Forecasting Electricity Usage: Comparative Analysis of Prophet, LSTM, and Hybrid Models on Public Datasets

Research

Question. When evaluated under identical, leakage-safe rolling backtests, does Facebook Prophet match or beat Long Short-Term Memory networks for electricity-usage forecasting across multiple public datasets?

Goal. Reproduce head-to-head Prophet vs. Long Short-Term Memory results from recent literature and test a small hybrid that combines Prophet's trend/seasonality with Long Short-Term Memory residual modeling.

Brief summaries of the research papers

[1] T. Bashir, et al., "Short term electricity load forecasting using hybrid prophet-LSTM model optimized by BPNN", Energy Reports, Volume 8, P. 1678-1686, Nov 2022. [Online]. Available: *https://www.sciencedirect.com/science/article/pii/S2352484721015067?*

- ***Summary:*** Show a simple idea, let Prophet handle calendar patterns, then ask an LSTM to clean up the residuals. They test on Belgian load data with strong daily and weekly rhythm. The hybrid cuts MAE and RMSE compared to Prophet, ARIMA, and solo LSTM. It works very well with holidays and workday cycles which are most of structure. The takeaway is that you split the job and each model does what it's best at.

[2] J. Henzel, et al. "Energy Consumption Forecasting for the Digital-Twin Model of the Building", Energies, Jun 2022. [Online]. Available:https://www.researchgate.net/publication/361294260_Energy_Consumption_Forecasting_for_the_Digital-Twin_Model_of_the_Building

- ***Summary:*** Uses building energy data in the “digital twin”. They compare naive baselines, linear models, LSTM, and Prophet under different sets of inputs. Prophet sometimes beats LSTM when calendar effects dominate. Feature choices, like weather, matter as much as the model. You need the correct model for the type of data you have. The takeaway is that good inputs and a sturdy baseline can win.

[3] S. Arslan, “A hybrid forecasting model using LSTM and Prophet for energy consumption with decomposition of time series data” PeerJ Computer Science, Jun 2022. [Online]. Available:
https://www.researchgate.net/publication/361222023_A_hybrid_forecasting_model_using_LSTM_and_Prophet_for_energy_consumption_with_decomposition_of_time_series_data

- ***Summary:*** This paper builds two stages, break down trend and seasonality first, then model the remainder with LSTM. This keeps the deep net focused on truly nonlinear leftovers. The hybrid beats single models on MAE and RMSE. Variants like stacked or bidirectional LSTMs help but don’t change the core outcome. The structure then “nonlinearity” pattern is the best.

[4] Q. Dong, et al., “Short-Term Electricity-Load Forecasting by Deep Learning: A Comprehensive Survey” Engineering Applications of Artificial Intelligence, Vol 154, Aug 2025. [Online]. Available:
https://www.sciencedirect.com/science/article/abs/pii/S0952197625009807

- ***Summary:*** They survey modern short-term load forecasting with deep learning. They cover RNNs, CNNs, Transformers, data prep, “exogenous” features, and evaluations. One problem is data leakage and sloppy backtests that inflate the results. They also state that strong baselines can match complex models on some horizons.

[5] S. Albahli, “LSTM vs. Prophet: Achieving Superior Accuracy in Dynamic Electricity Demand Forecasting”, Energies, Volume 18, Jan 2025. [Online]. Available:
https://www.researchgate.net/publication/387913942_LSTM_vs_Prophet_Achieving_Superior_Accuracy_in_Dynamic_Electricity_Demand_Forecasting

- ***Summary:*** Proposes a model that shifts weight between Prophet and LSTM based on error signals. When the series is stable and strongly seasonal, the ensemble leans toward Prophet. When patterns are spiky, it increases the LSTM’s share of the data. This blending improves accuracy without adding heavy complexity. In short, it’s a practical between structure and memory.

[6] S. Chaturvedi, et al. "A comparative assessment of SARIMA, LSTM RNN and Fb Prophet models to forecast total and peak monthly energy demand for India" Energy Policy, Vol 168, Sep 2022. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0301421522003226

- **Summary**: This paper takes a look at India's national monthly power demand. They compare SARIMA, LSTM-RNN, CEA, and Prophet for totals and peaks. Prophet and CEA beat the others on error metrics and stability. Coarse time scales are better for simple, calendar structure over heavy sequence memory. For planning, they pick Prophet for clarity and reliability.

[7] X. Yang, et al., "An optimized short-term load forecasting method based on time and weather-fused features using a ConvLSTM-3D; Prophet used for temporal decomposition" Frontiers in Energy Research, Vol 12, Jan 2024. [Online]. Available: https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2024.1501963/full

- **Brief Description**: Fuse time patterns and weather with a ConvLSTM-3D model. They strip calendar effects out using Prophet, then let the deep model learn the residual dynamic. SHAP helps select the most useful weather features. The combined approach improves forecasts by splitting structure from noise. It's a blend of explainability and power.

Datasets

- **Electricity Transformer Temperature** (choose ETTh1 and or ETTm1). Official repository with data and documentation.
  - https://github.com/zhouhaoyi/ETDataset?utm_source=chatgpt.com
- **Electricity Load Diagrams** (household electricity demand). Official University of California Irvine repository page.
  - https://archive.ics.uci.edu/dataset/321/electricityloaddiagrams20112014
- **Highway traffic sensor occupancy** (for example, California Performance Measurement System). Official Caltrans data source; a University of California Irvine variant is also available.
  - https://dot.ca.gov/programs/traffic-operations/mpr/pems-source?utm_source=chatgpt.com