Energy Grid Forecasting and Anomaly Detection

Cody Hill
University of Colorado Boulder
Boulder, USA
cody.hill-1@colorado.edu

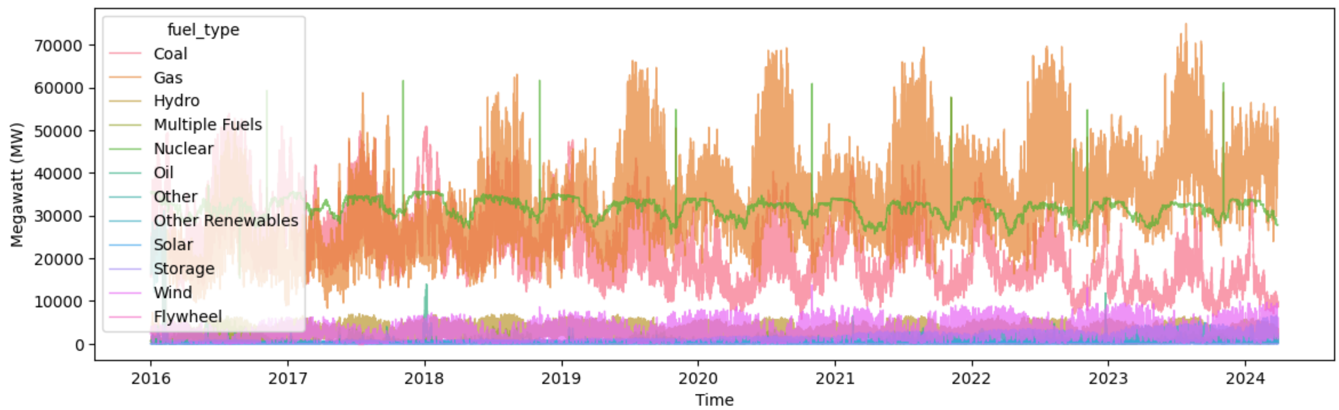

Figure 1: Energy Generated in megawatts (MW) by Fuel Type.

ABSTRACT

Energy forecasting is a vital part of today's energy markets and operations. It is used to set the cost of electricity, determine when power generation should be increased, or decide when power needs to be directed to or from other regions of the electrical grid. Electrical energy demand continues to rise, especially with the recent widespread adoption of electric vehicles. As a result, the electrical grid will need to become more robust and adaptive as this transition away from fossil fuels and into a more diversified energy market occurs. This makes forecasting predictions of electrical loads and generation that much more important and it is necessary to research the differences in forecasting applications and relevance in various short-term, medium-term, and long-term settings.

The time series forecasting techniques in use today each have their own strengths and weaknesses. Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing models tend to be useful in short-term forecasts, but machine learning techniques such as Gradient Boosting Machines or XGBoost and Long Short-Term Memory (LSTM) network models tend to excel at longer-term forecasts. By evaluating each forecasting technique's accuracy across various context windows and forecasting horizons, the results show which techniques are most relevant within the energy grid context and for what application domain.

CCS CONCEPTS

• **Mathematics of computing → Time series analysis**; • **Computing methodologies → Supervised learning by regression**; *Classification and regression trees*; *Support vector machines*; *Neural networks*; • **Applied computing → Forecasting**.

KEYWORDS

time series forecasting, anomaly detection, energy load, renewables, energy grid

ACM Reference Format:
Cody Hill. 2024. Energy Grid Forecasting and Anomaly Detection. In *Proceedings of Data Mining Project (Data Mining Project '24)*. ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX

Data Mining Project '24, April 2024, Boulder, CO
2024. ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Energy forecasting is an important task to not only determine the expected load on the energy grid and the requirements of energy providers to meet that demand by either ramping up production or diverting sources altogether. Identifying hourly, daily, and seasonal energy demands accurately has huge implications on not only the infrastructure of our energy grid but also a complex energy marketplace where prices dynamically shift based on generation supply, customer demand, and financial spot and derivative contracts which are based on energy forecasting.[x] In the past this predictive role was largely upheld with statistical point-estimate models, but with the growth of our predictive models along with the diversification of the energy grid more modern probabilistic interval forecasting techniques have begun to be used.[x] This diversification of the energy grid not only comes from a dismantling of the monopolistic qualities of the energy market in the 1990's,[x] but emerging cultural desires and improved technology in renewable energies are driving a need for a better electrical grid.[x]

Electricity generation has always been mostly an on-demand industry and largely continues to be, where the energy we generate

Table 1: Defining prediction forecast horizons. The forecast horizon determines how far into the future a forecast model predicts.

No.	Forecast Horizon Intervals	Duration
1	Very Short-Term	$X < 1$ Hour
2	Short-Term	1 Hour $\leq X < 1$ Week
3	Medium-Term	1 Week $< X \leq 1$ Year
4	Long-Term	$X > 1$ Year

needs to be transmitted in used within a short period of time because we lack efficient systems for energy storage at scale.[x] This fact presents additional complications when the majority of renewable energies only work in certain conditions, making its effective use more reliant on forecasting than traditional energy generation technologies. However, this means with accurate forecasting, we can reduce reliance on fossil-fuel energies by more efficiently filling in gaps of energy demands with renewables.

2 RELATED WORK

PLACEHOLDER

3 PROPOSED WORK

Models stemming from different statistical and machine-learning families will be trained to perform electricity load forecasting on the forecast horizons listed in Table 1. Each model's amount of compute required for training and evaluation scores will be compared on different forecast horizons. See evaluation section for more details on the specific metrics used.

3.1 Data

Data was collected by the Pennsylvania-New Jersey-Maryland Interconnection (PJM). PJM is a regional electrical transmission organization which coordinates electricity transmission in 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.[x]

Two sources of data from PJM are being utilized here, hourly load in megawatt-hours verified by the individual electric distribution companies and hourly electrical generation aggregated by fuel type, also in megawatts. The hourly load data is flagged with geographical location data which will be leveraged to include regional weather data as a covariate in the models.

3.2 Tools and Techniques

First these three datasets will be datetime aligned and analyzed for an outliers. The different generation sources and load data will then be analyzed and plotted for historical trends and to identify periodicity (cycles) at different context window intervals which will help set expected baselines for the forecast horizons. Once the data has been properly aggregating and aligned the data will then be split into a training, validation, and test split [METHOD TBD].

Models of varying complexity and methodology will be used to better represent the broad spectrum of techniques used today, and to attempt to capture a good representation of which models work best for the different forecast horizons.

Proposed models:

Baseline

- Naive Forecast Model

Statistical

- Autoregressive Integrated Moving Average (ARIMA)
- Vector Autoregression (VAR)

Machine Learning

- Support Vector Machine (SVM)
- XGBoost

Deep Learning

- Long Short-Term Memory (LSTM) Networks
- Transformer Networks

Pre-trained

- Time Series to Vector (TS2Vec)

4 EVALUATION

PLACEHOLDER

5 DISCUSSION

PLACEHOLDER

6 CONCLUSION

PLACEHOLDER

7 APPENDICES

PLACEHOLDER

If your work needs an appendix, add it before the "`\end{document}`" command at the conclusion of your source document.

Start the appendix with the "`appendix`" command:

```
\appendix
```

and note that in the appendix, sections are lettered, not numbered. This document has two appendices, demonstrating the section and subsection identification method.

ACKNOWLEDGMENTS

To Robert, for the bagels and explaining CMYK and color spaces.

A RESEARCH METHODS

A.1 Part One

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi malesuada, quam in pulvinar varius, metus nunc fermentum urna, id sollicitudin purus odio sit amet enim. Aliquam ullamcorper eu ipsum vel mollis. Curabitur quis dictum nisl. Phasellus vel semper risus, et lacinia dolor. Integer ultricies commodo sem nec semper.

A.2 Part Two

Etiam commodo feugiat nisl pulvinar pellentesque. Etiam auctor sodales ligula, non varius nibh pulvinar semper. Suspendisse nec lectus non ipsum convallis congue hendrerit vitae sapien. Donec at laoreet eros. Vivamus non purus placerat, scelerisque diam eu, cursus ante. Etiam aliquam tortor auctor efficitur mattis.

B ONLINE RESOURCES

Nam id fermentum dui. Suspendisse sagittis tortor a nulla mollis, in pulvinar ex pretium. Sed interdum orci quis metus euismod, et sagittis enim maximus. Vestibulum gravida massa ut felis suscipit congue. Quisque mattis elit a risus ultrices commodo venenatis eget dui. Etiam sagittis eleifend elementum.

Nam interdum magna at lectus dignissim, ac dignissim lorem rhoncus. Maecenas eu arcu ac neque placerat aliquam. Nunc pulvinar massa et mattis lacinia.