we compare the outcomes of our presumably more accurate representation of reality, the LMARX model, with other widely utilized approaches that have been employed in the literature.

Keywords— Prediction intervals, probabilistic forecasts, electricity prices, spikes, mixture models

1 Introduction

Forecast intervals and, more generally, probabilistic forecasts, are important when making decisions about future production and consumption of electricity. It allows the decision-makers in the market to determine not only which outcome they should expect from their decision but also the risks they are taking. This realization is probably the reason why the literature on probabilistic forecasts has grown significantly in the last decade, see, e.g. the extensive review by (Nowotarski, Weron, 2018).

A particular problem for probabilistic forecasting in electricity markets is the relatively frequent spikes. Several interesting questions arise because of these. Firstly, the ability to predict spike occurrence is, maybe, the golden grail of electricity price forecasting. A decision maker or trader with a superior ability to predict spikes will certainly be a very successful actor in the electricity market. A prominent example of this strand of the literature is (Christensen et al., 2012), which uses an autoregressive conditional hazard (ACD) model to estimate the conditional likelihood of a spike

*jonas.andersson@nhh.no
†samaneh.sheybanivaziri@nhh.no

occurrence. Other works focusing on spike predictions are (Amjady, Keynia, 2010) and (Amjady, Keynia, 2011). The latter two papers focus on both the probability and size of spikes.

Secondly, the issue of what a reasonable measure of risk is for a commodity with two such distinct regimes, regular and spike prices, surfaces. As an example, variance, often used in financial modeling, is not a proper measure of risk. There is certainly a big difference in risk profile between, say, a normally distributed, $N(2, 100)$, return distribution, and a return, distributed as a mixture of a $N(1, 3)$-variate and the value 100 with probability 0.01; even if they have almost identical expectations and the same variance.

Thirdly, and the main topic of this paper is that being able to model, or even predict spikes are of enormous use when probabilistic forecasts of electricity prices are needed. As part of their extensive review of probabilistic forecasting for electricity prices, (Nowotarski, Weron, 2018) present a study on the performance of methods to compute prediction intervals in this context. This was done using real-world data from the electricity forecast competition, GEFCom2014, that can be accessed in the supplementary data of (Hong et al., 2016). In this paper, the performance of some commonly used methods will, instead, be investigated by means of a Monte Carlo study.

In the past few years, research within the area of probabilistic electricity price forecasting has been quite active, as highlighted by (Nowotarski, Weron, 2018). (Hong et al., 2016) review advances in probabilistic energy forecasting methods. The paper also serves as an introduction to the Global Energy Forecasting Competition 2014 (GEFCom2014). The data from this competition is used in our paper. In the paper by (Dudek, 2016), the author suggests, and investigates, a method, based on a feed forward neural network to produce probabilistic forecasts. One argument for the method is that no preprocessing, such as detrending, is required. While dealing with probabilistic forecasts of another variable, wind power generation, the work by (Wan et al., 2013) is relevant to ours. Just like electricity prices, wind power generation does not have a smooth probability density. For electricity prices, a reason for this are the spikes and, to some extent, negative prices. For wind power production, the reason is, however, something else. The wind power turbines produce zero power for wind speeds below and, respectively, above certain thresholds. (Wan et al., 2013) use extreme learning machine (ELM), see (Huang et al., 2006), to produce prediction intervals for this application. The paper by (Botterud et al., 2012) introduces a model for optimal trading of wind power in the day-ahead market under uncertainty in wind power and prices. They highlight the importance of probabilistic prediction in the electricity market context. The paper also applies a kernel density method to compute the predictive distribution of the price. The application of probabilistic forecasts in optimal trading is, in our view, a topic that should be pursued further in the literature. It is, however, beyond the scope of our present work. In (Nowotarski, Weron, 2015), the quantile regression averaging (QRA) approach, is introduced. In QRA, the prices are regressed on a set of point forecasts in a quantile regression in order to obtain prediction intervals. (Maciejowska et al., 2016) expand QRA for cases with an exuberant number of forecast models as inputs. This is done by means of principal component analysis. (Gaillard et al., 2016) present their results from their participation in the GEFCom2014 competition. They found that a generalized additive model fitted by the fitting function from quantile regression was the most successful method for probabilistic forecasts for both load and price. In contrast to the methods described so far, possibly with the exception of (Botterud et al., 2012), the paper by (Ziel, Steinert, 2016) takes a more structural starting point to the problem. They use real auction data and arrive at probabilistic forecasts of the supply and demand curves. This approach is then extended to accommodate for long-term forecasts in (Ziel, Steinert, 2018). (Brusaferri et al., 2019) use a Bayesian neural network where a prior distribution is assumed on the weights of the network. A potential benefit of this is the possibility to evaluate model uncertainty.

(Muniain, Ziel, 2020) utilize autoregressive models with exogenous variables and modified error terms in order to capture price spikes. Their evaluation, considering interconnected peak and off-peak time series, identified the energy score (ES) as a measure for prediction accuracy. Using (Anderson, Davison, 2008)'s methodology, (Maryniak, Weron, 2020) model the probability of price spike occurrence as an increasing function of the demand-to-capacity ratio. They find a positive correlation between the likelihood of price spikes and an increase in the demand-to-capacity ratio, regardless of the spike detection method used. In addition to positive price spikes, probabilistic forecasting of extremely low prices has gained attention due to its relevance in market analysis and decision-making processes. Notably, the paper by (Bello et al., 2016), presents a method to predict the occurrence of extremely low prices and applies it to data from the Spanish wholesale market. The method combines different forecasting and spatial interpolation techniques with Monte Carlo simulation.

Our paper has two main contributions. Our first contribution is to add features, not included in the original LMARX model by (Wong, Li, 2001) and show how to estimate such a model. The second is to complement the study by (Nowotarski, Weron, 2018) by investigating coverage probability and length of prediction intervals by means of a Monte Carlo study where we simulate data from the logistic mixture autoregressive model with exogenous variables (LMARX), (Wong, Li, 2001). The benefit of doing this is the ability to study the performance of different models in a situation when the data-generating process is fully known. The values of the parameters in the Monte Carlo setup are guided by the parameter values obtained when the model is fitted to the GEFCom2014 data.

In the next section of the paper, we describe the model by (Wong, Li, 2001), how to estimate it, some properties of it, and why it is a good description of electricity prices. We also present extensions to allow for seasonality and lagged spike probabilities. We describe how to estimate the model using the R-package TMB, see (Kristensen et al., 2016). In Section 3 we fit the LMARX-model to the GEFCom2014 data using forecasted load as exogenous variables. In Section 4 we present other, commonly used, models used for forecasting electricity prices. This fitted LMARX-model, argued to be a realistic approximation for the data generating process of electricity prices, is then used in Section 5 where forecast intervals from the LMARX-model and models from Section 4 are evaluated by means of a Monte Carlo Study. A conclusion ends the paper.

2 An extension of the logistic mixed autoregressive model with exogenous variables (LMARX)

We extend the model by (Wong, Li, 2001) to capture the most important features of electricity price time series, namely price spikes, seasonality and explanatory variables, such as predicted load. The general model is specified by

$$p_t = \begin{cases} y_{0,t} \text{ if } z_t = 0 \\ y_{1,t} \text{ if } z_t = 1 \end{cases} \tag{1}$$

where z_t is a Bernoulli variable determining if the market is in a normal state or in a spike state. The probability of being in a spike state can potentially be predictable, given knowledge of a variable, $\boldsymbol{v}_t$, so that

$$P(z_t = 1) = \alpha(\boldsymbol{v}_t; \beta_0, \boldsymbol{\beta_1}, \delta_1) = \frac{1}{1 + \exp(-(\beta_0 + \boldsymbol{v}_t'\boldsymbol{\beta_1} + \delta_1 \alpha_{t-1}))} \tag{2}$$

where the variable $\boldsymbol{v}_t$ is observable at the time of prediction. α_{t-1} is a short notation for the lagged spike probability $\alpha(\boldsymbol{v}_{t-1};\beta_0,\boldsymbol{\beta_1},\delta_1)$. Note that lags of y_t can be included in $\boldsymbol{v}_t$. Within each of the two states, a seasonal ARX model, ARX$(p)(P)_s$, model

$$\phi_k(B)\Phi_k(B)(y_{k,t}-\mu_k)=\boldsymbol{x}_t'\boldsymbol{\gamma}+\varepsilon_{k,t}, \tag{3}$$

is governing the dynamics of the process, $k=0,1$. $\varepsilon_{k,t}\sim N(0,\sigma_k^2)$ are normally distributed white noise processes, B is the backshift operator, $\boldsymbol{x}_t$ is a vector of covariates,

$$\begin{aligned}\phi_k(B)&=1-\phi_{k,1}B-\phi_{k,2}B^2-...-\phi_{k,p}B^p \text{ and}\\ \Phi_k(B)&=1-\Phi_{k,1}B^s-\Phi_{k,2}B^{2s}-...-\Phi_{k,P}B^{Ps}.\end{aligned} \tag{4}$$

The model described above is an extended version of the model by (Wong, Li, 2001) in two ways. Firstly, we allow for a multiplicative seasonality and, secondly, for a lagged probability in the dynamics of the spike probability[1]. The EM-algorithm developed in (Wong, Li, 2001) is therefore not directly applicable. We estimate the model through the R-package (R Core Team, 2022), TMB (Kristensen et al., 2016), which employs an automatic differentiation algorithm, thus allowing the computation of exact gradients. The conditional density function for y_t is

$$\begin{aligned}f(y_t|\mathcal{F}_{t-1},\boldsymbol{x}_t,\boldsymbol{v}_t,\boldsymbol{\phi},\boldsymbol{\Phi},\boldsymbol{\gamma},\boldsymbol{\sigma},\beta_0,\boldsymbol{\beta_1})=(1-\alpha(\boldsymbol{v}_t;\beta_0,\boldsymbol{\beta_1},\delta_1))\phi\left(\frac{y_t-\mu_0(\mathcal{F}_{t-1},\boldsymbol{x}_t)}{\sigma_0}\right)\\ +\alpha(\boldsymbol{v}_t;\beta_0,\boldsymbol{\beta_1},\delta_1)\phi\left(\frac{y_t-\mu_1(\mathcal{F}_{t-1},\boldsymbol{x}_t)}{\sigma_1}\right)\end{aligned} \tag{5}$$

where $\mu_k(\mathcal{F}_{t-1},\boldsymbol{x}_t)$ are the conditional means of y_t for $k=0,1$, $\mathcal{F}_{t-1}=\{y_{t-1},y_{t-2},...\}$ and ϕ is the standard normal density function. The conditional log-likelihood can then be computed as

$$\log L(\boldsymbol{\phi},\boldsymbol{\Phi},\boldsymbol{\gamma},\boldsymbol{\sigma},\beta_0,\boldsymbol{\beta_1}))\sum_{t=Ps+1}^{T}\log f(y_t|\mathcal{F}_{t-1},\boldsymbol{x}_t,\boldsymbol{v}_t,\boldsymbol{\phi},\boldsymbol{\Phi},\boldsymbol{\gamma},\boldsymbol{\sigma},\beta_0,\boldsymbol{\beta_1}) \tag{6}$$

which is then maximized using TMB. The model, with around 1000 observations, is estimated in ca 0.2 seconds with a standard 2023 desktop computer.

The predictive probability density of the model, $f(y_{T+h}|\mathcal{F}_T,\boldsymbol{x}_T,\boldsymbol{v}_T)$, is obtained by simulating $y_{T+1},y_{T+2},...,y_{T+h}$ from the model . The values of $\boldsymbol{x}_T$ and $\boldsymbol{v}_T$ must be known or predicted. Modeling a potential endogeneity, so that variables, such as load, could be dependent on previous price or spike probability, would be very interesting but is beyond the scope of this paper.

3 The seasonal LMARX-model fitted to the GEF686Com2014 data

We first fit the model to the price data from the Global Energy Forecasting Competition 2014 and use seasonally adjusted forecasts of zonal and total loads as variables in both $\boldsymbol{x}$ and $\boldsymbol{v}$. v_{1t} is forecasted zonal load and v_{2t} is forecasted total load at time t. The processes $\{y_{k,t}\}$, $k=0,1$ are both seasonal ARX(1)-models. The purpose of the data analysis here is to choose realistic parameter values for the model used in the Monte Carlo study, performed in the next section.

Since the prices are set once a day and we consider this as a 24-dimensional daily time series rather than one hourly time series. An example of one such price series, for the hour 8 am, is given in Figure 1. The plot shows that there are

[1] By setting $\delta_1=0$ and $\Phi_k(B)=1$, we obtain the model by (Wong, Li, 2001).

spikes in the prices at some times, which deviate a lot from most other prices.

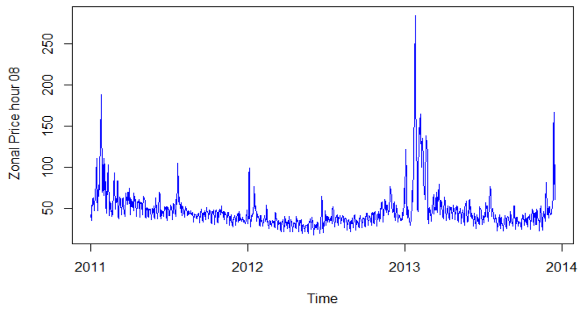


Figure 1: Zonal marginal prices for hour 8 am for the GEFCom2014 data

We also investigate the unconditional distribution of the data for hours 13 and 17, see Figure 2. These two examples show that the unconditional distribution of the price can be close to uni-modal, as for hour 13, or not, as for hour 16. For the uni-modal case, there is a possibility that a model based on a uni-modal distribution can capture the price distribution. Even a model based on a symmetrical distribution, such as the normal distribution, could work after a simple transformation, such as the logarithm, of the data. The bi- or multi-modal case, however, is a strong argument for using a model based on a mixture of distributions. In terms of electricity prices, a bi-modality can be thought of as a process determining prices in a "normal state" and another process determining the prices in a "spike state". This interpretation should not be seen as anything close to a structural model for electricity prices but only as a vehicle to capture the predictive distribution in a more realistic way.

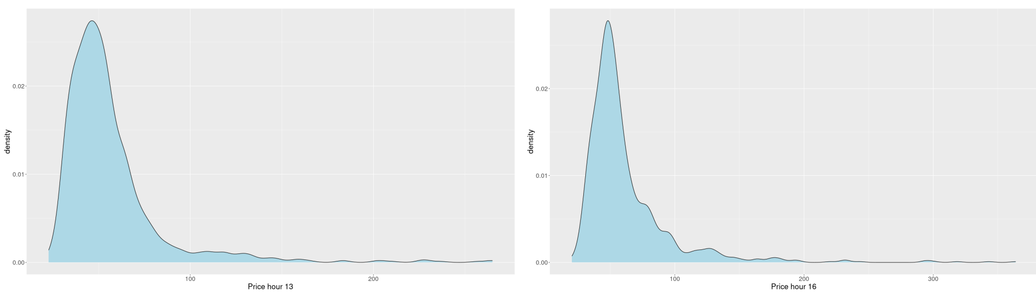


Figure 2: Kernel density estimates for hours 13 and 16.

We now fit the following seasonal LMARX(1) model to the series in Figure 1.

$$\begin{cases} y_{0,t} = \phi_{00} + \phi_{01} y_{t-1} + \Phi_{01} y_{t-7} - \phi_{01}\Phi_{01} y_{t-8} + \gamma_{01} v_{1t} + \gamma_{02} v_{2t} + \varepsilon_{1,t} \\ y_{1,t} = \phi_{10} + \phi_{11} y_{t-1} + \Phi_{11} y_{t-7} - \phi_{11}\Phi_{11} y_{t-8} + \gamma_{11} v_{1t} + \gamma_{12} v_{2t} + \varepsilon_{2,t} \\ y_t = \begin{cases} y_{0t} \text{ if } Z_t = 0 \\ y_{1t} \text{ if } Z_t = 1 \end{cases} \\ \alpha_t = P(Z_t = 1) = \dfrac{1}{1 + \exp(-(\beta_0 + \beta_1 v_{1t} + \beta_2 v_{2t} + \beta_3 y_{t-1} + \delta_1 \alpha_{t-1}))} \end{cases} \tag{7}$$

where y_t is zonal price, v_{1t} is forecasted zonal load and v_{2t} is forecasted total load. The result is shown in Table 1.

	coef	se	tval	pval
ARIMA-parameters				
ϕ_{00}	6.425	0.533	12.047	0.000
ϕ_{10}	14.417	4.251	3.391	0.001
ϕ_{01}	0.615	0.026	23.935	0.000
ϕ_{11}	0.781	0.050	15.471	0.000
Φ_{01}	0.590	0.024	24.461	0.000
Φ_{11}	0.147	0.083	1.764	0.078
γ_{01}	-0.004	0.002	-1.914	0.056
γ_{11}	-0.063	0.027	-2.381	0.017
γ_{02}	0.005	0.001	5.905	0.000
γ_{12}	0.031	0.009	3.314	0.001
σ_0	1.500	0.031	48.937	0.000
σ_1	3.188	0.062	51.476	0.000
Spike parameters				
β_0	-4.892	0.053	91.478	0.000
β_1	-0.006	0.001	4.796	0.000
β_2	0.002	0.001	-3.121	0.002
β_3	0.037	0.000	-96.710	0.000
δ_1	5.366	0.097	-55.319	0.000

Table 1: The model fitted to zonal prices for hour 8 am

As can be seen in Table 1, most parameters are significant in both regimes. The only exceptions are that in the spike regime, seasonality is not significant and in the non-spike regime, the same is true of the forecasted zonal load. What is more interesting is that both forecasted zonal and total load is significantly and positively affecting the probability of a price spike on the next day. Furthermore, an increased price on day t increases the spike probability on day $t+1$, which is manifested in the parameter β_3 and an increased probability of observing a spike on the day t can increase the probability of a spike on the next day which is revealed by δ_1.

Finally, in Figure 3, we illustrate the one-step-ahead predictive distributions for two different days at hour 16. In Figure 3a, we see an example where the resulting distribution is uni-modal, and in Figure 3b, where it is not. This is a feature that the other models used in this paper cannot capture.

(a) (b)

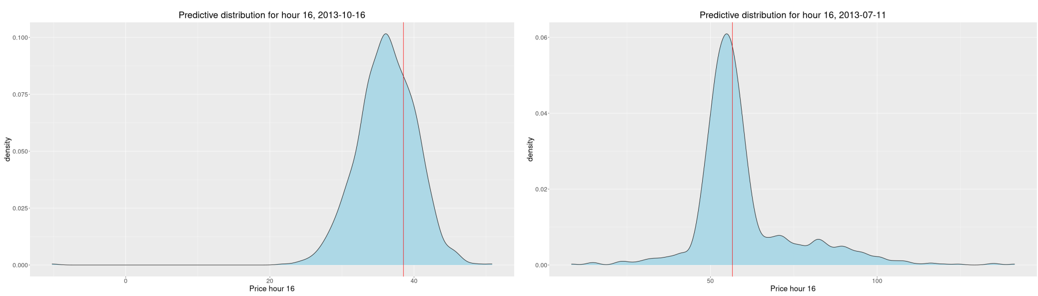


Figure 3: Predictive densities for two different days, hour 16. The vertical red lines are the actual prices.

4 Other common methods to compute forecast intervals

Seasonal autoregressive models were already mentioned in Section 2 when the LMARX model was described. A very common way to produce forecasts is to use an appropriate, special case from a class of models known as the seasonal autoregressive integrated moving average (SARIMA) models. The model can be written

$$\phi(B)\Phi(B)(1-B)^d(1-B^s)^D(y_t-\mu_k)=\theta(B)\Theta(B)\varepsilon_{k,t} \tag{8}$$

where $\{\varepsilon_t\}$ is a strict white noise and $\phi(B)$, $\Phi(B)$, $\theta(B)$ and $\Theta(B)$ are polynomials determining the lag structure of the model

$$\begin{aligned}
\phi(B) &= 1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p,\\
\Phi(B) &= 1-\Phi_1 B^s-\Phi_2 B^{2s}-...-\Phi_P B^{Ps},\\
\theta(B) &= 1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q, \text{ and}\\
\Theta(B) &= 1-\Theta_1 B^s-\Theta_2 B^{2s}-...-\Theta_Q B^{Qs}.
\end{aligned} \tag{9}$$

The model is often denoted as ARIMA$(p,d,q)(P,D,Q)_s$.

The last two polynomials add flexibility in the shape of the autocorrelation (ACF). Compared to the seasonal AR model, can be done in a parsimonious way, i.e., with fewer parameters. Since the model can be written in the so-called moving average (MA) form

$$y_t=\mu+\sum_{k=0}^{\infty}\psi_k\varepsilon_{t-k}, \tag{10}$$

where the coefficients ψ_k, $k=0,1,2...$, are determined by the coefficients in the polynomials above[2], it is straightforward to find the forecast distribution for y_{T+h} given information up until time T. If ε_t is normally distributed, $N(0,\sigma^2)$ so will be the forecast distribution.

$$(y_{T+h}|\mathcal{F}_T)\sim N(f_{T,h},v_h), \tag{11}$$

where

$$f_{T,h}=\mu+\sum_{k=h}^{\infty}\psi_k\varepsilon_{T+h-k} \tag{12}$$

is the conditional expectation of y_{T+h} given $\mathcal{F}_T=(y_T,y_{T-1},...)$ and

$$v_h=\sigma^2(1+\psi_1^2+\psi_2^2+...+\psi_{h-1}^2) \tag{13}$$

is the variance of the forecasting error $y_{T+h}-f_{T,h}$. Based on this result, it is straightforward to derive a forecast interval or the entire forecast distribution. We will now consider some special cases of particular interest to us. We will focus on the one-day-ahead prediction, i.e., $h=1$.

The model denoted **naïve** in this paper can be described by the following sentence: The most likely value in hour H of the next day is the value of hour H today. Framed in the ARIMA framework it would be an ARIMA$(0,1,0)(0,0,0)_7$ model; or more easily written

$$(y_{T+1}|\mathcal{F}_T)\sim N(y_T,\sigma_{\mathrm{N}}^2), \tag{14}$$

where σ_N^2 is the error term variance. A $100(1-\alpha)\%$ forecast interval based on this model is given by

$$y_T\pm z_{\frac{\alpha}{2}}\sigma_{\mathrm{N}} \tag{15}$$

[2] The first coefficient ψ_0 is always one.

where $z_{\frac{\alpha}{2}}$ is the $\alpha/2$ percentile of the $N(0,1)$-distribution. The **seasonal naïve** model can be described by the statement: The most likely value in hour in hour H of next day is the value of hour H on the same weekday last week. This corresponds to an ARIMA$(0,0,0)(0,1,0)_7$. The predictive distribution is then

$$(y_{T+1}|\mathcal{F}_T) \sim N(y_{T-6}, \sigma^2_{\text{SN}}) \tag{16}$$

yielding a $100(1-\alpha)\%$ forecast interval

$$y_T \pm z_{\frac{\alpha}{2}}\sigma_{\text{SN}}, \tag{17}$$

where σ^2_{SN} is the error term variance. We also use the $ARIMA(1,0,1)$-model in our study.

$$y_t = \mu + \phi_1 y_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{18}$$

The predictive distribution is here given by first computing the coefficients, ψ_k, $k = 1, 2, ...$ in (12) and then using them in (11). In Table 2 we summarize the alternatives to the mixture model. The AR(1) and SAR(1) models will be investigated both with and without explanatory variables. In Table 2 the point forecasts for the different models are given.

Forecast method	**One-step-ahead**
Naïve	$\hat{p}_{T+1} = p_T$
Seasonal naïve	$\hat{p}_{T+1} = p_{T-6}$
AR(1)	$\hat{p}_{T+1} = \hat{\mu} + \hat{\phi}_1 p_T$
ARMA(1,1)	$\hat{p}_{T+1} = \hat{\mu} + \hat{\phi}_1 p_T + \hat{\theta}_1 \hat{\epsilon}_T$
SAR(1)	$\hat{p}_{T+1} = \hat{\mu} + \hat{\phi}_1 p_T + \hat{\Phi}_1 p_{T-7} - \hat{\phi}_1 \hat{\Phi}_1 p_{T-8}$

Table 2: One-step-ahead forecasts for some other forecasting methods

To compute multi-step-ahead forecasts, the formulas in Table 2 are iterated. The observed values, i.e., the ones up and until time T are plugged in and the unobserved are substituted with the previous forecast. As an example, for the SAR(1)-model, the two-step-ahead prediction is

$$\hat{p}_{T+2} = \hat{\mu} + \hat{\phi}_1 \hat{p}_{T+1} + \hat{\Phi}_1 p_{T-6} - \hat{\phi}_1 \hat{\Phi}_1 p_{T-7} \tag{19}$$

The forecast intervals will be evaluated by their coverage probability and average length. Both those quantities are computed for both one-step-ahead and seven-step-ahead forecasts made repeatedly over the replications.

5 A Monte Carlo Study

As mentioned previously, the data generating process (DGP) in the simulation study is given by the model (7). The parameter values are guided by the estimates from the observed price series, such as exemplified for hour 08 in Table 1. We have done one such simulation study with 1000 replications for the parameter values corresponding to each of the 24 hours in addition to daily average prices. Since the results are quite consistent over the hours, in this section we include only ones for the daily average prices. Results for one peak hour, hour 18, and one non-peak hour, hour 03, is given in the appendix.

5.1 One-day-ahead forecasts

We first investigate the Monte Carlo results for the one-day-ahead forecasts for a DPG with parameter values corresponding to the ones estimated on daily mean prices. The results are presented in Figure 4 and Table 3.

In Figure 4a, the bias of the empirical coverage, i.e., the actual coverage minus the nominal coverage, is presented. We see that, not surprisingly, by specifying the correct model, we obtain intervals with coverage probabilities close to the nominal. Not equally obvious, the larger the nominal coverage probability, the less important the distributional assumptions seem to be. For small nominal levels, such as 50% and 60%, both the naive models and the ARIMA models with and without exogenous variables, give forecast intervals that are severely conservative, i.e., they cover a large part of the predictive distribution. For large nominal levels, on the other hand, this problem almost disappears. However, we see no reason to advise using these methods since the predictive distribution, seen in total, does not capture future behavior well. As seen in Figure 4b, the intervals for the correctly specified LMARX model, are the shortest, despite the fact that the other models are conservative. At the cost of longer intervals compared to a correctly specified model, bootstrapping, helps, to some extent, for all the miss-specified models.

A pattern that goes through all hours is that seasonality, while often significant in the estimated LMARX models, is not helping in a material way to reduce the length of the intervals for the other models. The length of the intervals is extremely long for the seasonal naive model in addition to, for small nominal levels, not having the correct coverage. When fitting an ARMA(1,1)-model, i.e., without seasonality, we observe that there is little autocorrelation left in the residuals. This is an indication that adding seasonal components will help little and benefits are likely to be outweighed by the disadvantage of more estimation uncertainty. This can also be seen by comparing the bootstrapped AR models with and without seasonality.

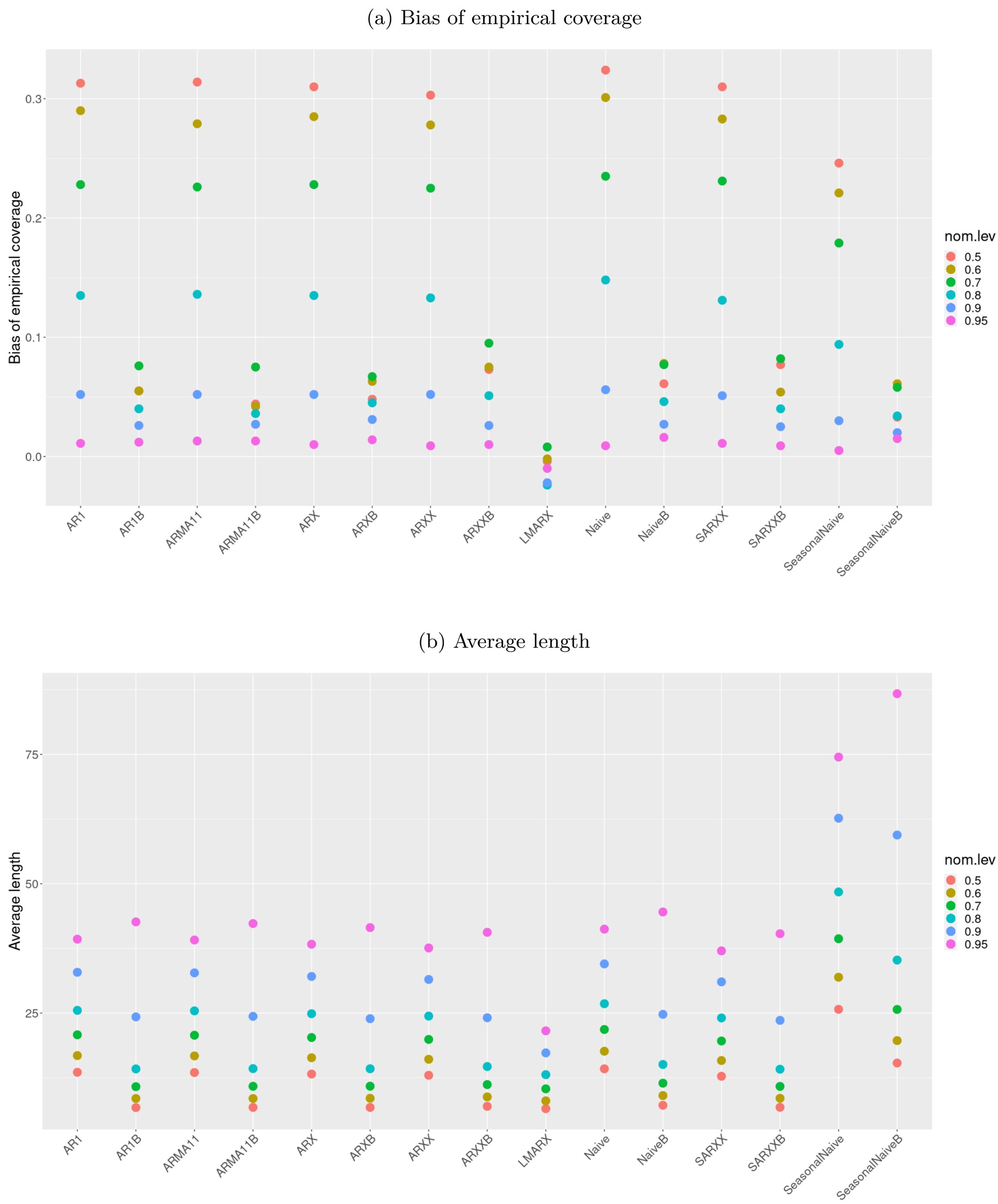


Figure 4: The bias of empirical coverage (a) and the average length of the confidence intervals (b) for the different models and nominal coverage levels, based on daily mean prices. One-step-ahead prediction.

Table 3: Empirical Coverage (EC) and Average Length of 50%, 60%, 70%, 80%, 90%, and 95% two-sided one-step ahead PIs for 15 models, daily mean, 1000 replications. Computation time 1 hour

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	**0.496**	6.52	**0.598**	8.04	**0.708**	10.36	**0.776**	13.10	0.878	17.31	0.940	21.57
Naive	0.824	14.23	0.901	17.63	0.935	21.83	0.948	26.81	0.956	34.51	0.959	41.23
Seasonal Naive	0.746	25.73	0.821	31.93	0.879	39.37	0.894	48.42	0.930	62.67	**0.955**	74.49
AR(1)	0.813	13.56	0.890	16.79	0.928	20.80	0.935	25.53	0.952	32.89	0.961	39.29
ARMA(1,1)	0.814	13.51	0.879	16.72	0.926	20.72	0.936	25.43	0.952	32.77	0.963	39.13
ARX	0.810	13.22	0.885	16.37	0.928	20.28	0.935	24.88	0.952	32.08	0.960	38.31
ARXX	0.803	12.98	0.878	16.07	0.925	19.91	0.933	24.42	0.952	31.50	0.959	37.59
SARXX	0.810	12.79	0.883	15.83	0.931	19.61	0.931	24.06	0.951	31.04	0.961	37.03
Naive-B	0.561	7.18	0.678	9.06	0.777	11.45	0.846	15.07	0.927	24.76	0.966	44.55
Seasonal Naive-B	0.533	15.35	0.661	19.68	0.758	25.71	0.834	35.25	**0.920**	59.42	0.965	86.74
AR(1)-B	0.555	6.74	0.655	8.49	0.776	10.78	0.840	14.21	0.926	24.27	0.962	42.63
ARMA(1,1)-B	0.544	6.77	0.642	8.49	0.775	10.85	0.836	14.27	0.927	24.37	0.963	42.31
ARX-B	0.548	6.77	0.663	8.54	0.767	10.87	0.845	14.24	0.931	23.92	0.964	41.53
ARXX-B	0.573	6.98	0.675	8.80	0.795	11.18	0.851	14.67	0.926	24.10	0.960	40.61
SARXX-B	0.577	6.79	0.654	8.53	0.782	10.83	0.840	14.15	0.925	23.60	0.959	40.35

While specifying the correct model seems like an obvious choice if one ever got the opportunity, it is also well known that estimation uncertainty will affect the quality of the forecasts for finite samples. If a feature of a model is sufficiently weak, it might be better to ignore it. In our case, this could, e.g., occur if the two states of the LMARX-model are sufficiently similar. We do, in fact, observe that some models perform better than the LMARX-models in quite a few cases in our simulation study.

A prominent example of this is the bootstrapped ARMA-models with and without seasonality and explanatory variables. For many hours and nominal levels, they obtained a correct empirical coverage and better or equivalent average length. At a 50% nominal level, LMARX demonstrates superior performance in 9 instances, including specific morning hours (e.g., 3, 9) and afternoon hours (e.g., 13, 17, 18, 20, 21), as well as the daily mean prices, presented in Table 3, in terms of empirical coverage. However, in the remaining hours, bootstrapped models have shorter average lengths. In terms of the average length of prediction intervals, the bootstrapped models consistently yield favorable results across most hours, except for hours 6, 7, and 8, and the daily mean price series, where LMARX generates narrower prediction intervals. These results also apply to the 60%, 70%, and 80% nominal levels for several parameter settings. For the 90 % level, the average lengths are mostly shorter for the LMARX-model and for the 95% level, this is always the case.

The LMARX-model is the only one, of the studied models, that can capture a bi-modal predictive distribution. Therefore, even though ignoring this feature *might* help in obtaining shorter prediction intervals, in the presence of many spikes, such a model would most likely not capture the entire predictive distribution.

5.2 Seven-step-ahead forecasts

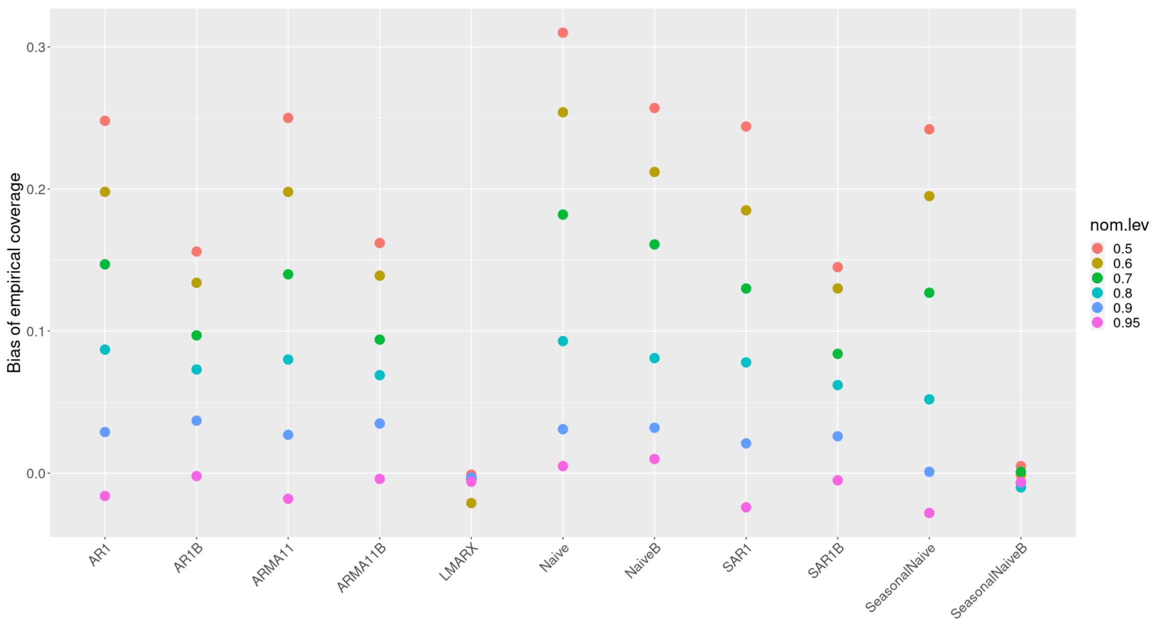
For the sake of avoiding the issue of having to predict covariates, models without covariates were fitted to the observed data in order to get realistic parameter values for studying seven-step-ahead forecasts. The results follow the same pattern as the one-step-ahead predictions in terms of the comparative performance of the different methods.

Again, not surprisingly, specifying the correct model yields superior results and the empirical coverage has the lowest bias, see Figure 5a. Among the models with a close to, correct coverage probability, the correct model has the shortest intervals. There is a tendency that the bootstrap does not correct the coverage probability as well as for the one-step-ahead forecasts. An exception to this is the seasonal naive model with the bootstrap which has correct coverage but much longer intervals, e.g., Figure 5b. This model is solely based on the weekly seasonality. The data is truly generated from the LMARX model so it is interesting to see that a very simplified model, like the naive seasonal, works well. This is not true just for the parameter values based on average prices but a result that goes through for most parameter values investigated (based on the estimated values for real data for different hours).

As the electricity market exhibits hour-specific characteristics, with each hour of the day reflecting distinct patterns influenced by factors such as fluctuating demand, diverse power generation sources, and market participants' activities, it is crucial to analyze each hour individually. When considering a seven-step ahead forecast at a 50% nominal level, the LMARX-model that incorporates the effects of both forecasted total load and forecasted system load demonstrates superior performance compared to all other models in 14 instances, including the daily mean series, see Table 4. However, during peak evening hours, the seasonal naive model with bootstrapped residuals works very well, producing comparable results, see for example Table 8 in the Appendix for hour 18:00. The analysis of average length prediction intervals over 1000 replications reveals a competitive performance between the LMARX and seasonal naive models, particularly in specific time blocks. In the early morning block, spanning six hours from midnight to 5 am, as well as in the afternoon block of seven hours from 12 pm to 6 pm, LMARX consistently produces the narrowest prediction intervals, as shown by the results of hour 03 in Table 7. However, during the second block of the day, encompassing hours 6 am to 11 am, and in the late evening hours from 7 pm to 11 pm, the seasonal naive model outperforms other models. Assuming the non-normality of residuals, the seasonal naive model demonstrates superior performance by effectively capturing the inherent seasonality and patterns present during these time blocks.

In the case of a 60% nominal level, for a 7-step ahead forecast, the empirical coverage of LMARX models is nearly on par with that of seasonal naive-B. When considering the average length of prediction intervals, LMARX outperforms in 17 out of the 25 cases we investigated, including the daily mean as shown in Table 4. LMARX demonstrates superior performance in most instances, except during the second block of the day from 6 pm to 10 pm and during hours 18, 21, and 22 and the seasonal naive model with bootstrapped residuals proves to be more effective during these specific hours which is interesting considering that the data used for analysis is generated from an LMARX model. These findings suggest that LMARX models exhibit competitive empirical coverage, even during peak hours, while outperforming in most cases regarding the average length of prediction intervals. However, during the second block of the day, the seasonal naive model with bootstrapped residuals proves to be more effective in producing narrower prediction intervals. These insights can be valuable in selecting the appropriate model based on the specific hour and desired forecasting performance.

(a) Bias of empirical coverage



(b) Average length

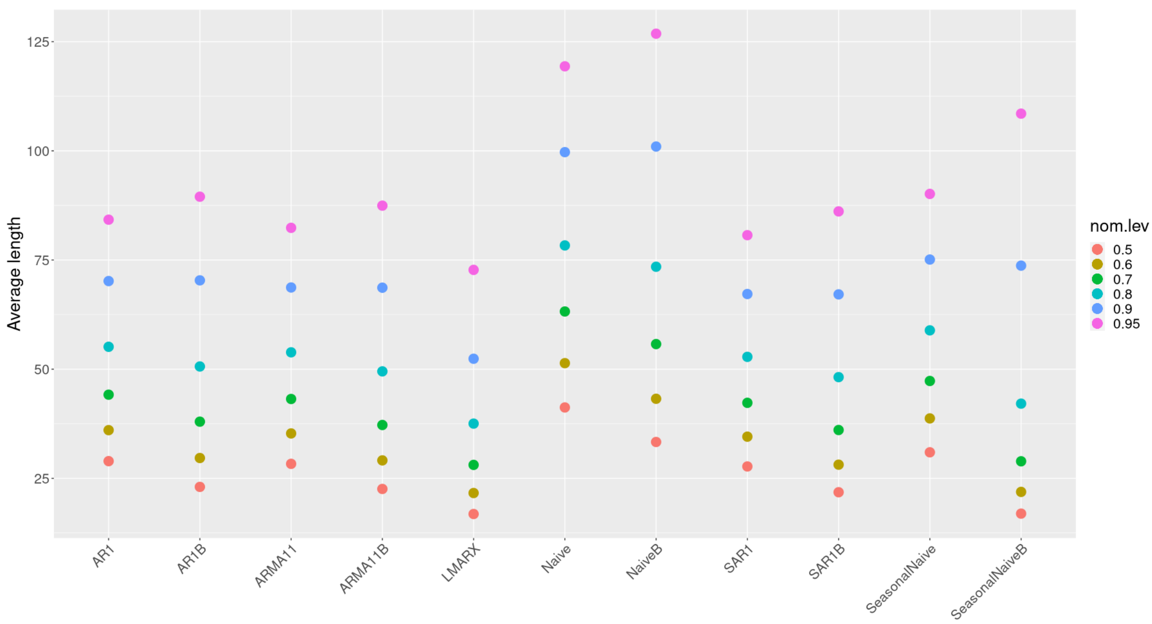


Figure 5: The bias of empirical coverage (a) and the average length of the confidence intervals (b) for the different models and nominal coverage levels, based on daily mean time series. Seven-step-ahead prediction.

At a 70% nominal level, both LMARX and seasonal naive-B models demonstrate effective performance in terms of empirical coverage for 7-step-ahead forecasts. Notably, there are instances, such as hours 4, 7, and 17, where both models achieve identical empirical coverage. However, when evaluating the average length of prediction intervals, LMARX consistently outperforms all other models in all hours except for hour 6, the seasonal naive-B model has a negligible better difference. The same findings persist when considering the nominal level of 80% for 7-step-ahead forecasts as demonstrated by Table 4 or Tables 7 and 8 in the Appendix.

The results also indicate that at the 90% nominal level for 7-step ahead forecasts, LMARX consistently produces the narrowest prediction intervals across all hours. This proves the model's precision and ability to capture the underlying patterns in the data. However, when considering the empirical coverage of point forecasts, LMARX performs effectively alongside seasonal naive models and seasonal naive models with bootstrapped residuals in most hours. Surprisingly, models such as AR(1)-B, seasonal AR(1)-B, and ARMA(1,1) also demonstrate efficient results in some hours, despite the data being generated from an LMARX model. These models showcase their capability to adapt and capture the essential features of the LMARX-generated data.

At a 95% nominal level, analyzing 7-step ahead point forecasts reveals that LMARX exhibits superior performance in terms of the average length of prediction intervals across all hours, except for hours 14 and 16, where seasonal AR(1) shows greater sharpness. However, when considering the empirical coverage of point forecasts, several other models emerge as strong competitors to LMARX. Among these competing models, we observe that naive, AR(1), naive-B, seasonal naive-B, AR(1)-B, ARMA-B, and seasonal AR(1) demonstrate competitive performance in terms of empirical coverage. Notably, AR(1) with bootstrapped residuals proves to be a proficient forecaster in 11 instances, specifically in hours 0, 1, 2, 6, 7, 8, 9, 11, 13, 16, 17, and also the daily mean series, see for instance Table 4. These findings emphasize that, despite the data being generated and simulated by a logistic mixture model with exogenous variables, simpler models can sometimes outperform more complex models in specific hours. While LMARX excels in terms of the average length of prediction intervals, other models demonstrate strong competition in capturing the empirical coverage of point forecasts. This highlights the importance of considering the strengths and limitations of different models in various forecasting scenarios, ultimately enabling the selection of the most appropriate model for each specific hour.

If the features of the electricity price for different hours are consistent over a long period of time and the LMARX-model is a good approximation to them, conclusions could be drawn on when one could, beneficially, use the different ARMA-models successfully. We leave this question for further research. For now, we conclude that the LMARX-model is a safe way to capture situations both with and without many spikes.

Table 4: Empirical Coverage (EC) and Average Length (AL) of 50%, 60%, 70%, 80%, 90% and 95% two-sided 7-step ahead PIs for 11 models, daily mean, 1000 replications. Computation time 2 hours

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	**0.499**	**16.83**	0.579	**21.63**	0.696	**28.07**	**0.796**	**37.53**	0.897	**52.39**	0.944	**72.74**
Naive	0.810	41.23	0.854	51.38	0.882	63.21	0.893	78.34	0.931	99.70	0.955	119.36
Seasonal Naive	0.742	30.96	0.795	38.72	0.827	47.30	0.852	58.89	**0.901**	75.11	0.922	90.14
AR(1)	0.748	28.94	0.798	36.04	0.847	44.15	0.887	55.12	0.929	70.17	0.934	84.23
ARMA(1,1)	0.750	28.32	0.798	35.28	0.840	43.17	0.880	53.86	0.927	68.71	0.932	82.37
SAR(1)	0.744	27.72	0.785	34.54	0.830	42.29	0.878	52.82	0.921	67.21	0.926	80.68
Naive-B	0.757	33.32	0.812	43.22	0.861	55.74	0.881	73.46	0.932	100.98	0.960	126.83
Seasonal Naive-B	0.505	16.92	**0.599**	21.89	**0.701**	28.90	0.790	42.11	0.893	73.71	0.944	108.52
AR(1)-B	0.656	23.04	0.734	29.65	0.797	37.98	0.873	50.61	0.937	70.34	**0.948**	89.50
ARMA(1,1)-B	0.662	22.57	0.739	29.11	0.794	37.21	0.869	49.50	0.935	68.66	0.946	87.45
SAR(1)-B	0.645	21.81	0.730	28.12	0.784	36.06	0.862	48.17	0.926	67.13	0.945	86.13

6 Conclusion

Computing forecast intervals for electricity prices is challenging since their distribution tends to be bi-modal. We investigate this by first modifying and then using a mixture model, the LMARX model, where one state represents price spikes and another has more regular prices. By fitting the model to the price series for different hours we get 24 parameter settings which we argue is a realistic representation of how electricity prices behave. Our results show, that using ARIMA models, together with the bootstrap, usually give a correct coverage probability and relatively short forecast intervals. The extreme values of electricity prices have the effect that, without the bootstrap, ARIMA models produce conservative, oversized, prediction intervals. This is particularly true for small nominal levels. Bootstrap only helps to some extent. Exploiting the benefits of a correctly specified LMARX model, however, yield huge benefits in terms of the length of the interval. An added benefit is also that the model produces an estimate of the spike probability for the next day.

If only a few unpredictable spikes, could seemingly be considered as outliers not worthwhile trying to model. However, in our view, the effect of spikes on probabilistic forecasts is a very interesting topic for further research. While few in numbers, badly predicted occurrences and values of the spikes might lead to huge losses, caused by bad decisions. This might be studied by means of a case study with a real, or hypothetical, decision problem yielding a specific loss function.

Statement: During the preparation of this work the authors used ChatGPT in order to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

Amjady N., Keynia F. Electricity market price spike analysis by a hybrid data model and feature selection technique // Electric Power Systems Research. 2010. 80, 3. 318–327.

Amjady N., Keynia F. A new prediction strategy for price spike forecasting of day-ahead electricity markets // Applied Soft Computing Journal. 2011. 11, 6. 4246–4256.

Anderson C L, Davison M. A Hybrid System-Econometric Model for Electricity Spot Prices: Considering Spike Sensitivity to Forced Outage Distributions // IEEE Trans. Power Syst. VIII 2008. 23, 3. 927–937.

Bello A., Reneses J., Muñoz A. Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case // Energies. 2016. 9, 3. cited By 19.

Botterud A., Zhou Z., Wang J., Bessa R.J., Keko H., Sumaili J., Miranda V. Wind power trading under uncertainty in LMP markets // IEEE Transactions on Power Systems. 2012. 27, 2. 894–903.

Brusaferri A., Matteucci M., Portolani P., Vitali A. Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices // Applied Energy. 2019. 250. 1158–1175. cited By 65.

Christensen T M, Hurn A S, Lindsay K A. Forecasting spikes in electricity prices // Int. J. Forecast. IV 2012. 28, 2. 400–411.

Dudek G. Multilayer perceptron for GEFCom2014 probabilistic electricity price forecasting // International Journal of Forecasting. 2016. 32, 3. 1057–1060.

Gaillard P., Goude Y., Nedellec R. Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting // International Journal of Forecasting. 2016. 32, 3. 1038–1050.

Hong T., Pinson P., Fan S., Zareipour H., Troccoli A., Hyndman R.J. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond // International Journal of Forecasting. 2016. 32, 3. 896–913.

Huang Guang-Bin, Zhu Qin-Yu, Siew Chee-Kheong. Extreme learning machine: theory and applications // Neurocomputing. 2006. 70, 1-3. 489–501.

Kristensen Kasper, Nielsen Anders, Berg Casper W., Skaug Hans, Bell Bradley M. TMB: Automatic Differentiation and Laplace Approximation // Journal of Statistical Software. 2016. 70, 5. 1–21.

Maciejowska K., Nowotarski J., Weron R. Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging // International Journal of Forecasting. 2016. 32, 3. 957–965.

Maryniak Paweł, Weron Rafał. What is the Probability of an Electricity Price Spike? Evidence from the UK Power Market // HANDBOOK OF ENERGY FINANCE Theories, Practices and Simulations. June 2020. 10, 231–245. (World Scientific Book Chapters).

Muniain Peru, Ziel Florian. Probabilistic forecasting in day-ahead electricity markets: Simulating peak and off-peak prices // Int. J. Forecast. X 2020. 36, 4. 1193–1210.

Nowotarski Jakub, Weron Rafał. Computing electricity spot price prediction intervals using quantile regression and forecast averaging // Comput. Stat. IX 2015. 30, 3. 791–803.

Nowotarski Jakub, Weron Rafał. Recent advances in electricity price forecasting: A review of probabilistic forecasting // Renewable and Sustainable Energy Reviews. 2018. 81. 1548–1568.

R Core Team . R: A Language and Environment for Statistical Computing. Vienna, Austria, 2022.

Wan Can, Xu Zhao, Pinson Pierre, Dong Zhao Yang, Wong Kit Po. Probabilistic forecasting of wind power generation using extreme learning machine // IEEE Transactions on Power Systems. 2013. 29, 3. 1033–1044.

Wong C S, Li W K. On a logistic mixture autoregressive model // Biometrika. X 2001. 88, 3. 833–846.

Ziel F., Steinert R. Electricity price forecasting using sale and purchase curves: The X-Model // Energy Economics. 2016. 59. 435–454.

Ziel F., Steinert R. Probabilistic mid- and long-term electricity price forecasting // Renewable and Sustainable Energy Reviews. 2018. 94. 251–266.

Appendix

Table 5: Empirical Coverage (EC) and Average Length of 50%, 60%, 70%, 80%, 90%, and 95% two-sided one-step ahead PIs for 15 models, hour 03, 1000 replications. Computation time 3.94 hours

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	**0.500**	4.98	0.618	6.80	**0.721**	8.10	**0.798**	10.75	0.903	14.98	**0.950**	18.59
Naive	0.758	7.56	0.813	9.47	0.894	11.62	0.914	14.39	0.950	18.58	0.948	21.87
Seasonal Naive	0.665	15.61	0.753	19.55	0.841	23.95	0.901	29.71	0.921	38.32	0.948	45.15
AR(1)	0.785	7.29	0.804	9.13	0.898	11.20	0.916	13.88	0.943	17.92	0.945	21.09
ARMA(1,1)	0.779	7.27	0.801	9.11	0.898	11.17	0.916	13.84	0.943	17.86	0.945	21.02
ARX	0.781	7.23	0.801	9.06	0.896	11.11	0.914	13.77	0.942	17.78	0.945	20.92
ARXX	0.774	7.16	0.800	8.96	0.896	11.00	0.913	13.62	0.942	17.59	0.945	20.70
SARXX	0.772	7.14	0.794	8.94	0.892	10.96	0.914	13.57	0.941	17.54	0.947	20.64
Naive-B	0.537	4.56	0.614	5.75	0.741	7.16	0.812	9.11	**0.900**	13.25	0.944	21.17
Seasonal Naive-B	0.509	11.01	0.629	14.09	0.744	17.82	0.829	23.29	0.909	35.05	**0.950**	49.11
AR(1)-B	0.533	4.36	0.610	5.48	0.742	6.85	0.811	8.74	0.908	12.81	0.948	20.53
ARMA(1,1)-B	0.529	4.35	0.618	5.48	0.735	6.86	0.808	8.75	0.902	12.83	0.947	20.43
ARX-B	0.516	4.28	**0.602**	5.40	0.723	6.73	**0.802**	8.61	0.898	12.64	0.947	20.32
ARXX-B	0.512	4.24	0.609	5.37	0.727	6.70	**0.802**	8.56	0.896	12.60	0.946	20.08
SARXX-B	0.520	4.28	0.604	5.40	0.726	6.75	0.806	8.64	0.898	12.67	0.945	19.94

Table 6: Empirical Coverage (EC) and Average Length of 50%, 60%, 70%, 80%, 90%, and 95% two-sided one-step ahead PIs for 15 models, hour 18, 1000 replications. Computation time 3.31 hours

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	**0.506**	10.15	0.581	13.93	**0.693**	15.96	**0.805**	20.47	0.876	27.44	0.943	36.74
Naive	0.856	21.73	0.878	27.16	0.925	33.38	0.939	41.34	0.941	53.13	0.960	63.14
Seasonal Naive	0.774	39.98	0.815	49.76	0.874	61.39	0.891	76.02	0.913	97.96	0.916	116.08
AR(1)	0.845	20.66	0.881	25.82	0.927	31.72	0.945	39.31	0.949	50.53	0.953	60.03
ARMA(1,1)	0.840	20.59	0.882	25.73	0.923	31.60	0.944	39.17	0.948	50.35	0.952	59.82
ARX	0.846	20.23	0.879	25.27	0.926	31.05	0.944	38.50	0.949	49.47	0.951	58.79
ARXX	0.839	19.90	0.878	24.86	0.919	30.54	0.942	37.87	0.946	48.68	0.951	57.83
SARXX	0.836	19.75	0.879	24.67	0.926	30.31	0.943	37.58	0.946	48.31	**0.950**	57.39
Naive-B	0.546	9.43	**0.609**	12.01	0.742	15.29	0.850	20.54	0.923	37.94	0.968	71.75
Seasonal Naive-B	0.544	21.76	0.619	28.18	0.750	37.09	0.824	53.24	**0.912**	95.29	0.946	138.85
AR(1)-B	0.512	8.86	0.612	11.26	0.734	14.33	0.844	19.42	0.920	36.81	0.961	67.57
ARMA(1,1)-B	0.511	8.91	0.620	11.31	0.730	14.43	0.834	19.48	0.922	36.93	0.960	67.19
ARX-B	0.555	9.07	0.634	11.56	0.734	14.68	0.853	19.76	0.924	36.36	0.959	66.11
ARXX-B	0.547	9.37	0.640	11.89	0.757	15.16	0.853	20.47	0.918	36.81	0.957	64.72
SARXX-B	0.548	9.29	0.649	11.80	0.765	15.01	0.858	20.35	0.921	36.67	0.956	64.28

Table 7: Empirical Coverage (EC) and Average Length (AL) of 50%, 60%, 70%, 80%, 90% and 95% two-sided 7-step ahead PIs for 11 models, hour 03, 1000 replications. Computation time 2.2 hours

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	**0.493**	12.05	0.611	15.38	0.716	18.68	**0.799**	24.57	**0.903**	34.05	0.944	44.86
Naive	0.732	22.41	0.798	27.87	0.867	34.30	0.901	42.27	0.933	54.14	0.945	64.62
Seasonal Naive	0.680	18.99	0.760	23.71	0.838	29.17	0.871	35.86	0.904	46.03	0.929	54.98
AR(1)	0.674	17.12	0.762	21.34	0.840	26.24	0.877	32.43	0.914	41.44	0.937	49.41
ARMA(1,1)	0.671	17.09	0.754	21.29	0.841	26.19	0.874	32.32	0.916	41.35	0.938	49.29
SAR(1)	0.671	16.94	0.754	21.13	0.841	25.96	0.871	32.07	0.913	41.01	0.936	48.89
Naive-B	0.658	18.35	0.744	23.45	0.837	30.10	0.888	39.22	0.927	54.51	**0.952**	68.73
Seasonal Naive-B	0.483	12.27	**0.599**	15.61	**0.704**	20.24	0.803	27.31	0.896	43.57	0.936	62.75
AR(1)-B	0.597	13.92	0.693	17.79	0.800	22.90	0.848	29.83	0.912	41.37	0.942	52.45
ARMA(1,1)-B	0.599	13.89	0.693	17.77	0.796	22.82	0.849	29.67	0.913	41.30	0.944	52.27
SAR(1)-B	0.588	13.79	0.682	17.65	0.796	22.58	0.853	29.50	0.916	40.90	0.942	51.95

Table 8: Empirical Coverage (EC) and Average Length (AL) of 50%, 60%, 70%, 80%, 90% and 95% two-sided 7-step ahead PIs for 11 models, hour 18, 1000 replications. Computation time 4.9 hours

	50% EC	50% AL	60% EC	60% AL	70% EC	70% AL	80% EC	80% AL	90% EC	90% AL	95% EC	95% AL
LMARX	0.482	24.53	0.601	31.47	**0.727**	40.08	**0.797**	54.45	0.883	79.71	0.942	109.19
Naive	0.803	61.53	0.846	76.62	0.893	94.58	0.906	117.05	0.939	151.02	**0.955**	179.82
Seasonal Naive	0.739	46.21	0.782	57.72	0.854	71.19	0.858	88.14	**0.899**	113.38	0.915	135.16
AR(1)	0.730	41.67	0.821	51.86	0.869	63.98	0.895	79.20	0.918	102.05	0.925	121.73
ARMA(1,1)	0.732	41.02	0.819	51.10	0.869	63.06	0.892	78.06	0.914	100.49	0.928	119.93
SAR(1)	0.721	40.41	0.811	50.33	0.863	62.10	0.889	76.93	0.916	99.04	0.921	118.12
Naive-B	0.746	48.95	0.803	63.99	0.884	83.82	0.892	110.61	0.940	152.95	0.958	191.07
Seasonal Naive-B	**0.510**	23.86	**0.600**	31.26	0.733	42.37	0.795	64.30	0.902	115.86	0.942	163.61
AR(1)-B	0.637	32.41	0.736	42.12	0.831	54.89	0.873	72.67	0.922	102.38	0.937	130.09
ARMA(1,1)-B	0.642	32.01	0.744	41.60	0.824	54.01	0.870	71.53	0.915	100.75	0.935	128.14
SAR(1)-B	0.625	31.15	0.736	40.48	0.826	52.79	0.867	70.25	0.917	99.34	0.932	126.45

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no