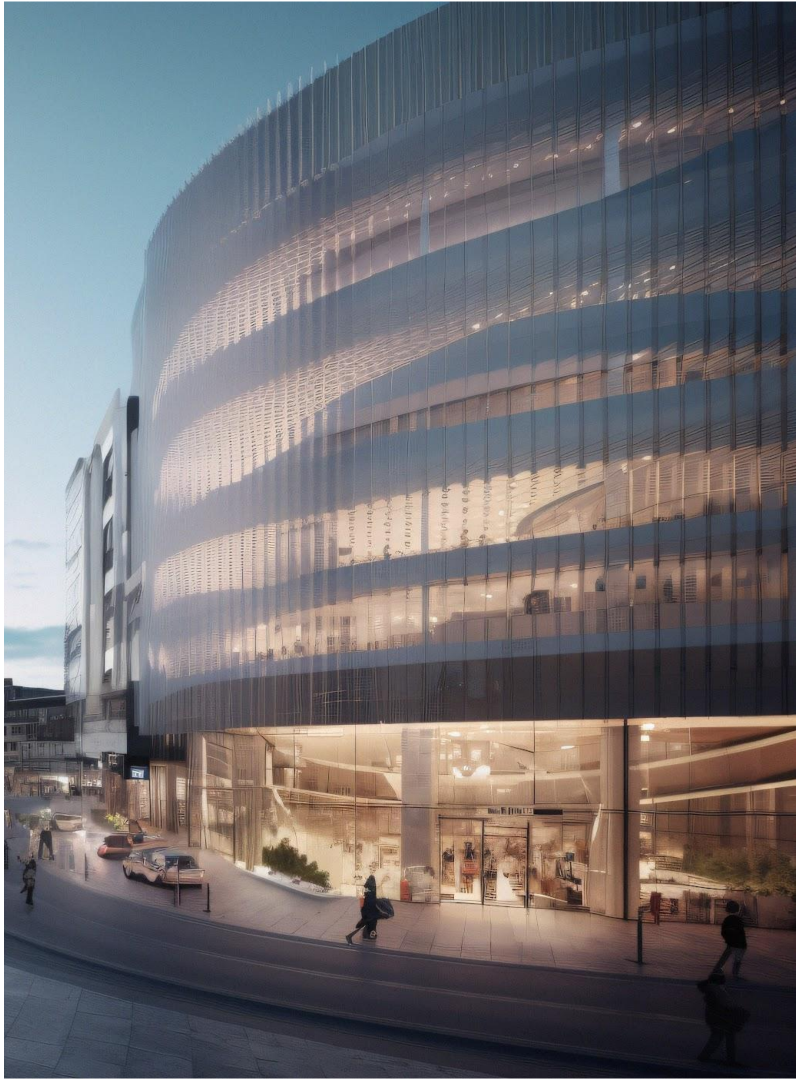
AI-Powered Measurement & Verification: Building Interpretable Counterfactual Models to Verify Energy Savings in Buildings

Dr. Benedetto Grillone
Lead AI Engineer at **Ento**
Author of **Reimagine Energy**

December 7, 2025
Tackling Climate Change with Machine Learning Workshop @ NeurIPS 2025



Intro

→ Benedetto Grillone, Lead AI Engineer @ **Ento**

→ Background in energy engineering

→ Ph.D. in Machine Learning applied to M&V

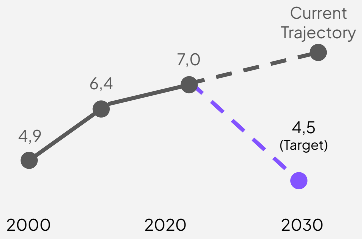
→ Main area of interest: AI + Energy

→ Reimagine Energy newsletter

A perfect storm is hitting the global buildings space

Emissions from buildings are massive and still rising rapidly

Accounting for 26% of global CO_2 emissions, energy in buildings is one of the few areas that has yet to improve.

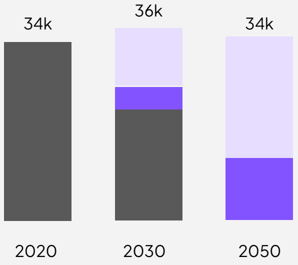


Energy-related emissions in buildings in $GtCO_2$

- What Happened
- What Needs to Happen

Energy efficiency is the single largest contributor to reaching climate goals

Immediate CO_2 cuts are essential to achieve net zero and limit global warming, and energy efficiency can deliver 1/3 of this.

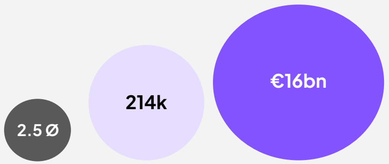


Net-zero roadmap in $GtCO_2$

- NZE Demand
- Avoided by Renewables and everything else
- Avoided by Energy Efficiency

But building owners lack budget, time and resources to adapt quickly

In EU municipalities alone, 214,000 new jobs will be needed over the next decade to transform the building stock.



Additional Costs and Labor Force in EU Countries from 2022-2030

- Additional FTE per Municipality per Year
- New Total Positions
- Extra Staff Costs for Municipalities Per Year

And are exposed to regulatory pressure and societal scrutiny

Energy management is now required, with ambitious targets depleting resources and risking reputational damage if unmet.

11.7%
reduction in EU energy consumption targeted by 2030.

1.3%
annual energy savings required by building owners, increasing to 1.9% by 2029.

ISO certifications
for energy or environmental management and climate assessments with energy review if above 85 TJ

Sources: International Energy Agency, Energy Cities, NYC.gov

Major barriers are causing a slow and ineffective optimization process

Costly and time-consuming to collect and manage data

Difficult to identify and plan energy optimization projects

Lack of resources for operating buildings and implementing projects

Limited follow-up and documentation of savings and reductions

Rate of reductions too slow and huge untapped savings potential!

Major barriers are causing a slow and ineffective optimization process

Costly and time-consuming to collect and manage data

Difficult to identify and plan energy optimization projects

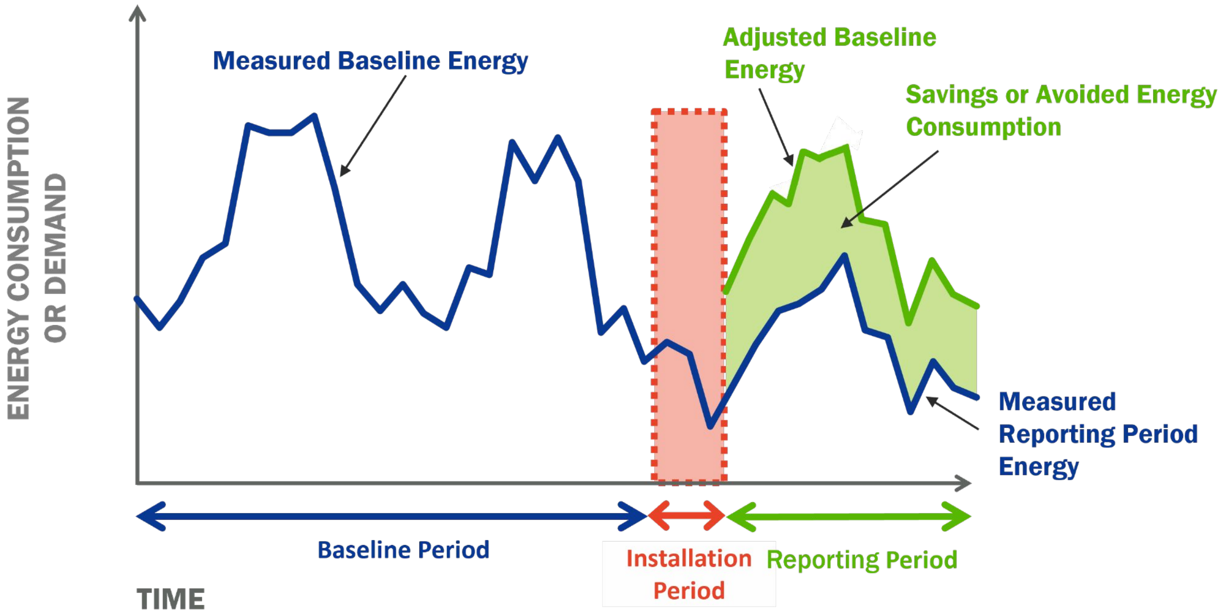
Lack of resources for operating buildings and implementing projects

Limited follow-up and documentation of savings and reductions

Rate of reductions too slow and huge untapped savings potential!

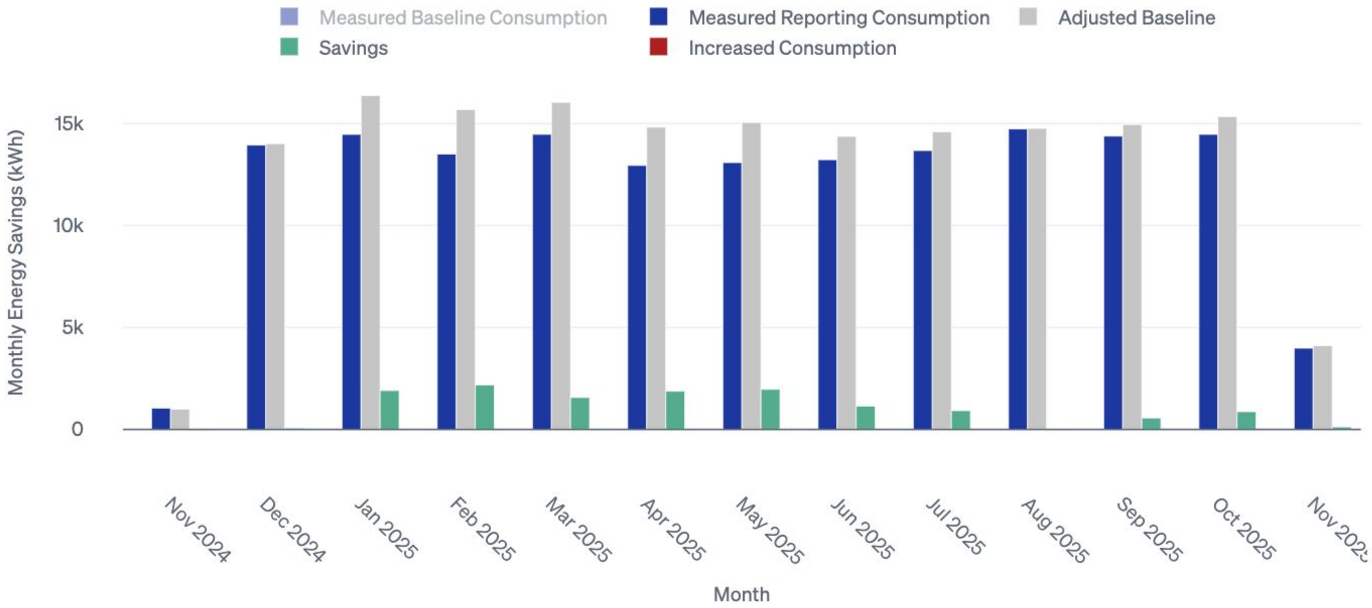
Measurement & Verification

The main goal of M&V is to create consensus around an unmeasurable number.



Source: EVO, International Performance Measurement and Verification Protocol

Historically a manual process using monthly billing data…



… today can be easily automated with ML

The M&V workflow at a glance

1. Define periods

Set baseline, blackout, and reporting windows capturing normal operations.

2. Build baseline

Train a driver-based model of pre-project "business-as-usual" use.

3. Detect non-routine events

Flag operational changes unrelated to the project and encode adjustments.

4. Predict counterfactual

Apply the baseline to post-period drivers to estimate no-project use.

5. Quantify uncertainty

Propagate model error to savings and report confidence intervals.

6. Validate the model

Check fit metrics, residuals, and backtests against guidelines and physics.

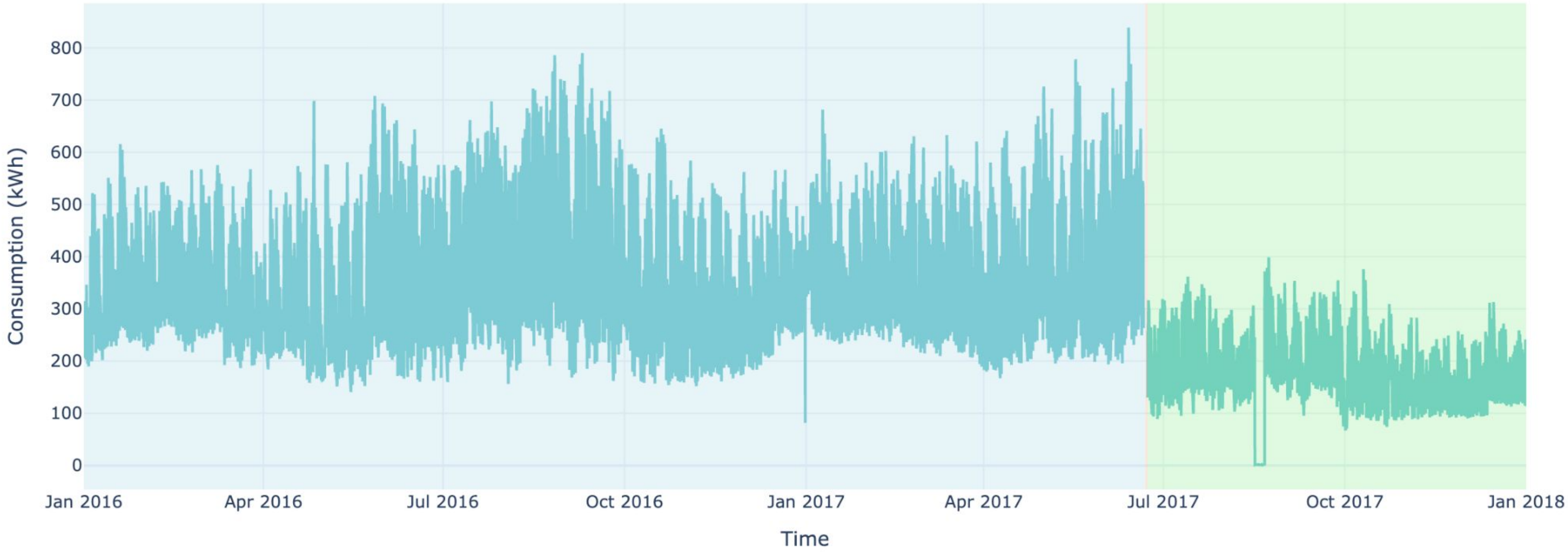
Dataset used in the analysis

BUILDING DATA GENOME PROJECT 2

```
building-data-genome-project-2
├─ README.md              <- BDG2 README for developers using this data-set
└─ data
|    ├─metadata           <- buildings metadata
|    ├─ weather           <- weather data
|    └─ meters
|         └─ raw          <- all meter reading datasets
|         └─ cleaned      <- cleaned meter data based on several filtering steps
|         └─ kaggle       <- the 2017 meter data that aligns with the Kaggle competition
├─ notebooks              <- Jupyter notebooks, named after the naming convention
└─ figures                <- figures created during exploration of BDG 2.0 Data-set
```

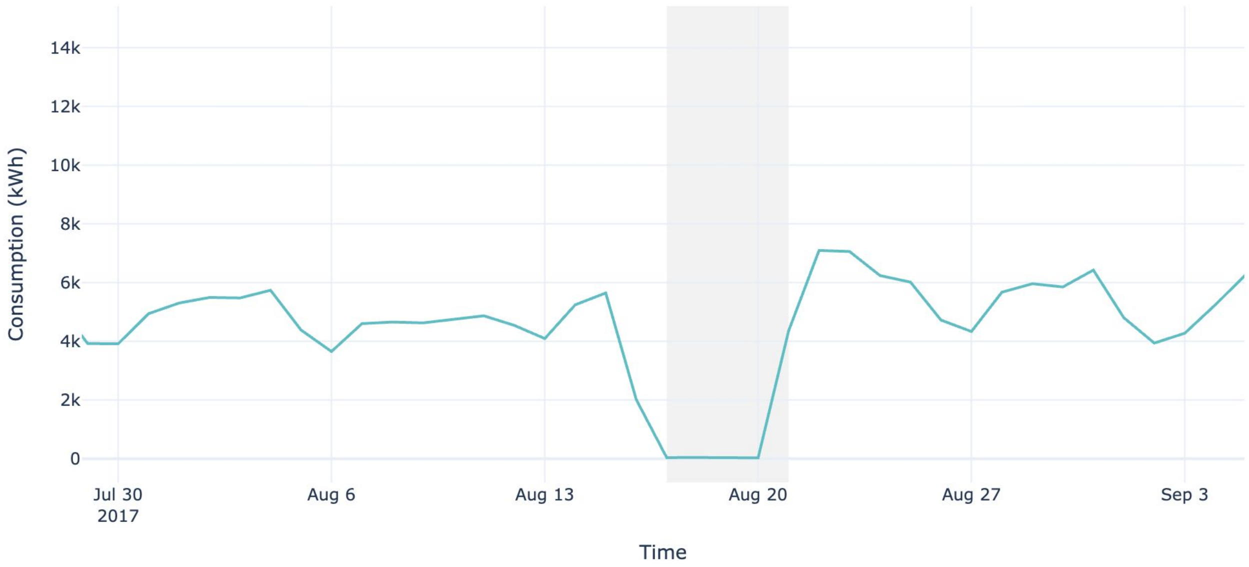
https://github.com/buds-lab/building-data-genome-project-2

Define periods for the chosen building



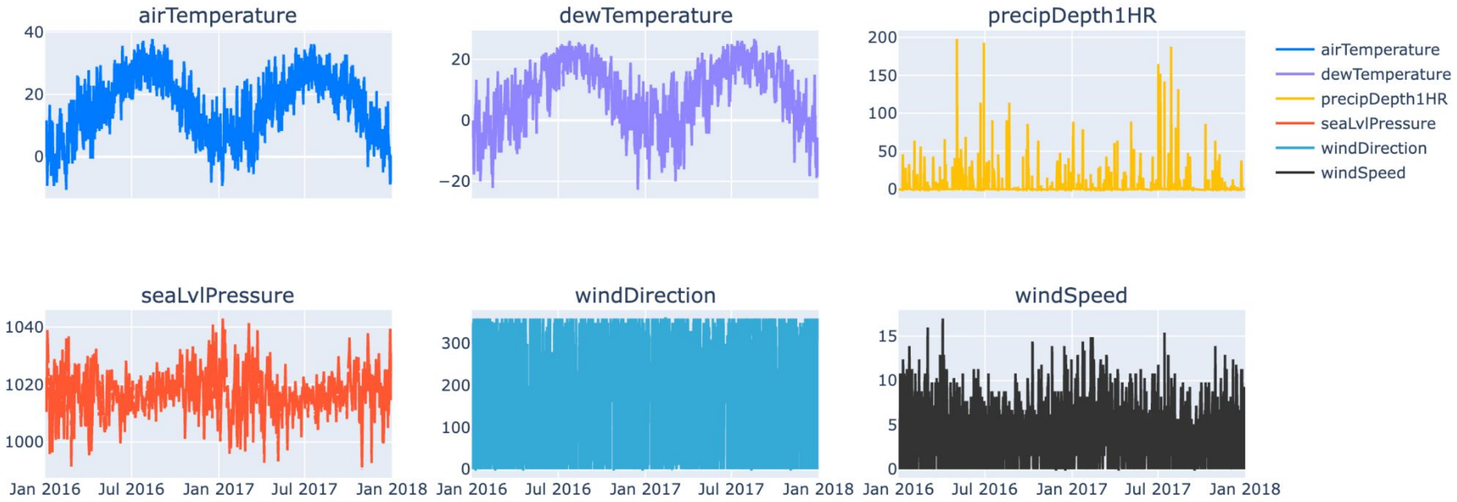
- → Blue: Baseline period
- → Pink: Installation period
- → Green: Reporting period

Detect and exclude Non-Routine Events



→ Exclude from the analysis periods with abnormal consumption

Validate weather data and select features



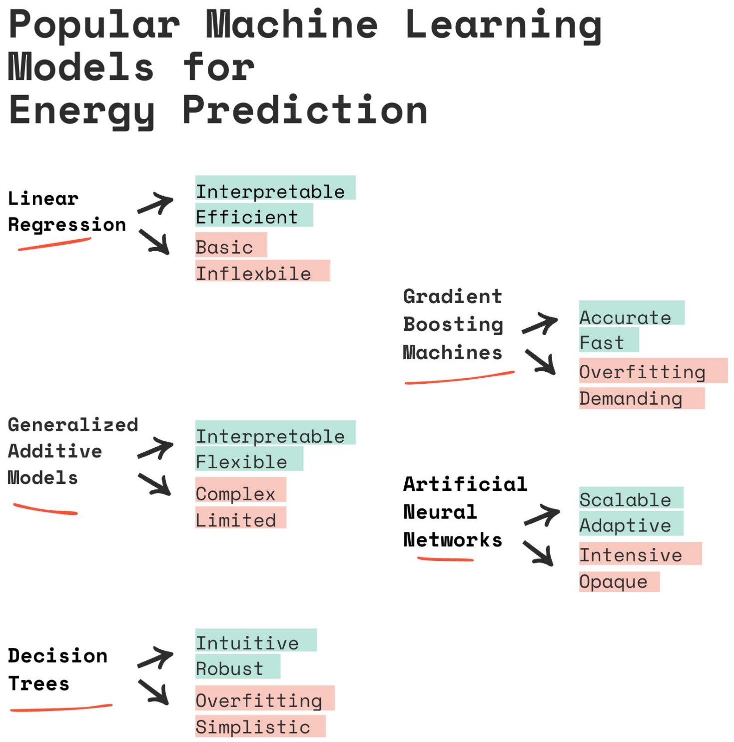
→ Commercial buildings' consumption is mainly driven by **weather** and **occupancy**

→ Weather data is available from the repo

→ Occupancy in commercial buildings is highly correlated with "**calendar features**"

→ In this case we will use "hour of day", "day of week", and "week of year"

Build an energy prediction model (LightGBM)

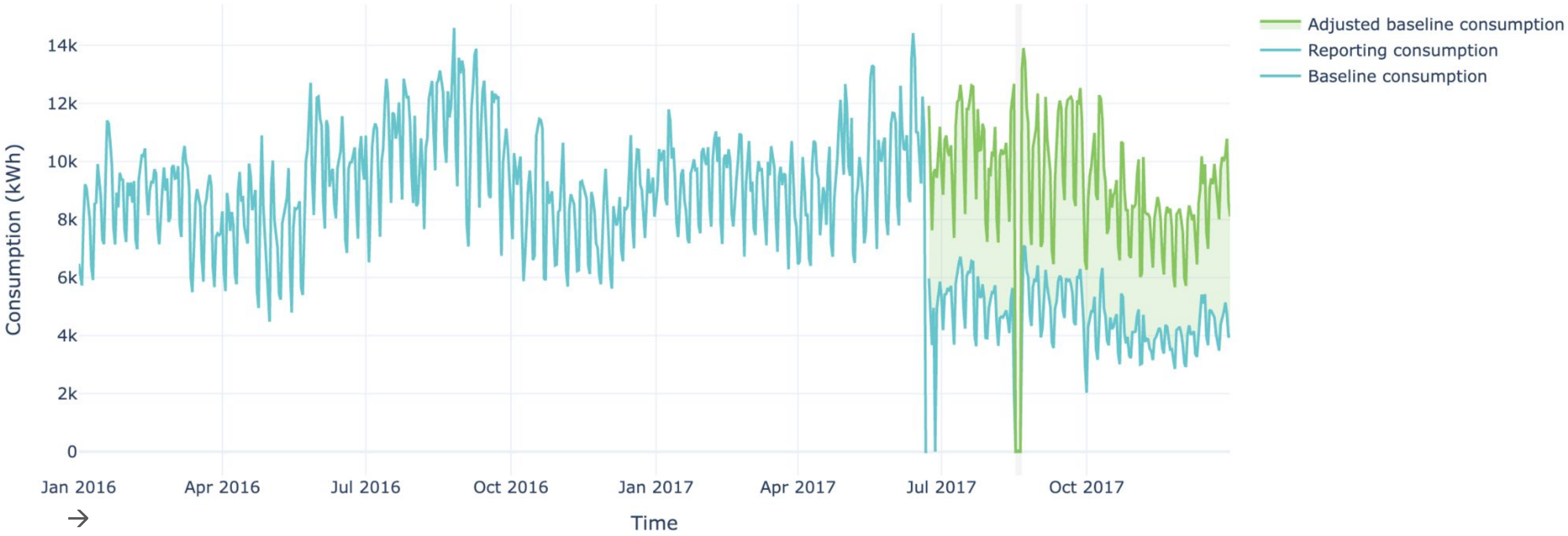


LightGBM strengths:

- → Strong on tabular time-series features
- → Captures non-linearities and interactions
- → Fast and stable with CV and early stopping

Source: https://www.reimagine-energy.ai/p/data-driven-efficiency-predicting

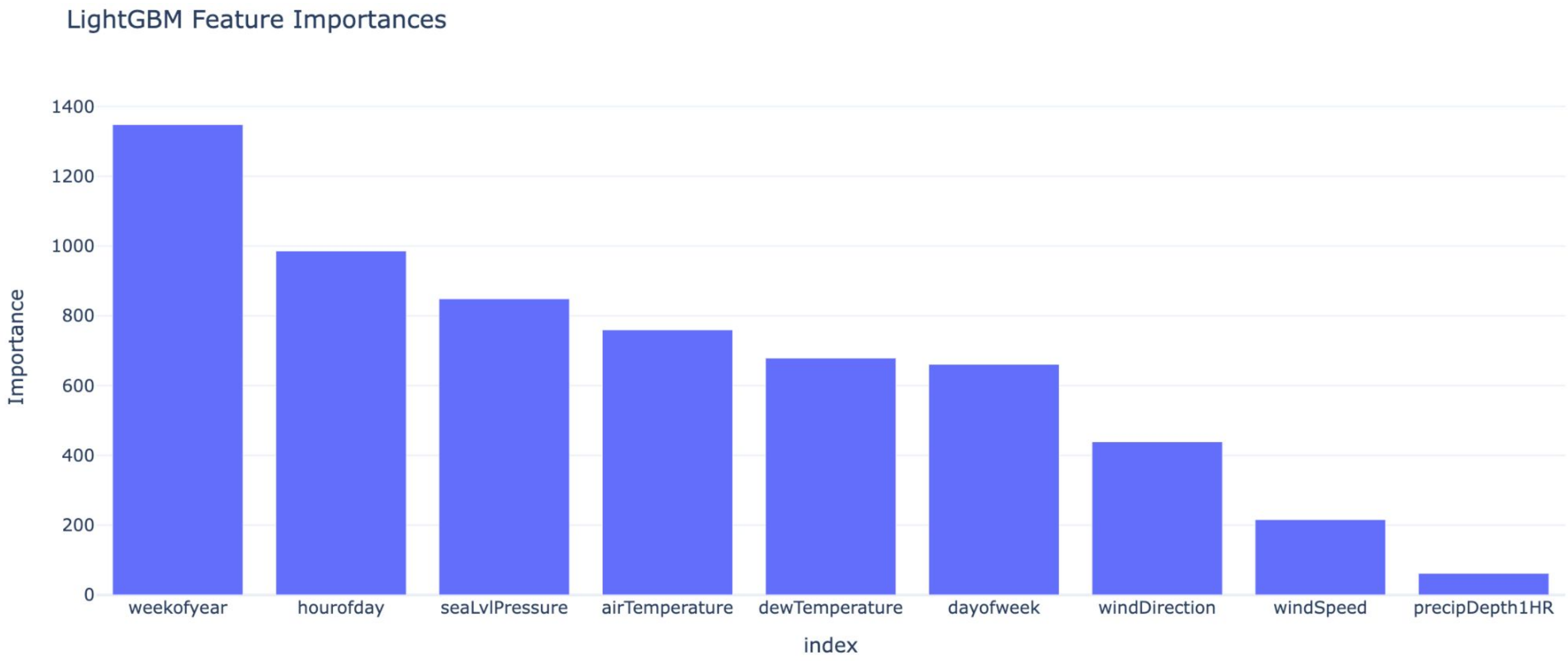
Use the counterfactual energy consumption model to estimate savings



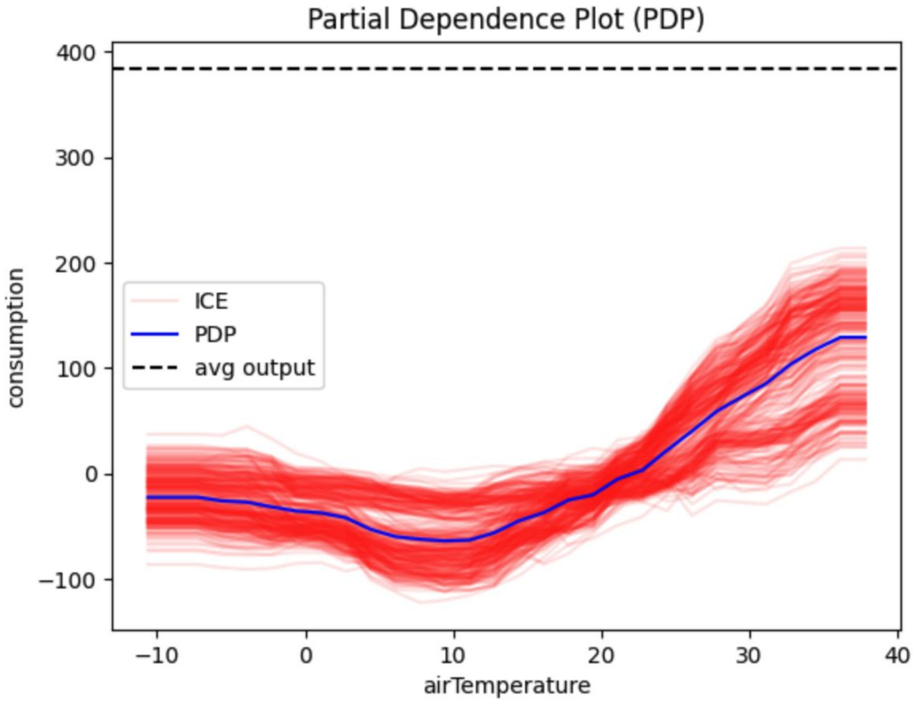
The savings from the energy efficiency project are equal to the difference between the adjusted baseline (counterfactual) consumption, and the metered consumption in the reporting period.

→ Total savings: **890 MWh**
→ Savings uncertainty: +/- **24.4 MWh**, 3% of the estimated savings (95% confidence)

Model interpretability: feature importances

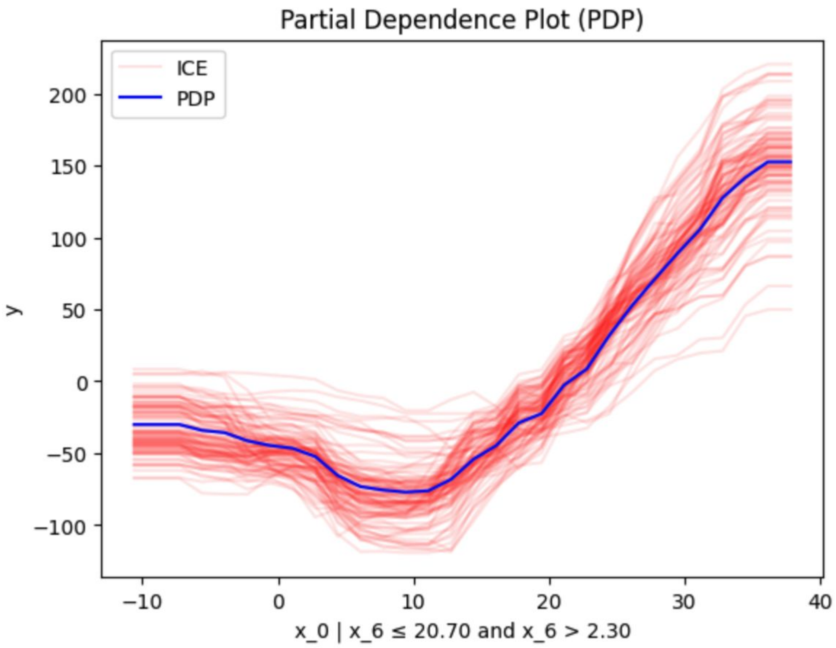


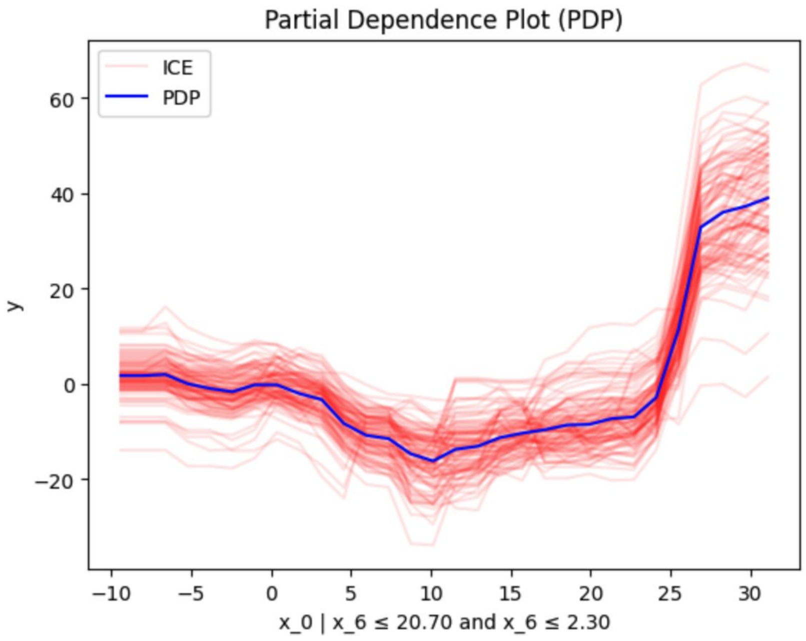
Model interpretability: PDP + ICE



→ **Global effect** plot (temperature)

Model interpretability: PDP + ICE





- → **Regional effect** plot (temperature varying by hour of the day splits)
- → During the **day** (occupied): crisp cooling slope after ~15–20 °C
- → During the **night** (unoccupied): weak effect, late kick-in

Real-world application

→ Methodology used daily to verify savings across **60,000 buildings** in Europe

→ Used to verify savings equivalent to more than **12,000 tCO2e** each year

Continue the conversation

- Try this yourself on Google Colab
- Find out more about **Ento**: ento.ai
- **Reimagine Energy** newsletter: reimagine-energy.ai
- Find me on LinkedIn: linkedin.com/in/benedetto-grillone/
- Reach out: benedetto@ento.ai

Thank you!

Found this presentation interesting? Let's connect