

Energy Grid Forecasting and Anomaly Detection

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

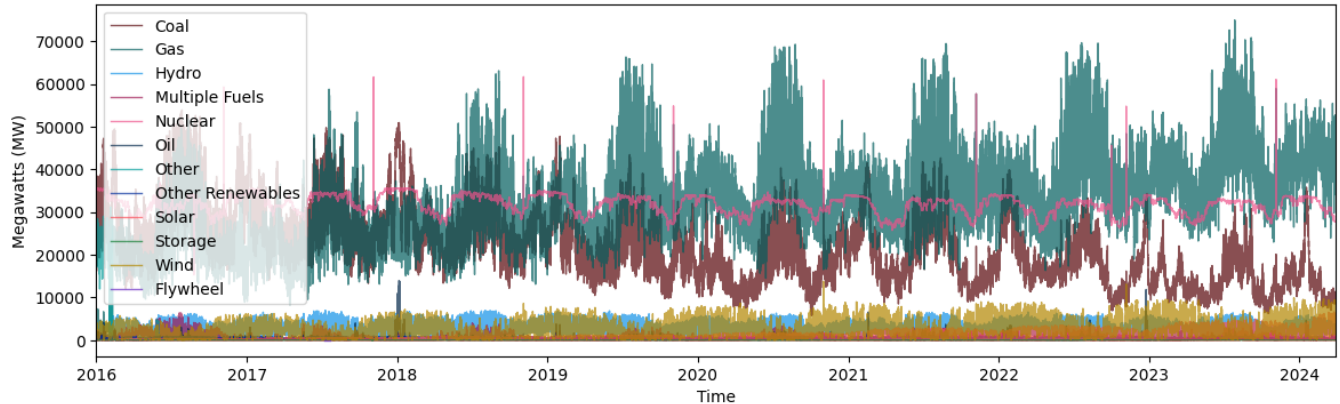


Figure 1: Energy Generated in Megawatts (MW) by Fuel Type.

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, and 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 and heat pumps. As a result, the electrical grid will need to become more robust and adaptive as this increase in demand comes in tandem with a transition away from fossil fuels and into a more diversified energy market and with it, more complex cycles. These circumstances make accurate forecasting periodicity 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.

Many of the statistical time series forecasting techniques of the past are still in use today, each with their own strengths and weaknesses, often trading compute efficiency with lesser accuracy in long-term forecast horizons. More modern machine learning models and neural network-based architectures tend to generalize through forecasting horizons better, but at the cost of compute resources. Though because of the potential for complexity of energy load patterns, both categories see benefits in task-specific tuning. By evaluating each forecasting model's accuracy across various context windows and forecasting horizons, this project presents a thorough comparison on which techniques are most relevant in the task-specific energy grid context and forecast horizons.

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, 5 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 AUTHOR'S NOTE

This document is to be used as a project proposal for the beginning planning stages of the project as well as a living-document as the project progresses. Up until project completion when a new version will be created to fill in the results.

2 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 [7]. Identifying hourly, daily, seasonal, or even yearly energy demands accurately has substantial impact on not only the logistics of the infrastructure of our energy grid, but also the complex energy marketplace. These intervals, or forecast horizons, in the context of predictive models are defined in Table 1. Short-term forecasts can be useful for players in power markets

Table 1: Defining prediction forecast horizon intervals. 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	$1 \text{ Hour} \leq X < 1 \text{ Week}$
3	Medium-Term	$1 \text{ Week} < X \leq 1 \text{ Year}$
4	Long-Term	$X > 1 \text{ Year}$

where prices dynamically shift based on generation supply, customer demand, and financial spot and derivative contracts which are all based on energy forecasting [6]. Very short-term forecasts can be important for power distributors and operators that want to meet customer demands more efficiently. Long-term forecasts are important in infrastructure planning and energy provider and distributors long-term planning.

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 [7, 12, 15]. This diversification of the energy grid not only comes from energy reforms by way of deregulating the national electric market in the 1990's, but emerging cultural desires and improved technology in renewable energies, driving a need for a better electrical grid [10].

Electricity generation has always been mostly an on-demand industry and largely continues to be, where the energy we generate needs to be transmitted and used within a short period of time because we lack efficient systems for energy storage at scale [11]. This fact presents additional complications when many renewable energies only produce energy in certain conditions, making its effective use more reliant on forecasting to navigate a more diverse and complex power grid. Additionally, with the cultural and regulatory push towards end-users adopting the use of electric appliances, electric vehicles, and heat pumps, energy demand will continue to change [5]. These circumstances mean accurate forecasting is becoming even more important with improvements having the potential to better reduce reliance on fossil-fuel energies by more efficiently filling in gaps of energy demands with a more diversified and efficient energy grid.

There is little consensus among researchers and professionals about what exemplifies state-of-the-art in time series forecasting. This seems to be due to many things including that time series forecasting is heavily task-specific, meaning one model may perform well with one type of data and forecast horizon, but might perform poorly when compared to other techniques with a different horizon or data type [16]. This is why it is still important to experiment and compare forecasting techniques on domain-specific time series data.

3 RELATED WORK

Autoregressive Integrated Moving Average (ARIMA) and Exponential Smoothing models tend to be useful in short and medium-term

forecasts but can experience limitations modeling complex periodicity and interactions [8, 13]. Newer, but certainly not state-of-the-art, machine learning techniques such as Support Vector Machines, XG-Boost, and a variety of Long Short-Term Memory (LSTM) network architectures tend to generalize better and excel across multiple forecast horizons [13, 14].

Hybrid models and ensemble methods, utilizing a variety of preprocessing techniques and combining models into a unified framework, are becoming more popular as of late. The process typically involves splitting the job of extracting different types of features from the data, say the temporal and spatial relationships of the time series, and feeding those outputs into yet another separate model to perform the forecasting [9]. This technique combines the strengths of multiple models resulting in improved results across all forecast horizons [8, 9]. Many of these still utilize LSTM layers as their attention mechanism [14].

Transformer-based models and other frameworks developed and popularized in the Large Language Model (LLM) space have found a foothold in time series forecasting in the last few years as well, but their use in time series forecasting is still being researched [17]. Similar to the LSTM, transformer-based models capture complex context and relationships in data, but across greater distances of the training corpus with increased efficiency because of their ability to be parallelized effectively [18]. This type of model benefits from pre-training on a large corpus of diverse data which simply does not exist in the public domain for time series data [1]. With more time and data, researchers believe transformer-based models can be effective at time series forecasting [1, 18].

4 METHODS AND PROPOSED WORK

Several models stemming from different statistical and machine-learning families will be trained to perform electricity load forecasting as well as electricity generation by fuel type across the different forecast horizon intervals listed in Table 1. Each model's amount of compute required for training and evaluation scores will be compared across different forecast horizons. See evaluation section for more details on the specific metrics used.

4.1 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 [2]. A map of this can be seen in Figure 2.

Table 2: Dataset size and date ranges.

Dataset	Size (Rows)	Date Range
Hourly Load	5,446,243	1993/01/01 - 2024/03/27
Hourly Generation	783,663	2016/01/01 - 2024/03/28
Avg Hourly Weather	TBD	TBD

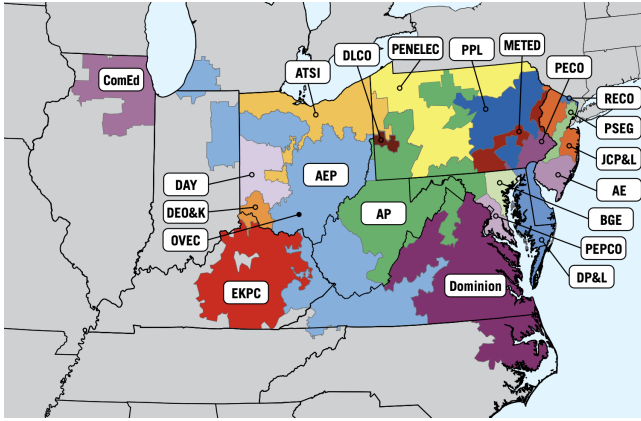


Figure 2: Regional Transmission Zones Labeled by Distributor IDs [4].

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 [3]. 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. An overview of the datasets can be seen in Table 2.

4.2 Tools and Techniques

First these three datasets will be datetime aligned and analyzed for 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. Some minor feature engineering will be performed on the datasets including encoding days of the week into the feature matrix since energy demands change throughout the week. Once the data has been properly aggregated and aligned the data will then be split into a training, validation, and test split.

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)

Some models may require data transformation or normalization to perform effectively. These changes will be listed here.

5 EVALUATION

5.1 Data Split

The data will be split into a 70/20/10 training, validation, testing split. A validation set is necessary here because several of the proposed forecasting models will require hyperparameter tuning. Once the data has been aggregated, cleaned, and finalized the exact ratio may be adjusted to ensure the test set is large enough to span the largest forecast horizon at least one or several times, depending on what is feasible with the data on hand.

During the splitting process extra considerations will be taken to ensure there are no temporal conflicts between sets (i.e. part of the test set is overlapping a part of the training set). Also, that each set contains a reasonable amount of periodicity.

5.2 Metrics

Symmetric Mean Absolute Percentage Error (sMAPE) will be the evaluation metric used to compare all models. This will ensure comparisons of error across different models and forecasting horizons are normalized.

$$\text{sMAPE} = \frac{200}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|} \quad (1)$$

where y_i and \hat{y}_i are the actual values and forecasted values respectively, and n is the number of points in the set.

One possible limitation of sMAPE is if a predicted value or actual value equals zero, the metric balloons to the maximum error value, and if both equals zero, the metric is undefined. Chosen evaluation metrics often have strengths and weaknesses, and it is yet unclear if this limitation will occur here – especially since power generation by different fuel types may have a periodicity that includes zero power generation.

Root Mean Squared Error (RMSE) will also be used in model evaluation to complement sMAPE. The value of including RMSE is that it is an easily interpretable quantitative metric resulting in a value that is in the same unit as the response. It also heavily penalizes predictions the further they are from the actual value. Alternatively, Normalized RMSE (NRMSE) will be considered if it is found that different models output at different scales.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

where y_i and \hat{y}_i are the actual values and forecasted values respectively, and n is the number of points in the set.

6 DISCUSSION

More discussions and analysis will be added here when the project has been completed.

6.1 Project Timeline

Table 3 outlines the specific tasks and their projected schedule.

Table 3: A general framework onto which the project will progress. This timeline is subject to change.

Timeline	Agenda Item	Progress	Done
Days 1 - 5	Data Mining	X	
	Hourly Load Data	-	X
	Hourly Generation Data	-	X
	Weather Data	X	
	Cleaning/Preprocessing	X	
	Datetime Alignment	X	
	Aggregate by Zones	X	
	Outlier Inspection		
	Corrupted Data Removal		
	Exploratory Data Analysis		
Days 6-8	Train/Val/Test Split		
	Anomaly Detection		
	Forecast Model Setup		
	Training		
	Val. Optimizations		
Days 9-10	Build Eval. Functions		
	Test Results		
	Discussion		
	Conclusion		

6.2 Potential Challenges

A few challenges may arise impeding progress or causing the project to fall behind schedule are expected to be the following. Generally, properly fitting the data to the number of proposed forecast models within the timeline may pose a challenge. Specifically, the neural-network architectures can take considerable time to build from the ground up without intimate knowledge and experience of their specific architectures. Also, completely fleshing out the three goals of the project on schedule may be cause for concern and need reevaluation – anomaly detection, hourly load forecasting, and hourly energy generation forecasting.

The granularity of the data may post issues in accurately forecasting the shorter forecast horizons. Since all collected data has a time-step (sample rate) granularity of one hour, predictions less than one hour have little basis in reality. Data with smaller a time-step would be required for accurate shorter forecast horizons.

6.3 Alternate Approaches

If the project is running behind the timeline a few different approaches can be considered to make the target deadlines.

- (1) The number of models can be reduced but leaving at least one remaining from each model type.

- (2) In the case of some of the neural-network based architectures and machine-learning approaches, complexity can be reduced by utilizing more out-of-the-box approaches/models.
- (3) Instead of forecasting both hourly load usage and generation by fuel type, the scope of the project can be reduced to focus on just one of these domains.
- (4) Further reduce the scope to solely focus on anomaly detection.

7 CONCLUSION

By analyzing and modeling energy time series data we can efficiently and accurately forecast into the future different interval sizes, or forecast horizons, which correspond to real-world decisions. Modern techniques in time series forecasting have reduced the errors in all forecast horizon intervals and continue to generalize those prediction abilities across many domains. This type of research and technology is necessary for uninterrupted energy generation and distribution due to the ever-increasing demands on our energy grids.

7.1 Key Findings

To be completed upon project completion.

7.2 Future Work

To be completed upon project completion.

8 APPENDICES

PLACEHOLDER

ACKNOWLEDGMENTS

Sqyd, for always brewing the morning coffee – I know it's 8 cups of water, now.

Luna, for breaking up the hyper-focused monotony with mind-clearing afternoon walks.

REFERENCES

- [1] 2024. <https://huggingface.co/blog/autoformer>
- [2] 2024. <https://www.pjm.com/about-pjm/who-we-are/territory-served>
- [3] 2024. https://dataminer2.pjm.com/feed/hrl_load_metered/definition
- [4] 2024. <https://www.pjm.com/library/maps.aspx>
- [5] James Archsmith, Erich Muehlegger, and David S. Rapson. 2022. Future Paths of Electric Vehicle Adoption in the United States: Predictable Determinants, Obstacles, and Opportunities. *Environmental and Energy Policy and the Economy* 3 (2022), 71–110. <https://doi.org/10.1086/717219> arXiv:<https://doi.org/10.1086/717219>
- [6] Derek W. Bunn. 2004. *Modelling Prices in Competitive Electricity Markets*. Wiley.
- [7] Tao Hong and Shu Fan. 2016. Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting* 32, 3 (2016), 914–938. <https://doi.org/10.1016/j.ijforecast.2015.11.011>
- [8] Konstantin Hopf, Hannah Hartstang, and Thorsten Staake. 2023. Meta-Regression Analysis of Errors in Short-Term Electricity Load Forecasting. In *Companion Proceedings of the 14th ACM International Conference on Future Energy Systems (Orlando, FL, USA) (e-Energy '23 Companion)*. Association for Computing Machinery, New York, NY, USA, 32–39. <https://doi.org/10.1145/3599733.3600248>
- [9] Heng Hua, Mingping Liu, Yuqin Li, Suhui Deng, and Qingnian Wang. 2023. An ensemble framework for short-term load forecasting based on parallel CNN and GRU with improved ResNet. *Electric Power Systems Research* 216 (2023), 109057. <https://doi.org/10.1016/j.epr.2022.109057>
- [10] Paul Joskow. 2001. U.S. Energy Policy During the 1990s. *National Bureau of Economic Research, Inc, NBER Working Papers* (01 2001).
- [11] Chun Sing Lai, Giorgio Locatelli, Andrew Pimm, Xiaomei Wu, and Loi Lei Lai. 2021. A review on long-term electrical power system modeling with energy

- storage. *Journal of Cleaner Production* 280 (2021), 124298. <https://doi.org/10.1016/j.jclepro.2020.124298>
- [12] Eunjung Lee and Wonjong Rhee. 2021. Individualized Short-Term Electric Load Forecasting With Deep Neural Network Based Transfer Learning and Meta Learning. *IEEE Access* PP (01 2021), 1–1. <https://doi.org/10.1109/ACCESS.2021.3053317>
- [13] Shahzad Muzaffar and Afshin Afshari. 2019. Short-Term Load Forecasts Using LSTM Networks. *Energy Procedia* 158 (2019), 2922–2927. <https://doi.org/10.1016/j.egypro.2019.01.952> Innovative Solutions for Energy Transitions.
- [14] Lichao Sun, Hang Qin, Krzysztof Przystupa, Michał Majka, and Orest Kochan. 2022. Individualized Short-Term Electric Load Forecasting Using Data-Driven Meta-Heuristic Method Based on LSTM Network. *Sensors* 22 (10 2022), 7900. <https://doi.org/10.3390/s22207900>
- [15] Rafal Weron. 2014. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting* 30, 4 (2014), 1030–1081. <https://doi.org/10.1016/j.ijforecast.2014.08.008>
- [16] Zhihan Yue, Yujing Wang, Juanyong Duan, Tianmeng Yang, Congrui Huang, Yu Tong, and Bixiong Xu. 2021. TS2Vec: Towards Universal Representation of Time Series. In *AAAI Conference on Artificial Intelligence*. <https://api.semanticscholar.org/CorpusID:237497421>
- [17] Ailing Zeng, Muxi Chen, Lei Zhang, and Qiang Xu. 2022. Are Transformers Effective for Time Series Forecasting? [arXiv:2205.13504](https://arxiv.org/abs/2205.13504) [cs.AI]
- [18] Vincent Zhihao Zheng and Lijun Sun. 2024. Multivariate Probabilistic Time Series Forecasting with Correlated Errors. [arXiv:2402.01000](https://arxiv.org/abs/2402.01000) [stat.ML]