

Explainable Data-Driven Smart Meter Analysis

Marius Führer, FH Technikum Wien
Anahid Wachsenegger, Austrian Institute of Technology (AIT)
Lukas Durnwalder, optivice.energy
Matthias Blaickner, FH Technikum Wien
Torsten Priebe, FH St. Pölten
Rosana De Oliveira Gomes, FH Technikum Wien

Abstract:

Energy communities are gaining importance as decentralized energy supply structures, but they also pose new challenges for grid operators in terms of data transparency and forecasting capabilities. This paper presents a FAIR-AI pipeline for an Austrian energy community consisting of three modules: (1) an explainable consumption forecasting model based on gradient boosting methods with SHAP attributions for a 7-day horizon; (2) a physics-based probabilistic estimation of behind-the-meter PV self-consumption using clear-sky modeling and Monte Carlo simulation; and (3) an efficiency assessment to estimate the potential of battery storage using linear optimization.

Keywords: energy community, explainable AI, smart meter data, battery energy storage, load forecasting

1. INTRODUCTION

FAIR-AI is an Austrian flagship project, translating trustworthy AI and EU AI Act requirements into operational methods, tools, and guidance. It targets technical, organizational, and socio-technical risks in real-world AI systems through a bottom-up, use-case-driven approach for compliant deployment.

In this context, energy communities provide a relevant use case for applying trustworthy AI in data-driven energy system analysis. Smart meter data supports key decisions such as demand forecasting, PV generation estimation, and battery planning, which can influence how resources and benefits are distributed among participants. This work integrates an explainable consumption forecast, a physics-informed PV analysis, and a battery potential estimation within a unified FAIR-AI pipeline, enabling transparent and interpretable assessment of energy flows in energy communities.

2. EXPLAINABLE FORECASTING¹

2.1. Problem Description and Forecasting Objective

We forecast hourly electricity consumption in an energy community for metering points of solar-participating entities using historical meter readings and exogenous weather variables. The target is weighted measured consumption, where weighting effects are rare, meaning that only few exceptions participated in other energy communities at the same time. Forecasts are produced for a fixed horizon of 7 days (hourly step) and are evaluated via a rolling hold-out on the final 90 days to mimic operational deployment.

2.2. Data Quality Check

Smart meter readings are aggregated sum values from 15-minute to hourly resolution, which defines the available lag structure and affects peak visibility and seasonality strength. Preprocessing aligns consumption and generation meters across participants and propagates the data-quality flag quality (L1 measured, L2 replaced) so that imputations (using backward filling followed by forward filling) are explicitly accounted for in modelling and in explanations. Exploratory analysis is hypothesis-driven and uses Mann–Whitney U tests with rank-based effect sizes to identify practically relevant regime differences (e.g., daylight vs. night) and translate them into feature choices and interpretation priors.

General procedure (applied per variable):

1. Define a binary grouping variable (e.g., daylight indicator)
2. Split the hourly samples into two sets
3. Compute Mann–Whitney [1] and corresponding p-value; compute effect size (rank-biserial or equivalent).
4. Interpret results jointly: significance + effect size + domain plausibility.

Hypothesis H1: Daylight vs. night-time differences: Comparing daylight (07:00–18:00) to night shows only a minor consumption drops (~4%) but strong daylight-driven allocation and near-zero night PV generation, motivating time/irradiance features and the expectation that explanations attribute generation effects primarily to these drivers.

Hypothesis H2: Weekday vs. weekend differences: Weekday–weekend comparisons show only negligible consumption differences but slight shifts in coverage (self-consumption vs. grid) and a modest weekend increase in generation (~5%, with surplus more sensitive), implying calendar features are secondary to daylight/temperature yet should be enriched with holiday indicators to avoid confounded explanations.

Hypothesis H3: Temperature regime effects and seasonality: Quantile-based cold–warm comparisons show much higher consumption in cold periods (~+65 kWh/h, consistent with heating load) and nearly triple generation with higher surplus in warm periods, making temperature and irradiance key forecasting covariates and suggesting SHAP [3] temperature effects should be read as regime indicators rather than simple linear sensitivity.

¹ Anahid Wachsenegger

Covariate definitions: To match the forecasting objective, 15-minute smart-meter data are aggregated to hourly resolution and split by availability at prediction time. The target is the measured electricity consumption; past covariates include community-level photovoltaic potential and coverage indicators, measured electricity generation, surplus generation, and weather observations (temperature, wind speed, and solar irradiance). Future covariates are limited to the weather variables over the 7-day forecasting horizon. Preprocessing is kept minimal (no normalization or feature engineering) to maintain a transparent baseline for later explainability comparisons.

Baseline and tree-based forecasters: Baselines provide technical anchors for accuracy and interpretability over the 7-day horizon. We use persistence forecasts to quantify how much predictive power comes from short-term autocorrelation, and seasonal naïve baselines (daily/weekly) to capture dominant periodicities without exogenous inputs. As an interpretable parametric reference, we include an ARX model [14] with a small lag set (e.g., 1 h, 24 h, 168 h) and meteorological covariates; its coefficients provide a sanity check for later attribution results. We then employ a higher-capacity tree-based regression forecaster with time-aware encodings (cyclic hour/day patterns, calendar encodings, and horizon position). Evaluations (Table 1) indicate that this model adapts slowly to abrupt changes and systematically underestimates peaks, implying that explanations must be interpreted in the context of these error modes, particularly during rare high-impact events.

Table 1. Overview of the tree-based forecasters performance.

model	MAE	RMSE	MAPE_%	sMAPE_%	R2
CatBoostModel	8.12	11.15	12.36	11.96	0.61
LightGBMModel	8.45	11.24	12.59	12.67	0.60
XGBModel	8.79	12.10	13.07	12.91	0.54

Explainability: Figure 1 illustrates the forecasting results together with the model inputs and their SHAP-based explanations [3]. This visualization method is motivated by [2], where the top panel compares the observed power consumption with the model forecast, where the grey line represents historical observations and the dotted black line the predicted values. The shaded region indicates the forecast window.

The middle panel shows a heatmap of the input features used by the model over time. Periodic patterns are visible in several variables, particularly solar radiation, reflecting daily environmental cycles. The highlighted region corresponds to the forecast horizon where some covariates are unavailable.

The bottom panel displays the SHAP explanations for the most influential features. Positive SHAP values indicate features that increase the predicted consumption, while negative values reduce it. The visualization suggests that environmental variables, especially solar radiation and its lagged values, are among the main drivers of the forecast, while operational signals often reduce the predicted consumption.

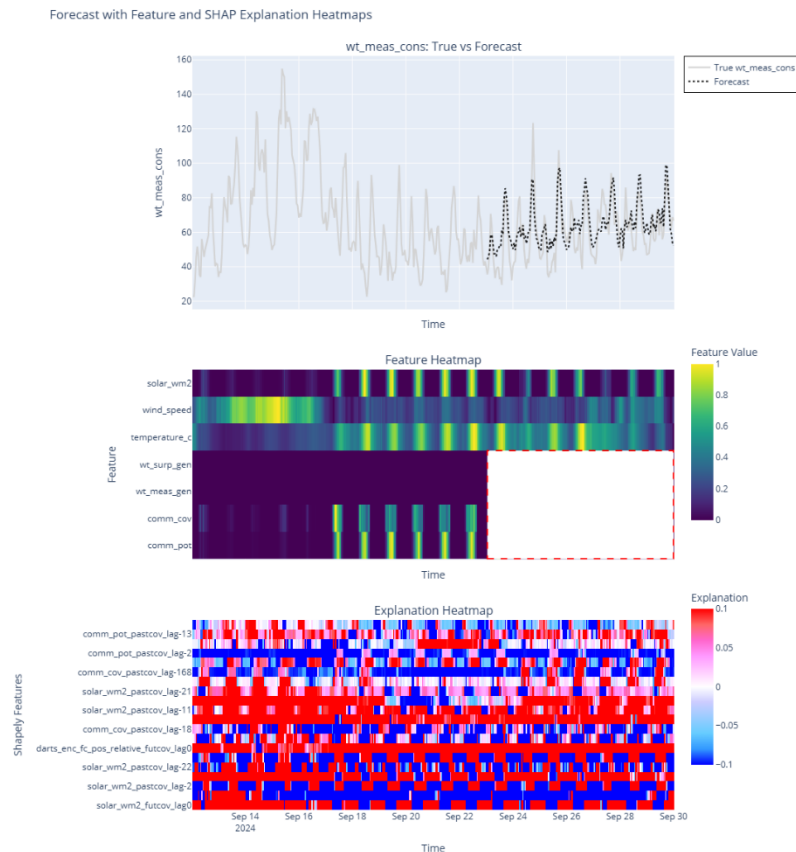


Figure 1. Explanation of the forecast with three layers: past data, forecast, covariate values heatmap, and SHAP values per lagged feature.

3. PHYSICS-INFORMED PROBABILISTIC ESTIMATION OF PV SELF-CONSUMPTION²

The proliferation of rooftop photovoltaic (PV) installations and behind-the-meter (BTM) battery storage systems poses a fundamental challenge for distribution system operators: PV generation is rarely directly metered, meaning that the energy a prosumer self-consumes is structurally unobservable from smart meter data alone [4,5]. Self-consumption rates directly determine the net load a prosumer places on the distribution network, influence the residual load curves that operators must balance, and affect the design of dynamic tariffs and grid reinforcement strategies [6]. This section presents a physics-informed, probabilistic methodology for estimating BTM PV self-consumption from 15-minute smart meter interval data and frames this estimation problem as the first stage of an explainable machine learning pipeline for prosumer segmentation.

² Marius Führer

3.1. Problem Formulation

Let each prosumer p be characterised by a set of metering points recording gross import from the grid and gross export to the grid at 15-minute resolution. Under a standard net-metering or injection-tariff regime, the observable quantities are import $I(t)$ and export $E(t)$. The latent quantity of interest is PV generation $G(t)$, from which self-consumption $SC(t)$ is defined as:

$$SC(t) = G(t) - E(t).$$

In the absence of battery storage, export closely tracks surplus solar generation, and a simple proxy $G(t) \approx E(t) + [\textit{correction}]$ is reasonable [7]. However, with a co-located battery, surplus PV is first routed to storage, strongly decoupling export from irradiance [8]. A prosumer with a large battery and time-of-use tariff incentives may export very little at midday even in full sun, making export-based proxies systematically biased downward.

3.2. Physics-informed PV Generation Proxy

Rather than relying on export as a PV proxy, we construct a physics-based estimate of generation using the `pvlb` library [9]. For each unique timestamp in the dataset, we compute the clear-sky global horizontal irradiance $GHI_{clear}(t)$ under the Ineichen Perez model [10]. This model accounts for solar geometry (zenith angle, azimuth) and atmospheric turbidity as a function of site coordinates and date, producing a deterministic upper envelope for solar irradiance.

The clear-sky signal is then normalised to a fractional factor $f_{cs}(t)$:

$$f_{cs}(t) = \frac{GHI_{clear}(t)}{\max(GHI_{clear})}.$$

To account for cloud cover—which can suppress actual irradiance well below the clear-sky envelope—we derive a community-level cloud attenuation factor $k_{cloud}(t)$ by comparing aggregate community export across all prosumers against the clear-sky envelope, similar in spirit to clear-sky index estimation used in solar forecasting [11]. Because individual battery effects cancel at sufficient scale, the community's export signal serves as a reliable cloud proxy.

The per-user PV estimate is then scaled as:

$$\bar{G}(t) = \alpha_p \cdot f_{cs}(t) \cdot k_{cloud}(t)$$

where α_p is a user-specific scaling coefficient estimated by minimising the residual between $\bar{G}(t)$ and observed export during periods where battery interference is expected to be minimal. This hybrid approach follows the broader paradigm of physics-informed machine learning [12].

3.3. Battery-Aware Bounds and Monte Carlo Uncertainty Quantification

Given $\bar{G}(t)$ and observed $E(t)$, we define a deterministic feasibility envelope for self-consumption. The upper bound assumes that all surplus PV not exported is self-consumed or stored:

$$SC_{max}(t) = \max(0, \bar{G}(t) - E(t)).$$

The lower bound accounts for the possibility that a battery is charging at maximum rate C_{max} , absorbing PV that would otherwise serve the load directly:

$$SC_{min}(t) = \max(0, \bar{G}(t) - E(t) - C_{max}).$$

To move beyond point-in-time bounds and obtain a distribution over plausible self-consumption trajectories, we employ a Monte Carlo simulation with 300 iterations per user over the unknown battery parameter space, which is derived from typical battery sizes. Figure 2 shows an example of the bounds of self-consumption (green shade) for different days.

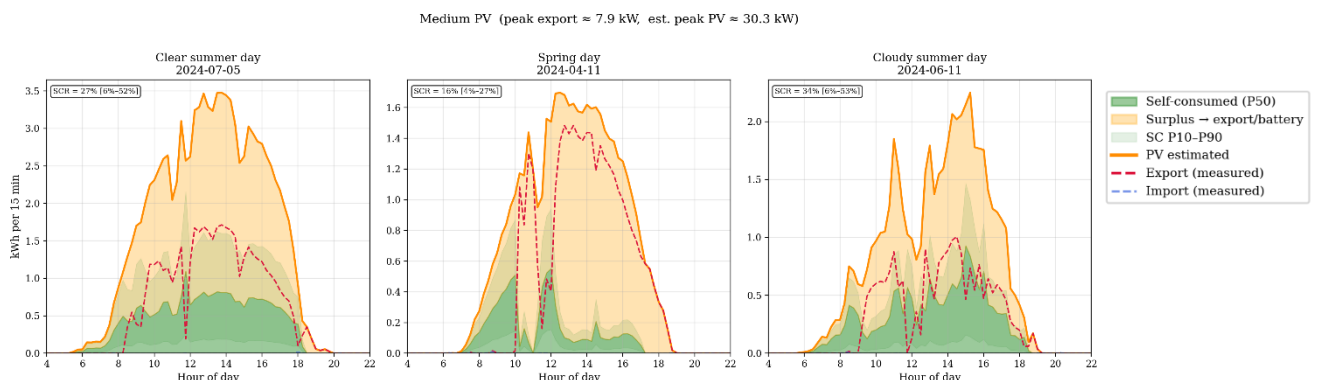


Figure 2. Daily average of estimated total production, measured export and self-consumption boundaries inferred from the Monte Carlo simulation result.

4. BATTERY POTENTIAL ESTIMATION³

4.1. Problem Description

Solar electricity production occurs only during daylight hours. Consequently, electricity generation in the energy community is concentrated during the day, while a substantial share of consumption occurs during evening and nighttime periods. Battery storage can mitigate this temporal mismatch by storing excess daytime generation for later use. This raises the practical question of appropriate battery sizing.

³ Lukas Peham

4.2. Methodology and Results

We define an efficiency space spanned by two metrics: self-sufficiency and energy utilization. Self-sufficiency measures the share of demand covered by community generation, while energy utilization quantifies the fraction of produced energy that is consumed within the community.

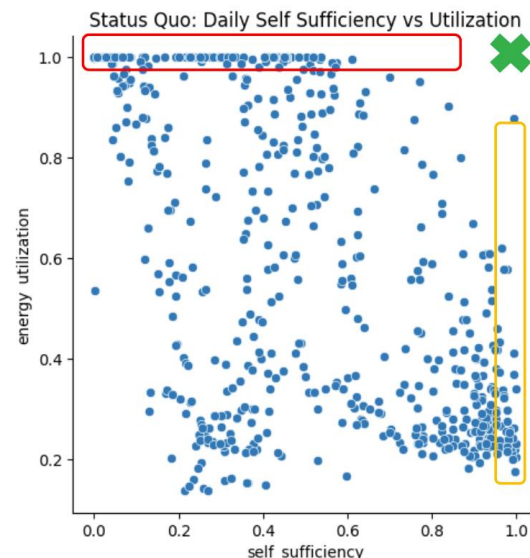


Figure 3. Energy Utilization and Self Sufficiency. The red area indicates that no energy leaves the community, the yellow area indicates that the demand is fully covered by the community. The green X indicates the best-case scenario.

In this space visualized in Figure 3, the upper-right boundary represents the theoretical efficiency frontier and thus the maximum achievable improvement through energy storage. The efficiency gap is defined as the shortest distance between the current operating point and this frontier.

Smart meter readings of community members are aggregated into community-level consumption and generation time series. To estimate battery potential, we simulate different battery configurations using a linear optimization model that determines optimal charging and discharging decisions over time [13]. The overall improvement potential is obtained by accumulating daily efficiency gaps over approximately one and a half years. Results indicate that a battery capacity of about 500 kWh already captures most of the achievable efficiency gains, leaving only minor additional improvement potential for larger systems.

Literaturverzeichnis

- [1] Mann, H. B., and D. R. Whitney. "On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other." *The Annals of Mathematical Statistics* 18, no. 1 (1947): 50–60. <http://www.jstor.org/stable/2236101>.

- [2] Wachsenegger, Anahid, Anita Graser, Axel Weißenfeld, and Melitta Dragaschnig. "CrowdSense: Interpretable and Efficient Multivariate Crowd Forecasting with Active Learning." <https://ceur-ws.org/Vol-4064/SENTIS-paper3.pdf> (2025), <https://ceur-ws.org/Vol-4064/SENTIS-paper3.pdf>.
- [3] Scott M. Lundberg and Su-In Lee. 2017. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777, doi: 10.5555/3295222.3295230.
- [4] Beck, T., Kondziella, H., Huard, G., & Bruckner, T. (2017). Optimal operation, configuration and sizing of generation and storage technologies for residential heat pump systems in the spotlight of self-consumption of photovoltaic electricity. *Applied Energy*, 188, 604–619.
- [5] Weniger, J., Tjaden, T., & Quaschnig, V. (2014). Sizing of residential PV battery systems. *Energy Procedia*, 46, 78–87.
- [6] Morales, J. M., Conejo, A. J., Madsen, H., Pinson, P., & Zugno, M. (2014). Integrating Renewables in Electricity Markets: Operational Problems. Springer.
- [7] Luthander, R., Widén, J., Nilsson, D., & Palm, J. (2015). Photovoltaic self-consumption in buildings: A review. *Applied Energy*, 142, 80–94.
- [8] Parra, David & Walker, Gavin & Gillott, Mark. (2014). Modeling of PV generation, battery and hydrogen storage to investigate the benefits of energy storage for single dwelling. *Sustainable Cities and Society*. 10. 1–10. 10.1016/j.scs.2013.04.006.
- [9] Holmgren, W. F., Hansen, C. W., & Mikofski, M. A. (2018). pvlib python: a Python package for modeling solar energy systems. *Journal of Open Source Software*, 3(29), 884.
- [10] Ineichen, P., & Perez, R. (2002). A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*, 73(3), 151–157.
- [11] Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T. C., & Coimbra, C. F. M. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review. *Solar Energy*, 168, 87–111.
- [12] Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. *Nature Reviews Physics*, 3, 422–440.
- [13] Temiz, M., & Javani, N. (2022). Optimization of PV and battery energy storage size in grid-connected microgrid. *Applied Sciences*, 12(16), 8247. <https://doi.org/10.3390/app12168247>
- [14] Ljung, L. (1999). System identification: Theory for the user (2nd ed.). Prentice Hall. doi: 10.1109/MRA.2012.2192817.