Bridging the Temporal Gap

From Monthly Invoices to Hourly Energy Forecasting for Sustainable Operations

Prathamesh Pawar[1], Alec Hewitt[1], William Schuerman[1], Seyma Gunes[1], Will Sorenson[1]

[1]**Amazon Web Services**

Tackling Climate Change with Machine Learning workshop at NeurIPS 2025

The Climate Challenge

Industrial Electrification Requires Granular Energy Insights

The Urgency

- Electricity's share of total energy must rise from 20% to 27% by 2030 for NetZero [1]
- Example: 800MW electrolyzer delayed due to grid stability concerns [2]
- Hourly forecasts needed for carbon accounting, demand response, and renewable integration

The Data Gap

- Most facilities only have monthly utility invoices
- Smart meter installation requires expensive outages, permitting, and coordination
- Without hourly data, sustainability planning is severely limited

Our Approach

Invoices Are All You Need” - A 3-Step ML Framework

Step 1: Monthly Forecasting: ML models predict 1-3 months ahead

Step 2: Pattern Learning: Bayesian model learns hourly distributions from training sites

Step 3: Disaggregation: Convert monthly forecasts to hourly predictions for test sites

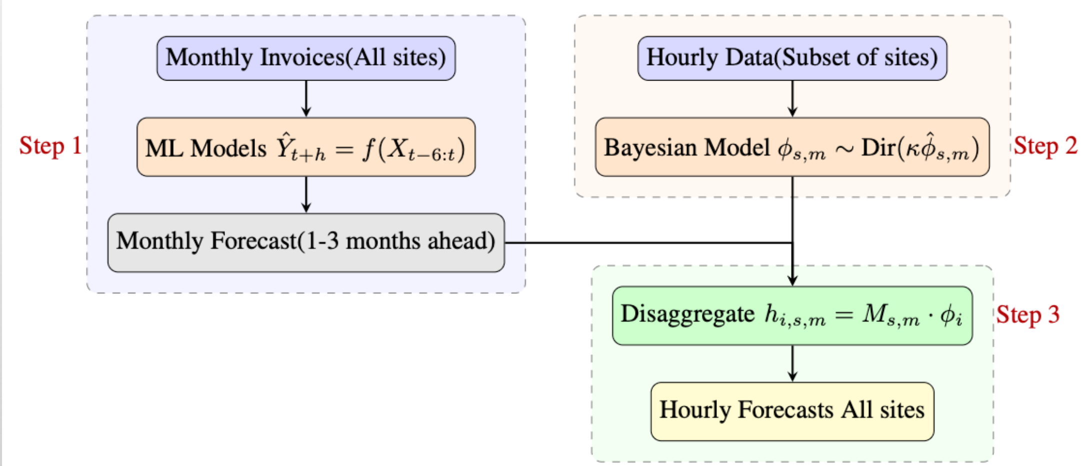


Figure 1: The proposed 3-step methodology

Dataset

Industrial site load data

- UCI Electricity Consumption & Load (ECL) dataset [3] containing 15-minute resolution data from 370 Portuguese industrial facilities (2012-2015). Note: *19 sites dropped due to missing data.*
- All sites used for monthly load forecasting (1-3 months horizon)
- For hourly disaggregation, we use a 50/50 split: 175 sites with hourly meters train the disaggregation model, while 176 held-out sites test the full pipeline using only monthly invoices.

Step 1: Monthly Forecasting Results

TabPFN Achieves Stable Performance Across Horizons

- TabPFN [6] maintains consistency: MAE stays <70 MWh across all horizons
- Baseline degrades sharply: 50 → 128 MWh
- 24-month training window, expanding cross-validation

Model	**H1 (1mo)**	**H2 (2mo)**	**H3 (3mo)**
Baseline	50	99	128
AutoGluon [4]	73	76	78
LightGBM [5]	69	69.2	68.9
TabPFN [6]	**68**	**65**	**65**

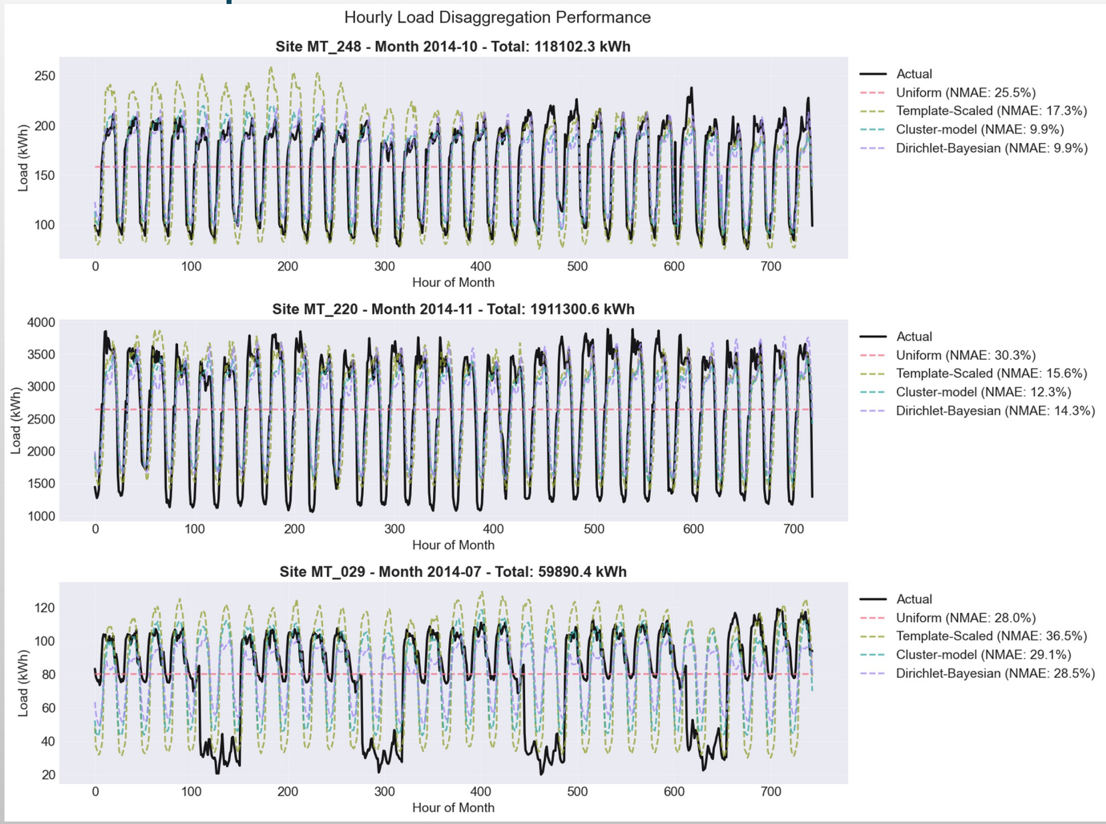
Table: Monthly Forecasting Performance (MAE in MWh)

Step 2-3: Monthly to Hourly Disaggregation

Hourly disaggregation performance on Representative Sites

We compared our method against standard approaches for converting monthly data to hourly predictions:

- Uniform Model: Distributes total monthly load evenly across all hours
- Template-based Model: Scales historical average hourly patterns by monthly totals
- Cluster-based Model: Groups sites by monthly consumption features. Applies cluster-specific templates
- Dirichlet-Bayesian Model:
    - Robust prior estimation via trimmed means
    - Day-of-week corrections
    - Cluster-aware template mixing



Step 2-3: Monthly to Hourly Disaggregation

Full Pipeline Evaluation

- 28-30% error reduction vs. uniform baseline
- Stable performance across 176 test facilities (50% of dataset)
- Works across different industrial load profiles

Model	Baseline			Forecast			Actual
	H1	H2	H3	H1	H2	H3	Observed
Uniform	178.8	185.7	189.3	174.2	176.7	178.3	157.6
Template-Scaled	152.4	160.6	164.9	159.2	160.1	159.1	145.6
Cluster-Based	132.2	143.6	150.7	131.4	133.5	134.0	112.7
Dirichlet-Bayesian	**130.5**	**142.5**	**150.2**	**129.7**	**132.1**	**133.1**	**109.0**

Table 2: Performance comparison of disaggregation models across forecasting scenarios and horizons

Note: H1, H2, H3 represent 1-, 2-, and 3-month ahead forecast horizons. Mean MAE across 1232 test site-month pairs (176 x 7 rolling test windows) in kWh. Baseline uses lagged monthly totals, Forecast uses ML predicted totals, and Actual uses observed monthly totals. Bold values indicate best performance in each column. We choose MAE primarily because it is most relevant for use cases like Demand Response, Integrated Resource Planning, etc.

Climate Impact Applications

Democratizing Hourly Energy Insights

Immediate Applications

- Carbon Accounting: Real-time emissions tracking vs. 3-month delayed insights
- Demand Response: Accurate load curtailment commitments for grid stability
- Renewable Integration: Precise solar and storage sizing for net peak reduction

Grid-Scale Impact

- Grid operators can improve load forecast accuracy
- Enables better demand response and renewable integration
- Supports large-scale industrial electrification
- No new metering infrastructure required on urgent basis

Conclusion & Impact

From monthly invoices to Climate Action

Key Contributions

- 28-30% improvement in hourly energy forecasting accuracy over uniform baseline
- No additional metering instrumentation required
- Enables immediate sustainability applications
- Supports grid stability during energy transition

Next Steps

- Sub-hourly predictions for granular grid balancing
- Incorporating weather patterns and grid conditions
- Automated facility clustering techniques
- Deployment in diverse geographic regions

Thank You

Questions?

Contact:

prathavp@amazon.com

Acknowledgments:

NeurIPS 2025 CCAI Workshop

AWS

NEURAL INFORMATION PROCESSING SYSTEMS

References I

[1] International Energy Agency. Net Zero by 2050: A Roadmap for the Global Energy Sector. Tech. rep. Paris: IEA, 2021.

[2] J. Burgess. Sweden's H2 Green Steel signs 14-TWh PPA to power planned electrolyzer. S&P Global Commodity Insights. 2022. url: https://www.spglobal.com/commodityinsights/en/market-insights/latest news/electric-power/060822-swedens-h2-green-steel-signs-14-twh-ppa-to-power-planned-electrolyzer.

[3] Trindade, A. (2015). ElectricityLoadDiagrams 20112014 [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C58C86.

[4] N. Erickson, J. Mueller, A. Shirkov, H. Zhang, P. Larroy, M. Li, and A. Smola. "AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data". In: arXiv preprint arXiv:2003.06505 (2020).

[5] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: Proceedings of the 31st International Conference on Neural Information Processing Systems. 2017, pp. 3149–3157.

[6] N. Hollmann, S. M¨uller, K. Eggensperger, and F. Hutter. "TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second". In: arXiv preprint arXiv:2207.01848 (2023).

[7] M. Lamagna, B. Nastasi, D. Groppi, et al. "Hourly Energy Profile Determination Technique from Monthly Energy Bills". In: Building Simulation 13 (2020), pp. 1235–1248.