Forecasting Hourly Prices in Indian Spot Electricity Market

Mukherjee, Paramita and Coondoo, Dipankor and Lahiri, Poulomi

International Management Institute Kolkata, Former Professor, Indian Statistical Institute, Rabindra Bharati University

2019

Forecasting Hourly Prices in Indian Spot Electricity Market

Paramita Mukherjee[†$] , Dipankor Coondoo[††] and Poulomi Lahiri[†††]

[†]Professor, International Management Institute Kolkata, 2/4C, Judges Court Road, Alipore, Kolkata
India 700027
Email: oparmita@hotmail.com, Phone: 94331 20454

[††]Former Professor, Indian Statistical Institute, Kolkata, India
Email: dipankor@gmail.com

[†††]Guest Faculty, Rabindra Bharati University, Kolkata,
Email: poulomilahiri@gmail.com

[$] Corresponding author.

ABSTRACT

In this paper, an attempt has been made to forecast the hourly electricity spot prices in India as this is very important for the bidders in the energy exchange for participating in the day-ahead market. Forecasting high frequency data is a challenging task. In forecasting, different variants of ARMA, ARMA-GARCH models are applied in different contexts, but no unequivocal dominance of a particular model exists. In this paper, based on hourly data for several years for all the regions in India, several variants of ARMAX models are estimated, by combining static and dynamic forecasts. Along with ARMA, intra-day, inter-day and hourly variations in prices as well as seasonalities on weekdays, holidays and festive days are incorporated. ARMAX models in this context performed quite well for forecasting horizons of hourly prices of upto 5 days. Interestingly, the ARMAX models provide reasonably good forecasts for day-ahead-market and the simple structure can be quite easily implemented. Such forecasts are not only essential for the players in the spot market, but also provides insights for policymakers as it reveals several aspects of Indian electricity market including the different dimensions of seasonality in demand.

Keywords: Forecasting, electricity, hourly data, energy, spot price, ARMAX model, day-ahead market

JEL Classification: C53, Q47

Forecasting Hourly Prices in Indian Spot Electricity Market

1 Introduction

Electricity prices exhibit unique behaviour owing to its special characteristics and forecasting electricity prices in the spot market has been very important across countries as the price discovery mechanism that works in the power exchanges is crucial for bidding for market participants. Electricity price forecasts have become a fundamental input to the decision-making process of energy companies (Weron, 2014). India, being a developing country and starting the reforms in the sector quite recently, such mechanisms of price discovery is still evolving and less explored.

Indian power sector has undergone significant transformation in the last two decades. With the implementation of Electricity Act, 2003 that seeks to introduce competition, protect consumers' interests and provide power to all, several changes are brought in like open access in transmission, competitive bidding for procurement of power, power trading from one region to another through the power exchange, competitive bidding for ownership and establishment of inter-state transmission schemes etc. One notable development is that the contribution of the private sector to additional capacity of electricity generation has gone up now from 42 per cent of total addition during 2007-12 to nearly 60 per cent of total addition during 2012-15 [Central Electricity Regulatory Commission (CERC), 2017].

There have been significant qualitative changes in transmission of power, too. By the end of '80s a strong regional networks have come into existence and the focus of planning and development of the transmission system has shifted from the State Grid system to the Regional Grid system. Gradually, the emphasis on the generation and transmission system has shifted from the orientation of regional self-sufficiency to the concept of optimization of resource utilization on an all India basis. Facilitated by the advancement in technology, the five regions, viz. North, East, North-East, West and South have now virtually become two markets called the "N-E-W grid" and the "SOUTH grid". The NEW grid is a leap forward from a regional grid operation to a national grid operation. In 2014, synchronization of the south grid with the national grid has been accomplished. Now a single grid, one of the largest grids of the world, is operational in India which accommodates a huge power capacity. During the last decade, the inter-regional transmission capacity of this grid has increased substantially from 14050 MW at the end of the 10th plan to 75650 MW at the end of 12th plan, i.e. as of 2016-17. The International Energy Exchange Country Report (2020) states that the Central Energy Regulatory Commission (CERC) in India is on the track of implementing several reforms and the progress made towards improved real-time markets.

In India, short term trading of electricity takes place through bilateral contracts, power exchanges and Deviation Settlement Mechanism (DSM). In January 2017, share of total volume of transaction of electricity out of the total electricity generation done through long term contracts was 89 per cent, while the shares of bilateral transactions, power exchanges and DSM were 4.7, 3.4 and 2.96 per cent, respectively. During the same month, out of a total short term trading of 10509 MUs electricity, the shares of bilateral transactions, power exchanges and DSM were 42.4, 30.8 and 26.8 per cent, respectively[1].

At present there are two power trading platforms in India, viz., Indian Energy Exchange (IEX) and Power Exchange India Limited (PXIL); and around 99.8 per cent of all power exchange-based electricity transactions are done through IEX, which started its operations in 2008. IEX, which provides a power trading platform for physical delivery of electricity, has more than 6000 participants located across states and union territories. The participants

[1] CERC (2017).

registered to trade electricity contracts include 54 distribution companies, over 450 electricity generators and over 3,900 open access consumers. As of March, 2018, participants registered to trade RECs included over 1,050 renewable energy generators and over 3,000 industry and corporate customers. The benefiting open access consumers belong to various industries such as metal, food processing, textile, cement, ceramic, chemicals, automobiles, information technology industries, institutional, housing and real estate and commercial entities (IEX website, accessed on June 5, 2018).

The trading takes several forms, viz., Day-Ahead Market (DAM), Term-Ahead Market (TAM), Renewable Energy Certificates (REC) and Energy Saving Certificates (ESCerts). In DAM, participants transact electricity on a *15-minutes block basis* a day prior to the actual delivery of the electricity transacted; the price and quantum of electricity to be traded being determined through a *double-sided closed auction bidding process*. Under TAM, contracts cover a range for buying/selling electricity for duration up to 11 days. Four types of contracts fall under this category, viz., Intraday, Day Ahead Contingency, Daily and Weekly. REC market facilitates transaction in environmental attributes. ESCerts are the tradable certificates under the Perform, Achieve, Trade (PAT) Scheme of Bureau of Energy Efficiency (BEE), a market-based mechanism to incentivise energy efficiency in large energy-intensive industries. The volume of electricity transaction through DAM and TAM has increased substantially over the years from 6214 MUs in 2009-10 to 46214 MUs in 2017-18.

Given the above marketing set up, the price discovery mechanism that works in the power exchanges is crucially important. From the point of view of the market participants, it is important because having forecast of electricity price may facilitate participation in the bidding. Electricity exhibits unique price behaviour due to its distinct characteristics compared to other commodities. In particular, electricity, being a flow cannot be stored in large scale and therefore smoothening of demand/supply shocks by inventory management is not possible. That is why electricity price exhibits high volatility and hence involves significant price risk for market participants. Apart from this, other characteristics of electricity include multiple levels of seasonality, price spikes, mean reversion etc. (Knittel and Roberts, 2005; Longstaff and Wang, 2004). Presence of such dynamism in the price formation makes forecasting of electricity price a challenging task.

Raviv *et al* (2012), based on a study of electricity price data of the Nordic Power Exchange, observe that a multivariate model for the full panel of hourly prices significantly outperform the corresponding univariate model for daily average price. Taylor *et al* (2006), based on a time series data of hourly demand for Rio de Janeiro, and a time series of half-hourly demand for England and Wales, compare forecasting performance of several univariate models for short-term electricity demand forecasting for lead times up to a day-ahead and find the model based on exponential smoothing to give the best forecast. Meng (2010) forecasts the hourly wind power production in Sweden using time series models. In this study the seasonality of the given data is examined alternatively by using ARARMA and seasonal ARIMA (SARIMA) models separately for cold and warm seasons and it is found that whereas for the warm season the ARARMA outperforms SARIMA, for the cold season both models give comparable forecasting results.

Studies focusing on models related to electricity price forecasting are mostly based on daily average price taken as a proxy for the spot price of electricity and also as a reference price for forward and futures contracts and other derivative contracts. The daily average price is obtained in the so-called day-ahead market for most electricity markets of the world and relates to an auction market where participants typically trade electricity for delivery during individual hours of the next day. Quotes for all hours (or regular intervals) are submitted simultaneously. Market prices are then determined by the intersection point of the underlying aggregate demand

and supply curves. The daily average price is basically the average of the 24 individual hourly prices. However, since the hourly electricity prices contain useful predictive information for the daily average price, modelling hourly prices may lead to better forecasting and that, in turn, may help market participants develop efficient trading and bidding strategies. Intraday price forecasting has not received much attention in the literature. Cuaresma *et al.* (2004) and Kristiansen (2012) use different autoregressive time series models for forecasting individual hourly prices. However, forecasting hourly prices is nothing short of a challenging task as it potentially contains different kinds of variations.

This paper makes an attempt to forecast the hourly spot electricity price that arises out of electricity transactions at the power exchange of India. Very few attempts have been made in this regard, so far for India. The available studies are either dated [Girish (2016)] or suffer from severe data limitations. These studies mostly focus on modelling of volatility of electricity prices [Kanjilal and Ghosh (2014)]. In contrast, the study we report here builds up a forecasting model for electricity price for India by incorporating relevant exogenous factors that are likely to affect spot electricity prices into a basic ARMA model specification. So, it not only provides an econometric model that may be useful for a market participant for bidding in the market, but also provides significant insight about the possible drivers of prices in the electricity market in India in each region. This will enable the policymakers to rightly assess the possible demand-supply mismatch on different hours, days and months and accordingly design reforms for each region separately. Needless to mention, forecasting of a variable of hourly frequency is in itself interesting as well as a challenging task. The next section provides an overview of the Indian electricity market scenario. Section 3 describes data and methodology followed by section 4 presenting the analysis and results. Finally, section 5 concludes the paper.

2. Indian Electricity Market: Facts and Studies

This study is on short term forecasting of hourly electricity price. Leaving aside a few studies that belong to the time series econometrics literature, studies on short term spot electricity price forecasting mostly belong to the electrical engineering literature. Models used for electricity price forecasting are of various types – viz., game-theoretic models, simulation models, time-series models etc.[Aggarwal, Mohan and Kumar (2009)]. Time-series models may further be classified into regression analysis based models, models of artificial intelligence, parsimonious stochastic models etc. The model developed here is a time series econometric forecasting model. For a review of the literature on spot electricity price forecasting, see Murthy *et al* (2014)].

There are a few studies on short term forecasts of electricity prices for India. For example, Girish (2016) does a region-level forecasting exercise for hourly electricity price and finds an ARIMA-PARCH (1,1) model, an ARIMA-GARCH (1, 1) model and an ARIMA-EGARCH model to perform best for northern, eastern and north-eastern regions, for western region and for southern region, respectively. An hourly data based analysis of Kanjilal and Ghosh (2014) finds an MSARIMA-EGARCH model to marginally outperform an MSARIMA model in terms of in-sample forecasting performances.

As the present study is on forecasting of hourly price based on ARMA/ARMAX (Auto Regressive Moving Average with Exogenous Inputs) model using price data originating in power exchange, a description of how the price evolves may help understand the price-setting mechanism in operation. Thus, in the Day-Ahead-Market (DAM) of IEX, the prices and quantity of electricity traded at individual time intervals are determined through a double sided closed auction bidding process. However, some time congestion may arise in this process – viz., an area having an inadequate transmission system may need power import from a surplus area to meet its power requirement and the supply-demand balance problem is resolved by adopting a market splitting mechanism. If the power flow exceeds capacity at the common price for the entire market area, it is split into surplus and deficit parts. The price is reduced for surplus part (where sale > purchase) and increased for deficit part (where purchase > sale). This will reduce sale (purchase) and increase purchase (sale) in surplus (deficit) part of the market. The flow is thus reduced to match available transfer capability (IEX website, accessed on October 25, 2018). Prices in such a case may differ across bid areas and the problem is resolved by obtaining Area Clearing Price (ACP). At present there are 13 bid areas covering 35 states and union territories of the country. The volume traded in DAM has grown substantially over the years and the market clearing price has declined during 2008-09 to 2017-18 [**Figure 1**].

Here we try to build up an adequate ARMAX model capable of generating reliable forecast of hourly electricity price for individual regions, taking into account the exogenous determinants of electricity use of individual regions, if any. Such a study is quite new in the Indian context.

3. Data and Methodology

We build up an econometric model for 10 days-ahead forecast of hourly electricity spot prices on a rolling basis for individual regions across India. The sample period of price data is January 1, 2011 to April 30, 2015. It may be noted that since a bunch of regulations about working of power grids in India came in force in 2010[2] and also that the southern region power grid got connected to the national power grid in December 2013, we take data well beyond 2013 covering

[2] For example, Enforcement and implementation of Indian Electricity Grid Code Regulations, Central Electricity Regulatory Commission Power Supply Regulations, Central Electricity Regulatory Commission Power Market Regulations etc.

the present wider data base. However, since older electricity price data may not be useful for forecasting, we use a dataset from January 1, 2012 relating to hourly electricity transaction price recorded in the IEX every day in the DAM segment. There are 29184 hourly observations in this data set, giving hourly price for each of the 12 electricity regions of India, viz., A1, A2, E1, E2, N1, N2, N3, S1, S2, W1, W2 and W3 [**Table 1**]. The data set is sourced from the spot electricity market of IEX.

An hourly electricity price data set as the present one may contain different types of variations such as intra-day hourly seasonality, inter-day weekly seasonality, monthly seasonality, seasonality due to day of the week effect, holiday effect, week-end effect etc., apart from regional variations. For generating efficient forecast, the analysis needs to take care of these variations in the forecasting model to the extent possible.

A preliminary data analysis is done to understand the nature of the data and identify the kinds of variations that may be present in the data set. Results of these preliminary analyses are summarized in subsection 3.1. The methodology followed for developing the econometric forecasting model is described next in subsection 3.2, followed by details on forecast generation and validation of the models in subsection 3.3.

3.1 Preliminary Examination of Price Variations

A preliminary analysis of the hourly price data for January 1, 2011 to December, 2014 sample period indicates presence of (1) regional clusters and (2) systematic hourly, weekly, monthly and other kinds of variations. These are as summarized below:

Regional Clusters: Table 2 gives the correlation matrix of hourly electricity prices for pairs of region. This correlation analysis shows eastern regions E1, E2, north-eastern regions A1 and A2, northern regions N1 and N2, western regions W1 and W2 and finally southern regions S1 and S2 form regional clusters in terms of similarity of patterns of hourly electricity price variation.

Further, whereas the correlation coefficient for S1, S2 pair is high, each of S1 and S2 has much smaller correlation with any region out of A1, A2, E1, E2, N1, N2, W1 and W2. Effectively thus there are two regional clusters in the data set, viz., the South Cluster (SC) comprising S1 and S2 and the Non-South Cluster (NSC) comprising the remaining ten regions.

Hourly Variation: Hourly variations of electricity price are of two kinds – intra-day and inter-day variation. We have examined these variations separately for individual regions using factor analysis to identify hour blocks that show systematic variation patterns. For illustration, we summarize in Table 3 the factor analysis results for regions A1 and S1 belonging to NSC and SC clusters, respectively. As this Table shows, there are three hour blocks (i.e., presence of three factors) for A1 and two hour blocks for S1. For both regions, price for evening hour block is different from those of the remaining hour blocks. To eliminate hourly seasonal effects present in the data, if any, we de-seasonalize the data[3]. The hourly seasonality of the data however is found to be weak. To check further, we fit region-specific ARMA model to the de-seasonalized data and obtain corresponding fitted values. These fitted values are multiplied by corresponding seasonal indices and the time series thus obtained is graphically matched with the corresponding observed times series[4]. Figure 2 shows this graph for A2 and similar result is obtained for other regions.

Monthly and other Seasonal Variations: To explore the nature of hourly variations across months, we have done Lowess Smoothing of the scatter diagram of hourly prices for individual regions separately for two relatively recent years, viz., 2013 and 2014. Graphs for 2014 for selected regions are presented in **Appendix 1**. The results are summarised in **Table 4.** These

[3] The deseasonaliztion procedure used is as follows: Let y_{it}: observed price for *i-th* hour of *t-th* sample day; *i=1,24*; $\bar{y}_i$: sample mean price for *i-th* hour; and $\bar{y}$: overall sample mean price. The deseasonalized observed price for *i-th* hour of *t-th* sample day is then $y_{it}^* = \left(\frac{\bar{y}}{\bar{y}_i}\right). y_{it}$; for all i, t.

[4] The ARMA model with hourly seasonal indices is also compared with an ARMA model estimated using the non-deseasonalized original actual data. The predictions turn out to be close.

graphs suggest presence of monthly and other types of seasonality in the hourly electricity price data used which should be taken care of in the forecasting model. For further look at the variation of hourly prices throughout the year, the trend in prices for the entire sample period for each of the regions are analysed by taking into account representative hour blocks for day and night, viz. hours 12 and 24. The plots of prices reveal that for hour 24, except S1, the trend in prices are more or less similar, exhibiting a surge during the monsoon months and festive months (July - October). The volatility is also not too high. Though similar pattern is observed for hour 12, it is not as distinct as hour 24.

3.2 Econometric Modelling for Forecasting

To build a satisfactory forecasting model for hourly electricity price capable of capturing the systematic variations present in the data, we first estimate an ARMA model as a benchmark specification. The AR component of the fitted ARMA model is known to capture the autoregressive elements present in the price data, while the corresponding MA component of the model captures the underlying error-correcting mechanism. Our region-specific ARMA model-based exercises show that AR and MA terms of fairly large lag order often turn out significant. For example, as Table 5 shows, the estimated coefficient of the AR term of lag 96 and that of the MA term of lag 120 for region A2 and E1, respectively, are statistically significant.

Such results suggest a parametrically parsimonious modelling for efficient forecast generation. In the next stage, we therefore try ARMAX specification of ARMA model, which incorporates exogenous explanatory variables into an ARMA model to improve the model's explanatory/predictive power. Guided by the preliminary data analysis results, we include qualitative exogenous explanatory variables in dummy variable form in the ARMAX model specification. The dummy variables used are - a Sunday Dummy – Sdum = 1, if sample day is

Sunday, 0 otherwise; a Holiday Dummy – Holiday = 1, if sample day is a holiday, 0 otherwise; and eleven Month Dummies to capture monthly seasonality for the months January to November (December being taken as the reference month).

The ARMAX model has the following stochastic specification (Hamilton, 1994). Consider the time series econometric model

$$y_t = \alpha_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots.. + \beta_b x_{bt} + w_{t,} \tag{1}$$

$x_{1t}, x_{2t}, \ldots.., x_{bt}$ being exogenous explanatory variables and the random disturbance term w_t following an ARMA (p,q) process, viz.,

$$\begin{aligned}(1 - \phi_1 L - \phi_2 L^2 - \ldots\ldots - \phi_p L^p) w_t \\ = (1 + \theta_1 L + \theta_2 L^2 + \ldots\ldots + \theta_q L^q)\, a_t;\ a_t \sim iid\ \Phi(0, \sigma^2),\end{aligned} \tag{2}$$

symbols having their usual meaning.

Substituting $w_t = y_t - (\alpha_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots.. + \beta_b x_{bt})$ in (2), the ARMAX *(p, q)* model is obtained as

$$\begin{aligned}(1 - \phi_1 L - \phi_2 L^2 + \ldots\ldots + \phi_p L^p) y_t \\ = (1 - \phi_1 L - \phi_2 L^2 + \ldots\ldots + \phi_p L^p)(\alpha_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \cdots.. + \beta_b x_{bt}) \\ + (1 + \theta_1 L + \theta_2 L^2 + \ldots\ldots + \theta_q L^q)\, a_t;\ a_t \sim iid\ \Phi(0, \sigma^2).\end{aligned} \tag{3}$$

For the present exercise, ARMAX model is estimated using STATA commands. Following the estimate of ARMAX model, corresponding k-step ahead dynamic forecasts are obtained. Because the predictions cannot update the changing structure of the MA error term of the ARMAX model in successive iterations, the forecasts tend to be smoother when k is large.

3.3 Forecast Generation based on Estimated ARMAX Model

As already mentioned, ARMAX Model is estimated at region level. Before this estimation, some data adjustments are done for regions having data gaps[5]. Validation of an estimated model is done through out-of-sample forecasting as follows: First, ARMAX model is estimated with 28944 observations out of the total number of 29184 hourly observations of the data set. Hourly forecast values for the remaining 240 sample hours (i.e., for 24 hours of each of the last ten sample days) are then generated and matched with the corresponding observed values.

To improve forecast accuracy, static and dynamic forecasts are combined[6]. For generating these forecasts, ARMAX model is estimated following four different approaches and forecast performances based on these approaches are compared as follows:

In the first approach, ARMAX model is first estimated on 28944 observations and using that dynamic forecast for last 240 sample hours are generated. Next, pretending these 240 forecast values to be observed values, ARMAX model is re-estimated using 29184 observations (i.e., on full sample size) and corresponding static forecast is obtained for the last 240 sample hours;

The second approach involves five iterations. First, ARMAX model is estimated on 28944 observations and dynamic forecast for next 48 hours are obtained. These 48 observations are added to the data set and ARMAX model is estimated on (28944 + 48) observations and dynamic forecast for 96 hours (i.e., last 48 hours plus next 48 hours) are obtained and added to the data set. This process is repeated till 240 dynamic hourly forecasts are obtained;

The third approach is a marginal modification of the second one. Here, using the dynamic forecasts of the entire 29184 observations, ARMAX model is re-estimated and static one-step-

[5] Such data gaps are especially large for the southern regions – e.g., for one region as much as 12 per cent out of a little over 29000 observations are found missing. We have used the method of filling up data gap of STATA for this.

[6] Dynamic forecast takes previously forecasted values while static forecast takes actual values to predict next step forecast. For one-step ahead forecast the two types of forecasts are equal.

ahead forecast for the last 240 sample hours is obtained. Finally, in the fourth approach, first an ARMA (1, 1) with all the dummy variables is estimated based on the 28944 observations and then dynamic forecasts for the last 240 sample hours are made.

Using the ARMAX model, prediction of price for specific hours of the day has also been attempted. These hourly forecasts however are not always superior to the corresponding 240 hourly forecasts by the previous approach. For example, whereas for Hour 1 both forecasts turn out to be similar, for Hour 2 and Hour 3, forecasts by the previous approach are found better for first 4-5 days, but hourly forecasts are found better for the 5^{th} - 6^{th} days onwards. For Hour 20, from the 3^{rd} day onwards, forecasts based on previous approach are often better than the corresponding hourly forecasts. In fact, prediction based on estimated hourly forecast model in which the price is predicted separately by hour, does not capture well all the variations present in the data and possibly therefore the hourly forecast model fails to forecast well.

4. Results

As mentioned, the present forecasting exercise is done separately for each of the 12 regions in which the electricity market of the country is divided. We present here illustrative results for the three selected regions A2, E1 and N1. **Table 5 and Table 6** present the region-specific estimated ARMA model and the corresponding estimated ARMAX forecasting models, respectively. These results show how parametric parsimony can be achieved by using the ARMAX model in place of the corresponding ARMA model – viz., by using a few exogenous explanatory dummy variables, AR terms of the order of 120 required in the ARMA model can be conveniently dispensed with. It should however be mentioned that for many individual regions (including those for which results are not presented here) the estimated ARMAX model leaves a small but significant residual autocorrelation unexplained.

Coming to the results of forecasting exercise with the fourth approach of ARMAX specification that provides dynamic forecasts only, the Mean Absolute Percentage Error (MAPE) of generated forecast are reported in the last three rows of Table 6. The first three forecasting approaches which use combinations of dynamic and static forecasts produce similar forecasts. In all the three approaches, the prediction is quite good for about 90 to 100 out of 240 observations and in all cases the Mean Absolute Percentage Error (MAPE) is within 30% [**Figure 3**]. Since forecast performances of approaches 1 to 4 are similar, only the forecast performance results for approach 4 are presented in **Table 6**. It should be noted that for all regions, Sunday dummy is significant. Also for some regions, prices are significantly less in holidays compared to weekdays. Moreover, compared to the December prices (base category), February, March and November experience less prices. This probably indicates the dominant influence of supply side compared to the demand side in determination of prices.

From the estimates of the model with 4^{th} approach, too, as expected, the forecasts are smoother beyond around 90 to 100 observations [**Figures 4.1, 4.2, 4.3**]. This implies that the model forecasts are good for hourly prices for 4 to 5 days, but beyond that it fails to perform well. This is reflected in MAPE which is around 40% to 47% for 240 observations, but falls to 20% to 30% for forecasts of 120 observations and further to 10% to 16% for forecasts of 96 observations. The possible explanation is that the estimated model has failed to capture autocorrelation of residuals of some large order i.e. the long memory autocorrelation. Therefore for forecasting hourly prices for next 3-5 days, models 1-3 as well as 4 perform equally well, but for hourly forecasting beyond 5 days, these models do not perform well.

5 Conclusion

The paper focuses on estimation of an appropriate forecasting model for hourly electricity prices in the context of Indian market. It differs from other studies in this context in terms of the methodology and treatment of the variables. Unlike the existing studies, it takes into account not only the data generating process through ARMA, but also incorporates relevant variables that has serious impact on electricity prices. Four variations of ARMAX models with combinations of dynamic and static forecasts are estimated. For hourly forecasts of prices till 5 days, the models perform quite well whereas for shorter horizons, all the models perform very well. Interestingly, the combinations of dynamic and static forecasts (in models 1 – 3) and only dynamic forecast (model 4) have shown comparable performance in forecasting for the short and the long horizons.

The forecasting models estimated here are not only dealing with high frequency data, but have tried to check whether a combination of dynamic and static forecasts produce the same results. The results are quite insightful in more number of ways than one. First, this forecasting model for electricity prices captures the intra-day as well as inter-day variations and the seasonal changes across months (through month dummies), and also special variations on Sundays and other holidays (through Sunday and Holiday dummies). This becomes particularly relevant in bidding in the energy exchange for short term. This also helps policymakers to understand the price variations on different hours, days and months in a particular region and undertake measures accordingly. *Second*, Interestingly, the characteristics of regions are reflected in these dummies. For each region, the price variations can be understood distinctly. *Third,* the results that summer month prices are less compared to winter month prices, point to the possibility of dominance of supply related factors over demand related factors in determining prices. The study points out that for high frequency data like hourly one, though the long memory of autocorrelations could not be captured properly in ARMA or ARMAX models, one may achieve a reasonably good and accurate forecast at least upto four days. Therse forecasts are easy to implement and may be of great use to the traders/ players in the day ahead market in the energy exchange.

References

Aggarwal, S. K., Mohan, S. L., Kumar, A. (2009), Electricity price forecasting in deregulated markets: A review and evaluation. *Electrical Power and Energy Systems*, 31(1), 13-22.

Central Electricity Regulatory Commission (2017), *Growth of Electricity Sector in India from 1947- 2017*, Ministry of Power, Governement of India

Cuaresma, J. C., Hlouskova, J., Kossmeier, S., Obersteiner, M. (2004), Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*, 77(1), 87-106.

Girish G. P. (2016), Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models, *Energy Strategy Reviews*, 11-12, pp.52-57

International Energy Exchange (2020), *Country Report on India*, January.

Kanjilal, Kakali and Sajal Ghosh (2014), Modelling and Forecasting day ahead electricity price in Indian Energy Exchange: Evidence from MSARIMA-EGARCH model, *International Journal of Indian Culture and Business Management,* January, pp. 413-423.

Knittel , C.R. and M.R. Roberts (2005), An empirical examination of restructured electricity prices, *Energy Economics* 27, pp. 791-817.

Kristiansen, T. (2012), Forecasting NordPool day ahead prices with an autoregressive model. *Energy Policy*, 49, 328–332

Longstaff, F.A., and A.W. Wang (2004): Electricity forward prices: A high-frequency empirical analysis, *Journal of Finance*, 59, 1877-1900.

Meng, Xiangli (2010), Modeling and Forecasting Hourly Wind Power Production in Sweden with Time Series Models, School of Economics and Social Sciences, Högskolan Dalarna, Sweden, May, mimeo

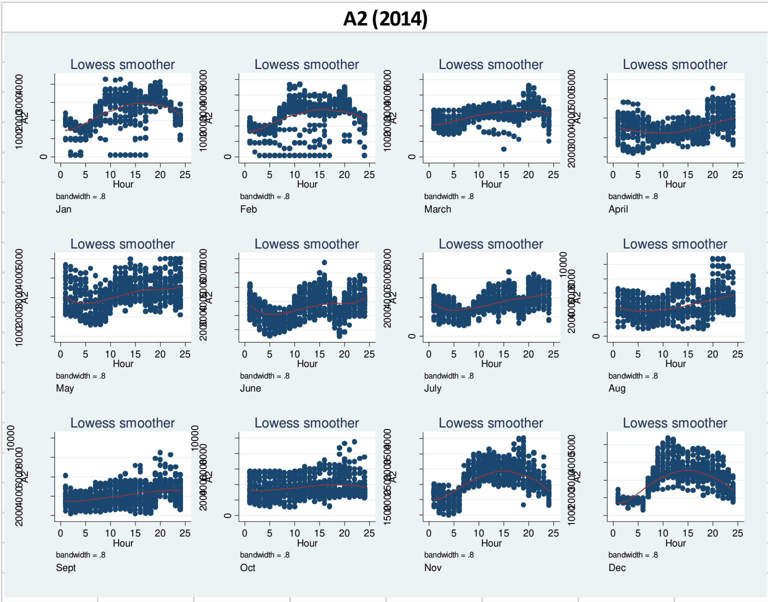
Murthy, Girish G. P., V. Sedidi, A. K. Panda and B. N. Rath (2014), Forecasting Electricity Prices in Deregulated Wholesale Spot Electricity Market: A Review, *International Journal of Energy Economics and Policy*, Vol. 4, No. 1, 2014, pp.32-42

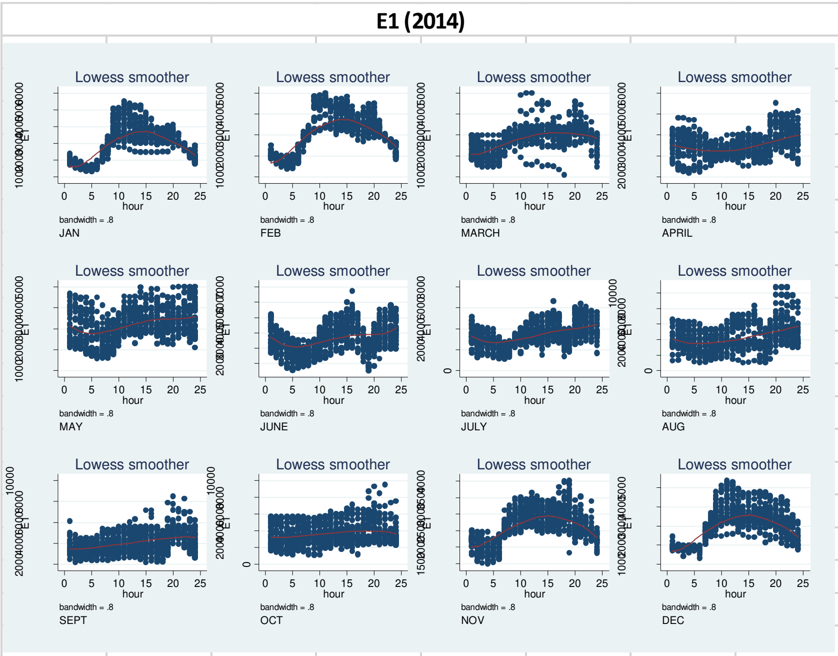
Raviv, Eran, Kees E. Bouwman1 and Dick van Dijk.(2012), Forecasting Day-ahead Electricity Prices: Utilizing Hourly Prices, *Energy Economics*, Elsevier, vol. 50(C), pp. 227-239.

Taylor, James W. , Lilian M. de Menezes and Patrick E. McSharry (2006), A Comparison of Univariate Methods for Forecasting Electricity Demand Up to a Day Ahead, *International Journal of Forecasting*, 2006, Vol. 22, pp. 1-16

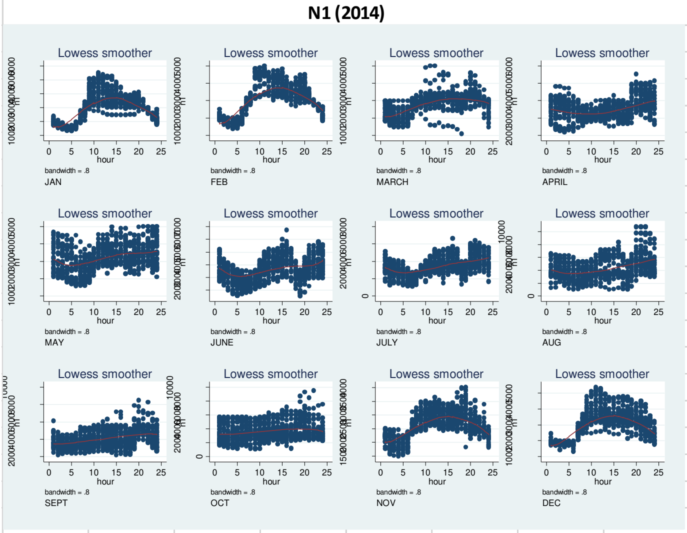
Weron, Rafal (2014), Electricity price forecasting: A review of the state-of-the-art with a look into the future, *International Journal of Forecasting*, Volume 30, Issue 4, October–December 2014, Pages 1030-1081

APPENDIX 1: Lowess Smoothing Graphs

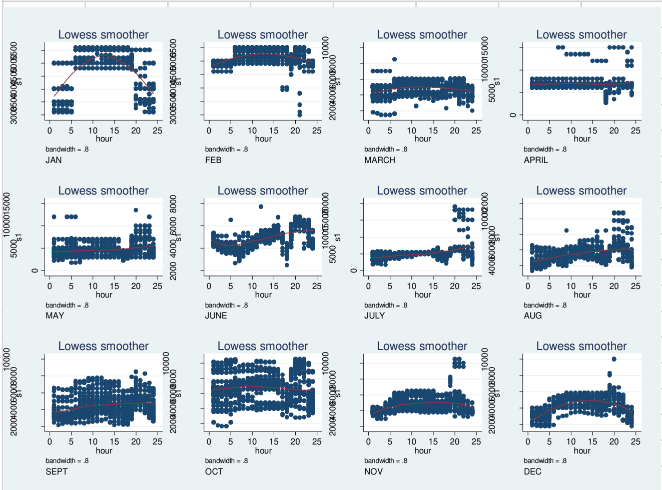




N1 (2014)



S1 (2014)



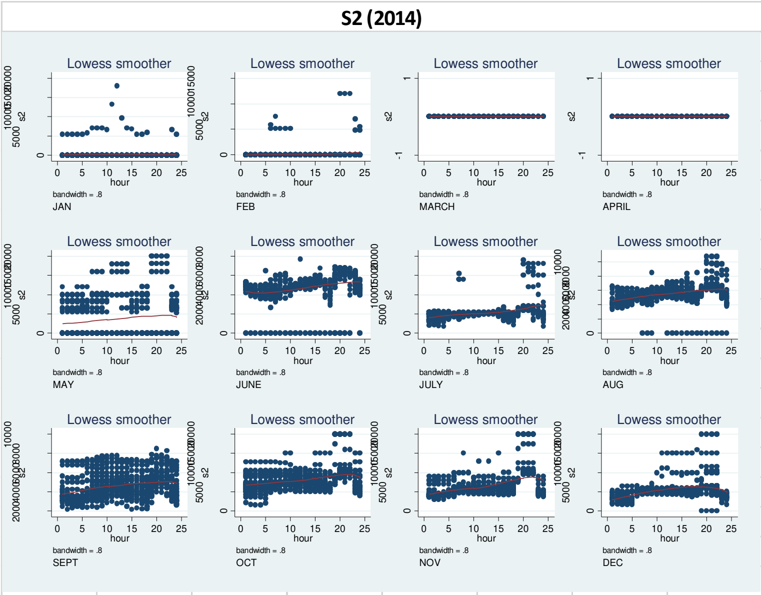

S2 (2014)
Lowess smoother
hour
bandwidth = .8
JAN
FEB
MARCH
APRIL
MAY
JUNE
JULY
AUG
SEPT
OCT
NOV
DEC

TABLE 1 : Details of the Regions			
Sr. No.	**Bid Area**	**Region**	**States covered under Bid Area**
1	N1	**North Region**	Jammu and Kashmir
			Himachal Pradesh
			Chandigarh
			Haryana
2	N2	**North Region**	Uttar Pradesh
			Uttaranchal
			Rajasthan
			Delhi
3	N3	**North Region**	Punjab
4	E1	**East Region**	West Bengal
			Sikkim
			Bihar
			Jharkhand
5	E2	**East Region**	Orissa
6	W1	**West Region**	Madhya Pradesh
7	W2	**West Region**	Maharashtra
			Gujarat
			Daman & Diu
			Dadar & Nagar Haveli
			North Goa
8	W3	**West Region**	Chhattisgarh
9	S1	**South Region**	Andhra Pradesh
			Telangana
			Karnataka
			Pondicherry (Yanam)
			South Goa
10	S2	**South Region**	Tamil Nadu
			Kerala
			Pondicherry (Puducherry)
			Pondicherry (Karaikal)
			Pondicherry (Mahe)
11	A1	**North East Region**	Tripura
			Meghalaya
			Manipur
			Mizoram
			Nagaland
12	A2	**North East Region**	Assam
			Arunachal Pradesh

Source: http://www.iexindia.com/bidareas.aspx?id=31&mid=2

Note: This table describes the different regions across India and the corresponding states falling under them.

TABLE 2: Correlation Matrix												
	A1	A2	E1	E2	N1	N2	N3	S1	S2	W1	W2	W3
A1	1.000											
A2	**.997**	1.000										
E1	.877	.877	1.000									
E2	.870	.870	**.993**	1.000								
N1	.886	.887	**.964**	.957	1.000							
N2	.886	.887	**.964**	.957	**1.000**	1.000						
N3	.649	.650	.720	.725	.745	.745	1.000					
S1	.238	.238	.204	.201	.210	.210	.110	1.000				
S2	.209	.210	.135	.132	.152	.152	.082	**.719**	1.000			
W1	.872	.873	.943	.938	.975	.975	.728	.210	.183	1.000		
W2	.872	.873	.943	.938	.975	.975	.728	.210	.183	**1.000**	1.000	
W3	.824	.826	.860	.856	.867	.867	.652	.186	.218	.886	.886	1.000

Note: This table describes the pairwise correlation co-efficients between the regions. The result shows that the southern regions are different from the rest of the regions. The electricity prices are very similar in A1 and A2; in E1, E2, N1 and N2 and they are exactly the same in W1 and W2.

TABLE 3: Hourly Pattern of Price Variation for Representative Regions			
Region A1		**Region S1**	
H7 to H18	Daytime	H19 to H22	Evening
H19 to H22	Evening	H23 to H18	Night & Day
H23 to H6	Night		

Note: This table shows that the pattern of variation in prices in A1 and similar regions are different from that of the Sounthern regions.

TABLE 4: Month-wise Trend in Hourly Prices		
Region	**2013**	**2014**
A1	Feb - May and Nov - flat Jun - Oct - rising towards night Dec - Jan - highest in middle of the day	Mar -Oct - flat Nov -Feb - highest in middle of the day
A2	Feb - May and Nov -flat Jun - Oct - rising towards night Dec - Jan - highest in middle of the day	Mar -Oct - flat Nov -Feb - highest in middle of the day
E1	Jan - Jun and Nov - flat Jul - Oct - slightly rising towards night Dec - highest in middle of the day	Mar - Oct -flat Nov - Feb - highest in middle of the day
E2	Jan - Jun and Nov -flat Jul - Oct - slightly rising towards night Dec - highest in middle of the day	Mar - Oct -flat Nov - Feb - highest in middle of the day
N1	Mar - Jun -flat Jul - Oct - slightly rising towards night Nov - Feb - highest in middle of the day	Mar - Jun - flat Jul - Oct - slightly rising towards night Nov - Feb - highest in middle of the day
N2	Mar - Jun - flat Jul - Oct - slightly rising towards night Nov - Feb - highest in middle of the day	Mar - Jun -flat Jul - Oct - slightly rising towards night Nov - Feb - highest in middle of the day
N3	Mar - May, Oct - flat Sep - slightly rising towards night Nov - Feb - highest in middle of the day	Mar - Oct - flat Nov - Feb - highest in the middle of the day
S1	Jan - Aug, Oct - flat	Feb - Nov - flat

	Sep - slightly rising towards night Nov - Dec - highest in middle of the day	Dec - Jan - highest in middle of the day
S2	May - Oct, Dec - flat Jan - Apr, Nov - slightly rising towards evening	May - Dec - very slightly rising towards night
W1	Mar - Jun - flat Jul-Oct - slightly rising towards night Nov - Feb - highest in middle of the day	Mar - Oct - flat Nov - Feb - highest in middle of the day
W2	Mar-Jun -flat Jul - Oct - slightly rising towards night Nov - Feb - highest in middle of the day	Mar - Oct - flat Nov - Feb - highest in middle of the day
W3	Mar - Jun - flat Jul - Dec - slightly rising towards night Jan - Feb - highest in middle of the day	Mar - Oct - flat Nov - Dec - highest in middle of the day Jan - Feb - mild rise towards night

Note: This table describes the monthly variation in electricity prices in different regions. For example, in 2014, regions vary in patterns of price and this leads to the consideration of different month dummies later in the regression analysis.

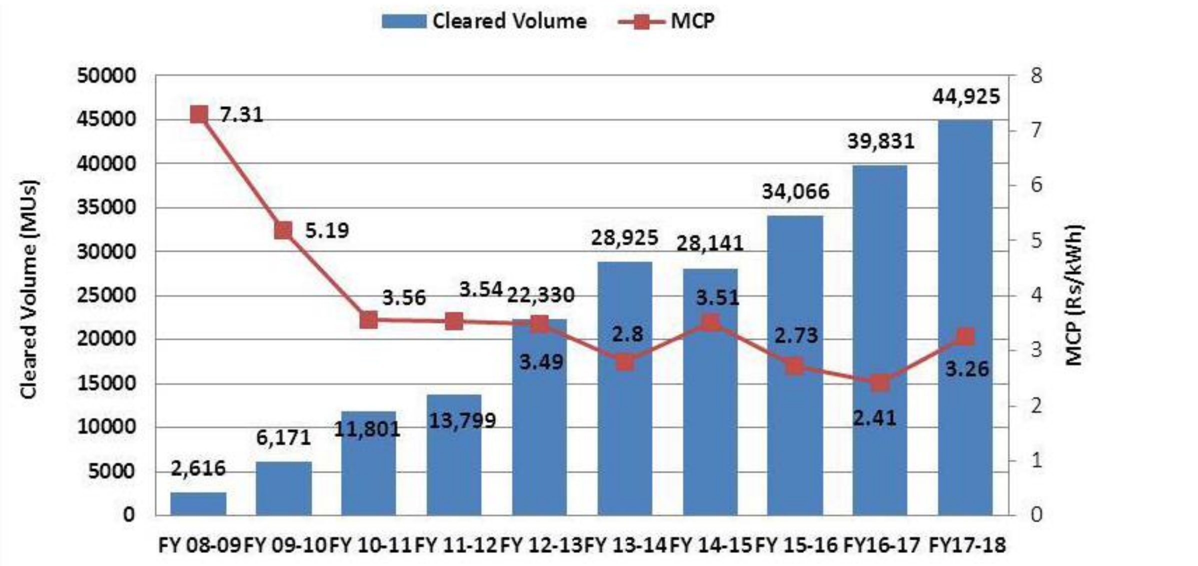
TABLE 5: ARMA ESTIMATION FOR SELECT REGIONS								
y_t : A2			y_t : E1			y_t : N1		
	Co-efficients	**P-value**		**Co-efficients**	**P-value**		**Co-efficients**	**P-value**
C	2957.60	0.00	C	2843.59	0.00	C	2947.37	0.00
y_{t-1}	0.124	0.000	y_{t-1}	0.912	0.000	y_{t-1}	0.074	0.000
y_{t-12}	0.022	0.004	y_{t-23}	0.007	0.000	y_{t-24}	0.640	0.000
y_{t-13}	0.018	0.016	y_{t-24}	0.979	0.000	y_{t-48}	0.099	0.000
y_{t-24}	0.442	0.000	y_{t-25}	-0.881	0.000	y_{t-72}	0.093	0.000
y_{t-48}	0.133	0.000	y_{t-26}	-0.006	0.002	θ_{t-1}	0.704	0.000
y_{t-72}	0.067	0.000	y_{t-48}	-0.014	0.000	θ_{t-2}	0.560	0.000
y_{t-96}	0.092	0.000	θ_{t-1}	-0.176	0.000	θ_{t-3}	0.472	0.000
θ_{t-1}	0.574	0.000	θ_{t-2}	-0.086	0.000	θ_{t-4}	0.413	0.000
θ_{t-2}	0.443	0.000	θ_{t-3}	-0.030	0.000	θ_{t-5}	0.371	0.000
θ_{t-3}	0.367	0.000	θ_{t-4}	-0.019	0.000	θ_{t-6}	0.308	0.000
θ_{t-4}	0.323	0.000	θ_{t-6}	-0.025	0.000	θ_{t-7}	0.270	0.000
θ_{t-5}	0.283	0.000	θ_{t-15}	0.009	0.035	θ_{t-8}	0.256	0.000
θ_{t-6}	0.230	0.000	θ_{t-23}	0.055	0.000	θ_{t-9}	0.239	0.000
θ_{t-7}	0.228	0.000	θ_{t-24}	-0.554	0.000	θ_{t-10}	0.201	0.000
θ_{t-8}	0.201	0.000	θ_{t-25}	0.145	0.000	θ_{t-11}	0.182	0.000
θ_{t-9}	0.168	0.000	θ_{t-26}	0.048	0.000	θ_{t-12}	0.155	0.000
θ_{t-10}	0.143	0.000	θ_{t-30}	0.019	0.002	θ_{t-13}	0.139	0.000
θ_{t-11}	0.125	0.000	θ_{t-48}	-0.085	0.000	θ_{t-14}	0.107	0.000
θ_{t-12}	0.098	0.000	θ_{t-72}	-0.077	0.000	θ_{t-15}	0.097	0.000
θ_{t-13}	0.069	0.000	θ_{t-96}	-0.030	0.000	θ_{t-16}	0.075	0.000
θ_{t-14}	0.034	0.000	θ_{t-120}	-0.033	0.000	θ_{t-17}	0.056	0.000
θ_{t-18}	0.011	0.028				θ_{t-18}	0.035	0.000
θ_{t-23}	0.083	0.000				θ_{t-23}	0.083	0.000
Adjusted R-squared	**0.882**		**0.933**			**0.946**		
Durbin-Watson stat	**2.003574**		**2.008758**			**2.003382**		
F-statistic	**9435.35**		**19231.10**			**22065.14**		
Prob(F-statistic)	**0**		**0**			**0**		

Note: This table shows that autocorrelations are significant even if a lag of 120 or 96 are included.

TABLE 6: Forecasting Models for N1, E1 and A2								
y_t : A2			y_t : E1			y_t : N1		
	Co-efficients	S.E.		Co-efficients	S.E.		Co-efficient:	S.E.
C	1285.379**	20.79	C	852.505**	14.47	C	677.436**	13.25
y_{t-24}	0.593**	0.002	y_{t-24}	0.7201**	0.002	y_{t-24}	0.787**	0.002
sdum	-154.547**	15.090	sdum	-152.171**	10.370	sdum	-114.066**	8.002
Holiday	-139.251**	33.300	Holiday	-59.514**	20.410	Feb	-101.114**	35.98
Jan	-167.877**	34.660	Feb	-142.998**	36.09	March	-87.331**	28.79
Feb	-81.394*	34.870	March	-103.273**	33.160	Oct	-67.994**	22.96
March	-137.558**	35.190	Aug	101.872**	25.370	Nov	-87.821*	36.29
Nov	-172.015**	61.760	Nov	-89.039*	41.450	y_{t-1}	0.855**	0.002
y_{t-1}	0.848**	0.002	y_{t-1}	0.858**	0.002	θ_{t-1}	-0.091**	0.004
θ_{t-1}	-0.143**	0.003	θ_{t-1}	-0.123**	0.004			
Wald Chi-square	**369487.18**			**397093.2**			**466914.8**	
Probability (Chi-square)	**0.0000**			**0.0000**			**0.0000**	
MAPE (%)	47.4			39.4			41.0	
MAPE (%, till 120 obs)	28.4			21.7			21.3	
MAPE (%,till 96 obs)	16.4			10.9			10.3	
Note: base dummy - Dec								
*5% los, **1% los								

Note: This table presents estimation of the ARMAX models for select regions. Apart from intra and inter-day variations, Sunday dummy, holiday dummy and several month dummies are significant implying significant seasonal variation.

Figure 1: Cleared Volume and MCP in DAM



Source: India Energy Exchange

Figure 2 : Model with Seasonal Indices and Original Prices of A2

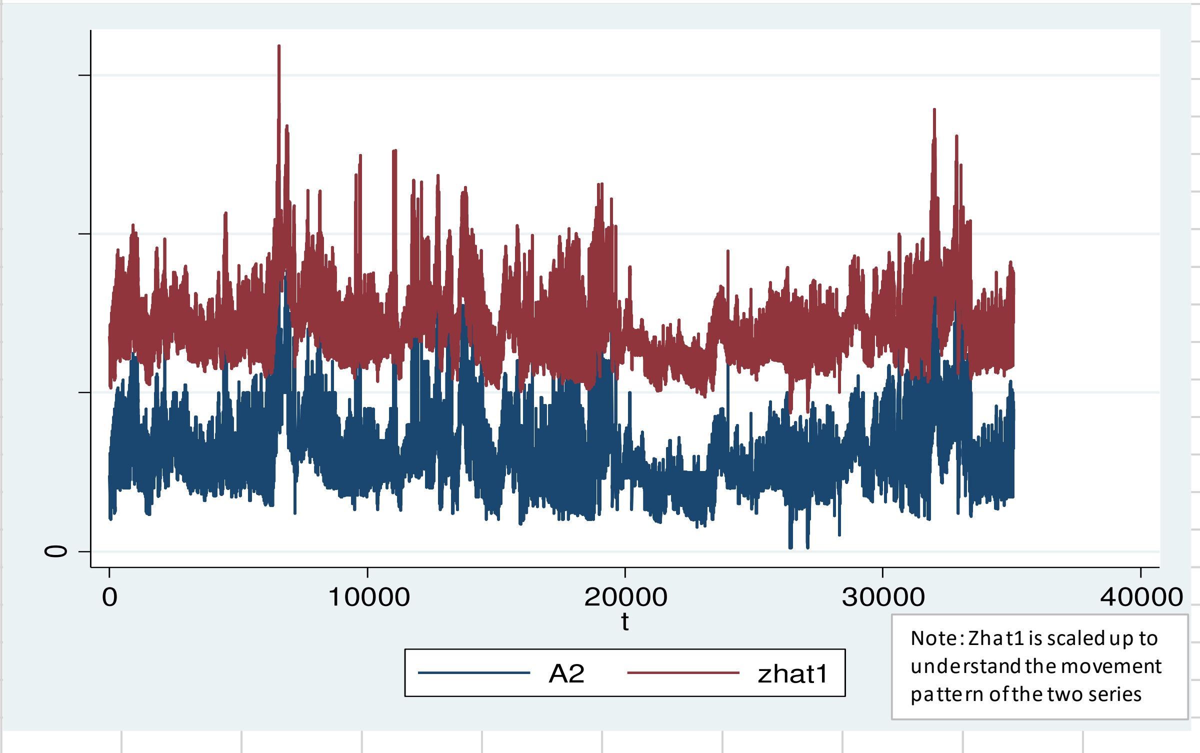


Figure 3: Model 3 – Actual and Predicted Prices for N1

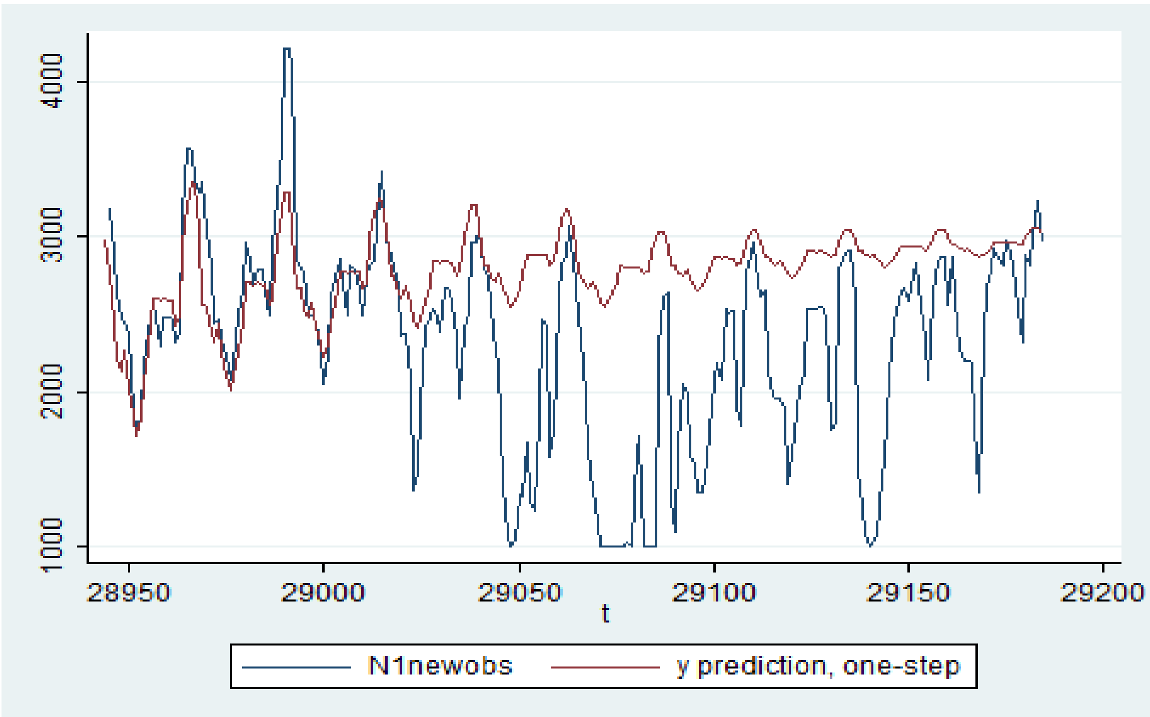


Figure 4.1: Model 4 Forecasts for N1

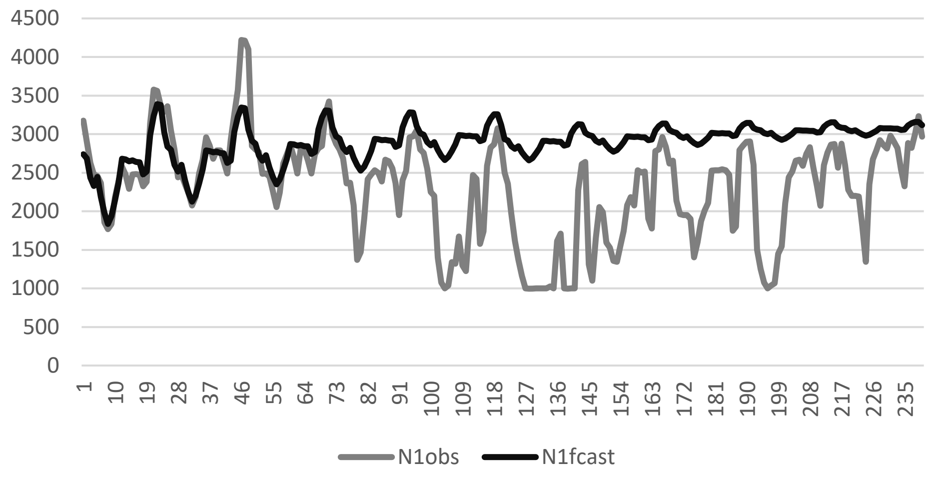
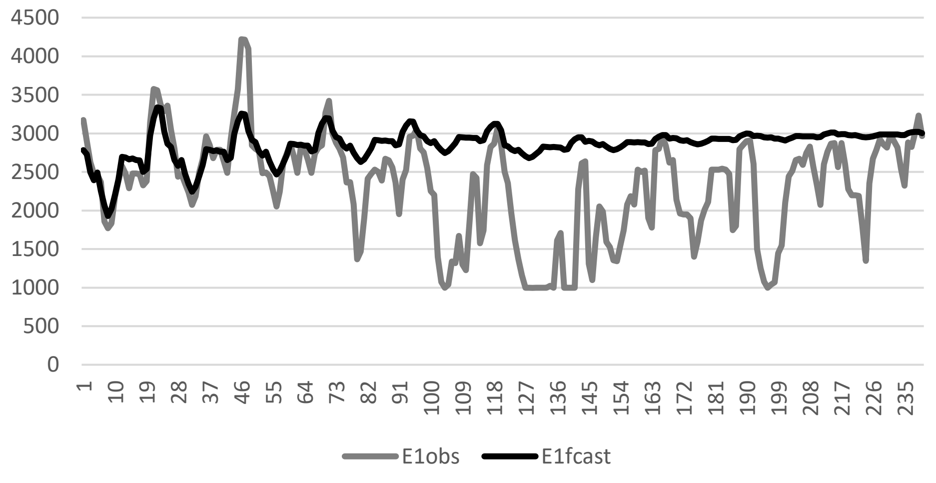


Figure 4.2: Model 4 FOrecasts for E1



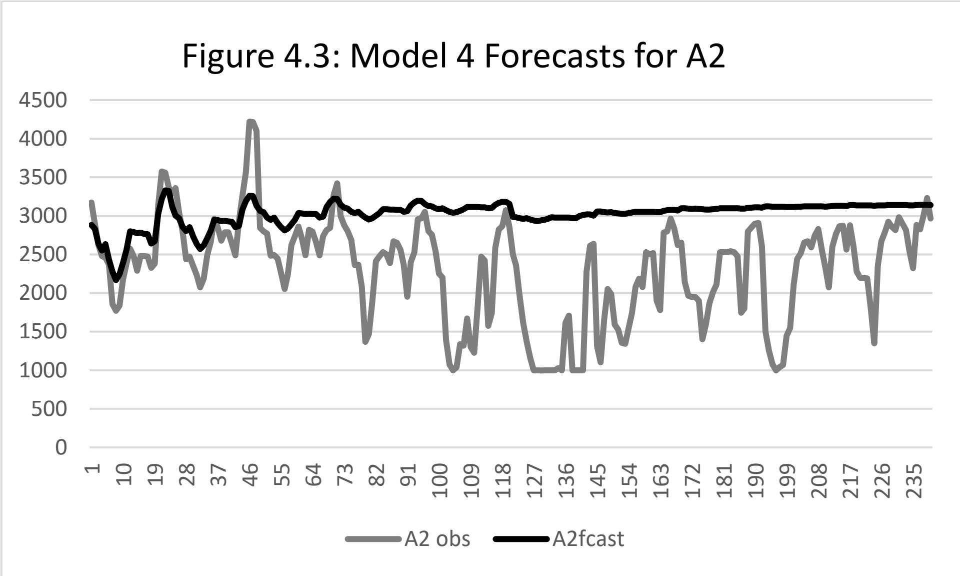


Figure 4.3: Model 4 Forecasts for A2
4500
4000
3500
3000
2500
2000
1500
1000
500
0
1
10
19
28
37
46
55
64
73
82
91
100
109
118
127
136
145
154
163
172
181
190
199
208
217
226
235
A2 obs
A2fcast