Two-Stage Load Forecasting for Residual Reduction and Economic Dispatch using PJM Datasets

Arun Sukumaran Nair, Mitch Campion, David Hollingworth, Prakash Ranganathan
Department of Electrical Engineering
University of North Dakota
Grand Forks, ND 58202
prakash.ranganathan@engr.und.edu

Abstract— **This paper discusses preliminary results obtained using auto-regressive integrated moving average (ARIMA), and exponential smoothing (ES) forecasting methods for loads in 11 regions of PJM Interconnection. The datasets used to predict day-ahead loads include demand values in both 24-hour and 30-day format for 2016 calendar year for multiple (e.g., 11 regions) areas. The accuracy of forecasting is evaluated using Mean Absolute Percentage Error (MAPE) and Mean Absolute Deviation (MAD) parameters. An economic dispatch was then carried using a linear programming formulation in Algebraic Mathematical Programming Language (AMPL) environment. The preliminary results indicate ARIMA outperforms ES for both 24-hour and 30-day to predict day-ahead forecasting.**

Keywords—Load forecasting, Economic dispatch, PJM, Smart grid

I. Introduction

Load forecasting is vitally important for the electric power industry in the deregulated economy. It has many applications including energy purchasing and generation, load switching, contract evaluation, and infrastructure development. A large variety of mathematical methods has been developed for load forecasting. Load forecasts can be divided into three categories: short-term forecasts which are usually from one hour to one week, medium forecasts which are usually from a week to a year, and long-term forecasts which are longer than a year.

In this work*, short-term load forecasting* of two well-known methods was conducted for PJM utility datasets that contained hourly historical data from 1/1/2016 to 11/16/2016 [1]. This work examines the accuracies involved in forecasting day ahead markets for 9 regions within PJM markets, so that the economic dispatch of this load can be evaluated. The 9 regions are PJM Mid-Atlantic Region (PJM-E), Allegheny Power (PJM-W), Dayton Power & Light Company (DAY), American Electric Power (AEP), Duquesne Light Company (DUQ), Dominion Virginia Power (DOM), American Transmission Systems, Inc (FE), and Duke Energy Ohio & Kentucky (DEOK).

II. Literature Review

Accurate models for electric power load forecasting are essential to the operation and planning of a utility company. As a result many algorithms and methodologies have resulted with varying levels of accuracy as part of a survey [2]–[6]. The work in [7] investigates ARIMA for day-ahead spot price forecasting. Additionally, [8] provides a great resource overviewing ARIMA, exponential smoothing (ES), and other statistical forecasting algorithms. To our knowledge, there are very limited, or no literatures exist that investigates accuracies on yearly and 30-day demand markets while combining this information with a perspective that includes examination of economic dispatch. Further, this work is of interest to the PJM market as it uses real data from the load areas of this market.

III. Methodology

A. Dataset

There are ten load areas in the PJM data. Each load area was had forecasts developed for the 24-hour period of 11/16/16 using each of the four forecasting protocols. Each individual forecast contains 24 data points corresponding to the 24 hours in which load occurs. In forecasting the day ahead power usage, forecasting each hour of the day is crucial as different demands occur at each hour of the day. This forecasting protocol resulted in a total of 40 individual forecasts, 4 for each of the ten areas. Once the forecasts for each load area were made, they were summed together to get a total PJM load forecast.

B. Forecasting

To perform meaningful power demand forecasting, multiple forecasting methods and protocols were attempted, and the accuracy of the methods were analyzed. In addition to these methods useful analysis of trend and seasonality of power demand data was analyzed. The methods of forecasting that were used and compared in this project were exponential smoothing [9] and auto-regressive integrated moving average (ARIMA) [7], [9], [10]. These are both well-known and well utilized traditional forecasting methods.

ARIMA is a forecasting method based on moving average but can account for seasonality and trend patterns in a time-series. There are 3 main parts of an ARIMA model formula: autoregressive portion, the integrated portion, and the moving average portion. The autoregressive part of the ARIMA model allows for a prediction to be influenced by the previous values of the prediction. In forecasting power demand this is important because the power demand of a given time interval is related and influenced by the demand of time intervals prior. The integrated component of ARIMA allows for seasonality trends that exist in a time series. This is an important aspect of power usage as there are seasonal trends to the usage of electric power. The integrated component of ARIMA models is accomplished

978-1-5386-5398-2/18/$31.00 ©2018 IEEE

by differencing and integrating a time series. The moving average component of and ARIMA model takes into consideration the moving average of a time series and the error that exists in fitting a real-time series with an auto-regressive model [11].

Forecasting for this project involved the use of different of amounts of historical data. The dataset used for this project contained hourly historical demand data ranging from 1/1/2016 to 11/15/2016. The goal of the load forecasting of this project was to create a power load forecast for 11/16/2016 so that the economic dispatch of this load could be simulated and examined. To perform this accurately different protocols were attempted. As mentioned above, ARIMA and exponential smoothing were used and compared. Additionally, different amounts of historical data were used in these forecasting methods to obtain different forecasts. The accuracy of these different methods was compared. ARIMA and exponential smoothing were applied to both the entire past year of power usage data as well as only the preceding 30 days of power demand data.

C. *Economic Dispatch*

The method of economic dispatch was used as the second step in our project. The forecasting methods deliver the next day load for 24 hours by evaluating the historical data and the results from forecasting was used for the economic dispatch process.

Economic load dispatch is a method to schedule the power generator outputs with respect to the load demands and to operate the power system most economically. In other words, the main objective is to allocate the optimal power generation of different units at the lowest cost possible while meeting all system constraints. The method of economic dispatch can be done subjected to different constraints of the system. Some of the constraints are generator constraints, tie line flow constraints, ramp-rate of the generator constraints etc.

Generator constraints limit the generation between the maximum and minimum value of a generator. We assume that all the generators in our system are thermal generators and for a thermal generator there will be minimum value of generation that should be maintained to keep the generator running. If the value of generation from the generator goes down below this value it can damage the generator and can cause it to shut down [12].

Tie line flow constraints are used to limit the flow between the generator and loads because each transmission line will be having a maximum amount of power flow it can handle. Ramp rate constraints are usually used in more complex systems to limit the amount of increase in power generation from a particular generator. The rate of change of increase in power a generator can handle can vary depending on the characteristics of the generator [12].

The economic dispatch model in this work considers only the generator constraints because it is assumed that there is a connection between all the generators to all the loads. Since we are considering each region as a load, we didn't consider tie line constraints as there is a need of large amount of power flow between the regions.

D. *Data Flow*

Fig. shows the data flow in the load forecasting – economic dispatch model we developed for this project. The historical load data values were obtained for the PJM system and different forecasting methods were applied to this data to obtain the day ahead data for the loads. The day ahead data implies the forecasted load values for the next 24 hours. The economic dispatch will be applied to the forecasted load values to obtain the hourly power generation values for the different generators.

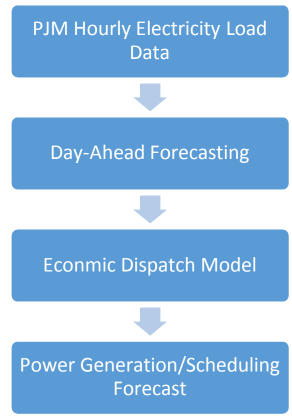

Fig. 1. *Graphical representation of the data flow in the forecasting- economic dispatch model*

E. *Forecasting and Economic Dispatch Methodology*

Using ARIMA, ES, and an ensemble ARIMA forecasting method, hourly-load for 11/16/2016 was forecasted for each region using the previous 11 months of hourly load data in [1]. The ensemble method forecasts the residuals of a first stage fitting model (ARIMA or ES) and uses ARIMA to forecast the residuals. The forecasted residuals are added to the forecast from the fitted model. The p,d,q parameters of ARIMA forecasts were determined using a method described in [8]. The accuracy of ARIMA and ES methods was then compared to actual hourly load data for 11/16/2016. The accuracy of the forecasts were quantified according to MAPE and MAD shown in equations (1) and (2) as follows:

$$MAPE = \frac{100}{n} * \sum^{n} |(A_l - F_l)| \,/\, A_l \quad (1)$$

$$MAD = \sum^{n} |(A_l - F_l)| \,/\, A_l \quad (2)$$

A_l is the actual load, F_l is the forecasted load, and n is the number of hours in the forecasted day (24). Additionally, accuracy of ARIMA and ES methods were compared by forecasting hourly load data of 11/16/2016 using the previous 30 days of hourly load data. After its comparison on accuracy on tested statistical methods, the output of best MAPE values of the method was fed as an input to an economic dispatch problem [13]. This is carried out using a linear programming formulation that does scheduling of generators for an optimal load dispatch cost. The combination of load forecasting, and optimal load dispatch gives a day-ahead forecast for optimal load dispatch cost.

IV. Load Data Analysis

Fig. 2 display a time-series plot of hourly demand for 2016 up to 11/15/16 for one of the PJM load areas called PJM-E. This raw time-series was decomposed into two component: 1) trends and 2) seasonality. They are shown in Figs 3 and 4. Trend and seasonality decomposition was performed using the loess algorithm as outlined in [14]. Fig. 3 display the yearly trend of load throughout the year of 2016. Fig. 3 shows a monthly seasonality. For example, 30 total demand peaks points are seen in Fig. 4 corresponding to 30 days for a month.

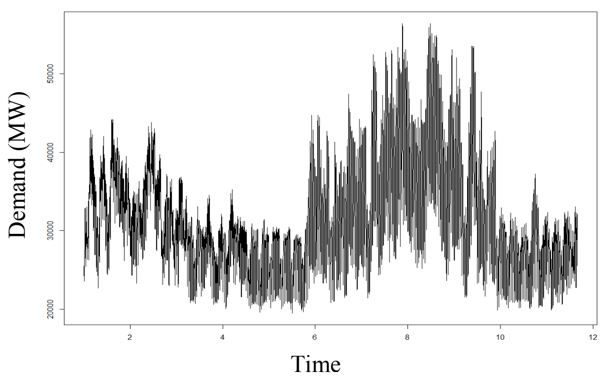

Fig. 2 *PJM-E demand levels for 2016.*

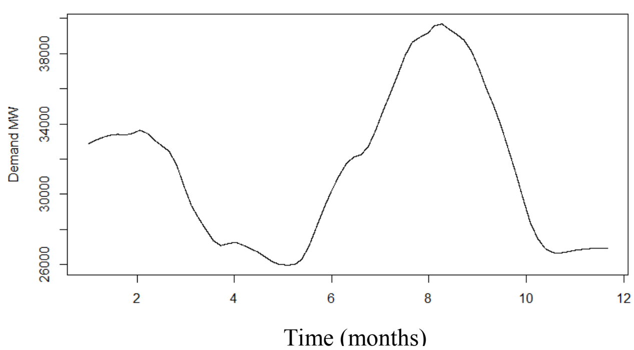

Fig. 3 *Yearly Seasonality by Month for PJM-E demand*

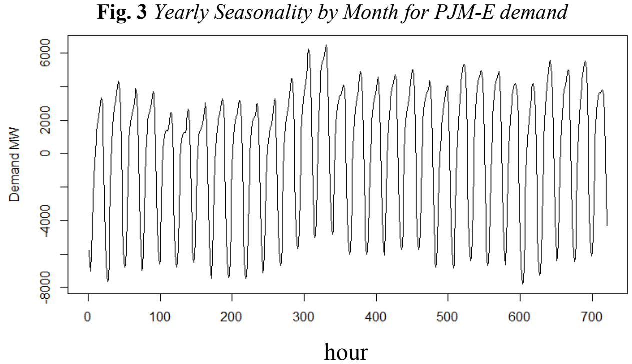

Fig. 4 *Additive monthly seasonality PJM-E region data*

V. Forecasting Results

Fig. 5 show the result of ARIMA and ES forecasting trained using 11-months of prior hourly-load data that are compared against the actual hourly-loads for 11/16/16. Fig. 6 display results of ARIMA (5, 1,5) and ES forecasting using the prior 30 days of hourly data as a training set data for PJM load areas. Fig. 7 compare several other methods of forecasting with the actual data.

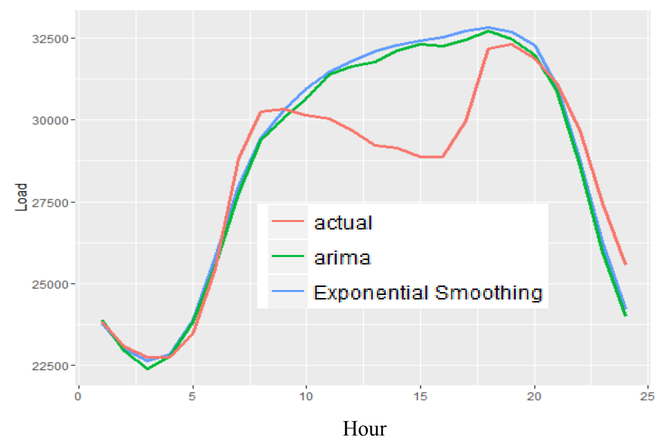

Fig. 5 *Day-ahead forecast (11/16/16) PJM-E using 11-month data*

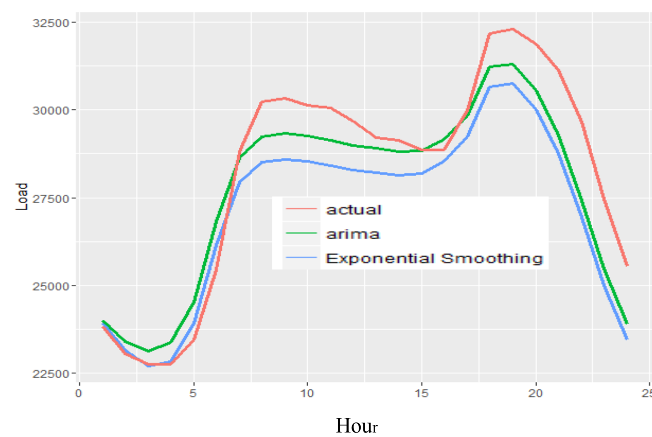

Fig. 6 *Day-ahead forecast (11/16/16) for PJM-E using 30 days data*

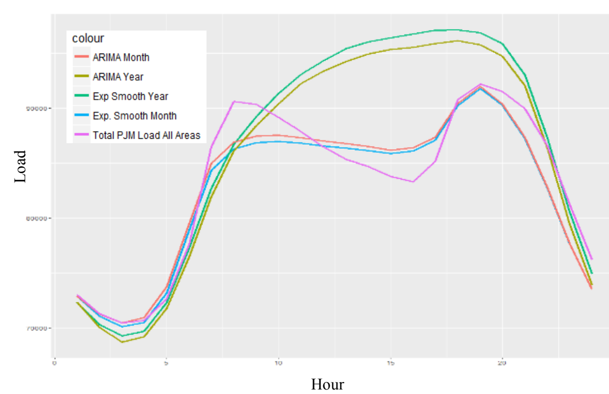

Fig. 7 *Day-ahead forecast comparison of all methods for all areas*

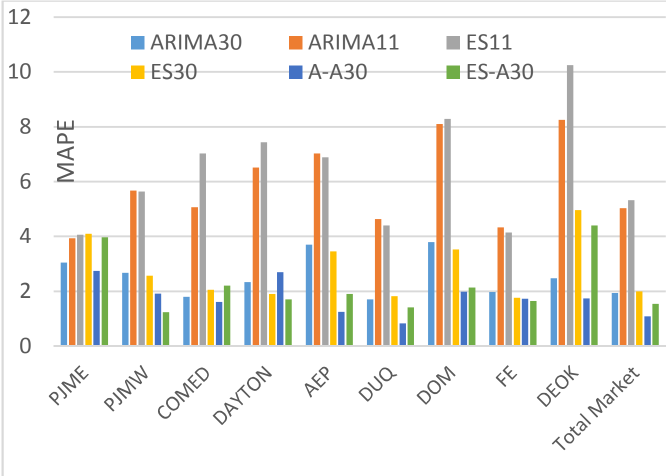

Fig. 8. *MAPE comparison for forecasts of each load area*

Table 1 and fig. 8 display 24-hour load forecasting accuracy metrics. It is observed that ARIMA forecasting using 30-day samples were found to be the most-accurate method, generating a mean absolute percentage error (MAPE) of 1.949% for the forecast of cumulative power demand of all PJM regions.

Fig. 5 shows that ARIMA and E.S. methods perform rather poorly when used in their traditional method when trained with one year of data. Fig. 6 suggests that training these methods with 30 days of data increases the accuracy and allows them to better find daily trends in the load data. The MAPE comparisons in fig. 8 show that ensemble methods perform most accurately in nearly all load areas in the PJM dataset. This suggests that ensemble methods are a more accurate method for forecasting daily electricity demand for these load regions. In general, ensemble methods consistently forecasted with MAPE values below 4%. Table 1 displays the MAPE values in a tabular form and shows that some ensemble methods performed forecasting with MAPE as low as 1%. This accuracy is very helpful for operations of utilities. Ensemble ARIMA-ARIMA is the most consistently accurate method examined in this work.

Table 1: *MAPE/MAD metrics from forecast for full PJM market*

Forecasting Method	MAD	MAPE
ARIMA-1 year	4287	5.316
Exp. Smoothing-1 year	4543	5.0314
ARIMA-30 days	1650	1.949
Exp. Smoothing- 30 days	1687	1.987
ARIMA-ARIMA-30 days	884	1.081
Exp. Smoothing-ARIMA-30	1782	1.544

VI. ECONOMIC DISPATCH RESULTS

Based on the results of the most accurate load forecasting (*e.g., ARIMA based on previous 30 days*), a simulated run of economic dispatch was carried for the forecasted hourly load of the following day, 11/17/2016. Fig. 9 shows the amount of power dispatched by the individual generators to meet the forecasted load demand. For example, at hour 0, GEN 1 to GEN5 should generate 13442 MW, 20754.1MW, 19869.7 MW, 7990.25MW and 10108.9MW respectively. Combining the forecasting scheme with an economic dispatch model allows for a day-ahead forecast for the scheduling of generators. This combination of techniques provides for operational insight in generator scheduling and planning from a utility perspective.

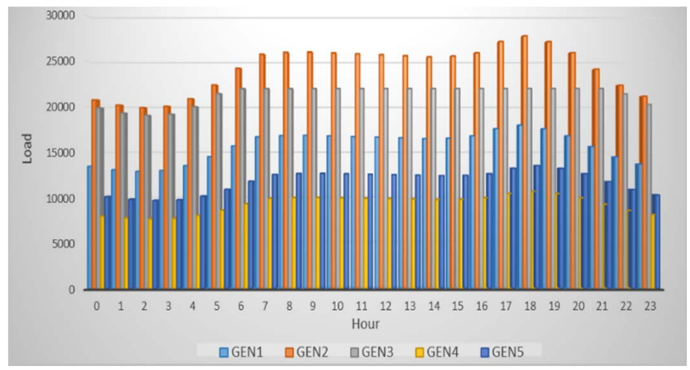

Fig. 9. *Economic dispatch of PJM generators*

VII. CONCLUSION

This work investigated the accuracies of ARIMA and ES implementation for PJM markets. The results indicate that ARIMA performs more accurately under both training scenarios. In addition, forecasts trained with 30 days prior to forecast day are more accurate. An economic dispatch was then performed. The combination of load forecasting and economic dispatch provides a simple framework for day-ahead optimal generation cost and generator scheduling. An ensemble forecasting method, specifically an ensemble ARIMA method in a two-stage combination that forecasts residuals allows for an increase in forecast accuracy. The ensemble ARIMA method produced a forecast accuracy of nearly 1%. This is a considerable increase in accuracy compared to traditional implementations of ARIMA and exponential smoothing. The increase in forecast accuracy is important for forecasting economic dispatch and the scheduling of generators.

VIII. ACKNOWLEDGEMENTS

This work is supported by National Science Foundation (NSF) under award number #1537565.

REFERENCES

[1] "PJM Hourly Estimated Hourly Load," 2017. .

[2] H. Alfares and M. Nazeeruddin, "Electric Load Forecasting: Literature Survey and Classification," *Int. J. Syst. Sci.*, pp. 23–24, 2010.

[3] K. Metaxiotix, A. Kagiannas, D. Askounis, and J. Psarras, "Artificial intelligence in short term electric load forecasting: a state-of-the art survey for the researcher," *Energy Convers. Manag.*, vol. 44, no. 9, pp. 1525–1534, 2003.

[4] T. Hong, T. Laing, and P. Wang, "Four Best Practices of Load Forecasting for Electric Cooperatives," in *2014 IEEE Rural Electric Power Conference*, 2014.

[5] E. Almeshaiei and H. Soltan, "A methodology for Electric Power Load Forecasting," *Alexandria Eng. J.*, vol. 50, no. 2, pp. 137–144, 2011.

[6] A. S. Nair, P. Ranganathan, and N. Kaabouch, "A Constrained Topological Decomposition Method for the Next-Generation Smart Grid," in *2017 International Conference on Electrical, Computer and Communication Technologies (IEEE ICECCT 2017)*, 2017, pp. 1–6.

[7] R. A. Chinnathambi and P. Ranganathan, "Investigation of forecasting methods for the hourly spot price of the day-ahead electric power markets," in *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*, 2016, pp. 3079–3086.

[8] R. Nau, "Statistical Forecasting: Notes on Regression and Time Series Analysis," 2017. .

[9] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting : the forecast package for R Automatic time series forecasting : the forecast package for R," *J. Stat. Softw.*, vol. 27, no. 3, pp. 1–22, 2008.

[10] R. J. Hyndman, "auto.arima," *RDocumentation*, 2017. .

[11] J. Contreras, R. Espínola, F. J. Nogales, and A. J. Conejo, "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1014–1020, 2003.

[12] D. Streiffert, "Multi-area Economic Dispatch with Tie Line Constraints," *IEEE Trans. Power Syst.*, vol. 10, no. 4, pp. 1946–1951, 1995.

[13] D. Streiffert, "Multi-area economic disptach with tie-line constraints," *IEEE Trans. Power Syst.*, vol. 10, no. 4, pp. 1946–1951, 1995.

[14] R. B. Cleveland, W. S. Cleveland, and I. Terpenning, "STL: A Seasonal-Trend Decomposition Procedure Based on Loess," *J. Off. Stat.*, vol. 6, no. 1, 1990.