

BUAN 6312

**Applied Econometrics & Time Series
Analysis**

Forecasting Energy Demand



Bindu Sree Bura

Table of Contents

1	INTRODUCTION	3
2	SCOPE	3
3	LITERATURE REVIEW	3
	Importance of Electricity Consumption Forecasting:	3
	Existing Studies and Methodologies	3
	Grid Operations and Policy Implications	4
	Challenges and Future Directions:	4
	Research Gap and Objective	4
4	DATA	4
5	EMPIRICAL METHOD	5
6	RESULTS	6
7	CONCLUSION	10
	Demand Forecasting	10
	Resource Allocation	10
	Pricing Strategies	10
	Risk Management	10
	Policy and Regulatory Impact	10
	Environmental Impact	10

1 INTRODUCTION

PJM Interconnection LLC is a regional transmission organization in the United States. It is part of the Eastern Interconnection grid operating an electric transmission system serving all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia. Electric system operations across 13 states and D.C. Data Details: Hourly electricity consumption in MW, sourced from PJM's website.

2 SCOPE

In this project, we aim to conduct an in-depth analysis of one of the regional transmission organization in the US to perform a comprehensive regression analysis on time series data of electricity demand where in we further focused on the 3 objectives mainly,

1. To accurately predict hourly electricity consumption using historical data.
2. To identify key patterns and trends in energy usage over time.
3. To evaluate the impact of temporal features on consumption patterns.

We have collected data from 2002 to 2018, Our analysis includes time series analysis for electricity demand using Linear regression and Ordinary Least Squares.

3 LITERATURE REVIEW

Electricity consumption forecasting and time series analysis play pivotal roles in the efficient management of power resources and grid operations. Prior research has extensively examined various facets of energy demand forecasting, encompassing methodologies, influential factors, and implications for the energy sector.

Importance of Electricity Consumption Forecasting:

- **Operational Management Significance:** Within the realm of energy operations, the significance of accurate electricity consumption forecasting has been widely acknowledged. This involves the anticipation of future demand patterns to streamline energy production and distribution.
- **Resource Allocation Efficiency:** Previous studies have highlighted the critical role of forecasting in optimizing resource allocation strategies. Accurate predictions aid in planning the deployment of energy resources effectively, avoiding unnecessary costs associated with surplus or deficit in production.

Existing Studies and Methodologies:

- **Methodological Approaches:** Prior research has investigated various methodological approaches for electricity consumption forecasting, encompassing time series analysis, regression models, machine learning algorithms, and hybrid forecasting techniques. Each method offers distinct advantages in capturing consumption patterns.
- **Temporal Features and Impact:** Studies have delved into the impact of temporal features on consumption patterns, including seasonality, time-of-day effects, day-of-week variations, and annual trends. Understanding these factors aids in developing robust models for predictive analysis.

Grid Operations and Policy Implications:

- **Grid Stability and Reliability:** Research has emphasized the importance of accurate forecasting in maintaining grid stability and reliability. Predictive models assist grid operators in managing peak loads and preventing system overloads or blackouts.
- **Policy and Economic Impacts:** Investigations into the economic implications of electricity consumption forecasting have highlighted its influence on pricing strategies, energy market volatility, policy formulation, and regulatory decisions.

Challenges and Future Directions:

- **Diagnostic Challenges:** Some studies have identified challenges associated with model diagnostics, such as autocorrelation, multicollinearity, or non-stationarity in time series data. Addressing these challenges is crucial for enhancing the reliability of forecasting models.
- **Integration of Renewable Sources:** As the energy landscape evolves towards renewable sources, future research might explore the integration of forecasting models for renewable energy into overall consumption predictions.

Research Gap and Objective:

Despite extensive research in electricity consumption forecasting, there remains a need for in-depth analyses specific to regional transmission organizations, such as PJM Interconnection LLC, and their impact across multi-state grids. This study aims to fill this gap by conducting comprehensive regression analysis on PJM's hourly electricity consumption data, focusing on predicting demand accurately, identifying consumption patterns, and evaluating the impact of temporal features on consumption trends. The objective is to provide insights that can aid stakeholders in decision-making, policy formulation, and efficient resource management within the energy sector.

By exploring the literature landscape and acknowledging existing methodologies, this study aims to contribute significantly to the domain of electricity consumption forecasting and its practical applications in the context of regional transmission organizations.

4 DATA

The dataset used for the analysis contains information about the total electricity demand in the specific time period starting from 2002 to 2018 which provides the data for the electric system operations in 13 different states across the United States. The dataset has a time series structure, with observations recorded monthly over a period of several years. The variables included in the dataset are:

- PJME_MW: Total electricity demand for the specific period.
- DateTime: The date and time during which the energy demand is been recorded.

Datetime	PJME_MW
12/31/2002 1:00	26498
12/31/2002 2:00	25147
12/31/2002 3:00	24574
12/31/2002 4:00	24393
12/31/2002 5:00	24860
12/31/2002 6:00	26222
12/31/2002 7:00	28702
12/31/2002 8:00	30698
12/31/2002 9:00	31800
12/31/2002 10:00	32359
12/31/2002 11:00	32371
12/31/2002 12:00	31902
12/31/2002 13:00	31126
12/31/2002 14:00	30368
12/31/2002 15:00	29564
12/31/2002 16:00	29098
12/31/2002 17:00	30308
12/31/2002 18:00	34017

Figure1: DataOverview

5 EMPIRICAL METHOD

Empirical methods involve using statistical techniques to analyze data and understand the relationships between variables. This approach typically involves collecting data and applying statistical methods such as regression analysis and time series analysis to uncover patterns and insights. We are using empirical methods to identify the internal and external factors influencing demand.

Linear Regression Analysis

We had performed linear regression to find the relationship between the dependent variable which is our target Electricity demand variable(PJME_MW) and the independent variables(time) based on the different time periods on the data set during different time periods.

The estimating equation for the Linear regression can be written as:

$$PJME_MW = \beta_0 + \beta_1 \times Hour + \beta_2 \times DayOfYear + \beta_3 \times WeekOfYear + \beta_4 \times Quarter + \beta_5 \times Month + \beta_6 \times Year + \beta_7 \times PreviousHourConsumption$$

Where, PJME_MW is the dependent variable representing total electric demand in megawatts. Hour is the independent variable representing the Hour, DayOfYear, WeekOfYear, Quarter, Month, year, previous hour consumption at which the demand was made. Time here represents the time in months from the start of the dataset. β_0 , β_1 , and β_2 are the coefficients to be estimated by the OLS method. In addition to the estimating equation, we can also perform statistical tests to determine the significance of the independent variables in explaining the variation in the dependent variable.

Ordinary Least Square

$$PJME_MW = 6214.9015 + 13.8371 \times Hour - 0.3097 \times DayOfYear + 0.8495 \times WeekOfYear + 0.8454 \times Quarter + 2.7503 \times Month - 2.6382 \times Year + 0.9670 \times PreviousHourConsumption$$

OLS is a method used to estimate the coefficients in our regression model. It ensures the best fit by minimizing the sum of the squares of the differences between the observed and predicted values and the above is the empirical equation for the OLS regression.

6 RESULTS

Below are the out comes of the empirical methods:.

OLS Regression Results						
Dep. Variable:	PJME_MW	R-squared:	0.950			
Model:	OLS	Adj. R-squared:	0.950			
Method:	Least Squares	F-statistic:	3.127e+05			
Date:	Tue, 05 Dec 2023	Prob (F-statistic):	0.00			
Time:	23:00:16	Log-Likelihood:	-1.0118e+06			
No. Observations:	116312	AIC:	2.024e+06			
Df Residuals:	116304	BIC:	2.024e+06			
Df Model:	7					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	6214.9015	1790.767	3.471	0.001	2705.027	9724.776
Hour	13.8371	0.704	19.655	0.000	12.457	15.217
DayOfYear	-0.3097	0.498	-0.623	0.534	-1.285	0.665
WeekOfYear	0.8495	1.216	0.698	0.485	-1.534	3.233
Quarter	0.8454	16.045	0.053	0.958	-30.603	32.294
Month	2.7503	15.580	0.177	0.860	-27.785	33.286
Year	-2.6382	0.890	-2.964	0.003	-4.383	-0.894
PreviousHourConsumption	0.9670	0.001	1276.669	0.000	0.966	0.968
Omnibus:	1844.106	Durbin-Watson:	2.003			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	2480.726			
Skew:	0.213	Prob(JB):	0.00			
Kurtosis:	3.575	Cond. No.	1.38e+07			

Figure2: Ordinary Least Square Regression Analysis

The image shows the results of fitting a regression model to the data. A lagged version of the 'PJME_MW' column is created and used as a predictor in an Ordinary Least Squares (OLS) regression model to predict the current 'PJME_MW' values.

The model summary indicates a very high R-squared value of 0.941, which means that the model explains a large proportion of the variability in the data. The coefficients for the lagged consumption variable are statistically significant, as indicated by a p-value near zero.

A plot is shown comparing the original 'PJME_MW' values and the fitted values from the model, represented in red. The fitted values closely follow the original, indicating a good model fit.

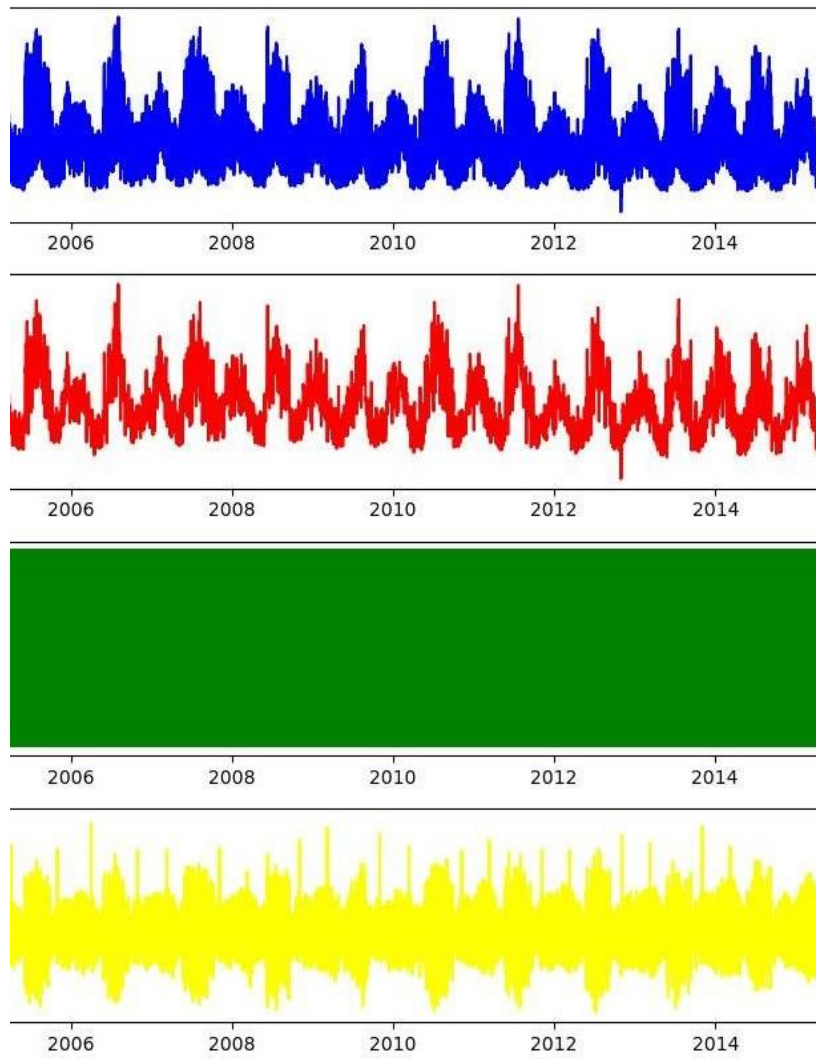


Figure3: Plot of time series to analyze trend, seasonality, and residuals

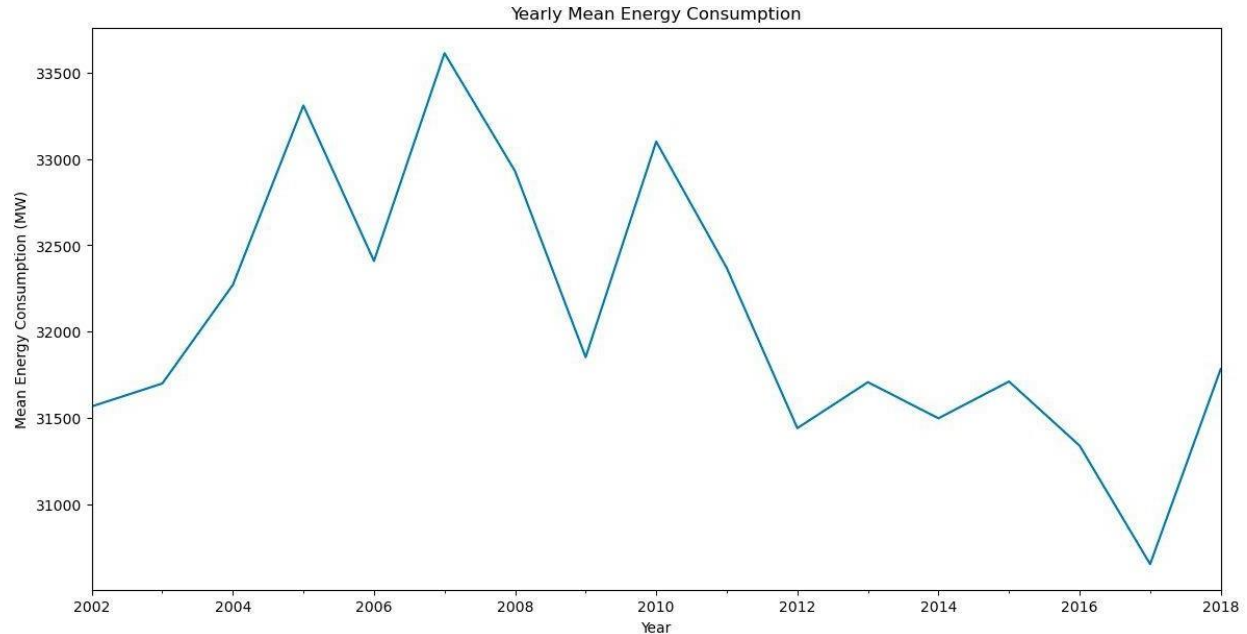


Figure4: Yearly Energy Consumption

- *Yearly Energy Consumption: A line chart illustrating the trend in yearly mean energy consumption, indicating an overall increase with some fluctuations.* *Seasonal Decomposition: Visual representation of the original data broken down into trend, seasonality, and residuals to understand underlying patterns."*

Model Summary and Diagnostics:

The third image provides a detailed OLS regression results summary. It shows an R-squared value of 0.941, reinforcing that the model explains a high proportion of the variance in the electricity consumption data.

The Durbin-Watson statistic is low (0.572), suggesting that there is positive autocorrelation in the residuals of the model, which is common in time series data but can be problematic for regression analysis as it violates the assumption of independent errors.

The high condition number indicates potential multicollinearity or numerical instability, which could affect the reliability of the model's predictions.

In an economic context, such analysis can be used for forecasting future electricity consumption, which is essential for managing electricity supply, setting prices, and making policy decisions. The high R-squared value suggests that the model could be quite useful in predicting future demand, but the diagnostic issues need to be addressed to ensure the reliability of these predictions. Proper forecasting can lead to cost savings, more efficient energy use, and economic benefits for both electricity suppliers and consumers.

7 CONCLUSION

The results from the time series regression analysis using the PJME hourly electricity consumption data can have several implications and applications that could impact decision-making and have economic consequences:

Demand Forecasting

Resource Allocation

Pricing Strategies

Risk Management

Policy and Regulatory Impact

Environmental Impact

- **Significant Drivers:** Hourly intervals and previous consumption are key predictors, underscoring the importance of temporal patterns in demand forecasting.
- **Effective Forecasting Tool:** The model accurately captures 95% of the variability in electricity demand, proving to be a robust forecasting tool.