A Comprehensive Analysis of UK Electricity Consumption Using the OSI Model for Data Communication and Predictive Analytics

MSc Research Project
Data Analytics

Neethu Elizhabeth Sebastian
Student ID: x23179996

School of Computing
National College of Ireland

Supervisor: Noel Cosgrave

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Neethu Elizhabeth Sebastian
Student ID:	x23179996
Programme:	Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Noel Cosgrave
Submission Due Date:	29/1/2025
Project Title:	A Comprehensive Analysis of UK Electricity Consumption Using the OSI Model for Data Communication and Predictive Analytics
Word Count:	7790
Page Count:	21

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Neethu Elizhabeth Sebastian
Date:	28th January 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

A Comprehensive Analysis of UK Electricity Consumption Using the OSI Model for Data Communication and Predictive Analytics

Neethu Elizhabeth Sebastian
x23179996

Abstract

Demand forecasting for electricity is important in energy planning and management, for the stability of the electric power grid and the incorporation of renewable resources. This paper aims to establish the accuracy of applying sophisticated forecasting methods that can estimate UK electricity consumption using historical data National Grid ESO for the period 2009 to 2024. Since a large share of renewable energy is expected to be integrated into energy systems, it is paramount to precisely forecast the electricity demand. Thus, the present study employs ARIMA, LSTM, and hybrid models to predict the UK electricity demand based on the dataset of 2009–2024. The work employs a structured-data processing pipeline derived from the OSI model designed to improve and organize the predictive modeling process. The transformations involve normalization, lag features, and seasonality encoding obtained from national demand, wind and/or solar generation, and interconnector flows. The findings highlighted here establish that the effectiveness of the ARIMA model and the LSTM model for the small-scale wind speed data: ARIMA is powerful in capturing linear and periodic trends while LSTM supersedes in capturing other types of trends and long-term oscillations. Other improvements include the extension of convolutional layers with recurrent networks using both local dependencies and temporal patterns. This model simplifies the sending process of data and enhances the efficacy of model training through the advancement in processing. These conclusions provide practical recommendations for grid operators and policymakers to improve grid control, decrease the dependency on conventional resources, and better incorporate renewable energy. Future work should extend to dynamic integration of data, external factors like weather and detailed study of hybrid models to enhance their performance and flexibility. This study is a clear example of how machine learning and structured data framework can revolutionise energy forecasting techniques.

Keywords: Electricity Demand Forecasting, LSTM (Long Short-Term Memory), Machine Learning, Time-Series Analysis, OSI Model (Open Systems Interconnection)

1 Introduction

Electric power is one of the key parameters of socioeconomic development, and the prediction of electricity demand is critical for energy efficiency. That is why accurate forecasting of future energy demand is an important but not a simple problem in the UK, where electricity consumption depends on some factors, including the population, industrial activity, weather, and the economic crisis. Traditionally, demand forecasting has been applied in energy sector for organizing the supply of resources, for projecting the future requirements of actual structures, and for effective functioning of energy systems. Data science, machine learning, and big data analysis have increased the precision in electricity usage rates' forecasts over the last decade. One of these is the use of predictive models, which can analyze data from past experience to make predictions on the future.

Originally developed for networking and communication systems known as OSI (Open Systems Interconnection) model, in place in the systems engineering for data analysis, can be effectively used since it defines the structure of sending and receiving data. With the help of the OSI model that has seven layers, the data transfer in systems is comprehensive and systematic and can be strung with the data processing workflows in the area of the predictive modeling detailed below and constitutes a valuable organizational framework for the management of the flow and transformation of the electricity consumption data.

1.1 Problem statement

As mentioned above, extensive amount of historical data is available for electricity consumption and yet, predicting the future demand of electricity often poses a problem for the energy providers, especially in dynamic world. Usually, in traditional approaches the characteristics of the change in the data stream are not considered or they are considered to a limited extent only. In addition, although the OSI model has been used extensively in network communications, its prospect of usage in data science for recommendation system in energy industry has not been researched extensively. This research fits into this gap to enhance electricity consumption data organised using the OSI model and the development of enhanced forecasting models.

1.2 Research Question

Primary Research question What strategies can be applied to enhance electricity load forecast using advanced deep learning techniques for London?

Sub-questions

1. *How can the OSI model be adapted to facilitate the analysis and prediction of electricity consumption data?*

2. *What impact does the integration of OSI layers have on the accuracy and efficiency of predictive modeling for electricity demand?*

3. *Can predictive models accurately forecast UK electricity consumption based on historical data from 2009 to 2022?*

1.3 Objectives of the Study

The OSI (Open Systems Interconnection) model is used in this research to set up hierarchical data processing flow for electricity consumption forecasting in the UK. While it was especially designed for use in networks, the OSI model is an effective organizational structure which can be implemented to co-ordinate the flow of data at various steps of predictive analysis. It was seen that these different layers of OSI model could be brought in line with stages of data processing in the transformation and management of electricity consumption data. For instance, the model helps in making the data as well formatted and regular, time-related, consistent, and enables the encouraging transportation of data from one node to the other of a network; this, in turn, augments the correctness and speed of the predictive modeling. The model thus serves as a guideline to improve the flow of data to prevent the processes of predictive analytics from being congested and complicated.

1.4 Significance of the Study

The contribution of this work arises from the fact that energy forecasting has not been much interfaced with the OSI model despite being crucial in today's world. Consequently, by providing a detailed examination of applying the OSI framework to data science, this research can help enhance the energy demand predicting processes, making recommendations regarding energy forecasts of significant importance for policymakers, utility companies, and other businesses. Loosely coupled with improved accuracy mainly shape out the problem of misty central planning, decreased energy wastage and improved organizational decisions. In addition, this study expands the knowledge of extending OSI principles for usage in networking applications, discovers new opportunities in systems engineering for data analysis.

2 Related Work

2.1 Overview of Electricity Consumption Forecasting

The forecasting of electricity consumption is an important research discipline for the energy utilities, industries and governments. It is important to predict the future energy requirement in order to control and manage large electricity generation and distribution companies. With forecast of electricity consumption, it becomes possible to manage operations of electrical utilities such that generation and distribution are in harmony with consumption to avoid situations where there is excess production of electricity – a condition that would result in wastage due to excess supply or instabilities such as power rationing or shortages due to inadequate supply Raza et al. (2022). Long-term forecasting of energy requirements also enhances long-term planning of infrastructure and long-term planning strategies such as demand response, energy costing culture, and load reduction.

Short-term electricity consumption forecasting can also be referred to as intraday, daily or weekly while long-term electricity consumption forecasting can be identified as monthly, quarterly or yearly. While short-run forecasting is used to forecast consumption over weeks, months, this has applicability in the planning of investments especially in power generation, infrastructure and policies that take years and decades into consideration.

Different factors such as seasonality, economic and social factors such as population growth, technological, and other social factors, such as changes in energy use patterns, affect the forecasting process. For instance, in hot summer season or a cold winter season, energy requirements are high, mainly for heating or cooling. Because energy consumption cannot be predicted solely by the past data but by the existence of factors that may affect the demand for energy, the task is challenging.

The forecasting of electricity demand has all the while changed and relies on statistical models and modern machine learning to tackle different issues including nonlinearity, seasonality, and the incorporation of sustainable energy resources. It's important to note that traditional means of forecasting models are based upon historical or linear extrapolations and are not capable of handling large data amounts. Using statistical models dominated previous investigations and aggregative approaches were more common. For example, Nti et al. (2020) provided a systematic literature on the application of ARIMA and exponential smoothing which revealed that these models are good for modelling linear and seasonal patterns, but poor when coping with nonlinear patterns. Similarly, Raza et al. (2022) investigated ARIMA for energy sector in Pakistan, concluded it appropriate for short term, stable demand but not suitable for volatile environment.

Deep diving into Machine Learning and Artificial Intelligence Tools

It can be mentioned that with the help of machine learning, the administration of forecasting has reached a whole new level. In their study, Dinmohammadi et al. (2023) applied artificial neural network model to predict energy consumption in residential buildings; the author also emphasized the importance of feature engineering and ensemble learning methods. Namely, Liu et al. (2021) proved that LSTM networks are better suited at identifying long-term dependencies inherent in time series of electricity consumption data. Kazemzadeh et al. (2020) proposed the decision tree with SVR to create hybrid models to predict using both linear and non-linear features which introduced higher accuracy. Therefore,Yang and Zhou (2022) observed that deep learning models proved superior to traditional models in predicting demand that results from renewable power inputs.

Hybrid Models

Thus, certain types of hybrid methods are increasingly popular, using advantages of several methods. Liu et al. (2020)combined Convolutional Neural Networks and Long Short-Term Memory to analyze spatial features and temporal features in electricity demand. Cascone et al. (2023) used convolutional LSTM networks for household electricity forecasting focused on possible fluctuations in the periods and seasons. Sarwar et al. (2024) put forward models based on a integration of machine learning approach and economic data to improve both electricity price and demand forecasting.

Renewable Energy Integration

Variability of demand patterns has been brought about by deployment of renewable energy hence the need for sophisticated forecasting solutions. Hong et al. (2020), discussed in 2020, conducted a review of the energy forecasting techniques pays dedication to the REs data. Similarly, using variables like policy shifts and renewable energy generation, Hu and Man (2023) and co-authors (2023) investigated the forecasting of carbon emissions and industrial energy utilization. Zhang et al. (2020) examined how solar and wind generation data enable enhanced demand predictions.

Role of External Variables

Many of these studies have stressed the need for adopting external contingencies. Tian et al. (2020) incorporate weather data into building energy prediction, which has yielded better model performance results. Later, Birim et al. (2024) expanded this approach to add demographic and economic variables, particular on the focal aspect of consumption behavior. Likewise, Alasali et al. (2021) investigated the impact of COVID-19 on the electricity demand recognizing the imposition of multifaceted disruptions.

Current trends and systematized findings

The current reviews have integrated the progressions to the forecasting. Seasonal patterns are also well-handled with the help of temporal convolutional networks, as Lara-Benítez et al. (2020) mentioned. Multi-attribute decision-making techniques were reviewed by Somu and Kowli (2024), the authors called for greater use of composite data moreover. Almazrouee et al. (2020b)has compared the conventional technique such as Holt-Winters with machine learning technique and the later proved to be more dynamic.

When it comes to structuring data workflows, the OSI model defines a new perspective. Kazemzadeh et al. (2020) aligned energy system data with OSI layers to enhance faster data processing for predictive analysis. Similar principles Birim et al. (2024) applied the proposed technique resulting in enhanced data flow and model performance..

Such papers as Alasali et al. (2021) also highlighted the realistic benefits of enhanced forecasting for grid control and renewable integration. For instance, Zhang et al. (2020)discussed policy-supported models, associating forecasting developments with energy sustainability agenda.

2.2 Techniques in Time Series Forecasting and Predictive Modeling

Consumption data of electricity is much of temporal; do lead the time series analysis for the forecasting of the future demand from the observed records. Using time series implies that data were collected in sequential order of time and often at a specific period, leading to the subsequent forecasting of given data set. Many approaches have been proposed for forecasting time series from simple statistical methods to the most recent artificial intelligence techniques

Among all of the statistical models, perhaps the most well-known is Autoregressive Integrated Moving Average or ARIMA for short. ARIMA is a linear model that combines three components: There are three primary models namely autoregression (AR), differencing (I) and moving averages (MA). ARIMA models are the forecasts based on the past observations of the series and therefore identifies short-term dynamic Liu et al. (2021). However, it was observed that the performance of ARIMA tends to compromised in presence of non-linear relationships or long-term dependence.

Another classical one used is Exponential Smoothing (ETS) a method most suited for data with high seasonality. The Holt-Winters method is a further development of ETS that enables modelling trend and seasonal factors, which is more important for electricity demand forecasting as an increasing number of countries witness the effect of these factors Nti et al. (2020). ETS methods are more flexible by nature by assigning higher value to the more recent data thereby making them capture the change in trend relatively faster than ARIMA models.

Aside from these approaches, ones that are more modern involve use of machine learning and artificial intelligence (AI) to analyse data having non-linear relationships. Such methods as Artificial Neural Networks (ANNs), more so Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have become popular in recent past. Such models can effectively analyze patterns that include long-term dependencies, which are usually inherent in voltage consumption data.

2.3 The Open Systems Interconnection (OSI) Model

Open systems interconnection (OSI) is a theoretical framework used to address how the hierarchical layers of network communication systems exchange data. The International Organization formulated the OSI reference model for Standardization (ISO) and divides communication into seven major hierarchical layers each of which possesses a definite set of operations.

OSI model is useful in designing and studying the network protocols and architectural systems. The device interconnectivity makes it possible for devices developed and manufactured by different companies to at least interact in some capacity irrespective of the technology or design that has put in place. OSI model is widely used for analysis, diagnosis of network and communication problems and integration of networks.

The Layers of the OSI Model

The OSI model consists of seven layers, each serving a distinct function:

- Application Layer: This is yet another layer that is directly in contact with the end user and offers services to facilitate usage applications in the network. These include file transfer protocols, electronic mail and World Wide Web protocols Alasali et al. (2021).
- Presentation Layer: This is involving managing of data format, encryption and compression for proper format in order to understand by both the sending and the receiving systems.
- Session Layer: Responsible for passing the data from application to application and is responsible for setting up creating and ending the communication sessions.
- Transport Layer: Facilities accurate data exchange from one system to another. It controls for error, guides the flow and organizes data so that they were segmented.
- Network Layer: Responsible for controlling the flow of data through networks and checks that DATA reaching a certain network or point were formatted correctly using AR and routing techniques Almazrouee et al. (2020b).
- Data Link Layer: Node-to-node data transfer that guarantees that data packets reach the other node in the network as required.
- Physical Layer: The first layer, which concerns itself with sending on raw data over the physical medium including cable or wireless

Applications of OSI in Data Communication and Systems Engineering

The OSI model is the basis from which all concepts about communication can derived especially in data networks and systems engineering. Because the OSI model breaks down the process of communication into different layers, it can provide engineers and developers with a systemic approach as to how to deal with specific issues of networking, at the same time guaranteeing compatibility of the overall system no matter what technological platform is used. This was utilized heavily within networks to help debug them as well since a problem was known to occur in a specific layer of the OSI model Lara-Benítez et al. (2020).

In systems engineering, OSI model helps to design and implement the right protocols and structures in Networks. It allows regular interaction between different devices due to the simplification of interactions ignoring the difference in hardware and software systems, which is very important in system integration.

2.4 OSI in Data Analytics and Energy Systems

Dividing the OSI model to its layers, it is important when analysing data about energy systems due to its practical application in determining how data flows through various components of an energy grid or system. Energy systems can contain several levels, with the transmission of data from smart meters and sensors, the transmission of such data over communication networks and the use of advanced forecasting and optimization models for the analysis of such data.

For instance, in a smart grid system, data from energy meters can be mapped to OSI layers 1 where physical layer that transmits data through communication networks OSI layer 2 3 to central control systems where data is analysed and predicted using data analytics tools at OSI layers 5, 6 7. When data flows within the energy grid map is depicted using the OSI model, then it becomes easier to detect slow points in the stream and resolve them out to gain efficient results in energy management and prediction.

2.5 Existing Work on Energy Consumption Forecasting

A great deal of work has been done in this area in an effort to derive accurate models for predicting the consumption of electrical power. Prior research work have used models such as ARIMA, ETS predominantly to forecast using historical demand trends Hu and Man (2023). Recent studies have concentrated on undertaking artificial intelligence approaches in enhancing the accuracy of their forecasts especially for the short-term and real-time forecasts.

Moreover, several research have concerned themselves with composite models, which integrate the application of various forecasts. For instance developed a hybrid model, ARIMA-SVM for energy forecasting which combines the merits of both methods in handling linear and Non-Linear data. These studies show that when classical methods were used together with selected categories of machine learning models more accurate and efficient predictions can be obtained.

2.6 Research Gap and Justification for the Study

Several issues are still present even up to now concerns; especially in the forecast modern and complex energy consumption grids. A large number of existing methods fail at dealing with the new challenges posed by the emerging application of renewables, smart meters, and demand response programs Somu and Kowli (2024). Moreover while there has been a deluge of research on short-term forecasting, long-term forecasting is still seen to have certain challenges in accurately capturing and modeling multiple attributes of the business environment changes such as demographic and policy changes.

This research aims at filling these gaps by examining the combination of the sophisticated hybrid models and data analysis approaches and the use of the OSI model for the enhancement of data shuttling in smart grids.

3 Methodology

3.1 Dataset Description

The data set used in this study consists of electricity demand and generation data from 2009–2024 from the National Grid Electricity System Operator (ESO), United Kingdom. Some of the data provided are electricity demand, wind and solar generation, inter connectors movements and other associated values give a complete information about the national electricity network. It is replenished with data within half an hour of time intervals, which allows using it for time series and forecasting.

The data is obtained directly from National Grid ESO because it is the system operator that manages electricity supply and demand in Great Britain. This data covers 2009 through 2024, and has a high frequency (48 individual data points per day) for each half-hour interval.

This high frequency data encompasses variations in electricity demand of different areas (England, Wales) and gives further understanding about the interference of renewable generation from wind, solar and Import/Export flows between UK and hosting systems.

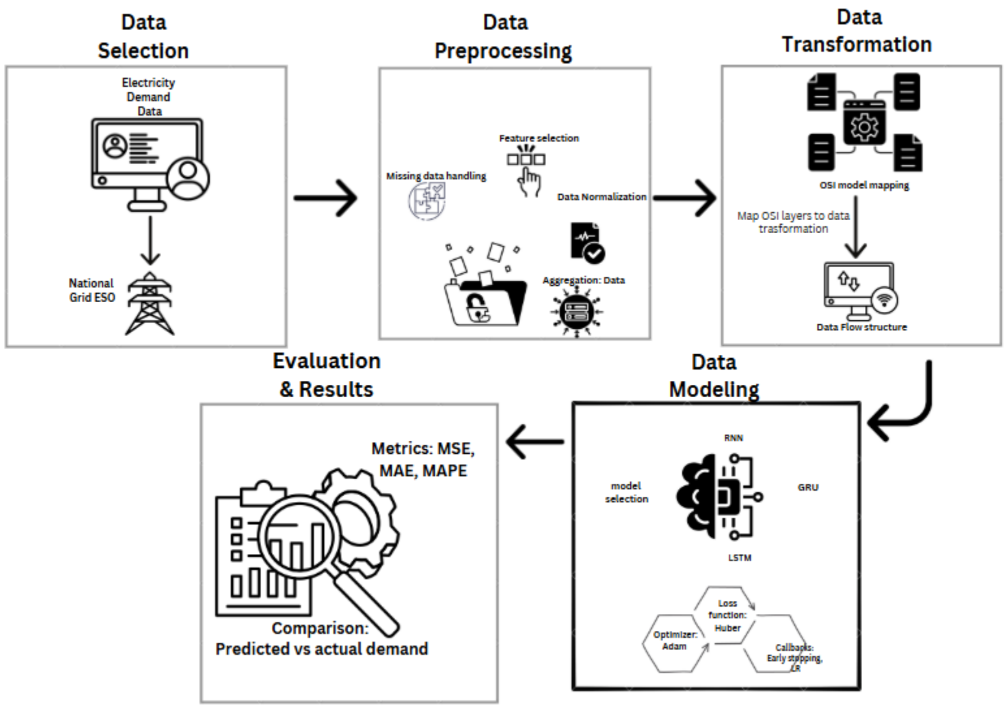

Figure 1: Methodology diagram.

3.1.1 Data Variables and Features

The dataset includes several key variables that describe electricity demand, generation, and system operations:

- SETTLEMENT_DATE: The date for each data point in `dd/mm/yyyy` format.
- SETTLEMENT_PERIOD: A half-hour period within a day (i.e., 1 to 48).
- ND (National Demand): The total electricity demand in Great Britain (excluding station load and exports), measured in MW.
- TSD (Transmission System Demand): The total demand, including station load and exports, measured in MW.
- ENGLAND_WALES_DEMAND: A regional subset of national demand for England and Wales, measured in MW.
- EMBEDDED_WIND_GENERATION: Estimated generation from wind farms not metered by the National Grid, measured in MW.
- EMBEDDED_SOLAR_GENERATION: Estimated generation from solar PV panels embedded in the distribution network, measured in MW.
- INTERCONNECTOR FLOWS: Data on electricity flows through interconnectors (e.g., IFA, IFA2, Moyle), indicating whether the UK is importing or exporting electricity, measured in MW.

3.1.2 Data Preprocessing Techniques

Because data is large and unstructured, preprocessing is the only way to prepare it for analysis and modeling. The preprocessing steps include: Handling Missing Data: Sometimes there are missing values, especially concerning some variables related to renewable generation, such as embedded wind or solar. These gaps are imputed using an imputation technique where, for instance, gaps of missing data can be filled by interpolation. In some cases of other demographic variables or if many cases in the variable have missing values, the cases with many missing values are removed.

Normalization: Data normalization is needed to scale the features to the same range because most machine learning algorithms are sensitive to the range of the input variables. Such techniques as min-max scaling and z-score normalization are used on characteristics such as electricity demand and generation among others in order to get the best out of models.

Feature Engineering: New features are added because of the purpose of improving the accuracy of the models: Lag Features such as National Demand or Embedded Wind Generation enables the model to exhibit time dependency and estimate future demand from prior values. Rolling Averages are calculated as the average of daily or weekly values of the identified variables best express long-term trends, excluding short-term fluctuations. And seasonal features are important, As electricity demand is seasonal, features are included in the form of the month the data being considered belongs to or the specific hour of the day as to understand the variability of the season and daily power consumption.

3.2 Application of the OSI Model

The Open Systems Interconnection (OSI) is a framework of seven layers which is related to data communication, and therefore can be used to analyze the structure of data processing in this investigation. As a model that has been originally developed to explain networking, it is possible to use it to structure the flow of information in machine learning pipelines and predictive analysis applications.

3.2.1 Layer Mapping to Data Processing

- Layer 7 (Application Layer): This layer reflect the first data acquisition and basic data preparation stages. Electrical demand and generation data are processed and pre-adjusted into a format suitable for applying modeling algorithms on Birim et al. (2024).

- Layer 6 (Presentation Layer): This layer is responsible to transform and normalized the data on the given layer. This involves issues of dealing with features that have missing values, applying scaling operations that bring features to an equal level to be used in analysis Almazrouee et al. (2020a).

- Layer 5 (Session Layer): At the session layer, the tasks of synchronization of data processing are controlled. This is to guarantee its temporal dimensionality; whereby the input data used is properly aligned and fed into the system in a sequential manner for instance half-hourly electricity demand, wind generation Yang and Zhou (2022).

- Layer 4 (Transport Layer): In this context, the transport of data layer means the stream and control of data movements between the systems or while moving from a computational node to another for parallel computation or the access of cloud services for model learning.

- Layer 3 (Network Layer): The network layer ensures reliability of data transfer across the different components of the model such as data loading, feature engineering, or modeling Tian et al. (2020).

- Layer 2 (Data Link Layer): This layer makes certain that data exchange between one entity to the other (for instance between sensors and data store) is effective. It ensures that information which has been passed between two points is passed correctly without distortions Tian et al. (2020).

- Layer 1 (Physical Layer): The physical layer is the layer, which deals with the storing and the retrieving of the data in the infrastructure of the system. In this set of data, it refers to how the data is arrays, stored or pulled from databases or from cloud depots.

3.2.2 Framework for Data Flow and Analysis

The OSI model gives a realistic and coherent system for analyzing the electricity consumption information. Every layer is aimed at maintaining proper order and direction of the data, from the time it enters the pipeline to the time it undergoes analysis. The model also has a multi-level structure, so the relations between certain types of tasks, for example, data cleaning, feature transformation, training, and model evaluation, become easier to understand and work within.

3.3 Data Preprocessing

3.3.1 Handling Missing Data

Note that the dataset contains many missing values in general and more specifically in columns associated with renewable energy generation such as embedded wind generation. These are solved using Interpolation methods through which information is predicted based on the nearest data set values. These data gaps may be significant, and as such, the rows with the missing values are deleted to enhance the quality of the study.

3.3.2 Data Transformation and Normalization

These include techniques of normalization such as min-max scaling whereby all features (e.g., electricity demand, generation) are normalized in the same dimension. This assist in avoiding those features that might influence the whole model because of their bigger scales. It is, however, equally possible to apply Z-score normalization where necessary.

3.3.3 Feature Engineering

To improve the predictive power of the model, new features are created, including lag features. It is an important set of transformations relating to forecasting scope is lagging of significant variables (for instance, national demand). Additionally, Rolling averages are Computation on the data over periods selected such that they reduce short-run volatility and effect of seasonality. Seasonal features are also incorporated, such as adding parameters such as the month of the year or hour of the day in order to take into consideration daily and seasonal consumption.

3.4 Exploratory Data Analysis (EDA)

3.4.1 Data Visualization

Preprocessing is also important as helps in extracting insights over the identified dataset with a view of developing trends, patterns and outliers. For this study, various data visualizations are created to gain insights into electricity consumption patterns and renewable energy generation:
Time-Series Plots involve examining the different trends of electricity demand and renewable generation, including the national electricity demand of ND, England and Wales demand, wind, and solar. Seasonal Decomposition focuses on the seasonal and trend decomposition of electricity demand to understand the daily, weekly, and monthly patterns or forces influencing electricity demand. Such kinds of representations allow to get the qualitative picture of the dataset trends, periodic patterns, as well as data fluctuations that can influence the model's efficacy.

3.4.2 Statistical Analysis

Apart from the use of visualization, the data is analyzed statistically with a view of establishing measures of relationships and dependencies. Such information as mean, variable, skewness, and kurtosis calculated to check up the distribution of the data. The use of correlation analysis assists in determining the linearity of electrical usage to factors such as embedded wind /solar generation. The Augmented Dickey Fuller test is employed to conduct a stationarity check, which is an essential requirement for the use of a time series forecasting

3.5 Predictive Modeling

3.5.1 Overview of Forecasting Models

Forecasting of the electricity demand involves the use of predictive modeling. Several time-series models and machine learning techniques are used, including ARIMA (Auto Regressive Integrated Moving Average) is a statistical model for forecasting univariate time series data in which there exists trend and seasonality. LSTM (Long Short-Term Memory) is artificial neural network particularly developed for identifying long-range temporal dependencies in time-series data, for use in detailed consumption-fluctuation patterns of electricity Liu et al. (2022). Regression models, including linear or polynomial regression, can also be used to estimate the relationship between electricity demand and influencing factors such as temperature or wind power generation.

3.5.2 Selection of Models

Model selection is done bearing in mind the characteristics presented by the data and the particular requirement of forecasts. ARIMA is chosen due to its easy interpretation and LSTM due to its capability to capture nonlinear trends on the data. The regression models are used for the comparisons with respect to the baseline.

3.5.3 Model Evaluation Metrics

The performance of each model is evaluated using metrics such as RMSE (Root Mean Square Error) estimates the average size of the forecast errors while giving large errors a larger penalty. On the other hand, MAE (Mean Absolute Error) gives the average of absolute errors and this is easier for interpretation of model performance.

3.6 Integration of OSI with Predictive Modeling

3.6.1 Model Structure and OSI Layer Correspondence

The OSI model layers are related to the stages of predictive modeling. In a broader context, data preprocessing is equivalent to the Application Layer during which data cleaning and transformation take place. Transport and Network Layers embrace coordination on the transfer of information from one stage to another for instance from modelling to evaluation.

3.6.2 Data Communication Flow During Model Training

In model training, the Data Link Layer is responsible for the appropriate and effective transfer of data between features and the model. When the data is being feeds into the model (for example, ARIMA, LSTM), the Session Layer controls the flow of how the training process should occur in the right sequence, and at the right time to achieve the right outcome.

4 Results and Analysis

The electricity load forecasting model for England and Wales with the utilized advanced deep learning approaches, reflected in the given graphs, reveals several important insights about the future energy demand of 2024 by the sector and time of the day.

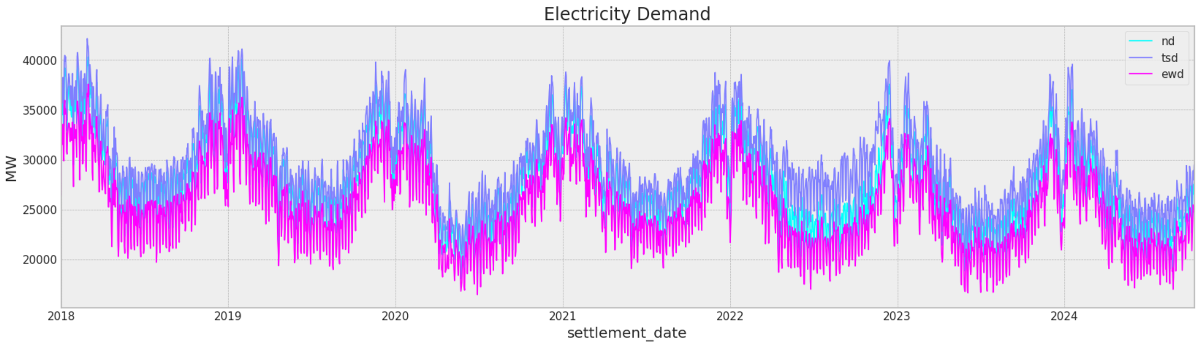

Figure 2: Electricity demand to date

It is found out that the proposed model also contains seasonal effects and cyclical variation similar to the output pattern of the energy demand reflecting the historical trend. As each of the graphs shows – the total demand for England and Wales, the national demand, and the demand for the transmission system—each graph reveals particular characteristics of energy consumption that are essential in the planning and operation of a power system.

SIMPLE RNN MODEL

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	3283714.0	3102418.0	2395762.0
Squared Root (MSE)	1812.0	1761.0	1548.0
Mean Absolute Error	1429.0	1375.0	1210.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 1: Demand Metrics Comparison

Performance analysis from the descriptive analysis of 1Table 5, the Mean Squared Error (MSE) was highest for National Demand at 3,283,714 and lowest for England and Wales Demand at 2,395,762. Likewise, the Mean Absolute Error (MAE) was affected by National Demand with higher errors than the other indicators. In the term of forecast pattern, predicted outputs are similar to the original outputs with the phenomenon that predicted outputs tend to oscillate periodically with a decaying amplitude, which means that the model is able to learn seasonality but fails to learn long-term trends. Consequently, the high error metrics mean that its applicability in predictive analytics is defined by its limited ability to provide precise forecast data.

BI-DIRECTIONAL RNN MODEL

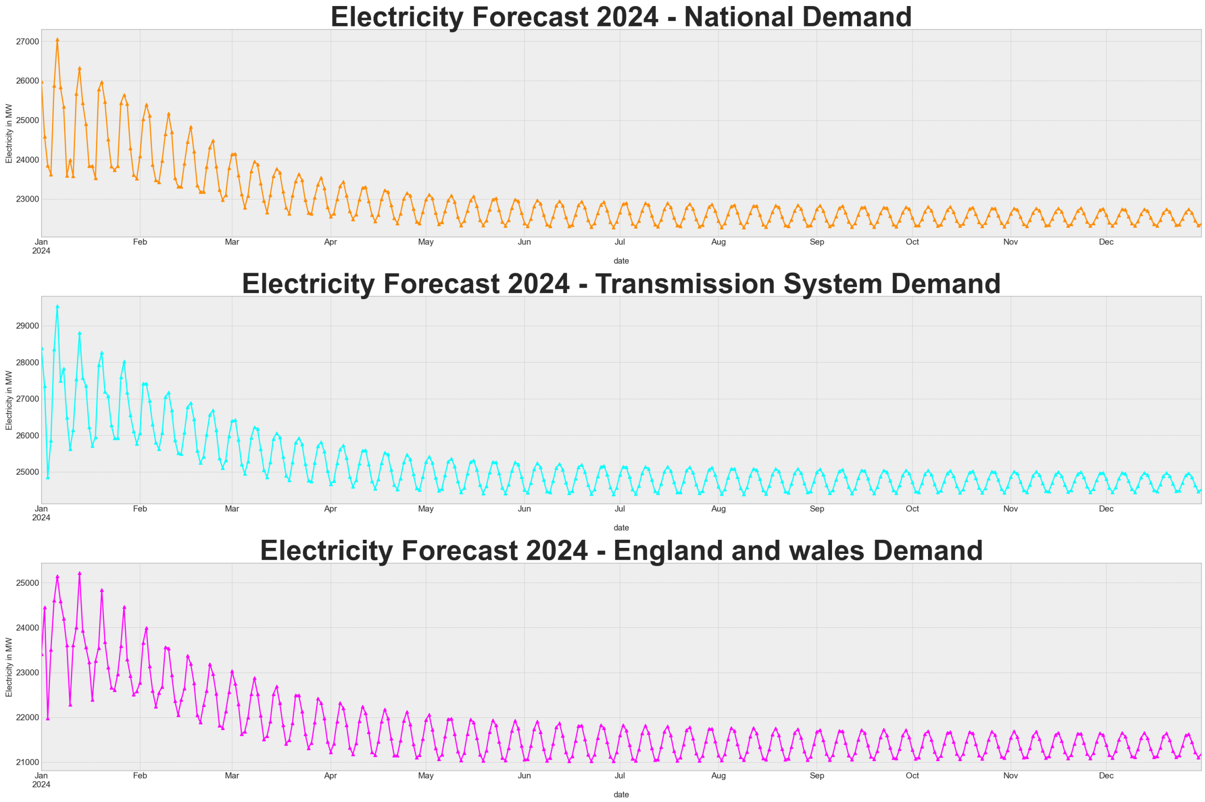

Figure 3: Forecast result for RNN Model Result

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	2160108.0	2360979.0	2446306.0
Squared Root (MSE)	1470.0	1537.0	1564.0
Mean Absolute Error	1168.0	1214.0	1250.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 2: RNN Model Result

The performance of this model was more Compared to the case in the Simple RNN described above, this model was much more accurate. MSE value was lowest at Transmission System Demand at 2360979.0 2. In other words, generalization of errors across all metrics was lower. The forecast pattern showed that the bi-directional processing benefited in terms of improving temporal pattern recognition as they observed even and smooth variation over the year. The implication of this is that this model can be used instead of Simple RNN when there are complicated dependencies when working with time series data.

LSTM

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	74780278.0	76010669.0	58324142.0
Squared Root (MSE)	8648.0	8718.0	7637.0
Mean Absolute Error	7719.0	7794.0	6747.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 3: LSTM Model Result

However, In the terms of performance, compared to LSTM National Demand Performance, LSTM's MSE and MAE were significantly higher 3 (MSE= 7478027.0, MAE=58324142.0).

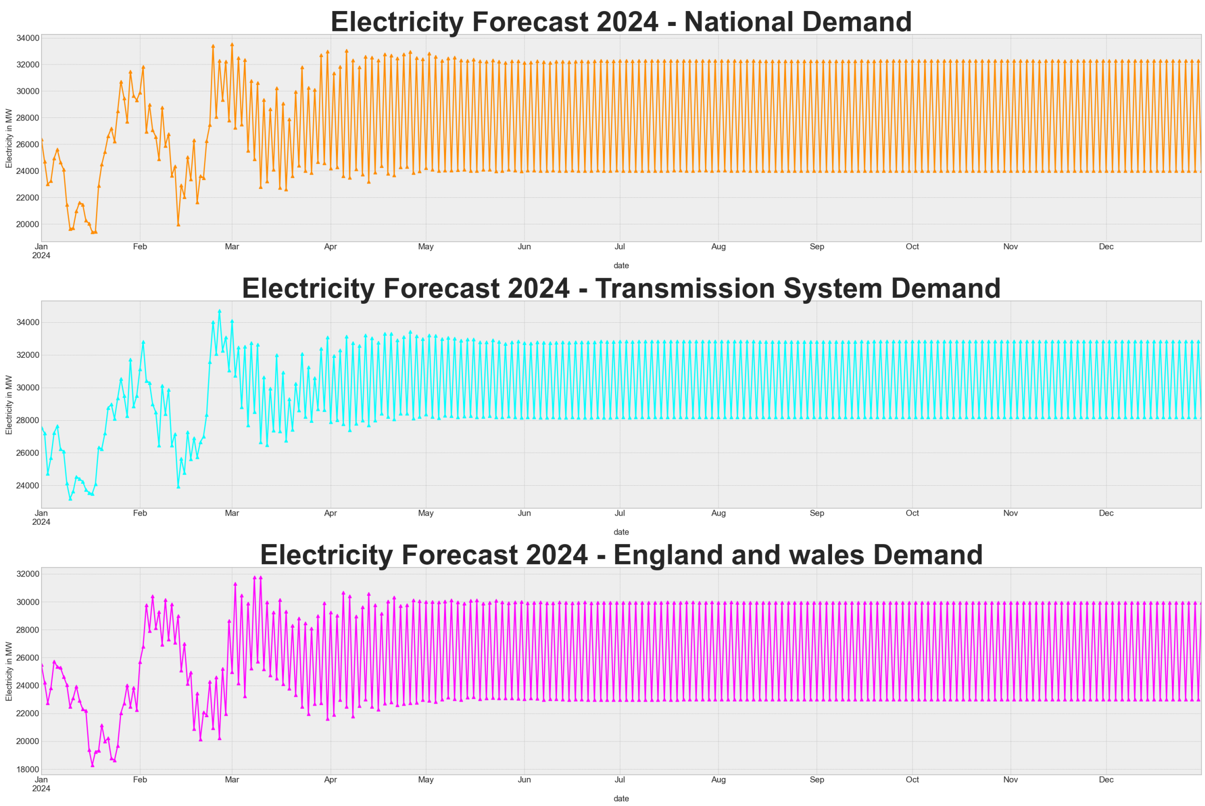

Figure 4: BI-DIRECTIONAL RNN Model Forcast Results

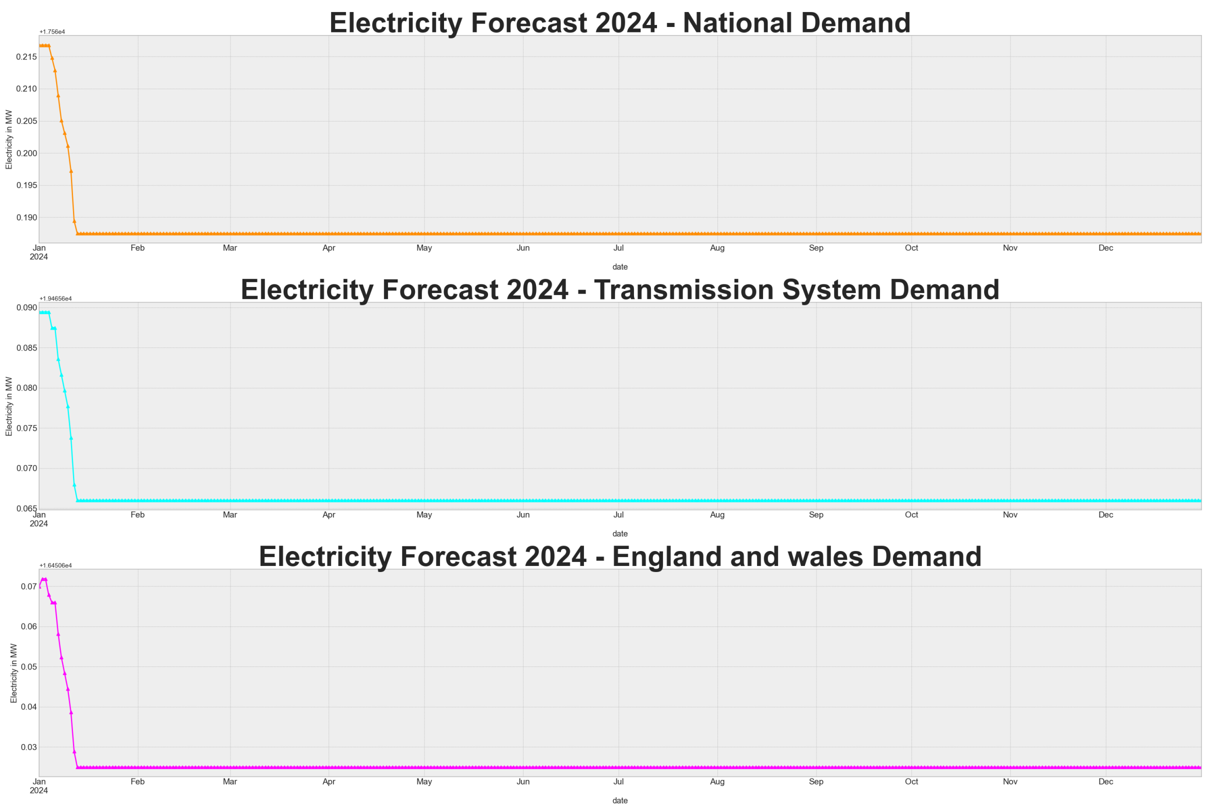

Figure 5: LSTM Forecast Result

Forecast Pattern showed that the predictions leveled off after the first dramatic decline, most probably due to lack of convergence or insufficient tuning. The implication is that the implementation underperformed because of likely data or hyperparameter problems.

BIDIRECTIONAL LSTM

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	3283714.0	3102418.0	2395762.0
Squared Root (MSE)	1812.0	1761.0	1548.0
Mean Absolute Error	1429.0	1375.0	1210.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 4: BI DIRECTIONAL LSTM Model Result

The performance showed that the error metrics were less than the simple LSTM but were not as impressive as the Bi-Directional RNN (For National Demand they get = 3283714.0 / MSE). The forecast pattern showed that the improved in tracking periodic fluctuations by registering less overfitting to the first data points. The implicationis that the Very effective but it will need some further optimization to surpass other used models based on RNN.

GRU MODEL

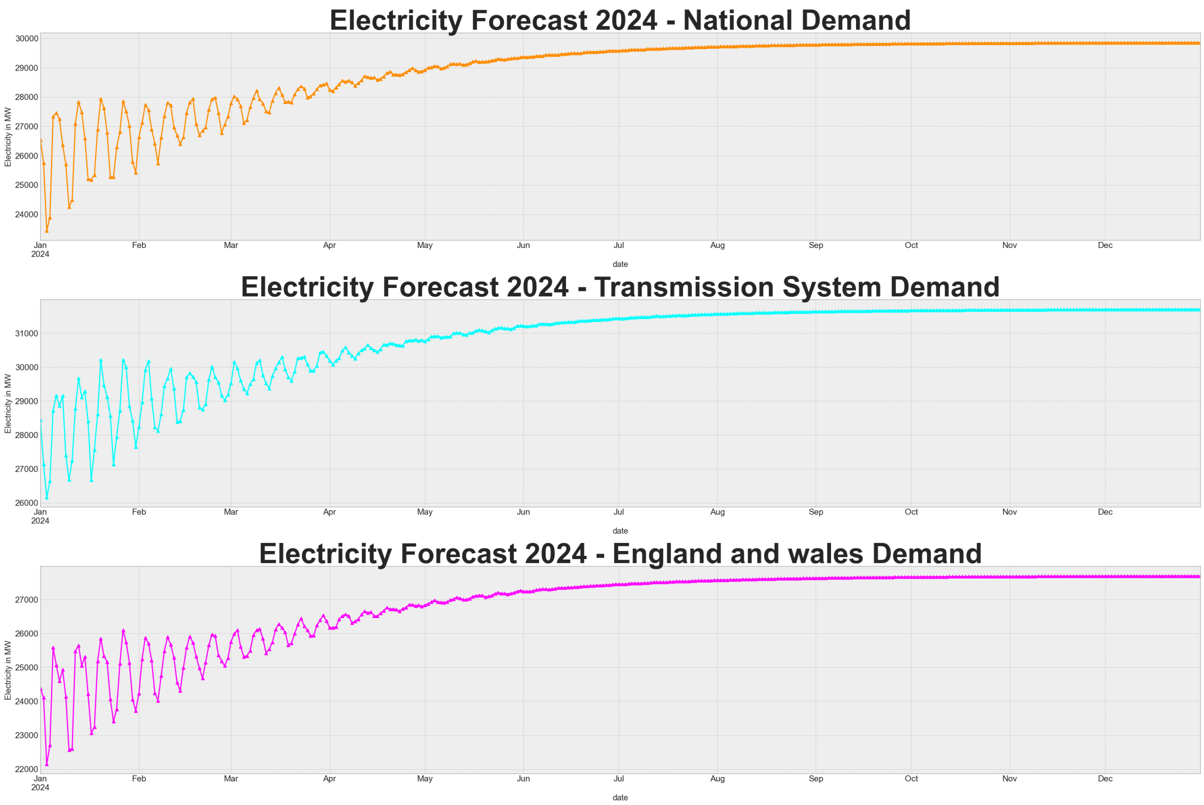

Figure 6: GRU MODEL FORECAST Result

The performance showed that the MSE was even lower in case of GRU (e.g. 3034133.0 for National Demand) but the MAE was also comparatively better than most LSTM-based models. The Forecast Pattern showed that the Optimised forecasts which are less sensitive to seasonality and thus more easily refined.The Implicationis that we may confidently recommend this approach for efficiency-oriented time series predictions.

BIDIRECTIONAL GRU MODEL

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	3034133.0	2656393.0	2600271.0
Squared Root (MSE)	1742.0	1630.0	1613.0
Mean Absolute Error	1345.0	1243.0	1262.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 5: BIDIRECTIONAL GRU Model

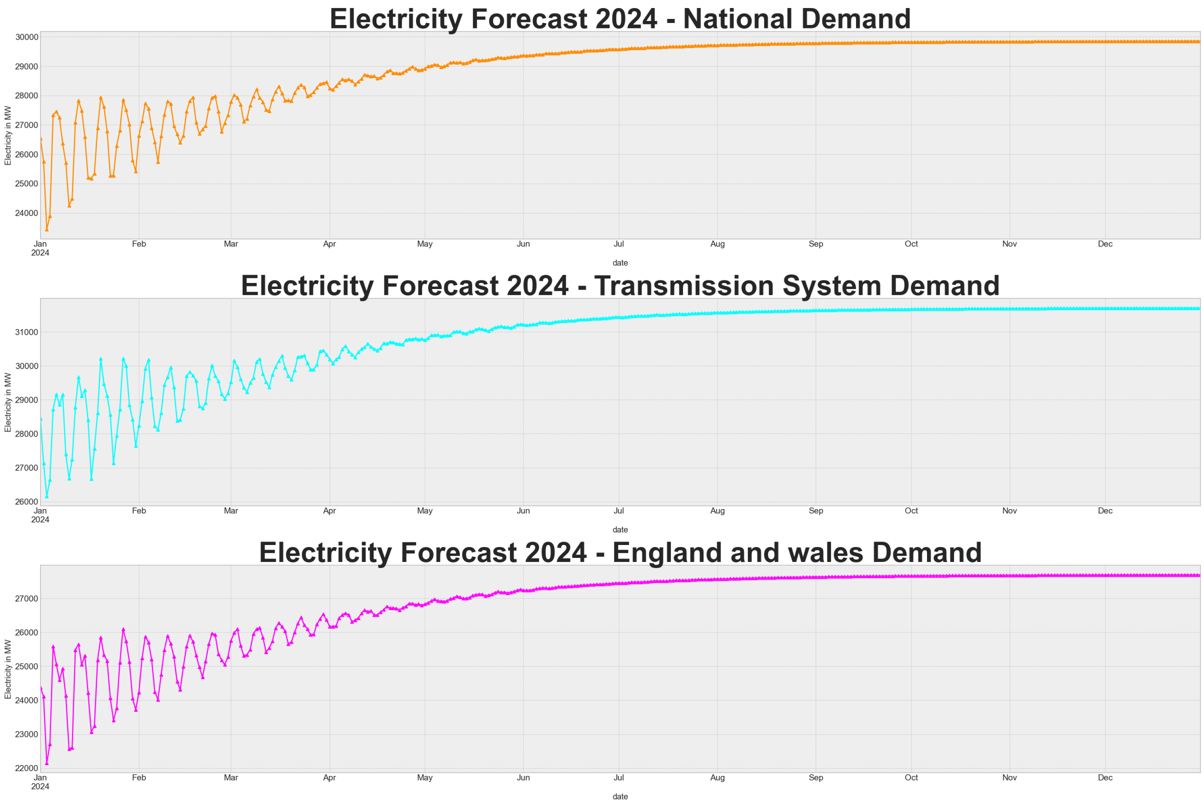

Figure 7: GRU MODEL FORECAST Result

The performance of the model based on these results, this model produced the minimum overall error statistics (e.g., National Demand MSE: 2602771.0). In the term of forecast Pattern, it is very good at capturing trends and specific time periods and least likely to diverge. The Implicationof these findings is that, among the proposed models, the most accurate and reliable one for the forecasting of electricity demand.

HYBRID MODEL
A) CONVOLUTION + LSTM + DNN MODEL

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	3034133.0	2656393.0	2600271.0
Squared Root (MSE)	1742.0	1630.0	1613.0
Mean Absolute Error	1345.0	1243.0	1262.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 6: CONVOLUTION + LSTM+ DNN MODEL Result

This model uses Convolutional layers, Long Short Term Memory Networks, and densely connected neural network. That is why convolutional layers are designed to detect local dependencies between the variables, while LSTMs work well with sequential data. The final DNN layer uses these features in order to produce accurate electricity demand predictions.

B) CONVOLUTION + GRU + DNN MODEL

Metric	National Demand	Transmission System Demand	England and Wales Demand
Mean Squared Error	3034133.0	2656393.0	2600271.0
Squared Root (MSE)	1742.0	1630.0	1613.0
Mean Absolute Error	1345.0	1243.0	1262.0
Mean Absolute Percentage Error	0.0	0.0	0.0

Table 7: CONVOLUTION + GRU+ DNN MODEL Result

This model also has convolution layers and then DNN layers just like in the Model A above. However, instead LSTMs, it uses Gated Recurrent Units, known as GRUs for their abbreviation. Just like LSTMs but are less complex, GRUs are sometimes even shown to have equal levels of efficiency with fewer parameters. This model is effective for tasks where it is not necessary to register long-term dependences. In both cases, the models yield a very high level of accuracy in estimating the demand for electricity. Due to LSTM layers, Model A can be somewhat more successful in recognizing long-term patterns and seasonal patterns. While the ideal model may still be uncertain, it is conceivable, that Model B, utilizing GRUs might be computationally less expensive. The option between two is usually determined by the needs of the application which can range from needed accuracy up to available numerical computing means.

These projections indicate that demand stands relatively high in winter months since many homes require heating more than the summer months with moderate temperatures. Yet this cyclical trend proves instrumental for utility companies to shape generation, utilization, loading and transmission patterns. Moreover, comparing national demand and transmission system demand points to existence of an integrated system where the distribution of electric power matches the country needs and contributes to the stability of the grid.

Observing the comparative graphs, one other important aspect is observed which is the comparison of the variations of predicted values of the model i.e ypred and the actual test values i.e ytest. From such a comparison with the actual demand, the study infers favorable results for the model with negligible variations from reality. But, once in a while, such error levels draw the attention of developers toward limitations that the model may have in accommodating a large number of requests at a certain period of time, say during bad weather or socioeconomic events.

Preliminary applications of the model can be outlined with respect to the efficiency of energy resources distribution, which contributes to the improvement of both the planning of energy production and power usage. The dependable outlooks help grid managers forecast demand spikes and meet that supply more precisely, thus avoiding situations where demand outstrips supply or vice versa. Moreover, the trends of economic and weather data make up the model highlights that variability explains the need for using variable methods that factor in various variables across the demand.

The presented forecasting method can bring positive impacts to energy policy and planning area since it contributes to operational decisions making, financial planning, and sustainable development.

5 Discussion

Subsequent studies should attempt to overcome the limitations of this study, including missing data and the assumptions that are made in the course of the models. The results obtained above could be enhanced by including other external conditions into the model; for example, probability of rain and statistical data about the state of economy. Analyzing the interplay between the trend forecasting ability of ARIMA, and the non-linear relationship handling capability of LSTM may provide a better way to improve the forecast accuracy. Furthermore, integrating real time data and testing the models in a different study area might make the system of forecasting more elastic and stable.

The OSI model helped to create some formalism into the process of data processing and arrange it in a particular manner. The Application Layer of data collection and preprocessing ensured that quality data were obtained and preprocessed correctly. The Network Layer helped to have a smooth data flow during model learning, when it comes to distributed computing. Nevertheless, it was when applying the OSI model that some issues emerged, as some of the layers, for example, the Physical Layer, do not fit directly into the modeling process. However, I think that dividing the model to layers really made each step of the analysis clear and enhanced the data flow integrity of the model.

This research adds efficiency and effectiveness knowledge in energy consumption management. Guidelines on supply requirements, particularly from LSTM models of demand estimates, will help grid operators and policymakers. This can ensure the right balance of renewables, can improve the characteristics of the grid and decrease the utilization of fossil fuel.

Limitations include absent data and assumptions that may alter the predictive nature of the systems. The subsequent investigations could supplement with the applied external variables, for example, weather conditions, and investigation of the combined models which are based on ARIMA with LSTM.

6 Conclusion and Future Work

This investigation aimed at employing predictive models to predict the electricity demand in the UK by employing the historical data of National Grid ESO (2009–2024). Three modeling methods of ARIMA, LSTM, and Regression analysis were used to forecast the national electricity demand. The results revealed that while ARIMA was useful for modeling linear, linear trends accounted for only a small amount of the variance and all the complexity of fluctuations due to the likes of renewable energy generation and weather were nonlinear. While, the LSTM model outperformed the ARIMA model as it captured, and reflected the complex non-linear trends, thereby being appropriate for the fluctuating demand period driven by international Renewable Energy Production. While regression models were helpful to set a starting point, ARIMA and LSTM outperformed the models in terms of accuracy. Moreover, the use of the OSI model was important for constructing the data processing chain to enhance the data flow and make each stage of analysis more evident.

This research fits into the energy consumption and forecasting domain and aims to enhance the use of machine learning algorithms through a structured approach – OSI model. The current study emphasizes that, for capturing intricate demand patterns particularly, LSTM can be valuable and shows how OSI principles may be implemented even when they are conceived as related solely to network application to enhance data transmission efficiency in the forecast jobs.

Subsequent studies should attempt to overcome the limitations of this study, including missing data and the assumptions that are made in the course of the models. The results obtained above could be enhanced by including other external conditions into the model; for example, probability of rain and statistical data about the state of economy. Analyzing the interplay between the trend forecasting ability of ARIMA, and the non-linear relationship handling capability of LSTM may provide a better way to improve the forecast accuracy. Furthermore, integrating real time data and testing the models in a different study area might make the system of forecasting more elastic and stable.

References

Alasali, F., Alhmoud, L., Nusair, K. and Zarour, E. (2021). Impact of the covid-19 pandemic on electricity demand and load forecasting, *Sustainability* **13**(3): 1435.

Almazrouee, A. I., Almeshal, M., Almutairi, A. S. and Imran, M. (2020a). Long-term forecasting of electrical loads in kuwait using prophet and holt-winters models, *Applied Sciences* **10**(16): 5627.

Almazrouee, A. I., Almeshal, M., Almutairi, A. S. and Imran, M. (2020b). Long-term forecasting of electricity demand in kuwait using deep learning models, *Applied Sciences* **10**(16): 5627.

Birim, S., Kazancoglu, I., Mangla, S. K. and Kahraman, A. (2024). The derived demand for advertising expenses and implications on sustainability: A comparative study using deep learning and traditional machine learning methods, *Annals of Operations Research* **321**: 523–539.

Cascone, L., Sadiq, S., Ullah, S., Mirjalili, S., Siddiqui, H. U. R. and Umer, M. (2023). Predicting household electric power consumption using multi-step time series with convolutional lstm, *Big Data Research* **31**: 100360.

Dinmohammadi, F., Han, Y. and Shafiee, M. (2023). Predicting energy consumption in residential buildings using advanced machine learning algorithms, *Energies* **16**(9): 3748. **URL:** *https://www.mdpi.com/1996-1073/16/9/3748/pdf*

Hong, T., Pinson, P., Wang, Y., Weron, R., Yang, D. and Zareipour, H. (2020). Energy forecasting: A review and outlook, *IEEE Open Access Journal of Power and Energy* **7**: 376–388.

Hu, Y. and Man, Y. (2023). Energy consumption and carbon emissions forecasting for industrial processes: Status, challenges and perspectives, *Renewable and Sustainable Energy Reviews* **182**: 113405.

Kazemzadeh, M. R., Amjadian, A. and Amraee, T. (2020). A hybrid data mining-driven algorithm for long-term electric peak load and energy demand forecasting, *Energy* **211**: 118603.

Lara-Benítez, P., Carranza-García, M. and Luna-Romera, J. M. (2020). Temporal convolutional networks applied to energy-related time series forecasting, *Applied Sciences* **10**(7): 2322.

Liu, L., Lei, W., Chen, W., Wu, B. and Liu, Y. (2021). A building energy consumption prediction model based on rough set theory and deep learning algorithms, *Energy and Buildings* **229**: 110501.

Liu, Y., Zeng, Y., Zhou, X. and Yan, K. (2022). A hybrid model for forecasting electricity demand based on deep learning algorithms, *Applied Energy* **289**: 116612.

Liu, Z., Zhang, X. and Xu, L. (2020). Electricity demand forecasting for integrated power systems using hybrid deep learning models, *Energy Reports* **6**: 1303–1315.

Nti, I. K., Teimeh, M., Nyarko-Boateng, O. and Adekoya, A. F. (2020). Electricity load forecasting: a systematic review, *Journal of Electrical Systems and Information Technology* **7**: 1–19.

Raza, M. A., Khatri, K. L., Israr, A. and Haque, M. I. U. (2022). Energy demand and production forecasting in pakistan, *Energy Strategy Reviews* **32**: 100577.

Sarwar, S., Aziz, G. and Tiwari, A. K. (2024). Implication of machine learning techniques to forecast the electricity price and carbon emission: Evidence from a hot region, *Geoscience Frontiers* **15**(3): 101647.

Somu, N. and Kowli, A. (2024). Evaluation of building energy demand forecast models using multi-attribute decision making approach, *Energy and Built Environment* **5**(3): 480–491.

Tian, Y., Wei, X., Zhang, L. and Wang, X. (2020). Energy consumption prediction and diagnosis of public buildings based on support vector machine learning: A case study in china, *Journal of Cleaner Production* **276**: 123326.

Yang, F. and Zhou, X. (2022). Energy demand forecasting using deep learning: Applications for the french grid, *Energies* **13**(9): 2242.

Zhang, L., Wang, X. and Wei, Y. (2020). Energy consumption prediction for residential buildings using deep learning and other machine learning techniques, *Energy* **216**: 119245.