This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/UPEC57427.2023.10294946, 2023 58th International Universities Power Engineering Conference (UPEC)

Investigating an Ensemble of ARIMA Models for Accurate Short-Term Electricity Demand Forecasting

Daniil Hulak
Department of Electronic and Electrical Engineering
Brunel University London
London, UK
Daniil.Hulak@brunel.ac.uk

Gareth Taylor
Department of Electronic and Electrical Engineering
Brunel University London
London, UK
Gareth.Taylor@brunel.ac.uk

This paper investigates the effectiveness of ensemble modelling for time series forecasting using Autoregressive Integrated Moving Average (ARIMA) models. In recent years, ensemble modelling has become a popular approach for improving forecasting accuracy by combining multiple models to achieve better performance than individual models. However, there is still limited research on the effectiveness of ensemble models for time series forecasting using ARIMA models. In this paper, we tested simple averaging of ARIMA models and investigate their performance in comparison to individual models. We conducted experiments using real-world datasets and evaluated the models' performance using metrics such as Mean Absolute Percentage Error, and Root Mean Squared Error. In this paper, we provided experiments on both short and long datasets to evaluate the performance of ensembled models compared to individual models. For the short datasets, our results clearly demonstrated the advantages of using ensembled models over individual models. The ensemble of models consistently outperformed the individual models in terms of accuracy. Our findings suggest that ensemble modelling can be a tool for time series forecasting and can provide improvements in accuracy. By leveraging the strengths of different models, ensemble models can effectively capture the underlying patterns in the data and make more accurate predictions.

Keywords — short-term forecasting, ARIMA, ensemble of models, open data, PJM, National Grid ESO

I. Introduction

Electric energy is an essential commodity that powers virtually all aspects of modern life. With the increasing global demand for electricity, it has become crucial to accurately forecast demand to ensure a stable and reliable energy supply [1]. Energy demand forecasting is critical in the energy industry, enabling utilities, grid operators and market traders to efficiently manage energy generation, transmission, and distribution systems. Accurate short-term energy demand forecasting is particularly important as it enables utilities to adjust their energy generation in real-time to match the demand, ensuring a stable and reliable energy supply.

In recent years, advanced statistical methods such as Autoregressive Integrated Moving Average (ARIMA) models have gained popularity for short-term energy demand forecasting due to their ability to capture complex temporal relationships in the data. However, since the energy demand data is stochastic and non-linear, the accuracy of a single ARIMA model can be limited.

In this paper, we investigate the use of an ensemble of ARIMA models for short-term energy demand forecasting using data from two major electricity markets, PJM, and National Grid ESO. We performed experiments on datasets of varying lengths to assess the performance of ensembled models in contrast to individual models.

We first evaluate the performance of single ARIMA models on both datasets and then demonstrate the effectiveness of ensemble techniques in improving forecasting accuracy. Specifically, we explore several variations of ensemble methods using simple averaging and compare their performance to that of a single ARIMA model.

Overall, our findings demonstrate the effectiveness of ARIMA models ensemble for short-term energy demand forecasting and provide insights into the best ensemble methods for improving accuracy. These results have important implications for utilities and grid operators, enabling them to improve their energy demand forecasting accuracy and better manage their energy generation and transmission systems.

II. Methodology and Data

A. Short-term and long-term energy demand forecasting comparison

Short-term energy demand forecasting, typically ranging from a few hours to a few days ahead, is critical for all market participants in the energy industry. Authors of [2] short-term energy demand forecasting refer to the hourly prediction of electricity demand or load for a lead time ranging from 1 hour to several days ahead, although some datasets are available at half-hour intervals, which can also be considered as short-term forecasting. Accurate short-term forecasting enables utilities and grid operators to adjust their real-time energy generation and distribution systems to match the demand, ensuring a stable and reliable energy supply. It also allows energy traders and market participants to make informed decisions about energy pricing and trading, minimizing the risk of financial losses due to inaccurate forecasting.

Long-term energy demand forecasting, on the other hand, typically ranges from several months to several years ahead and is primarily used for capacity planning and investment decisions [3]. Long-term forecasting is important for utilities and grid operators to make informed decisions about building new power plants, expanding transmission and distribution systems, and allocating resources. It also provides insights into future energy demand trends and can inform energy policy decisions.

While long-term forecasting is important for capacity planning and investment decisions, short-term forecasting is crucial for day-to-day operations in the energy industry [4]. Short-term forecasting provides real-time insights into energy

British Academy, Council for At-Risk Academics

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/UPEC57427.2023.10294946, 2023 58th International Universities Power Engineering Conference (UPEC)

demand patterns and enables market participants to respond quickly to changes in demand. This is particularly important in volatile energy markets, where sudden changes in energy demand can lead to price spikes and supply shortages.

For markets with high penetration of renewable energy sources, such as wind and solar, short-term energy demand forecasting becomes even more critical. This is because renewable energy sources are highly variable and difficult to predict, which can result in significant fluctuations in energy supply and demand.

B. General approach using ARIMA models

Today, the most advanced methods of forecasting energy demand involve the use of artificial intelligence (AI) [5] and deep learning [6] algorithms, which have demonstrated superior performance compared to traditional statistical models. However, these methods require significant computational power and can be economically expensive, making them less suitable for real-time or near-real-time applications. Additionally, the complexity of these methods can make it difficult to interpret their results, limiting their applicability in certain contexts.

Statistical methods of forecasting, such as ARIMA models, are often more appropriate than deep learning and AI methods when it comes to time series forecasting. This is because statistical methods are designed to capture the patterns and trends in time series data, whereas deep learning and AI methods are more suited to handling complex data structures such as images and natural language [7]. Furthermore, statistical methods have a proven track record in time series analysis, with established methodologies and frameworks that have been developed over many years.

In addition to their proven effectiveness in time series analysis, statistical forecasting methods typically require less computational power than deep learning and AI methods [8]. This makes them more accessible and practical for many forecasting applications, particularly those that involve large datasets or limited computing resources. While deep learning and AI methods have shown promise in certain areas of forecasting, such as natural language processing and image recognition, statistical methods remain a reliable and widely used tool for time series forecasting.

Accordingly, statistical models are useful for forecasting time series data as they are designed to analyse patterns and relationships within the data, which can be used to make accurate predictions about future values. These models can also be adjusted to account for changes in the data over time, allowing them to adapt to changing conditions and produce reliable forecasts.

Considering these challenges, it is important to mention that using simple and fast methods for forecasting energy demand is good because such methods can respond rapidly to changes in energy demand and supply conditions, making them useful in emergencies or when real-time or near real-time forecasting is required. Furthermore, simple methods can be more accessible and easier to use than advanced techniques, making them suitable for organizations with limited computational power or expertise.

For statistical models used in time series forecasting, accuracy is the most crucial factor as the goal is to produce reliable predictions that can help inform decision-making and planning. Even small errors in the forecast can have significant consequences, which is why the models must be carefully calibrated and validated to ensure their accuracy.

ARIMA models are a type of time series model that is widely used in energy demand forecasting. They are particularly useful when dealing with non-stationary time series data, where the mean and/or variance of the series changes over time. ARIMA models can capture both the long-term trends and short-term fluctuations in the data by considering the previous values of the series and the differences between these values. Another benefit of ARIMA models is their ability to handle missing data and data outliers. The description of the main components [9] of the ARIMA model stands as follows:

AR – autoregressive, which means that the model uses past values of the series to predict future values.

I – integrated, which means that the model uses differencing to make the series stationary.

MA – moving average, which means that the model uses past errors to predict future values.

In general, ARIMA model formula looks as follows:

$$Y_t = c + \phi_1 y_{dt-1} + \phi_p y_{dt-p} + \cdots + \theta_1 e_{t-1} + \theta_q e_{t-q} + e_t \quad (1)$$

Where:

ϕ is the AR characteristic polynomial.

θ is the MA characteristic polynomial.

e_t is an error term.

c is a constant.

However, there are also some disadvantages to using ARIMA models. One of the main challenges is selecting the appropriate parameters for the model. This can be a time-consuming process that requires expertise in time series analysis.

The p, d, and q coefficients of the ARIMA (p,d,q) model refer to the orders of the autoregressive, differencing, and moving average components of the model, respectively.

The autoregressive (AR) component involves the use of past values of the time series to predict future values and has an order of p. The moving average (MA) component involves the use of past error terms to predict future values and has an order of q. The integrated (I) component involves differencing the time series to remove trends or seasonality and has an order of d.

The values of p, d, and q are selected based on the characteristics of the time series being modelled. A higher value of p indicates a stronger dependence on past values, a higher value of q indicates a stronger dependence on past errors, and a higher value of d indicates a higher degree of differencing needed to remove trends or seasonality from the data. Choosing the appropriate values of p, d, and q is crucial for accurate modelling and forecasting of time series data using the ARIMA model.

In this paper, we use ARIMA models with p, d, and q coefficients equal from 0 to 2. For all the simulations and calculations, we use the programming and numeric computing platform MATLAB.

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/UPEC57427.2023.10294946, 2023 58th International Universities Power Engineering Conference (UPEC)

The goodness of fit of an ARIMA model can be evaluated using various measures, such as Akaike Information Criterion (AIC) [10]. To assess the accuracy of the forecasted results generated by the ARIMA model, commonly used measures include Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE). The formulas for these coefficients look as follows:

$$MAPE = \left(\sum_{t=T+h}^{T+h} \left|\frac{\widehat{y_t}-y_t}{y_t}\right| / h\right) * 100\% \quad (2)$$

$$RMSE = \sqrt{\sum_{t=T+1}^{T+h} (\widehat{y_t} - y_t)^2 / h} \quad (3)$$

Where:

$\widehat{y_t}$ is predicted (forecasted) value.

y_t is the real (observed) value.

h, in turn, is the number of forecasting steps ahead [11].

C. Datasets used in paper

Accurate energy demand forecasting relies on high-quality and reliable datasets that capture the complex and dynamic patterns of energy consumption. Energy demand datasets can be challenging to obtain due to the sensitive nature of energy consumption data and the complex data acquisition processes used in the energy industry.

Open data initiatives have emerged as a solution to these challenges, providing public access to energy demand data and enabling researchers and industry professionals to develop better forecasting models and tools. Open data initiatives offer a range of benefits, including improved transparency, collaboration, and innovation in the energy industry.

In this paper, we use energy demand datasets from two of the large energy markets of the world: PJM and National Grid ESO. PJM is a wholesale regional transmission system operator responsible for managing the transmission of electricity in 13 states in the United States, serving approximately 65 million people [12]. National Grid ESO is the national electricity transmission system operator for Great Britain, responsible for managing the flow of electricity across the country[13].

The energy markets served by PJM and National Grid ESO also differ in terms of their size, complexity, and the types of energy sources used. PJM serves a large and diverse market that includes a mix of energy sources, including coal, natural gas, nuclear, and renewable energy [12]. National Grid ESO serves a smaller market that is heavily reliant on renewable energy sources, such as wind and solar [13].

Both PJM and National Grid ESO provide open access to their energy demand data. The energy demand dataset from PJM contains hourly energy consumption data for the period from 2009 to 2023. On the other hand, the energy demand dataset from National Grid ESO contains half-hourly energy consumption data for the period from 2000 to 2023.

In this paper, we use both PJM's hourly energy demand dataset and National Grid ESO's half-hourly energy demand dataset to evaluate the effectiveness of an ensemble of ARIMA models for short-term energy demand forecasting. We use the most recent data from both datasets, focusing on the period from 2022 to 2023. This allows us to evaluate the effectiveness of forecasting using the ensemble of ARIMA models for short-term energy demand forecasting using the most up-to-date data. By using the most recent data, we can provide insights that are relevant to current market conditions and help improve the accuracy of short-term energy demand forecasting in real-world scenarios.

In the paper, we utilize a large datasets, but due to its extensive size, we only include a select few blocks of values to represent its content in the tab.I.

TABLE I. BREIF DESCRIPTION OF THE DATASETS USED

Value number	Energy demand for the National Grid ESO (every half hour), MW	Energy demand for the PJM (every hour), MW
1	21 940	130 754
2	22 427	125 644
3	21 896	121 109
4	20 693	119 146
…	…	…
…	…	…
…	…	…
8757	22 366	160 248
8758	21 955	151 878
8759	21 655	145 362
8760	21 332	135 924
…	…	X
…	…	X
…	…	X
17517	23 223	X
17518	22 627	X
17519	21 690	X
17520	21 229	X

D. Review of the forecasting methods, used in the other research papers

There has been lots of research in the field, with numerous papers exploring different techniques and methodologies for forecasting electric energy demand. In this paper, we will examine some of the latest research in this area, highlighting the strengths and weaknesses of different approaches. We will focus specifically on research papers that incorporate ensemble methods or combinations of techniques for energy consumption forecasting. A key requirement for these approaches is the use the statistical models.

The paper [10] proposes a new approach for predicting energy consumption using the Ensemble Empirical Mode Decomposition (EEMD) and Autoregressive Moving Average (ARMA) models. The EEMD-ARMA model is applied to time series data of energy consumption, which is decomposed into different Intrinsic Mode Functions (IMFs) using EEMD. The ARMA model is then used to predict the future values of each IMF and the final prediction is obtained by summing the predicted values of all IMFs.

The results of the study demonstrate the effectiveness of the proposed EEMD-ARMA model in predicting energy consumption. The model achieved high accuracy in terms of MAPE and RMSE compared to other models, such as the autoregressive integrated moving average ARIMA and Support Vector Regression (SVR).

Authors [11] proposes a new hybrid approach for mid-long-term electric energy consumption forecasting using bagging ARIMA and exponential smoothing methods. The bagging technique is applied to ARIMA and exponential smoothing models to reduce the variance of the forecasting error and improve the accuracy of the prediction.

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/UPEC57427.2023.10294946, 2023 58th International Universities Power Engineering Conference (UPEC)

Their results show the effectiveness of the proposed hybrid approach in accurately forecasting mid-long-term electric energy consumption. The model achieved high accuracy in terms of MAPE and Mean Absolute Error (MAE) compared to other models, such as the seasonal ARIMA and exponential smoothing methods. The study highlights the importance of considering both time series patterns and historical trends in mid-long-term electric energy consumption forecasting, which the proposed method can address.

In the work 'A Building Energy Consumption Prediction Method Based on Random Forest and ARMA' [14] authors propose a new approach for predicting building energy consumption using the Random Forest (RF) and ARMA models. The method combines the strengths of RF in handling high-dimensional and non-linear data with ARMA's ability to capture time series patterns and trends.

The results of the study demonstrate the effectiveness of the proposed method in accurately predicting building energy consumption. The model achieved high accuracy in terms of MAPE and MAE compared to other models, such as the artificial neural network and SVR. Paper emphasizes the importance of using a combination of machine learning and time series analysis in building energy consumption prediction, especially in scenarios with complex and dynamic data.

In conclusion, there are numerous papers in the literature that focus on forecasting energy consumption using a variety of techniques and models. Many of these studies use combinations of statistical models with other methods such as machine learning, ensemble learning, and time series analysis. The hybrid models tend to show promising results and are becoming increasingly popular.

In this chapter, we have reviewed a few examples of research papers that utilize different forecasting methods. However, it is important to note that these are just a few of the many approaches that have been proposed and evaluated in the literature. Future research may explore different combinations of models and methods or even introduce new techniques to further improve the accuracy of energy consumption forecasting.

E. This paper experiment details and performance algorithm

This chapter will detail the experiment and performance algorithm used in this paper. The experiment involves 2 cases of using the data. We used different types of the datasets to predict the first 40 values of 2023 based on the long and short type datasets. Long type dataset is limited to the entire previous year (2022) which is more than 17K values for the National Grid ESO and around 9K for the PJM. Short time dataset is limited to no more than 100 real values of 2022 which is in comparison to the long dataset evidence the lack of data. The datasets used for the experiment were obtained from public sources and we assume them to be accurate and error-free, accordingly no specific data preprocessing techniques were applied to them. The first forecast (N+1) is based on the entire dataset. The forecast is then compared to the real value, and MAPE with RMSE coefficients are calculated. The real value is then added to the dataset, and the N+2 value is forecasted, repeating the procedure for all 40 values.

For each step, all possible ARIMA (p,d,q) models with p,d,q ranging from 0 to 2 are built. Such values are quite standard for the forecasting. Using high values of p, d, and q coefficients in ARIMA models can lead to overfitting, increased computational complexity, and unreliable parameter estimates. It is generally recommended to choose lower values for these coefficients that strike a balance between model complexity and predictive performance. A comparative analysis is conducted to compare the best ARIMA model to an ensemble of 2 to 20 averaged ARIMA models. In the paper, we employed a simplified method, such as standard averaging, where we computed the sum of the forecasted values and divided it by the number of models utilized.

The algorithm used in our investigations is described on the fig1:

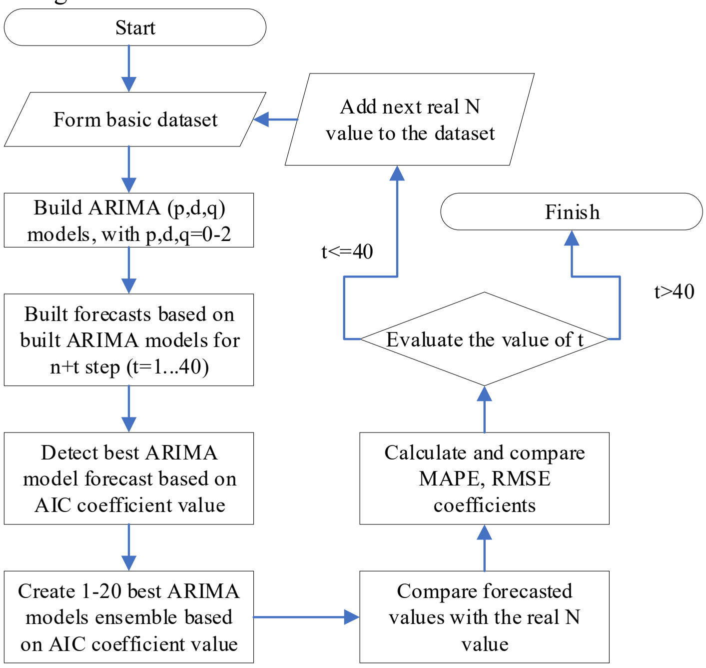

Fig. 1. Algorithm of building ARIMA models, and comparing them

To compare the accuracy of forecasts for two datasets, we would need to follow a similar process for each dataset. Once we have obtained forecasts for both datasets, we can compare the results to assess the accuracy of the models. The way to do this is to use MAPE and RMSE metrics to measure the difference between the predicted values and the actual values for the forecast period. While there are various metrics available for evaluating forecast accuracy, we chose to focus on MAPE and RMSE due to the limitations of the experiment and the scope of this paper. These metrics are widely used in forecasting literature and provide a standardized way of quantifying the prediction errors. However, it is important to note that other metrics, such as MAE can also be valuable in certain contexts. Given the specific goals and constraints of our experiment, we determined that MAPE and RMSE would provide sufficient insights into the performance of the models without overly complicating the analysis.

If these values are similar for both datasets, it suggests that the forecasting models are performing similarly for both datasets. However, if there is a significant difference in the accuracy of the forecasts for the two datasets, it may indicate

This article has been accepted for publication in a future proceedings of this conference, but has not been fully edited. Content may change prior to final publication.
Citation information: DOI: 10.1109/UPEC57427.2023.10294946, 2023 58th International Universities Power Engineering Conference (UPEC)

that there are underlying differences in the data or that one dataset is more complex than the other.

Overall, comparing the results of forecasting models for multiple datasets can help us to understand how well the models generalize to new data and can inform us about the strengths and limitations of the models or their ensembles.

III. Empirical Results

After building and comparing the forecasts for both datasets and 2 cases, we obtained the following results shown of MAPE and RMSE coefficients shown for working with National Grid ESO datasets in fig2 and the same for the PJM in fig.3.

We started by analyzing the performance of the ensemble of ARIMA models on long data series. The ensemble displayed behavior close to disarray, making it difficult to determine which model, either single or ensemble, would perform better. The chaotic behavior observed in the ensemble could be attributed to the complex interactions among the individual models. As a result, the predictions from the ensemble showed high variability, making it challenging to rely on them for accurate forecasting. The lack of stability and consistent performance in the ensemble indicated that using it for long data series may not be ideal.

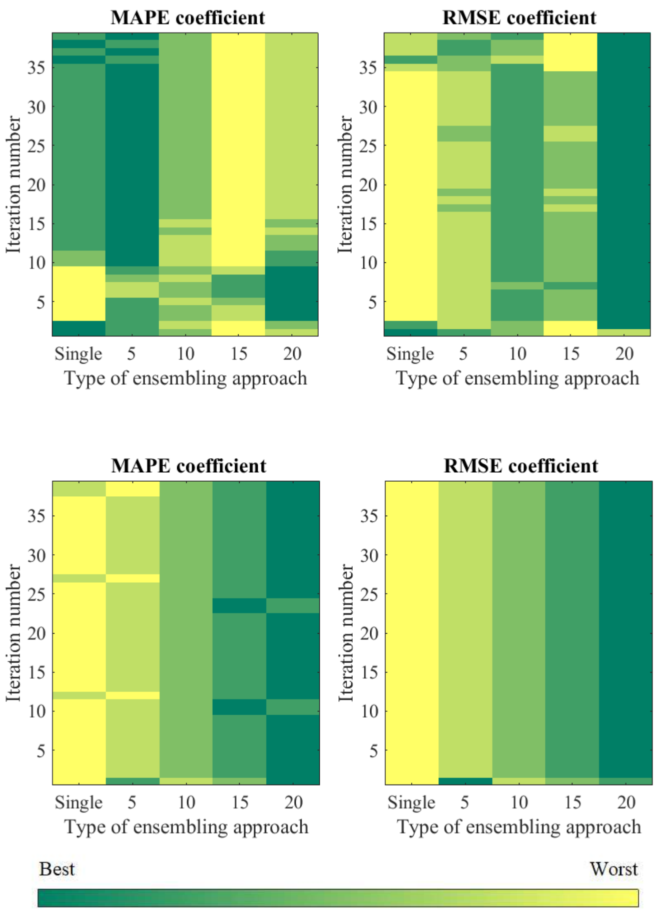

Fig. 2. Results comparing MAPE, RMSE coefficients for National Grid ESO with using long (upper graphs) and short (bottom graphs) datasets

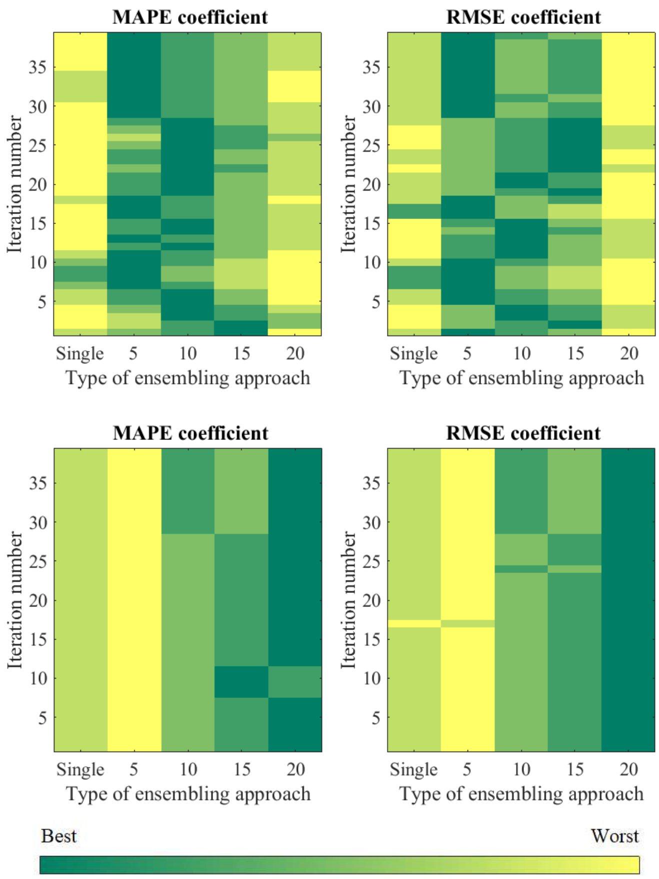

Fig. 3. Results comparing MAPE, RMSE coefficients for PJM with using long (upper graphs) and short (bottom graphs) datasets

Further we explored the potential of ensembles on the short data series. The results obtained with the ensemble of ARIMA models on short data series were more promising. We observed better stability and improved forecasting accuracy when compared to single models.

Comparative analysis of the coefficients obtained from the single ARIMA models and the ensemble approach, specifically focusing on the results obtained from the short-term dataset. However the difference in coefficients obtained from the ensemble were found to be very small. Comparison of performance of 20 ensembled ARIMA models on the last forecasting step to the single models shows the results which are reflected in the tab.II:

TABLE II. Comparison of the Models Performance Using the Coefficients Values

Dataset	Coefficient	Single ARIMA model	Ensemble of 20 ARIMA models	Difference, %
National Grid ESO	MAPE	1.299	1.268	2.4
	RMSE	437.1	395.7	9.5
PJM	MAPE	1.351	1.225	9.3
	RMSE	2492.9	2412.1	3.2

This finding indicated that the ensemble approach effectively reduced the individual model's reliance on specific coefficients, leading to a more balanced and robust forecasting model. By combining multiple models the ensemble minimized the impact of individual model and resulting in a more generalized and accurate forecast.

With the lack of data ensembles produced more reliable predictions, demonstrating their ability to capture the underlying patterns in the data. The ensemble approach leveraged the diversity of the individual models to provide a robust forecast. The increased stability of the ensemble made it a valuable tool for short-term forecasting tasks. These findings could be applied by energy companies of any types which can leverage ensemble modeling based on the results of our study, particularly in situations where data may be limited or time constraints require fast results. Ensemble models, By combining multiple ARIMA models, energy companies can benefit from improved forecasting accuracy and make more informed decisions regarding resource planning, demand management, and operational strategies.

For future papers, we plan also consider other open datasets in addition to the ones used in this study. This will allow us to test the effectiveness of ensemble modeling on different types of data and provide a more comprehensive understanding of its potential applications. Additionally, we aim to explore the performance of ensemble models on the same datasets but over different years to assess their robustness and stability over time.

IV. Conclusions

In conclusion, our paper tested the ensemble approach for ARIMA models and achieved good results. Working with the datasets lack of data the ensembled models outperformed individual models, indicating the effectiveness of combining multiple models for forecasting. However, we recognize that there is always room for improvement, and for future papers, we plan to explore even more advanced methods of working with data, such as stacking, boosting, and other ensemble techniques, in order to further improve the performance of our models. Overall, our findings suggest that the use of ensemble models can provide significant benefits for time series forecasting.

[1] M. A. Islam, H. S. Che, M. Hasanuzzaman, and N. A. Rahim, “Energy demand forecasting,” *Energy for Sustainable Development: Demand, Supply, Conversion and Management*, pp. 105–123, Jan. 2020, doi: 10.1016/B978-0-12-814645-3.00005-5.

[2] M. El-Telbany and F. El-Karmi, “Short-term forecasting of Jordanian electricity demand using particle swarm optimization,” *Electric Power Systems Research*, vol. 78, no. 3, pp. 425–433, Mar. 2008, doi: 10.1016/J.EPSR.2007.03.011.

[3] S. P. Filippov, V. A. Malakhov, and F. V Veselov, “Long-Term Energy Demand Forecasting Based on a Systems Analysis,” *Thermal Engineering*, vol. 68, no. 12, pp. 881–894, 2021, doi: 10.1134/S0040601521120041.

[4] N. Mohan, K. P. Soman, and S. Sachin Kumar, “A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model,” *Appl Energy*, vol. 232, pp. 229–244, Dec. 2018, doi: 10.1016/J.APENERGY.2018.09.190.

[5] E. Ofori-Ntow Jnr, Y. Y. Ziggah, and S. Relvas, “Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting,” *Sustain Cities Soc*, vol. 66, p. 102679, Mar. 2021, doi: 10.1016/J.SCS.2020.102679.

[6] J. Bedi and D. Toshniwal, “Deep learning framework to forecast electricity demand,” *Appl Energy*, vol. 238, pp. 1312–1326, Mar. 2019, doi: 10.1016/J.APENERGY.2019.01.113.

[7] M. Oussalah, “AI Explainability. A Bridge Between Machine Vision and Natural Language Processing,” in *Pattern Recognition. ICPR International Workshops and Challenges*, A. Del Bimbo, R. Cucchiara, S. Sclaroff, G. M. Farinella, T. Mei, M. Bertini, H. J. Escalante, and R. Vezzani, Eds., Cham: Springer International Publishing, 2021, pp. 257–273.

[8] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson, “The future of forecasting for renewable energy,” *Wiley Interdiscip Rev Energy Environ*, vol. 9, no. 2, p. e365, Mar. 2020, doi: 10.1002/WENE.365.

[9] Y. Yu, J. Wang, M. Song, and J. Song, “Network Traffic Prediction and Result Analysis Based on Seasonal ARIMA and Correlation Coefficient Network Planning and Optimization,” *2010 International Conference on Intelligent System Design and Engineering Application*, vol. 1, 2010, doi: 10.1109/ISDEA.2010.335.

[10] J. Li *et al.*, “Energy Consumption Data Prediction Analysis based on EEMD-ARMA Model,” in *2020 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2020, pp. 1338–1342. doi: 10.1109/ICMA49215.2020.9233741.

[11] E. M. de Oliveira and F. L. Cyrino Oliveira, “Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods,” *Energy*, vol. 144, pp. 776–788, 2018, doi: https://doi.org/10.1016/j.energy.2017.12.049.

[12] PJM, “About PJM.” https://www.pjm.com/about-pjm (accessed May 02, 2023).

[13] National Grid ESO, “What we do.” https://www.nationalgrideso.com/what-we-do (accessed May 02, 2023).

[14] B. Jiang, Z. Cheng, Q. Hao, and N. Ma, “A Building Energy Consumption Prediction Method Based on Random Forest and ARMA,” in *2018 Chinese Automation Congress (CAC)*, 2018, pp. 3550–3555. doi: 10.1109/CAC.2018.8623540.