GridVue

Smarter Grids, Greener Futures

Final Presentation | MIDS Capstone | Spring 2024

Jammy Chan, Jun Qian, Maiya Caldwell, Omar Al-Darwish

Berkeley
UNIVERSITY OF CALIFORNIA

The World Has Decided

However, Our Systems Are Not Ready

AP News
https://apnews.com › article › renewable-energy-climate...

Electrical grids aren't keeping up with the green energy ...

Oct 16, 2023 — Stalled spending on **electrical grids** worldwide is slowing the rollout of **renewable** energy and could put efforts to limit climate change at risk if millions ...

IEA – International Energy Agency
https://www.iea.org › news › lack-of-ambition-and-atte...

Lack of ambition and attention risks making electricity grids ...

Oct 17, 2023 — Lack of ambition and attention risks making **electricity grids** the weak link in clean energy transitions - News from the International Energy Agency.

The New York Times
https://www.nytimes.com › 2023/06/12 › climate › us-e...

Why the U.S. Electric Grid Isn't Ready for the Energy ...

Jun 12, 2023 — Already, a lack of transmission capacity means that thousands of proposed wind and solar projects are facing multiyear delays and rising costs to connect to the ...

Forbes
https://www.forbes.com › Innovation › Sustainability

Will Power Grids Keep Pace With Renewable Growth?

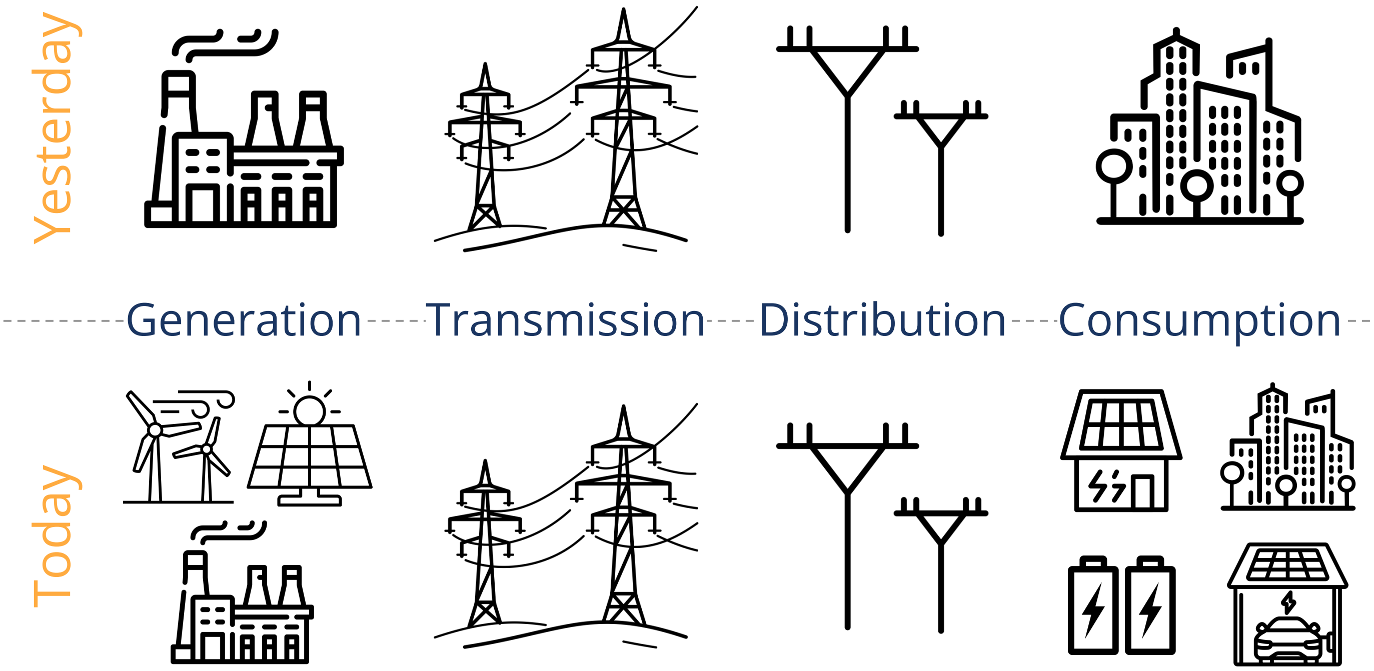
Jan 26, 2024 — The US **power grid**, for example, needs to be updated for transmitting modern **renewables**. Over 930 gigawatts of **renewable** energy, vital for an 80% **renewable** share ...

Reuters
https://www.reuters.com › business › energy › electric-...

Electric grids need major upgrades to aid global energy ...

Mar 30, 2023 — New transmission and distribution networks must also be able to accommodate sudden changes to **power** loads generated by **renewable** farms caused by reduced ...

Especially for local distribution grids ...

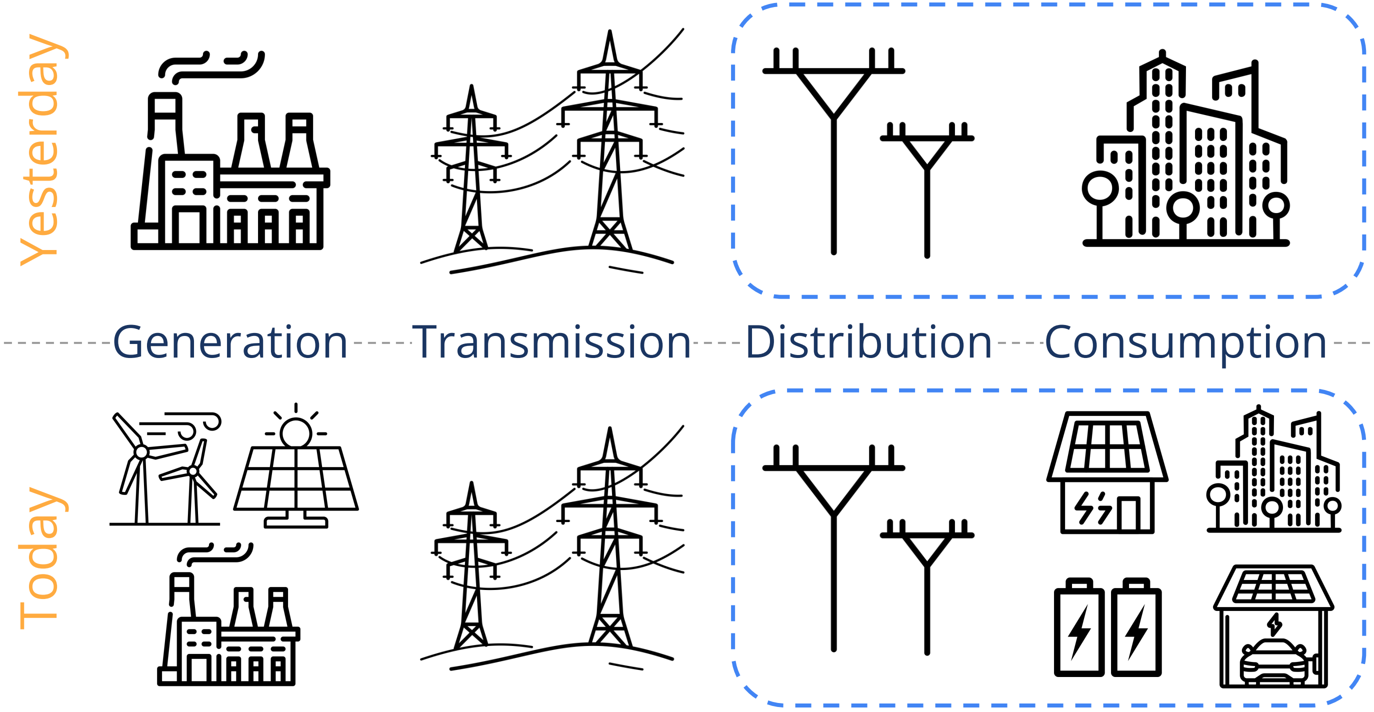


- Centralized Generation
- Low Renewable Share
- Unidirectional power flow
- Deterministic Control

- Distributed Generation
- High Renewable Share
- Bidirectional power flow
- Stochastic Control

Icons: thenounproject.com

Especially for local distribution grids ...

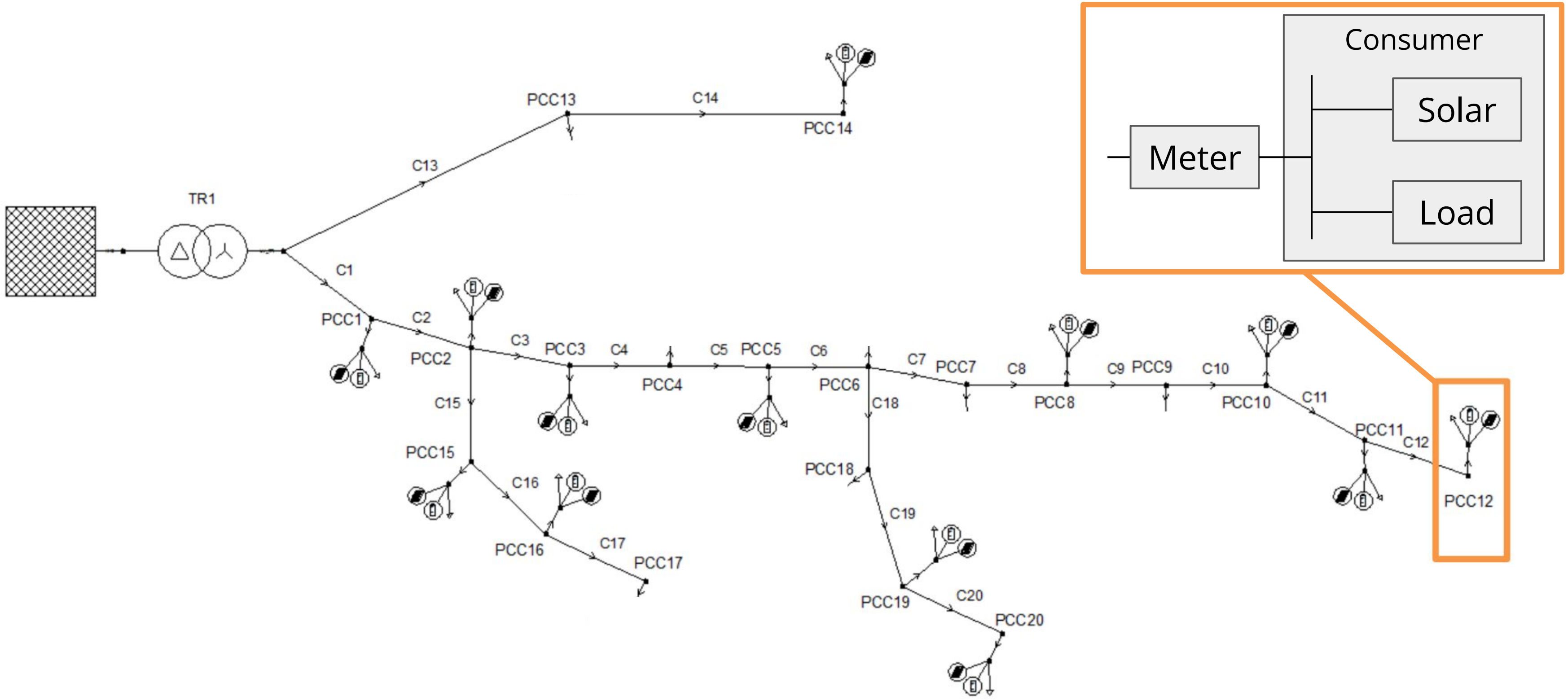


- Power from A to B
- Capacity Planning
- Network Maintenance
- System Protection
- Event Management

- ... all the above ...
- System Balancing
- Congestion Management
- Voltage Support
- Network Resilience

Icons: thenounproject.com

The Problem: Distributed and Hidden



Adapted from: Resch, Matthias & Ramadhani et al. (2015). Comparison of control strategies of residential PV storage systems. 10.13140/RG.2.1.3668.2084

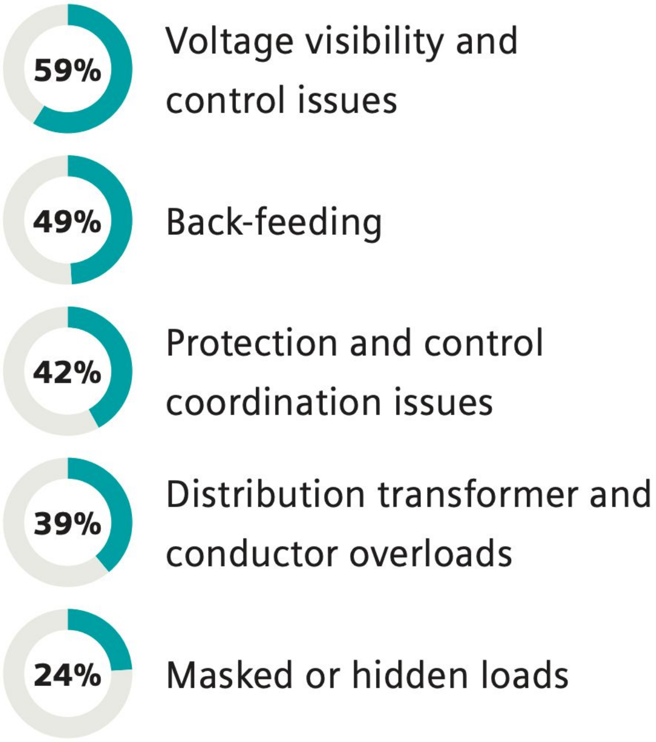
Utilities are already recognizing the problem …

“Nearly three-quarters of utilities say customer adoption of behind-the-meter DERs creates operational challenges.”

“Limited visibility and understanding of DER behavior creates operational challenges and impacts grid performance.”

“ (Traditional) Solutions Exist … data shows adoption to be slow.”

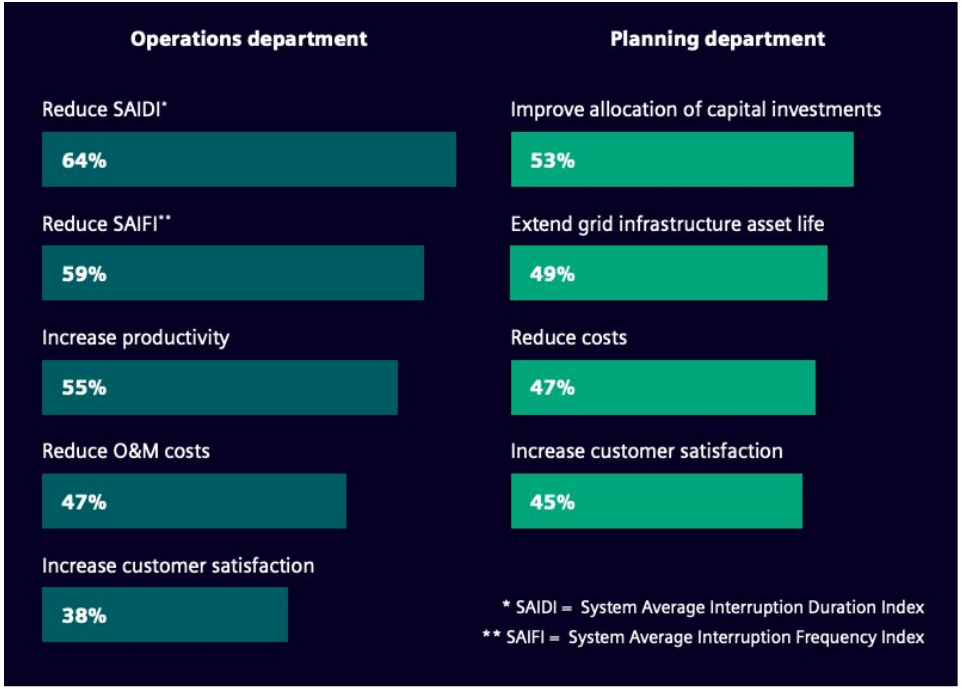
DERs: Distributed Energy Resources



Oxford Economics and Siemens, Seeing Behind the Meter: How Electric Utilities are Adapting to the Surve in Distributed Energy Resources (2024).

And the potential benefits of solving it …

“How would visibility into behind-the-meter DERs impact the following departments?”



Oxford Economics and Siemens, Seeing Behind the Meter: How Electric Utilities are Adapting to the Surve in Distributed Energy Resources (2024).

Addressing the problem using existing infrastructure …

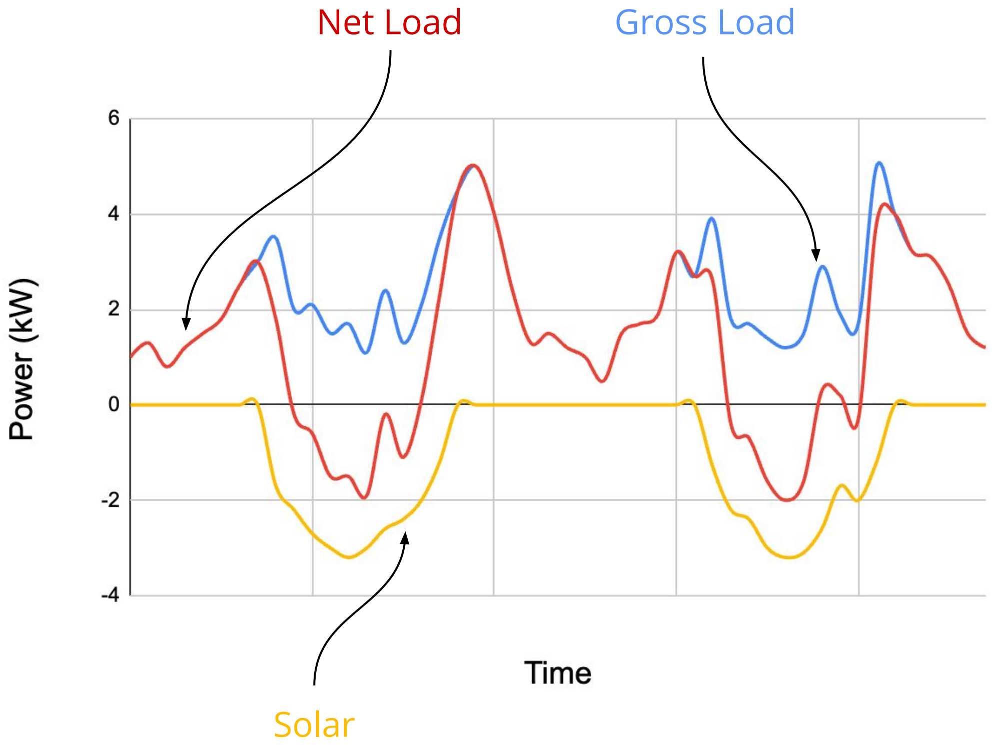
"In 2021, U.S. electric utilities had about 119 million advanced (*smart*) metering infrastructure (AMI) installations, equal to about 72% of total electric meters installations." - EIA

Number and percentage share of AMI installations by sector, 2021				
Residential	Commercial	Industrial	Transportation	Total
104,237,855 (73%)	13,908,481 (69%)	574,726 (68%)	1,879 (55%)	118,722,741 (72%)

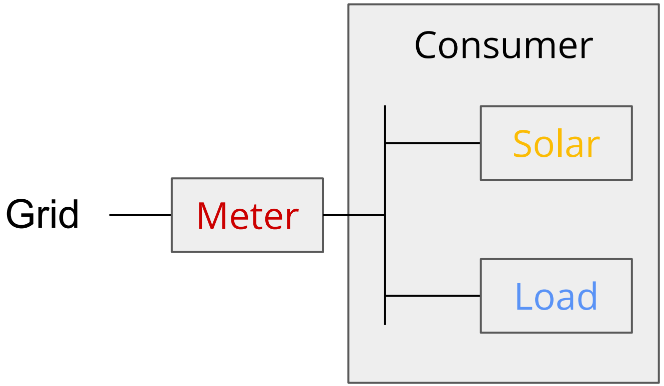
Current Focus

Energy Information Administration, How many smart meters are installed in the United States, and who has them? (2022). https://www.eia.gov/tools/faqs/faq.php?id=108&t=3

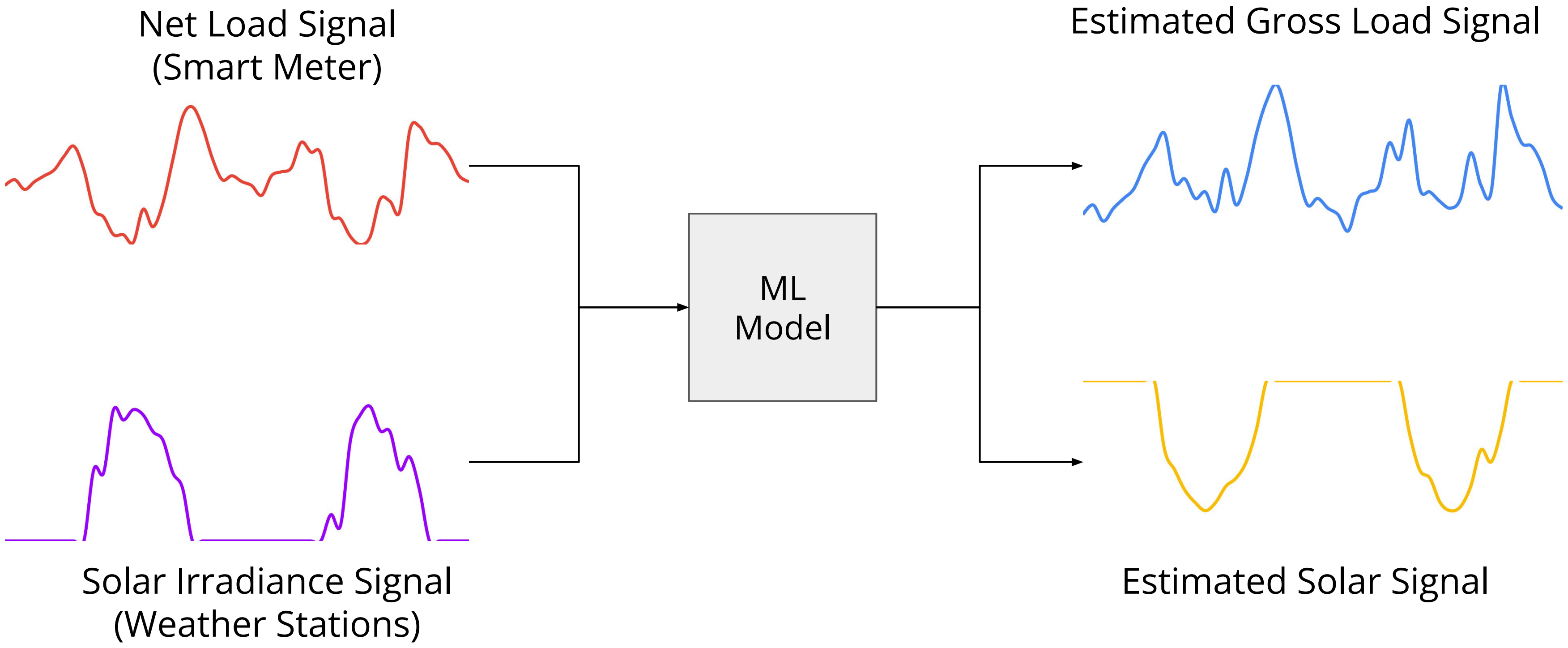
Addressing the problem using existing infrastructure ...



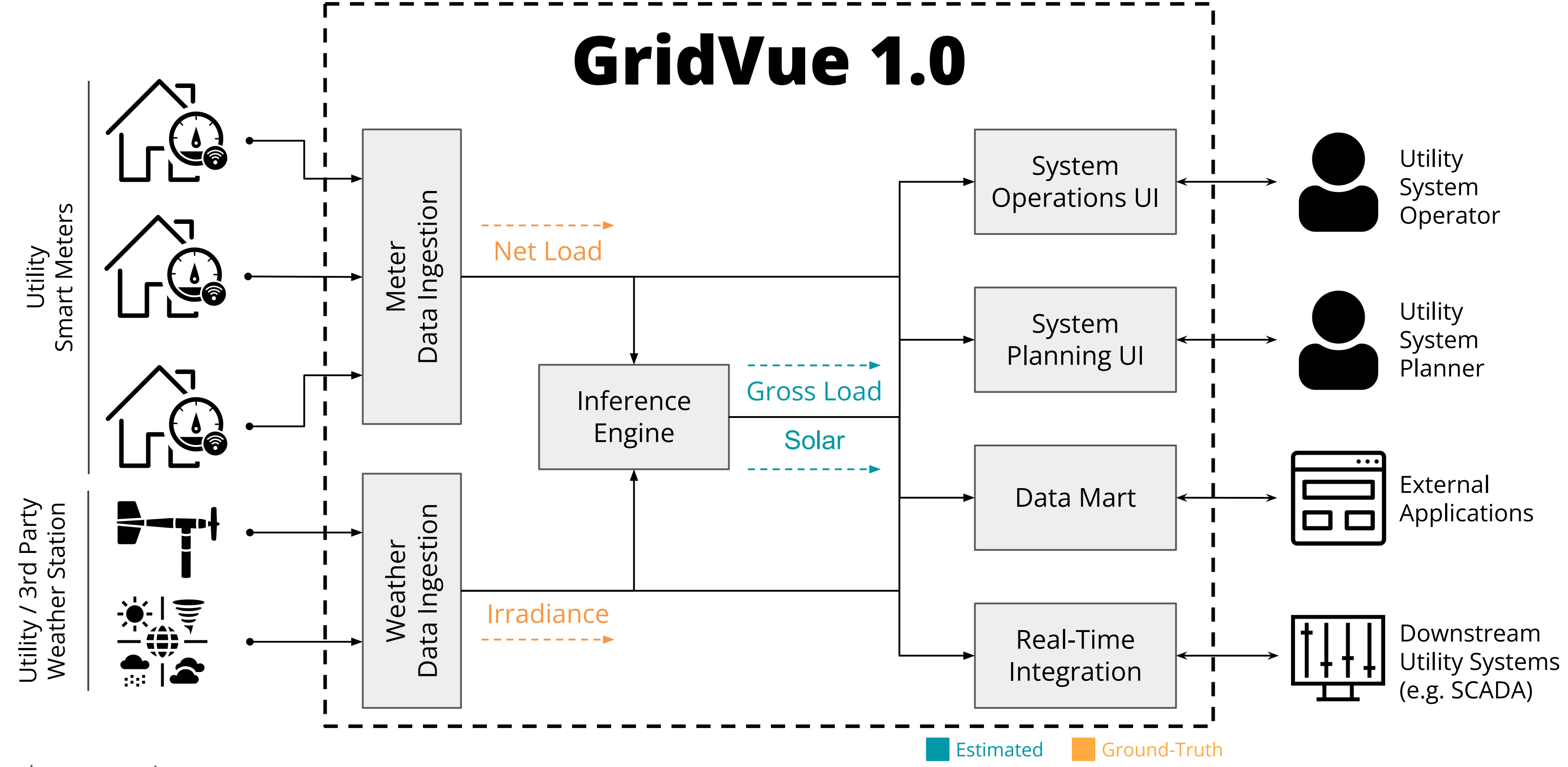
Net-Load = Gross-Load + Solar



Addressing the problem using existing infrastructure ...



Product



Icons: thenounproject.com

Data Sources

Training Data

Simulated
10k houses (1-year)

ResStock

NREL
NATIONAL RENEWABLE ENERGY LABORATORY

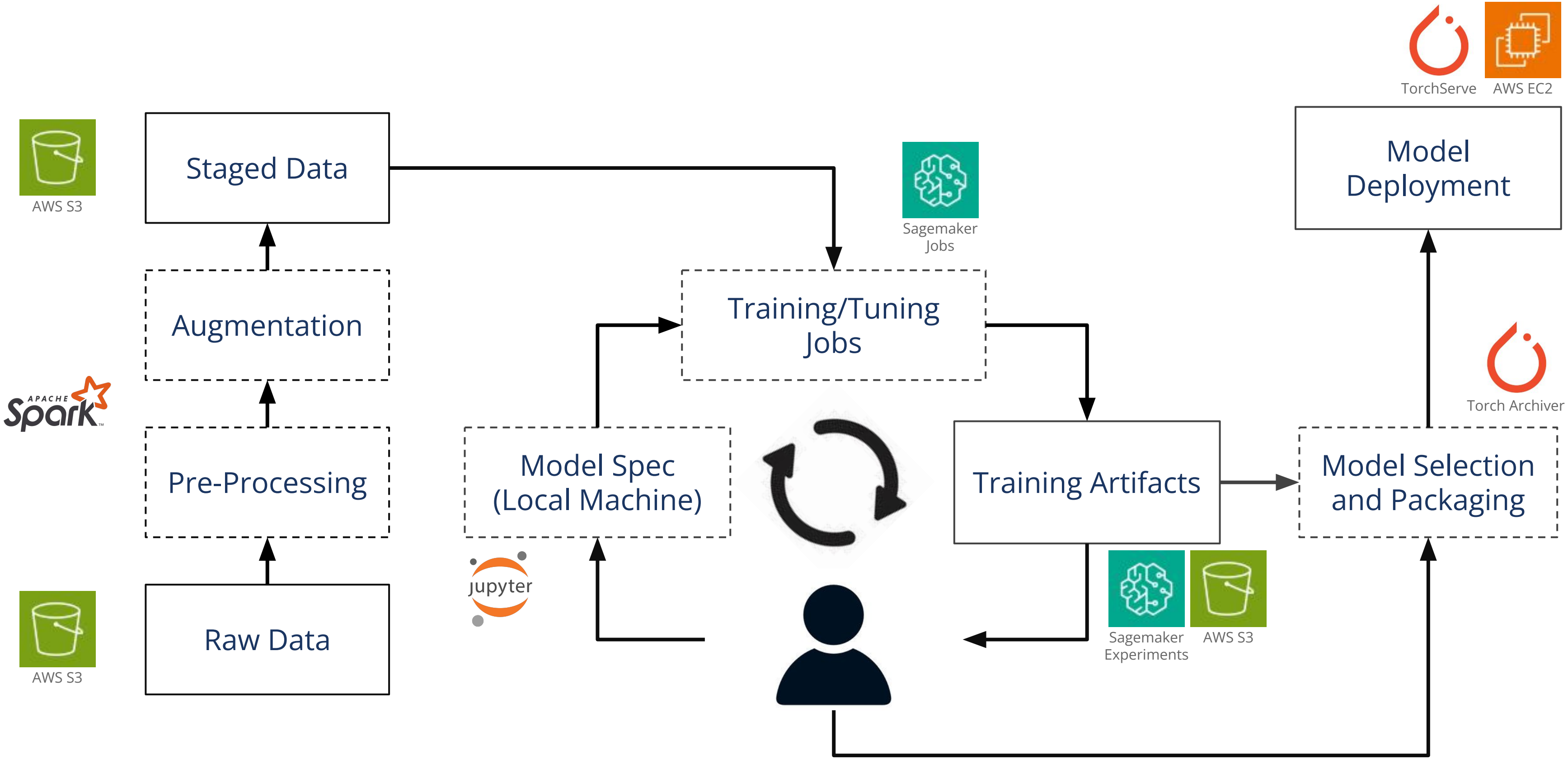
Test Data

Measured
68 houses (mixed)

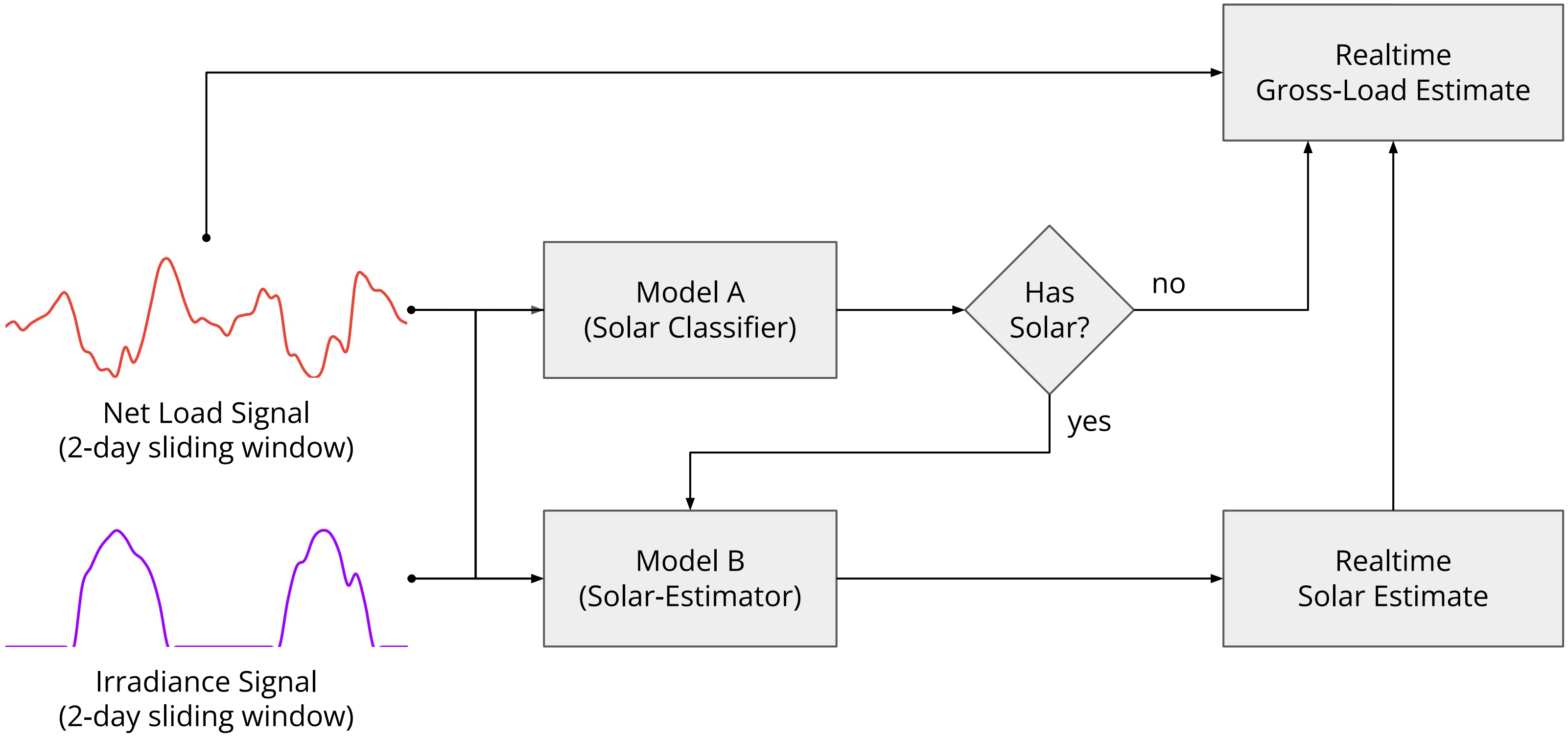
PECAN STREET

SOLCAST

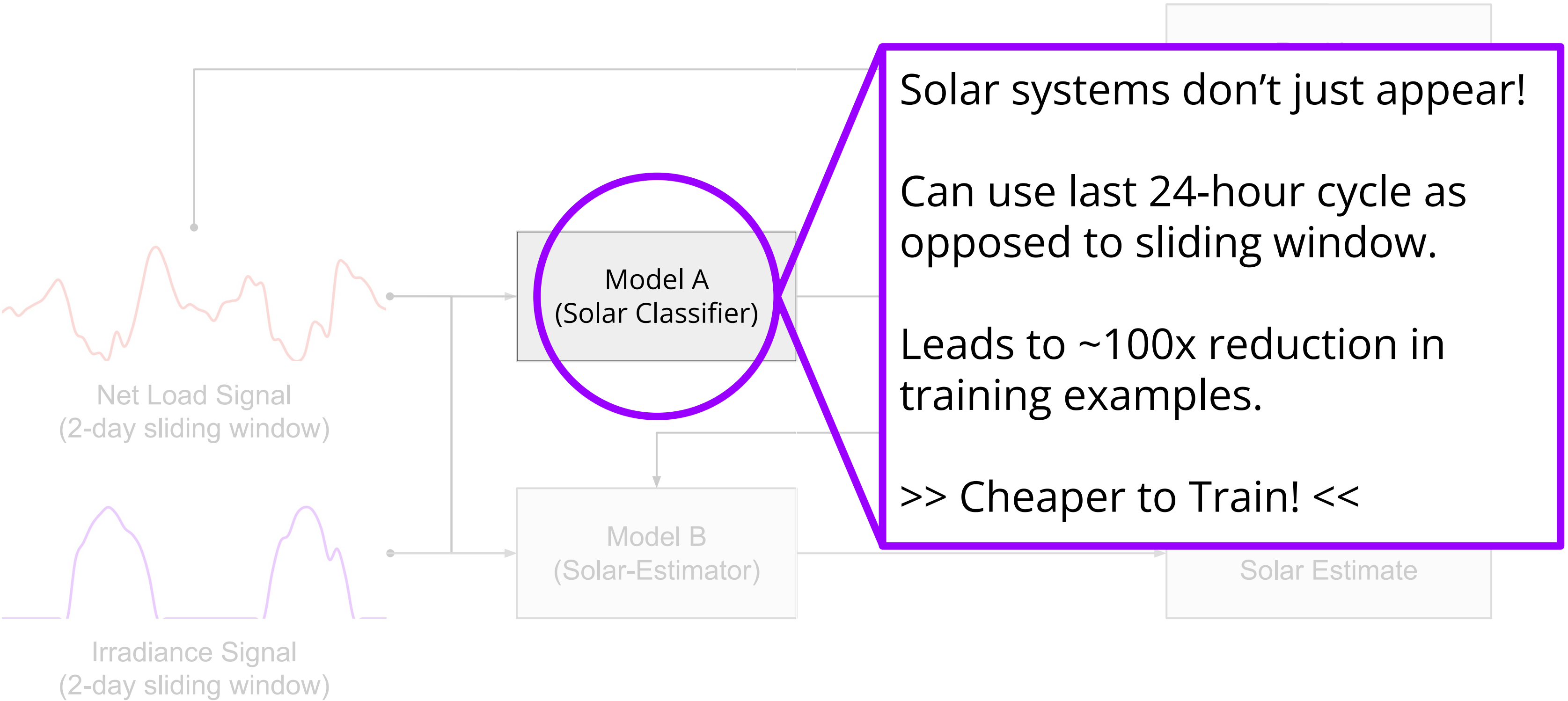
MLOps: Overall Setup



Multi-Model Approach



Multi-Model Approach: Why?



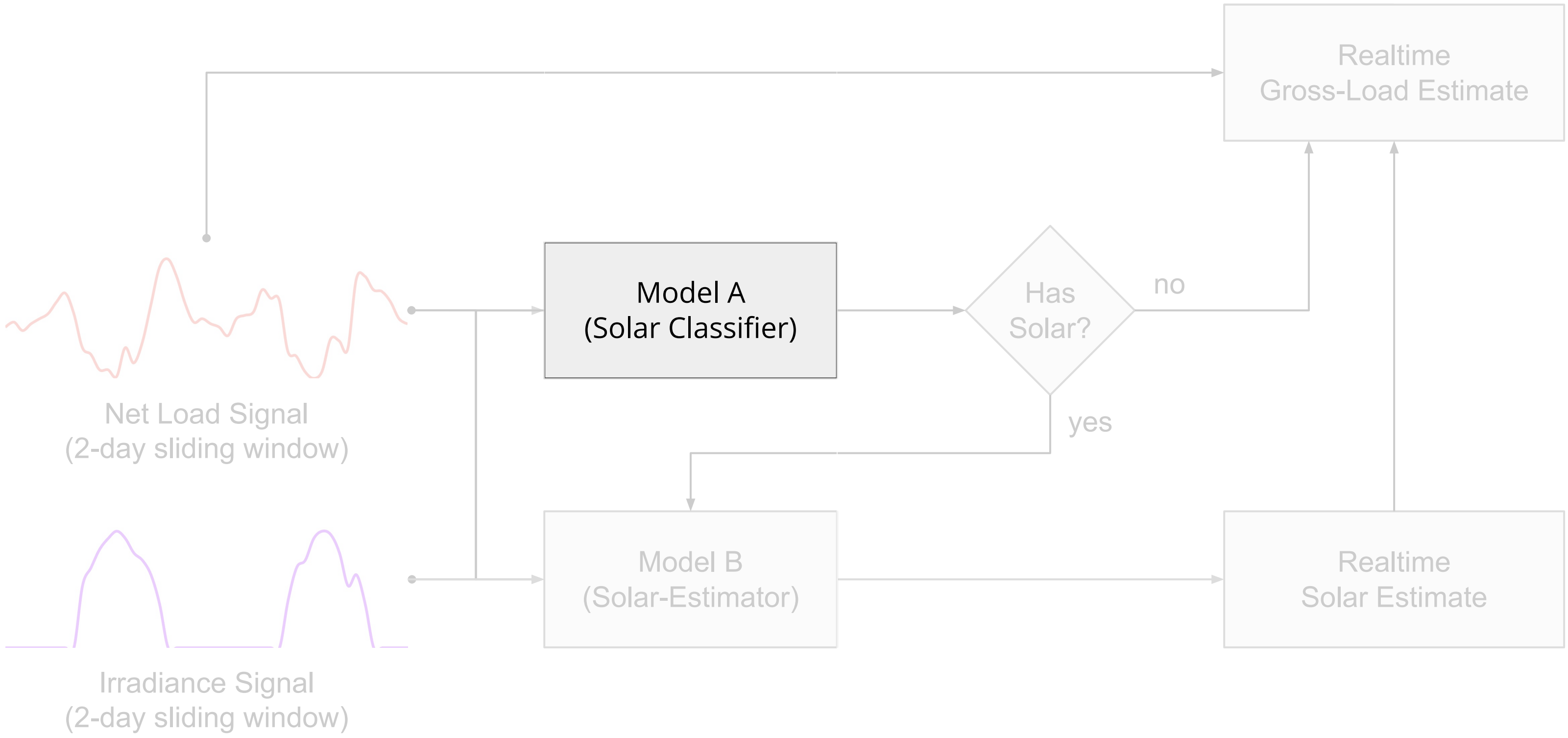
Solar systems don’t just appear!

Can use last 24-hour cycle as opposed to sliding window.

Leads to ~100x reduction in training examples.

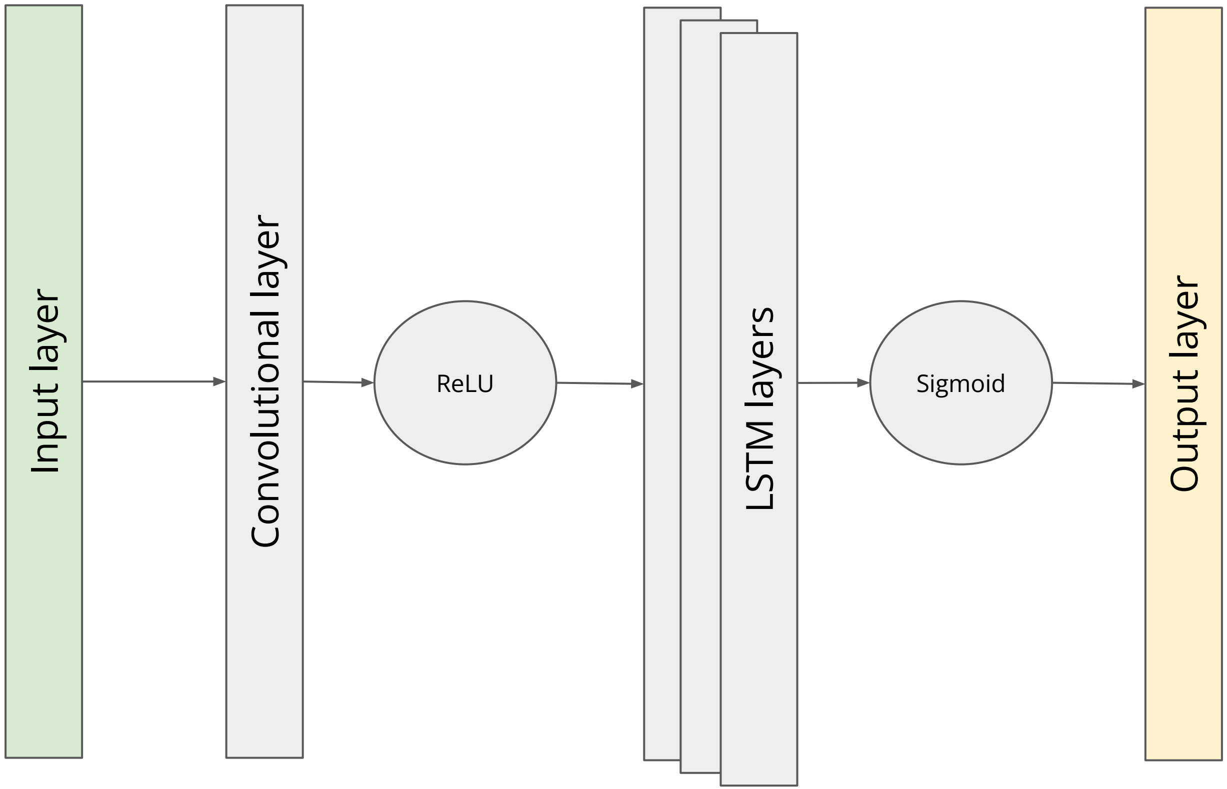
>> Cheaper to Train! <<

Models: Solar Classifier



Models: Solar Classifier | Architecture

- **Feature**: net electrical load ('x__net')
  - Window size: 96 (15-min increments) = 1 day
- **Label**: 0 or 1, house generates PV or not ('y__pv')
- **Hyperparameters**:
  - Batch size: 512
  - Learning rate: 0.005
  - Number of LSTM layers: 3
  - Hidden units: 64
  - Bidirectional: true

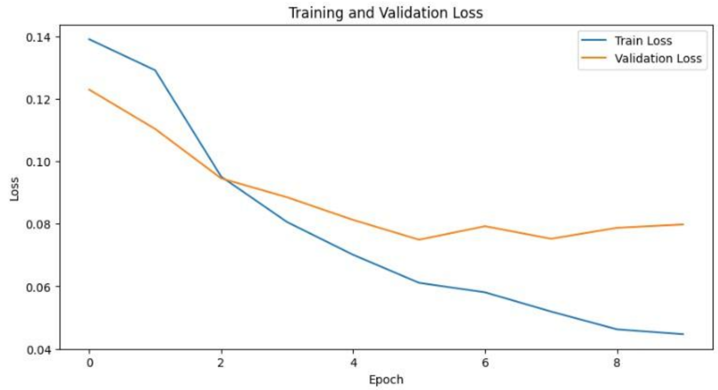


Models: Solar Classifier | Training and Performance

Validation Classification Report

ResStock data

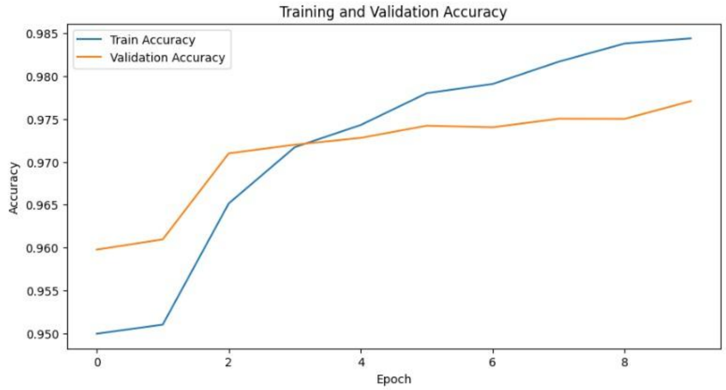
	precision	recall	f1-score	support
0.0	0.95	0.99	0.97	1839600
1.0	0.99	0.95	0.97	1839600
accuracy			0.97	3679200
macro avg	0.97	0.97	0.97	3679200
weighted avg	0.97	0.97	0.97	3679200



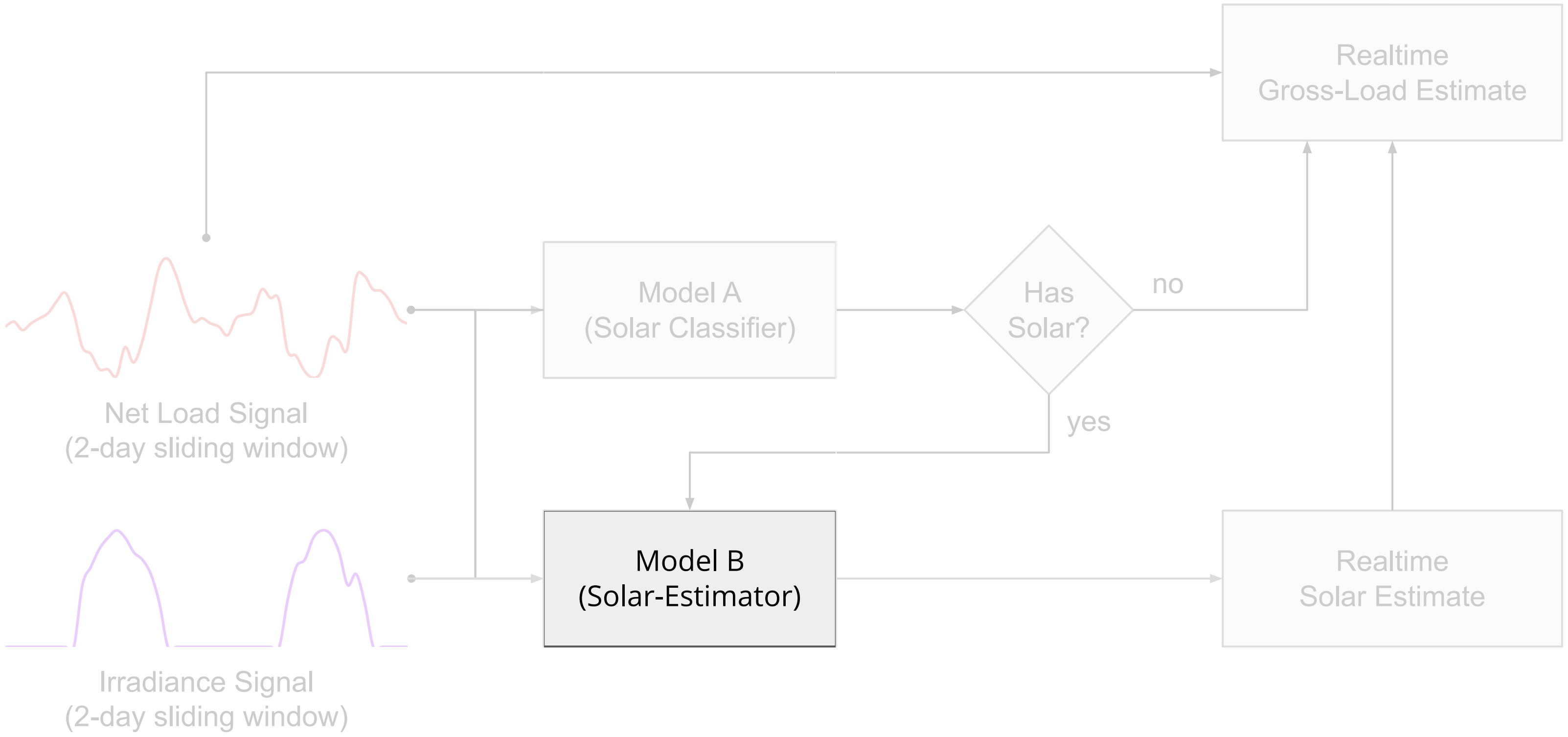
Test Classification Report

Pecan St data

	precision	recall	f1-score	support
0.0	0.74	0.93	0.83	98450
1.0	0.91	0.68	0.78	98450
accuracy			0.81	196900
macro avg	0.83	0.81	0.80	196900
weighted avg	0.83	0.81	0.80	196900

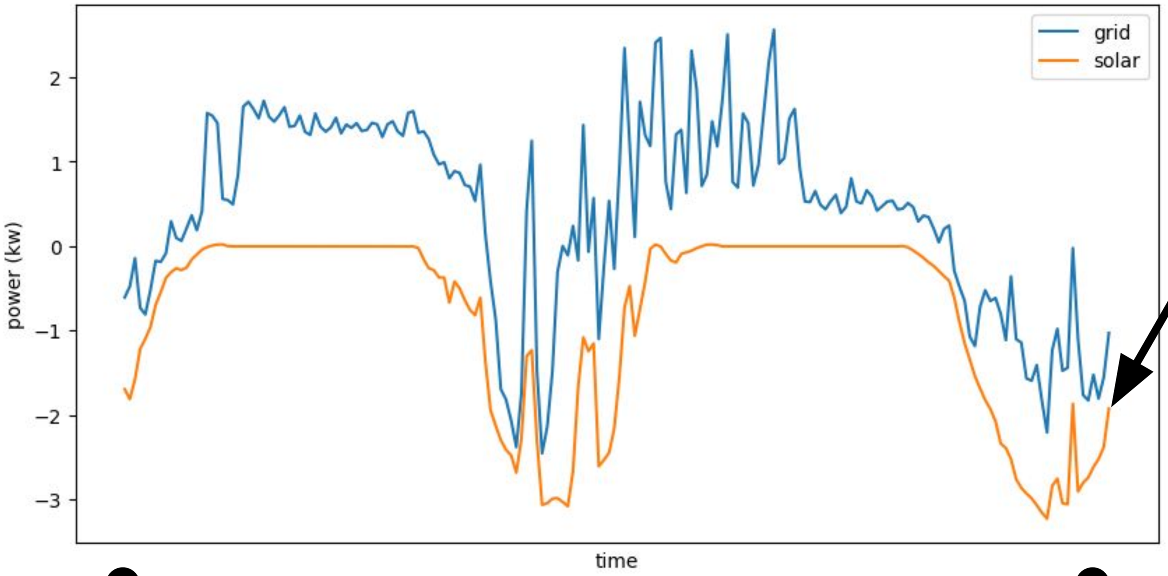


Models: Solar Estimator



Models: Solar Estimator | Function

Seq To Point

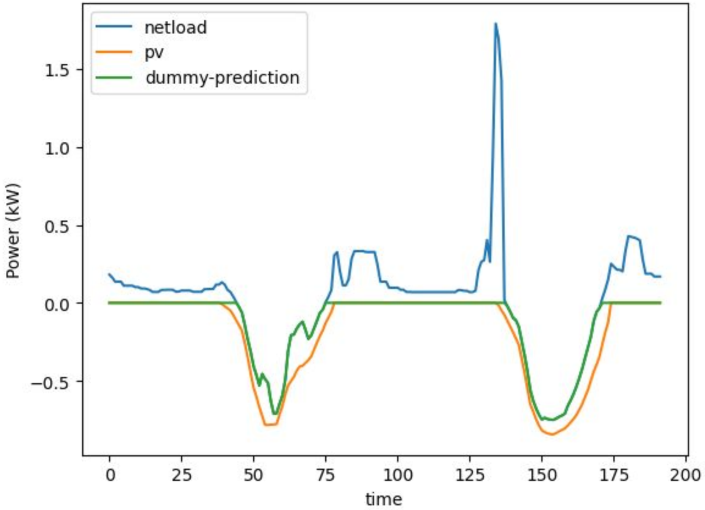


Predict the Solar output at T

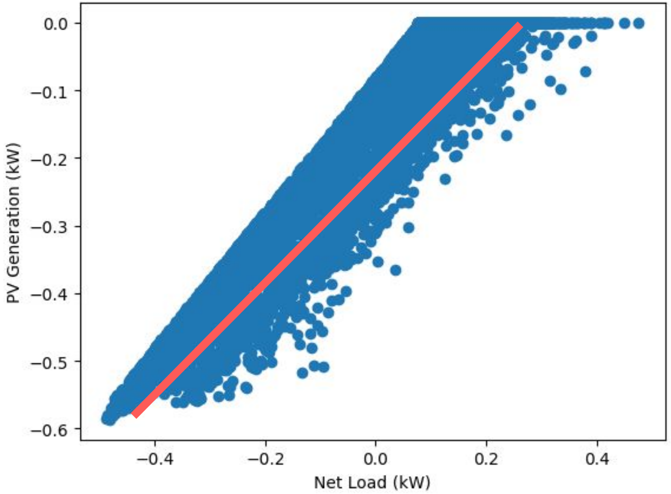
Given the net load signal's values from T to T-192

Models: Disaggregation | Baselines

Dummy (negative signal = solar)
MSE ~ 0.05

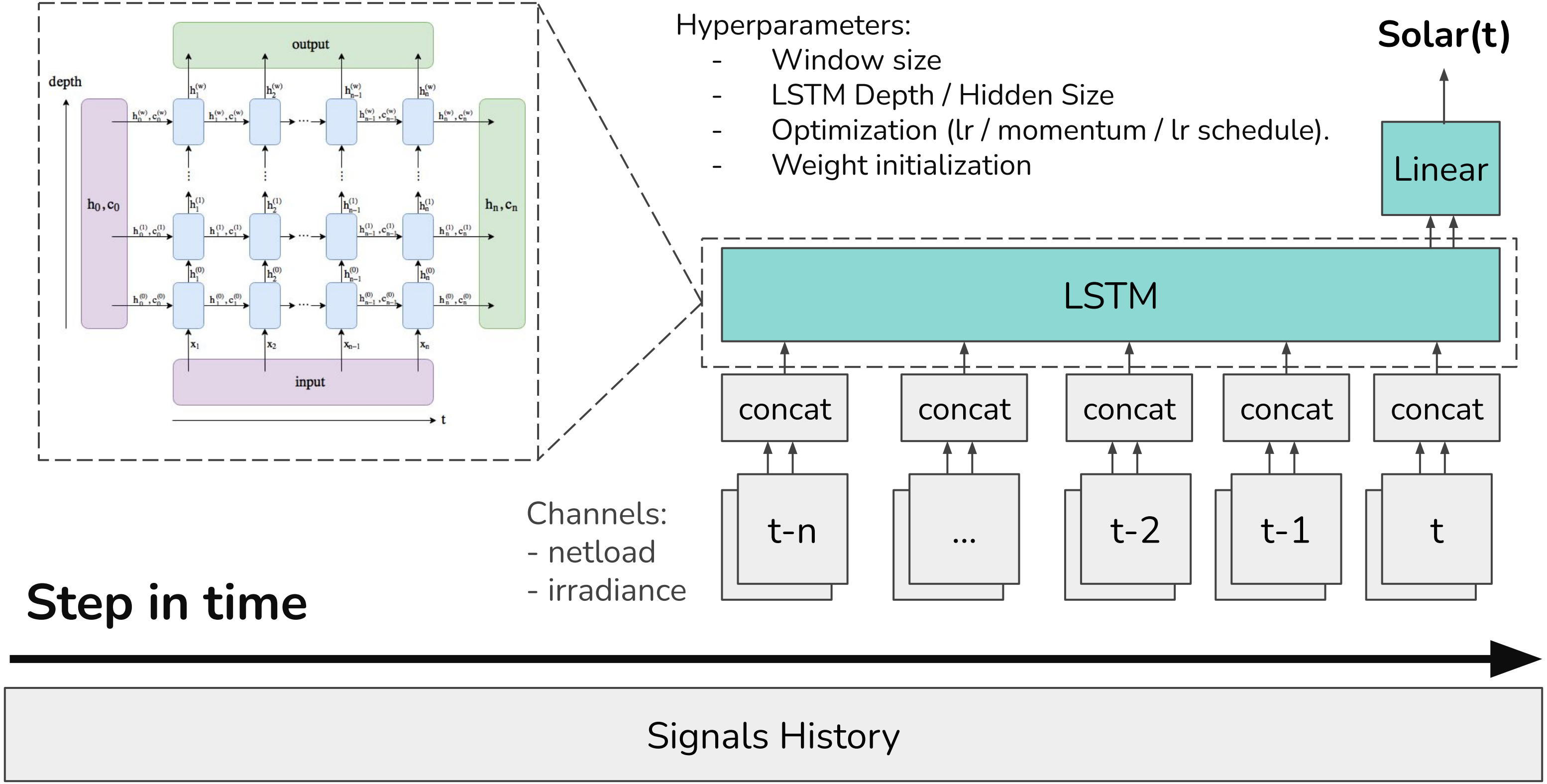


Linear Regression
MSE ~ 0.06

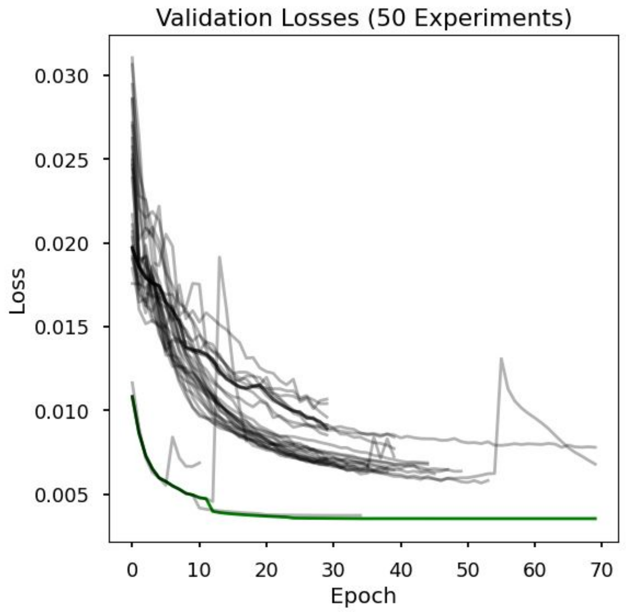


Models: Disaggregation | Model Architecture

Hyperparameters:
- Window size
- LSTM Depth / Hidden Size
- Optimization (lr / momentum / lr schedule).
- Weight initialization



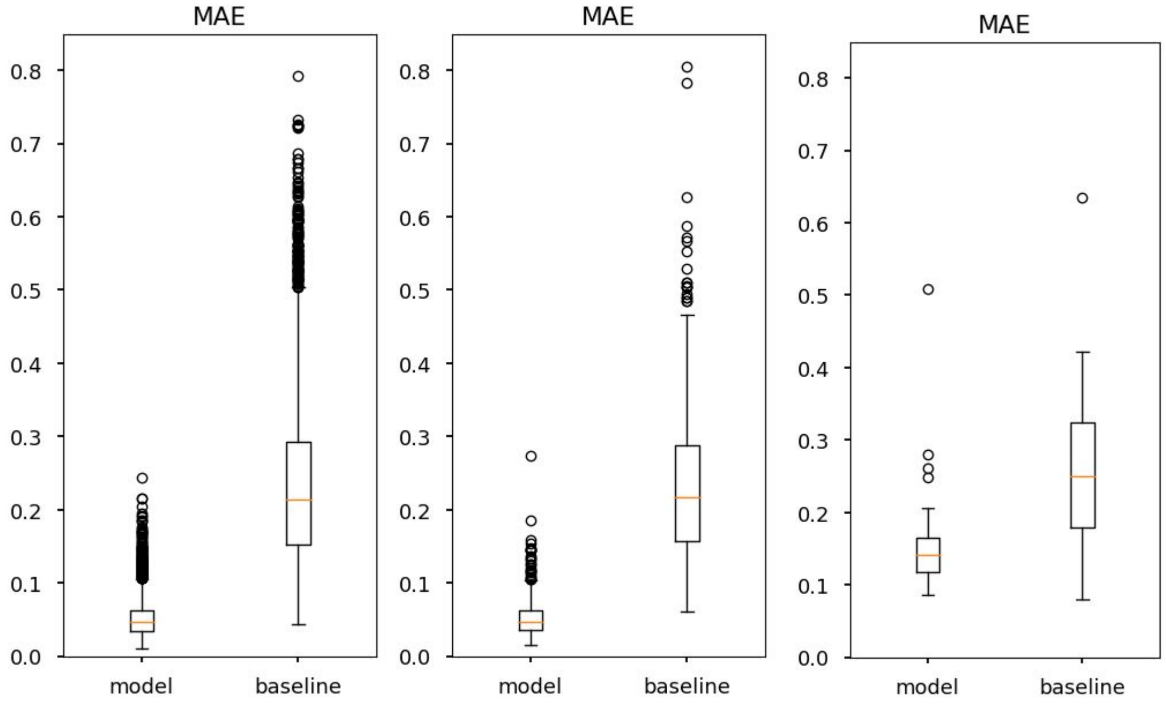
Models: Disaggregation | Training and Performance



Train Performance
4500 houses (Simulated)
~5x improvement

Validation Performance
500 houses (Simulated)
~4x improvement

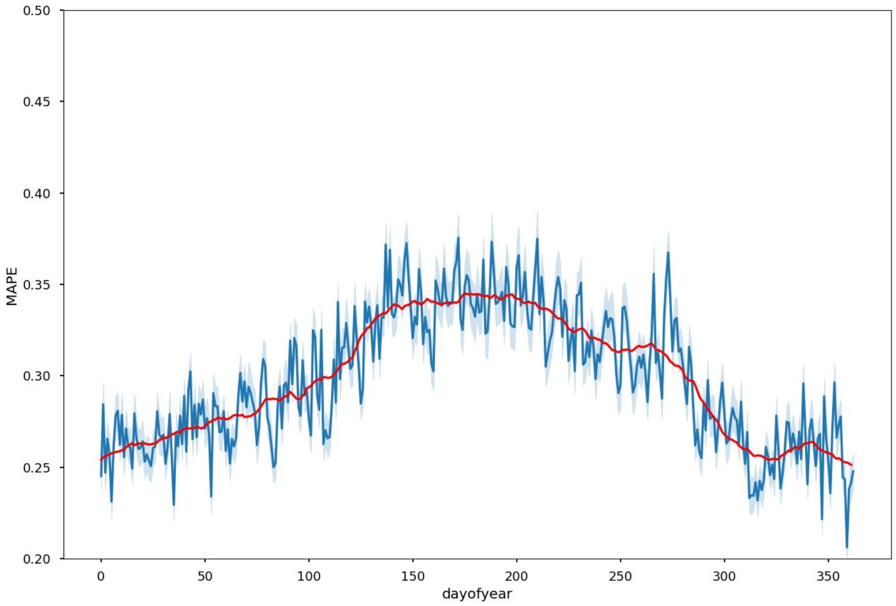
Test Performance
32 houses (Real-World)
~1.7x improvement



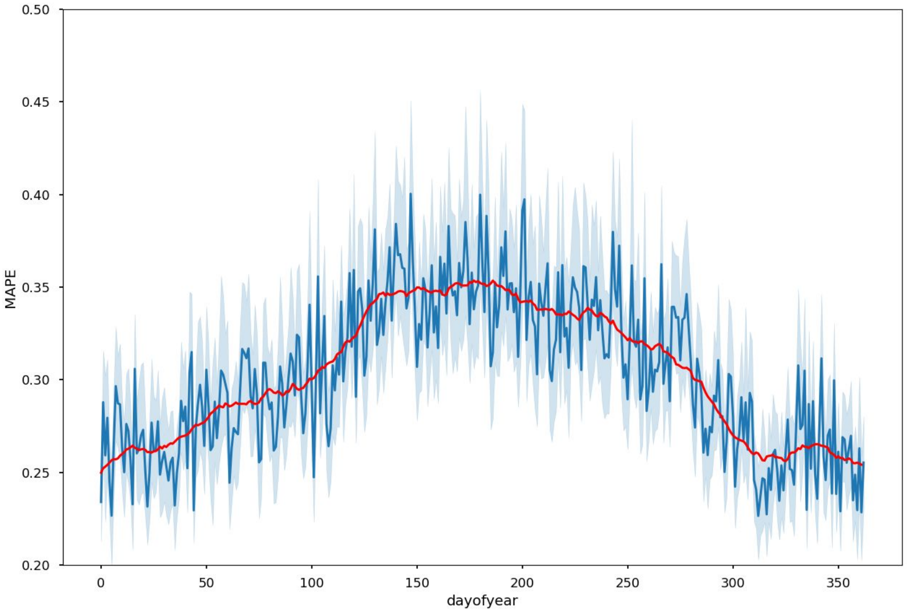
Models: Disaggregation | Post Mortem

Model performs better in winter time …

Training

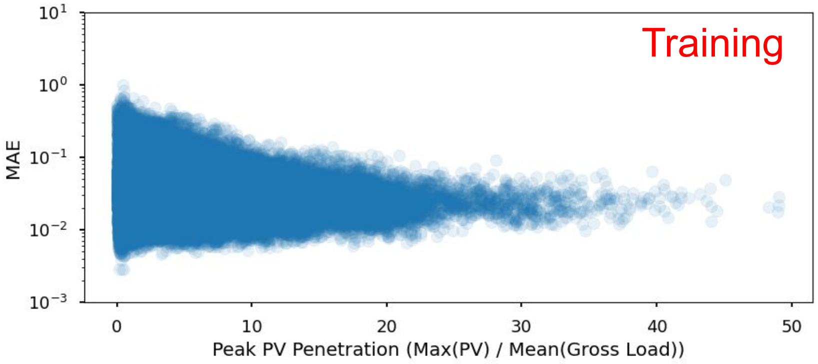


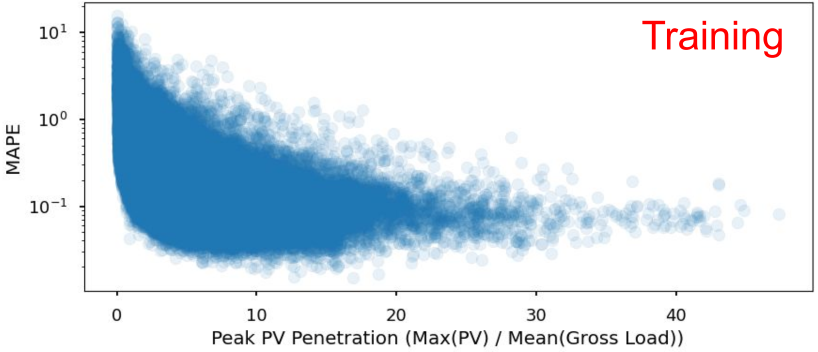
Validation

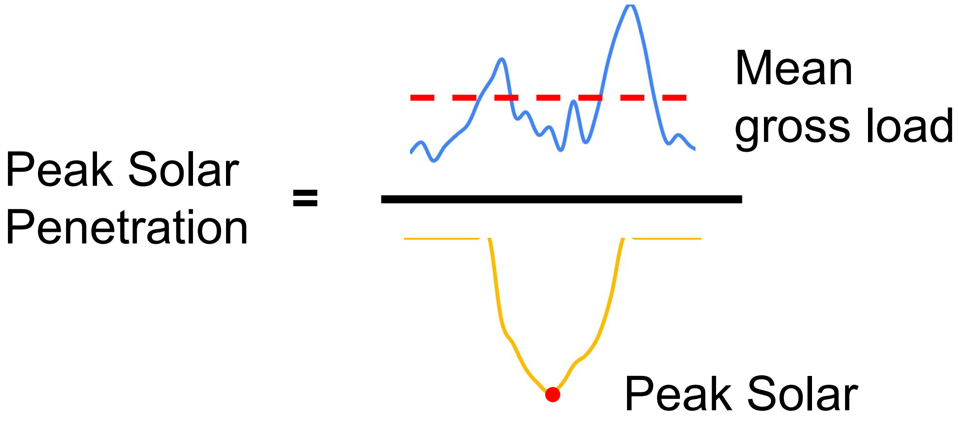


Models: Disaggregation | Post Mortem

Better performance / lower variance for higher pv penetration ...

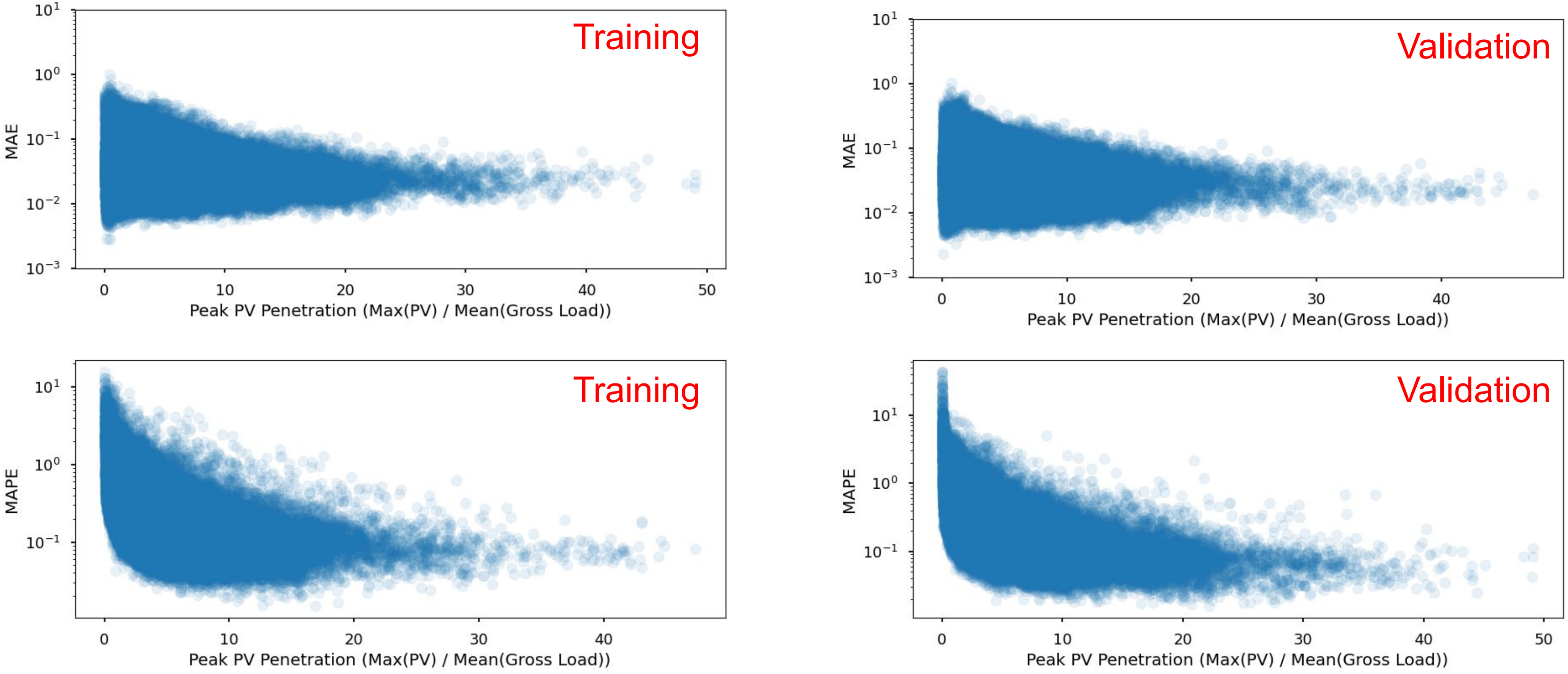




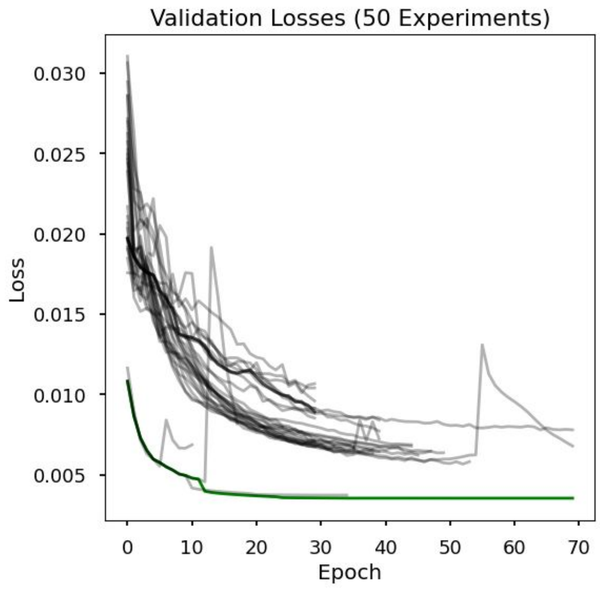


Models: Disaggregation | Post Mortem

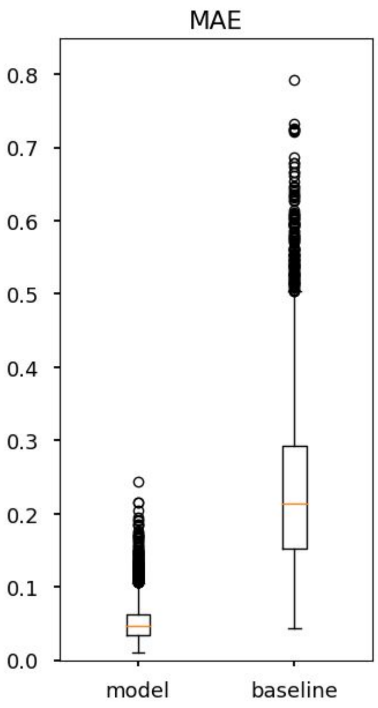
Better performance / lower variance for higher pv penetration …



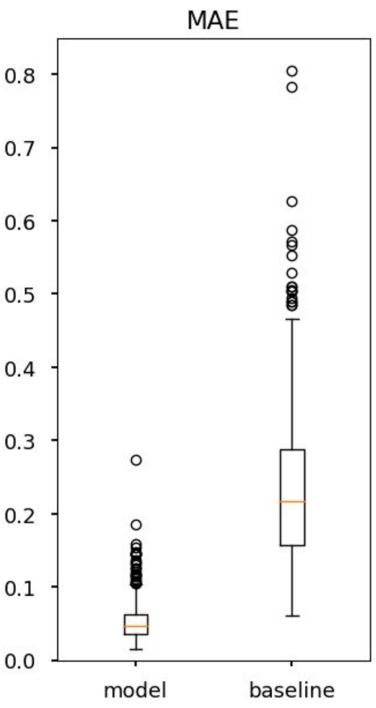
Models: Disaggregation | Post Mortem



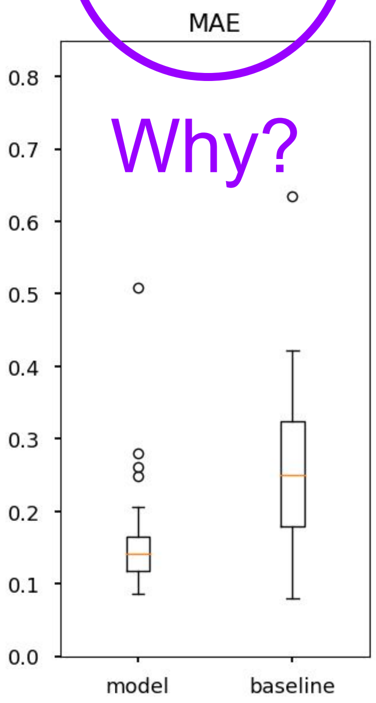
Train Performance
4500 houses (Simulated)
~5x improvement



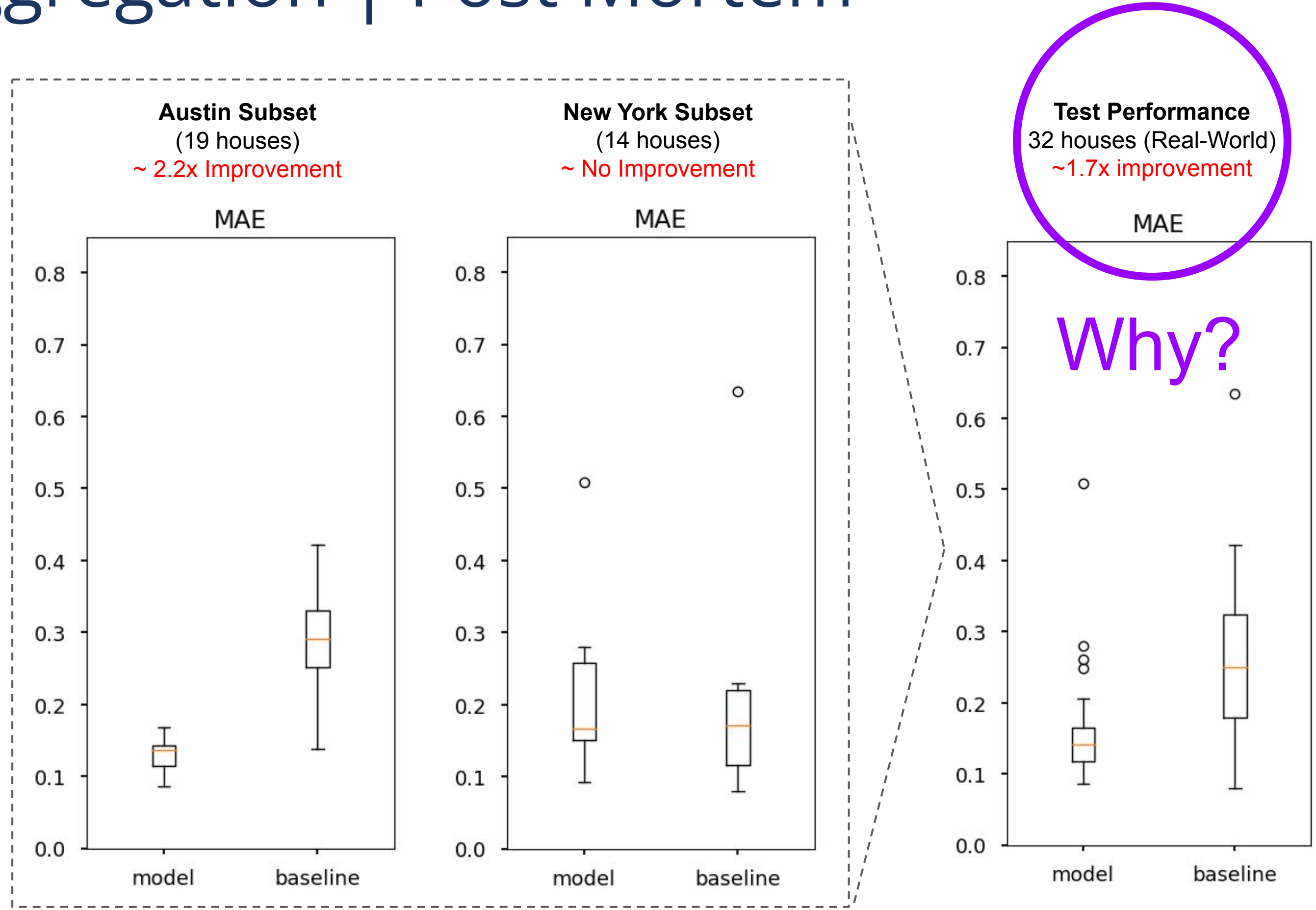
Validation Performance
500 houses (Simulated)
~4x improvement



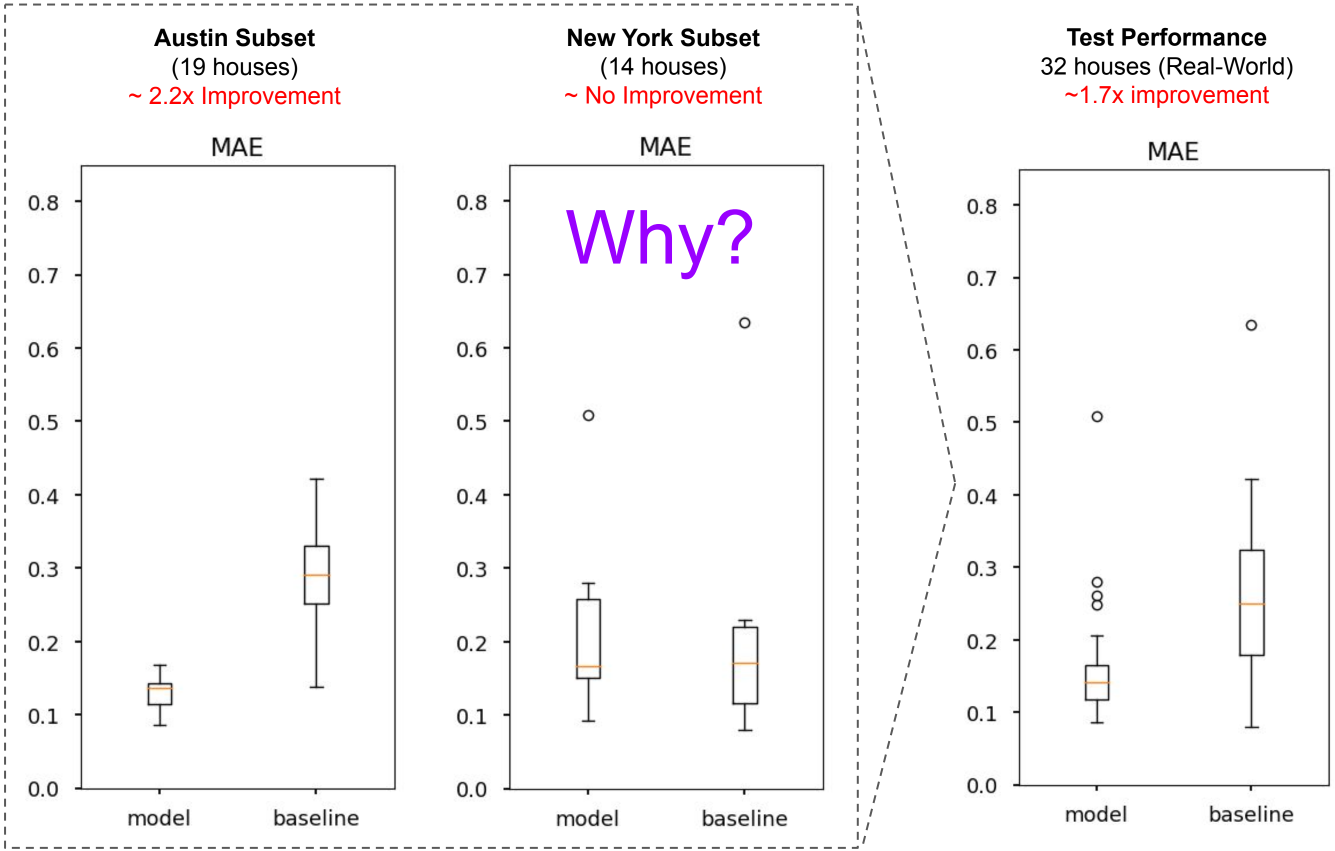
Test Performance
32 houses (Real-World)
~1.7x improvement



Models: Disaggregation | Post Mortem

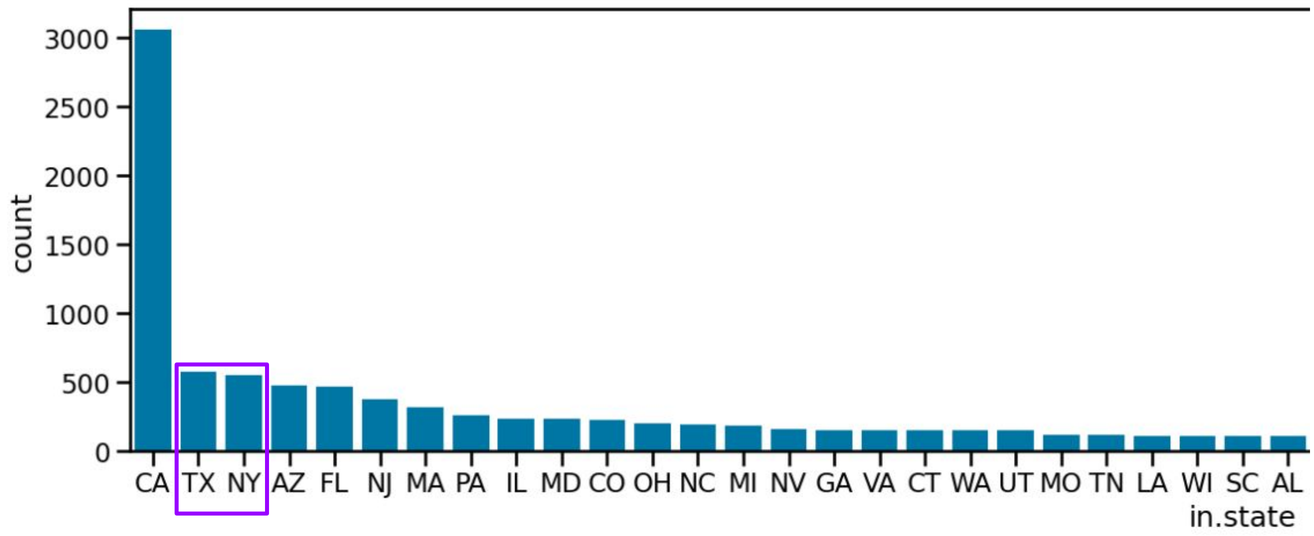


Models: Disaggregation | Post Mortem

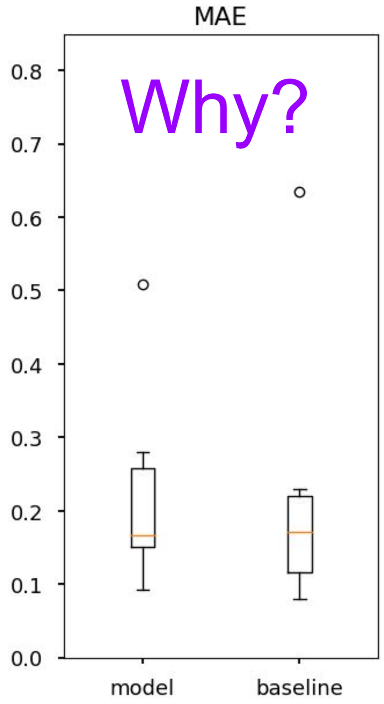


Models: Disaggregation | Post Mortem

Is NY underrepresented in training Data? No.



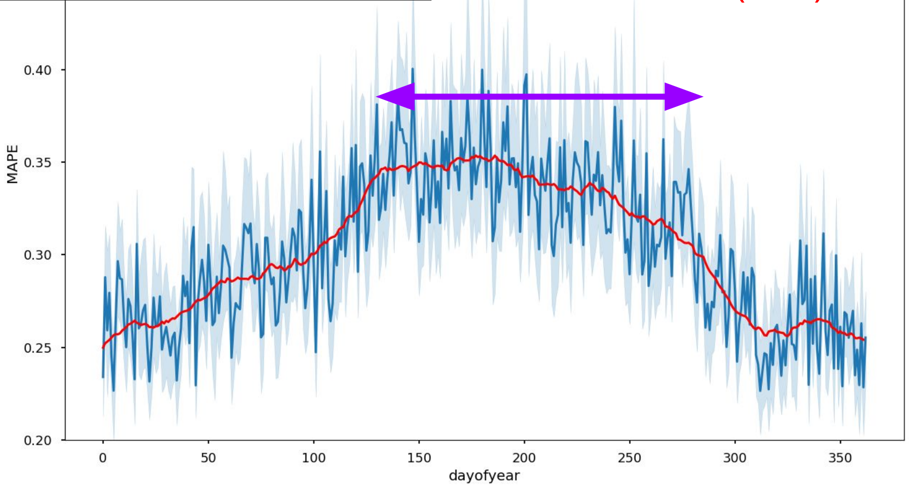
New York Subset
(14 houses)
~ No Improvement



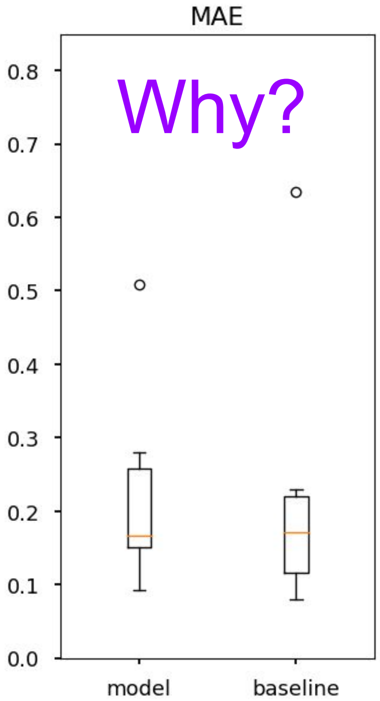
Models: Disaggregation | Post Mortem

NY data comes from primarily summer months …

	min	max
city		
austin	2018-01-01 00:00:00-06	2018-12-31 23:45:00-06
newyork	2019-05-01 00:00:00-05	2019-10-31 23:45:00-05

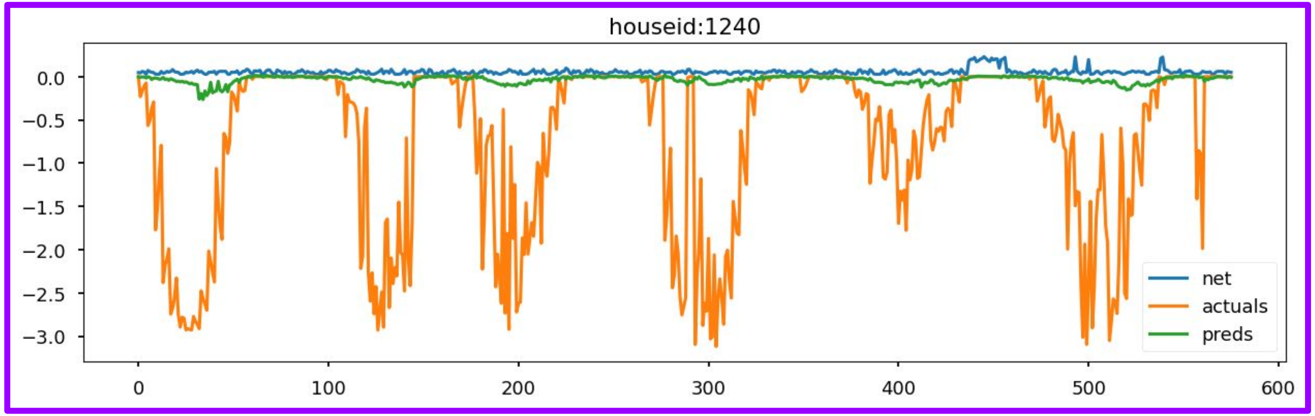


New York Subset
(14 houses)
~ No Improvement

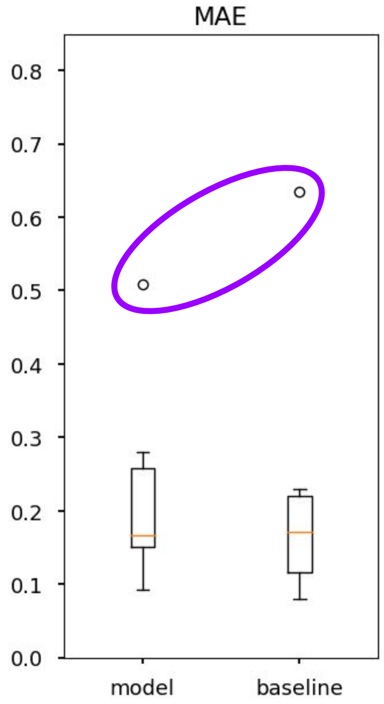


Models: Disaggregation | Post Mortem

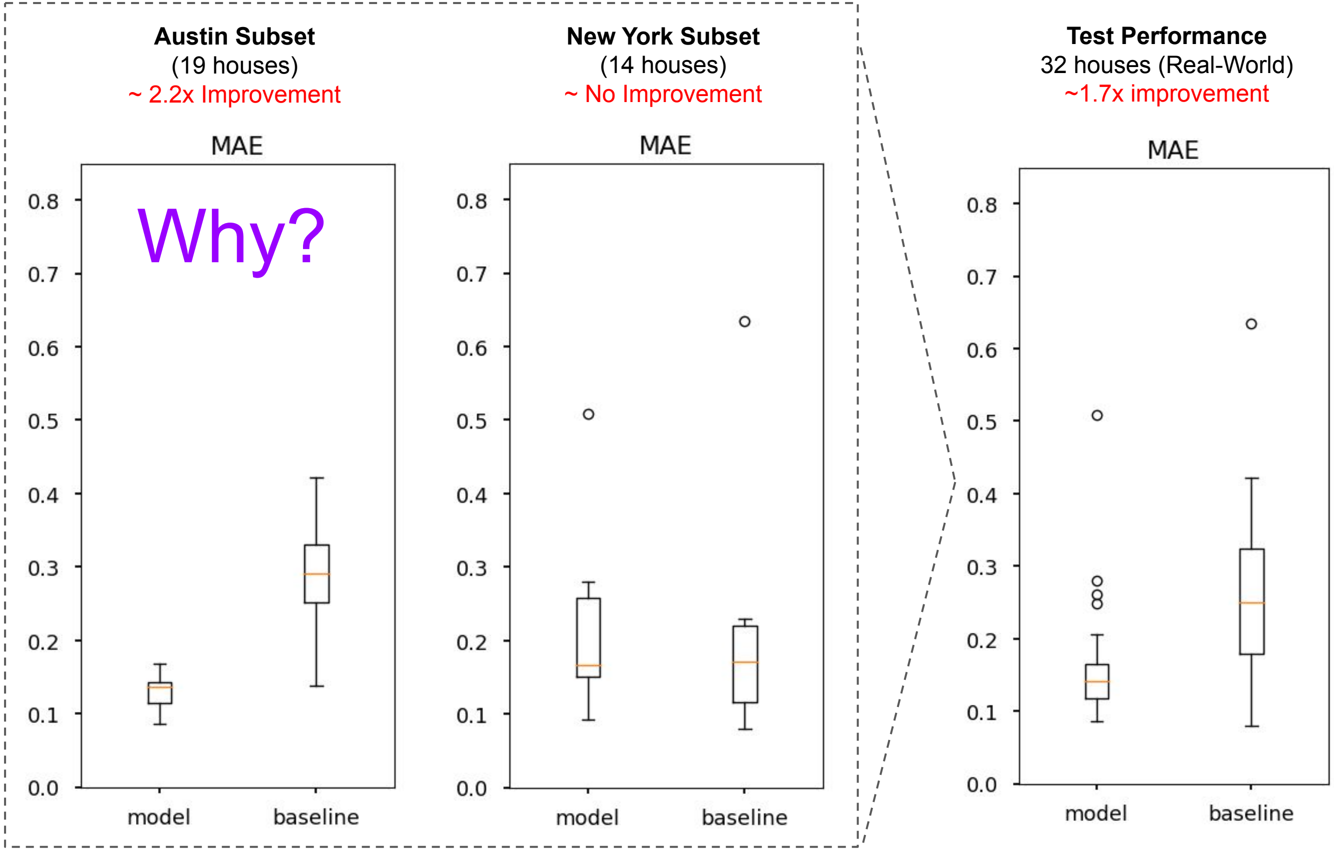
Behind the meter batteries!



New York Subset
(14 houses)
~ No Improvement

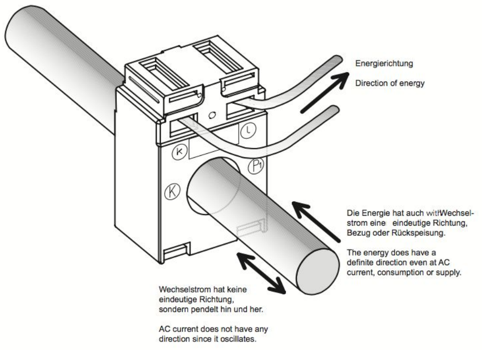


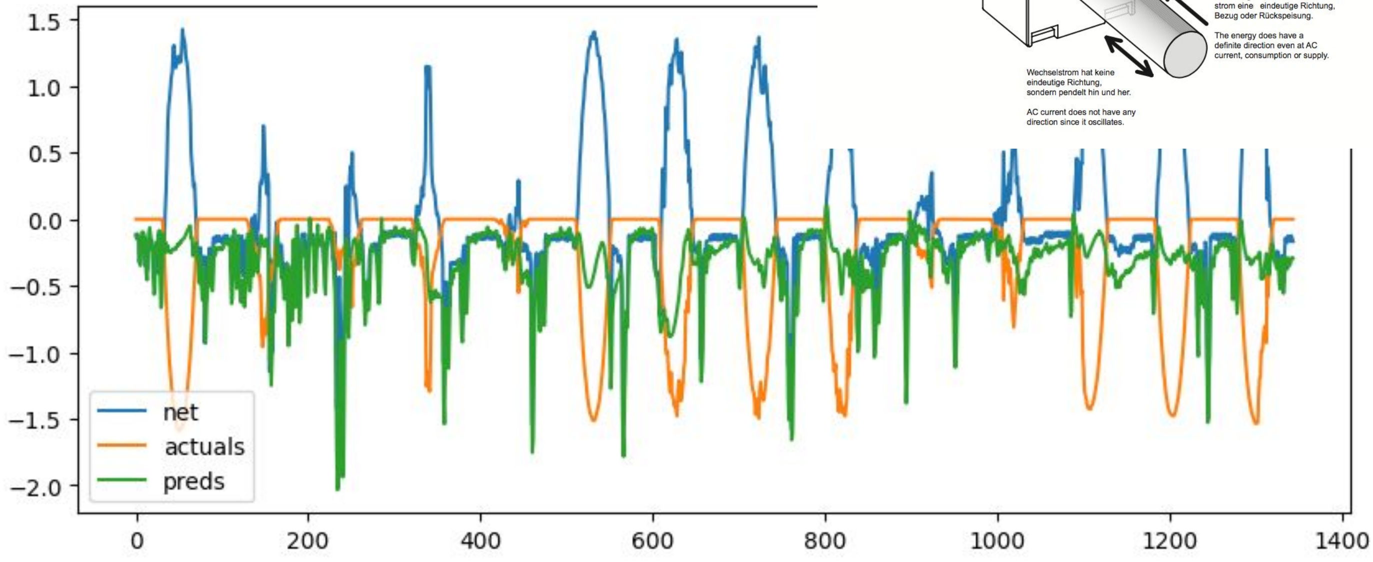
Models: Disaggregation | Post Mortem



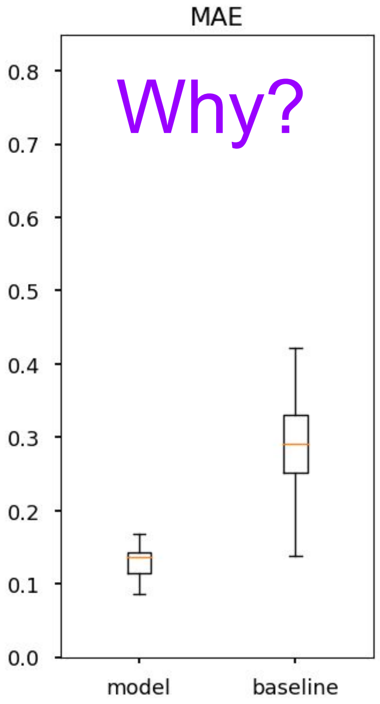
Models: Disaggregation | Post Mortem

Some bad apples!





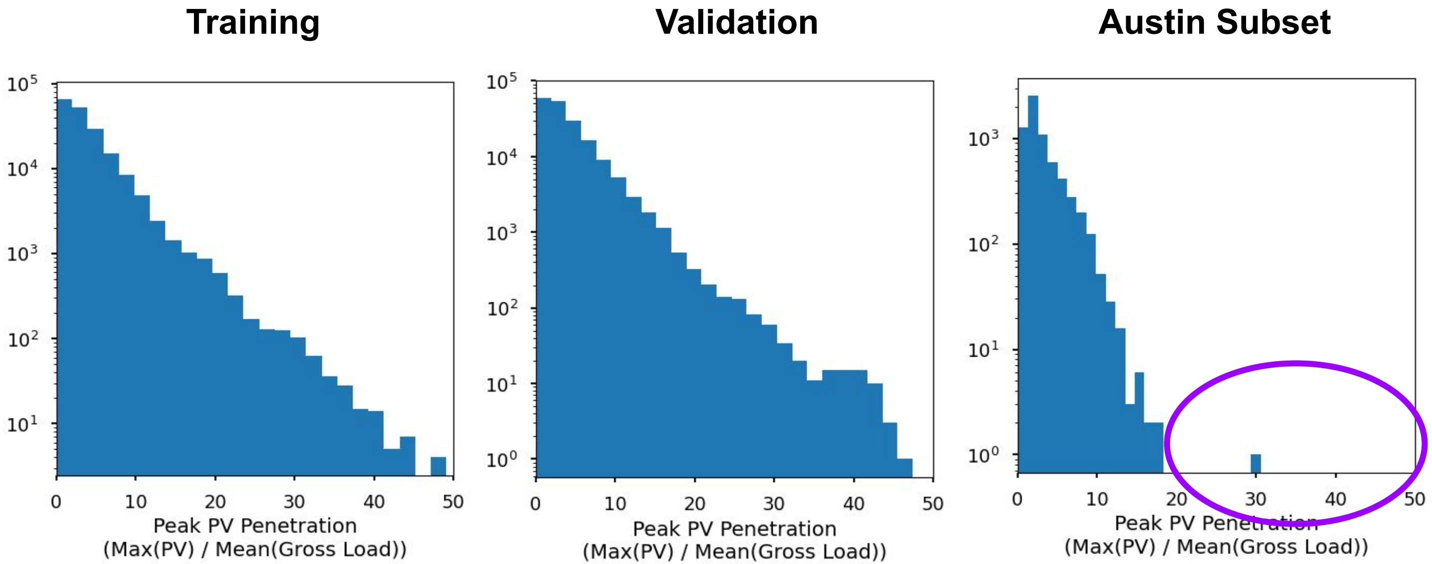
Austin Subset
(19 houses)
~ 2.2x Improvement



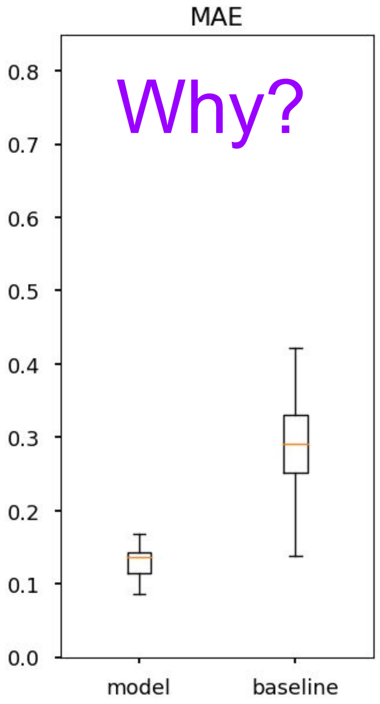
Why?

Models: Disaggregation | Post Mortem

Training data not representative …



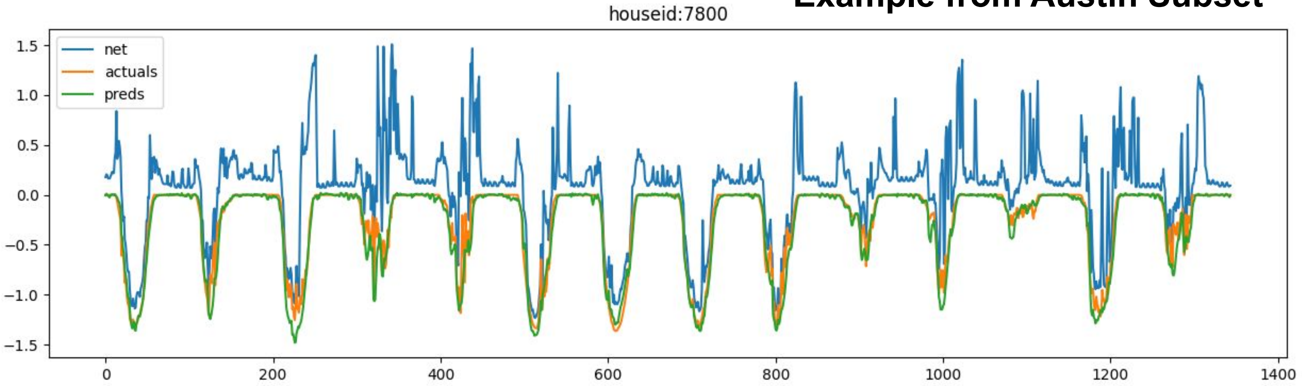
Austin Subset
(19 houses)
~ 2.2x Improvement



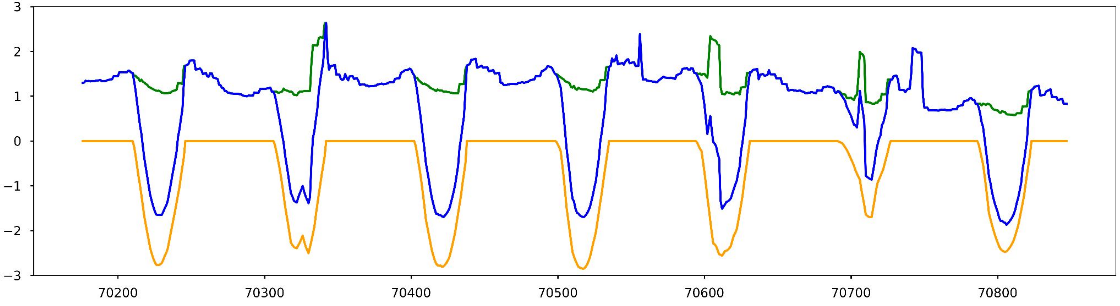
Models: Disaggregation | Post Mortem

Noisy Data!

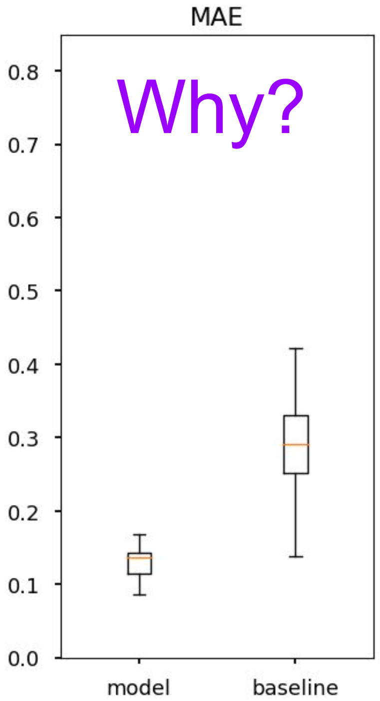
Example from Austin Subset



Example from Training Data

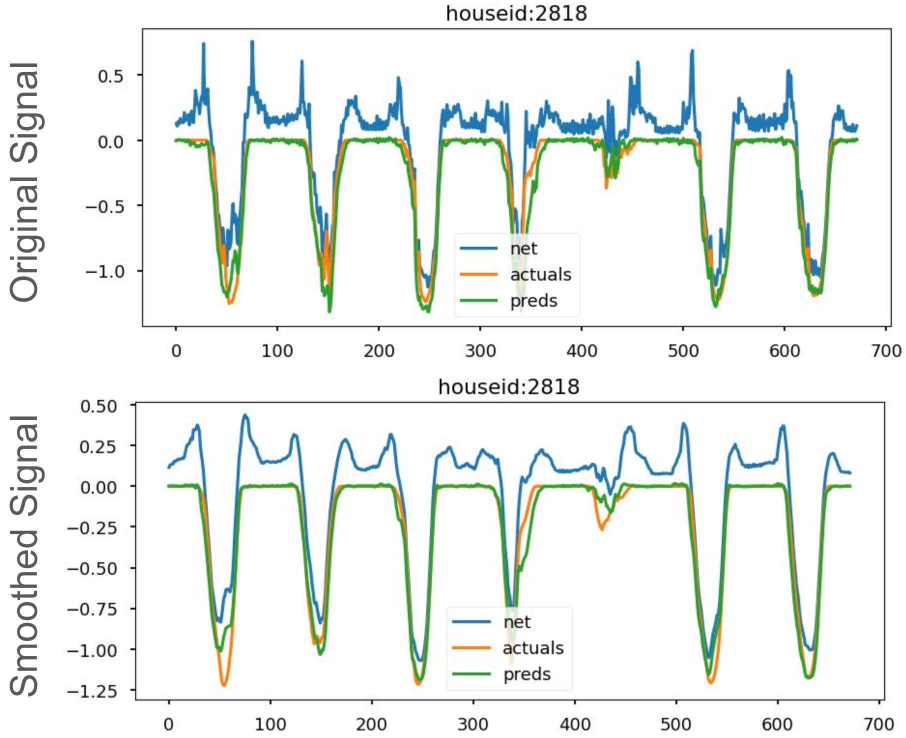


Austin Subset
(19 houses)
~ 2.2x Improvement

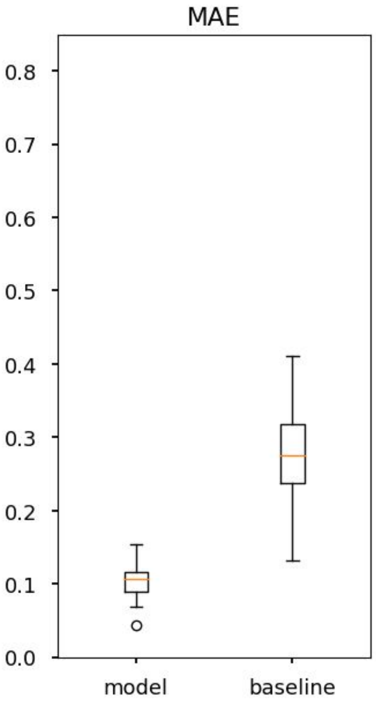


Models: Disaggregation | Post Mortem

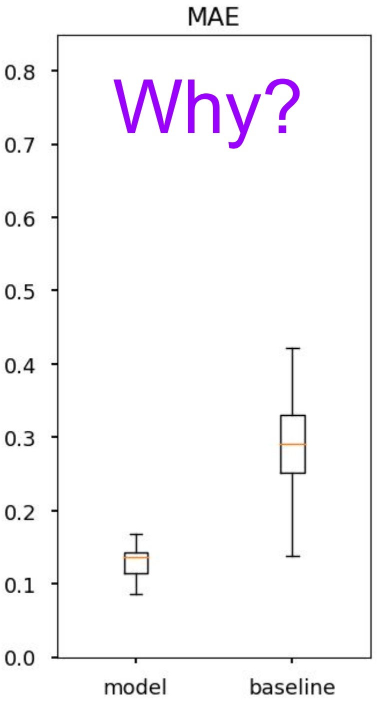
Noisy Data!



Austin Subset (smoothed)
(19 houses)
~ (1.25)(2.2)x Improvement



Austin Subset (Original)
(19 houses)
~ 2.2x Improvement



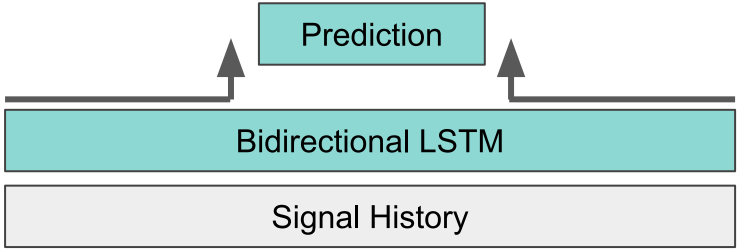
Models: Disaggregation | Possible Next Steps

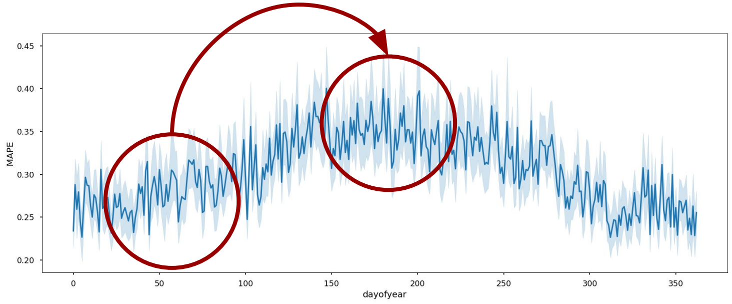
Retain existing architecture and:

1) Forcing Learned Behaviour: Weighted Loss, summer season / low penetration examples.
2) Data Augmentation: add noise to training data **?**
3) Feature Extraction: Convolutional / Pooling layers to counteract high-frequency noise.
4) Ensemble of different window sizes.
5) Hindsight is 20/20: Bidirectional LSTM

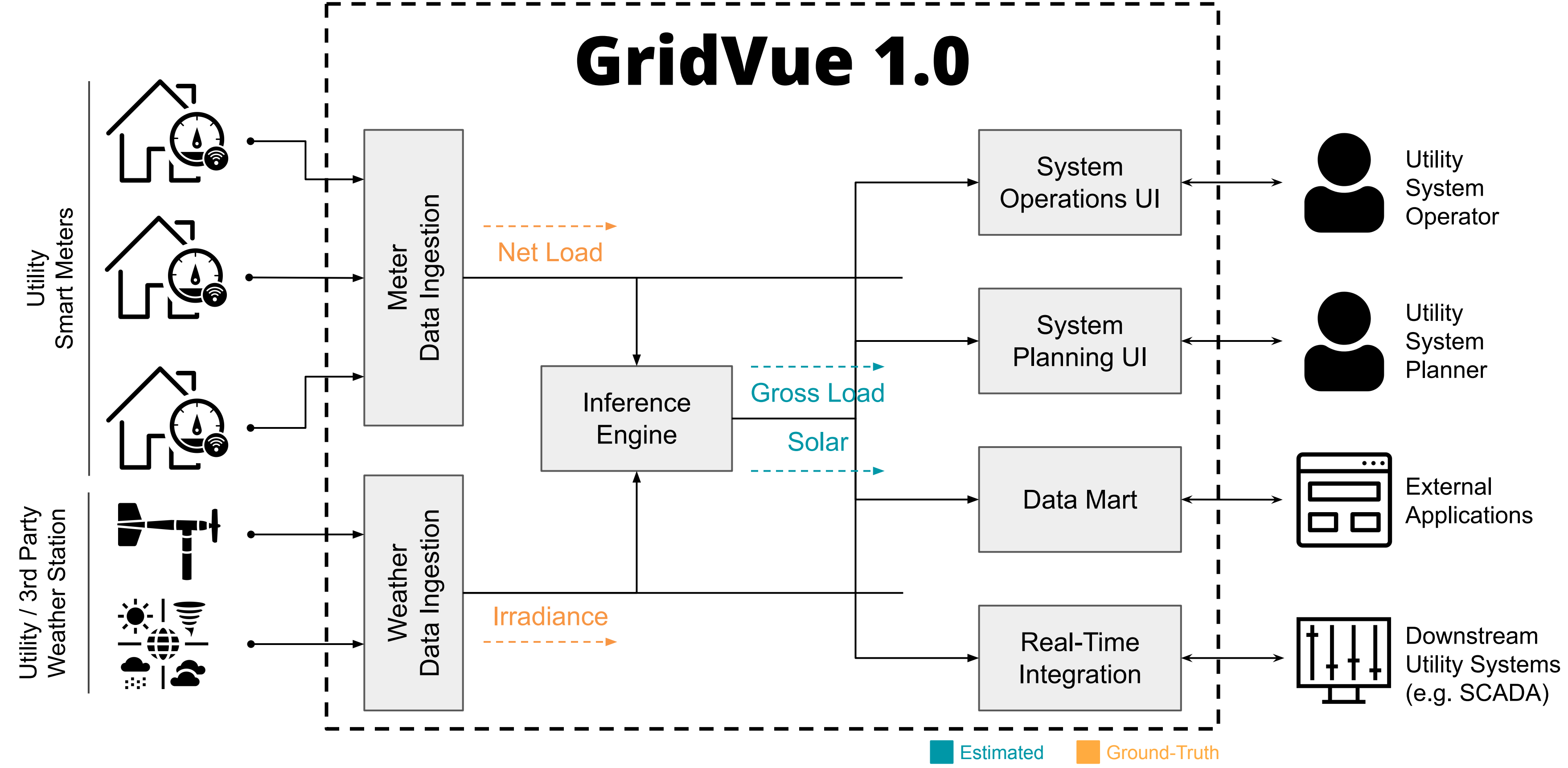
Or try a different approach:

6) Use physical model characterized by estimated parameters (e.g. capacity, tilt/azimuth angle, shading, soiling … etc.). Use data to constantly update belief of parameters.

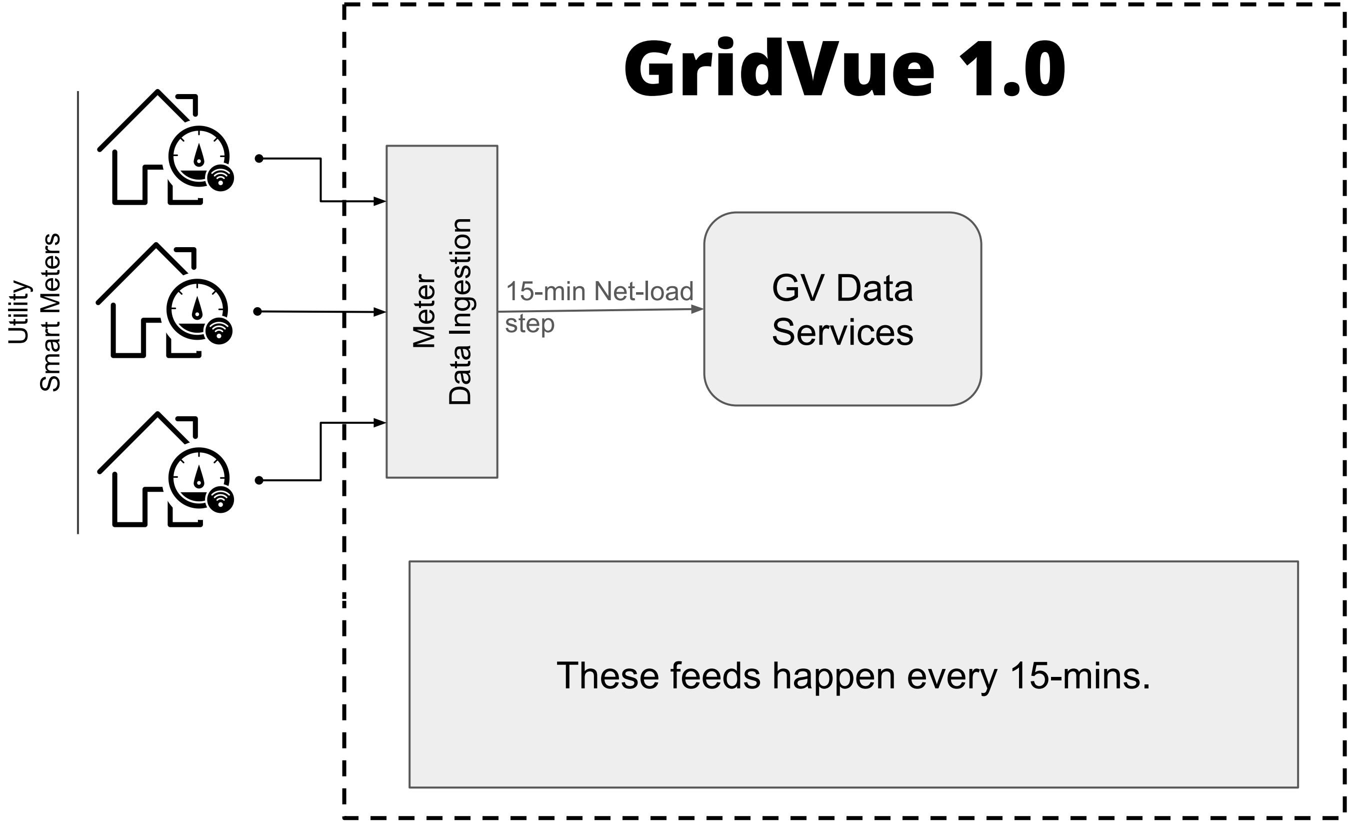




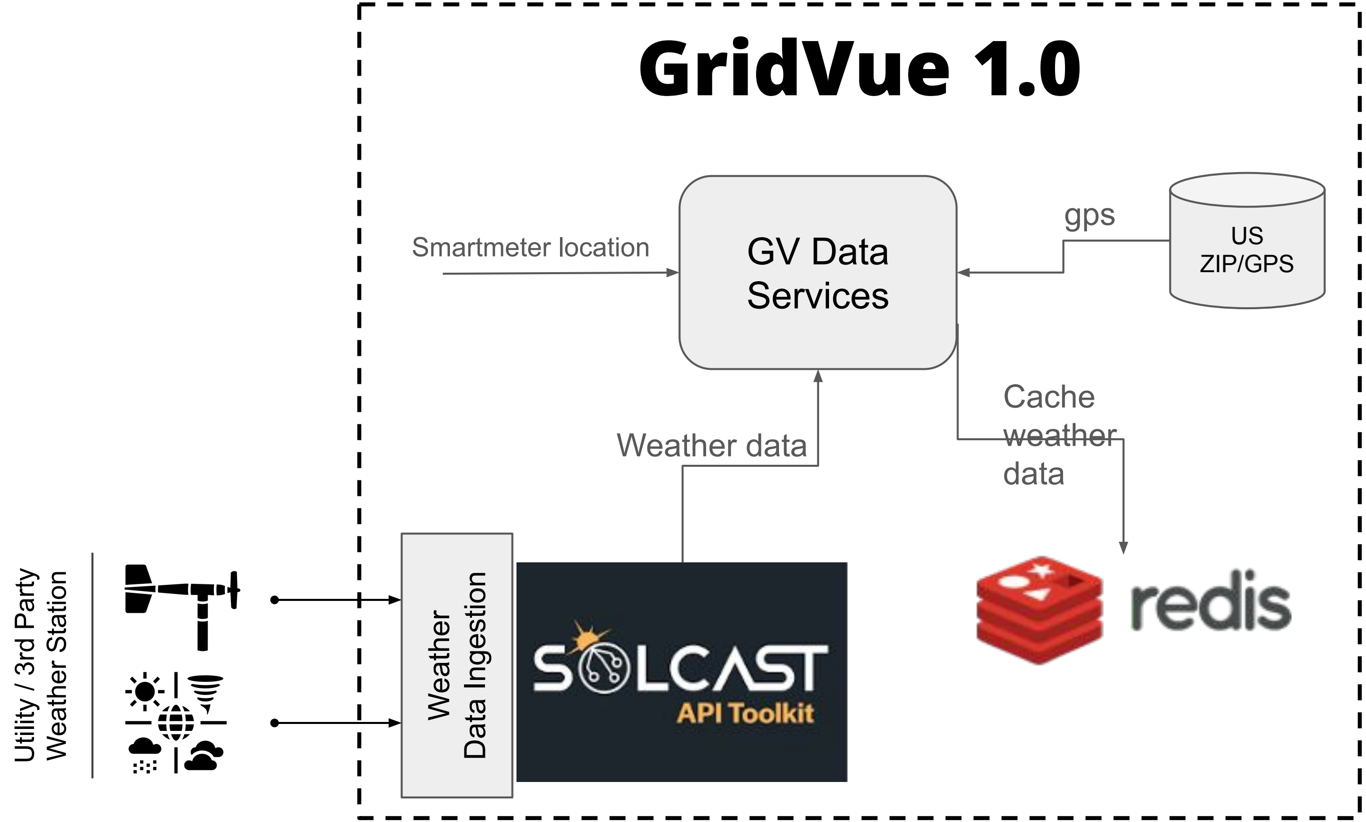
GridVue Infrastructure


GridVue 1.0
Utility
Smart Meters
Utility / 3rd Party
Weather Station
Meter
Data Ingestion
Weather
Data Ingestion
Net Load
Irradiance
Inference
Engine
Gross Load
Solar
System
Operations UI
System
Planning UI
Data Mart
Real-Time
Integration
Utility
System
Operator
Utility
System
Planner
External
Applications
Downstream
Utility Systems
(e.g. SCADA)
Estimated
Ground-Truth

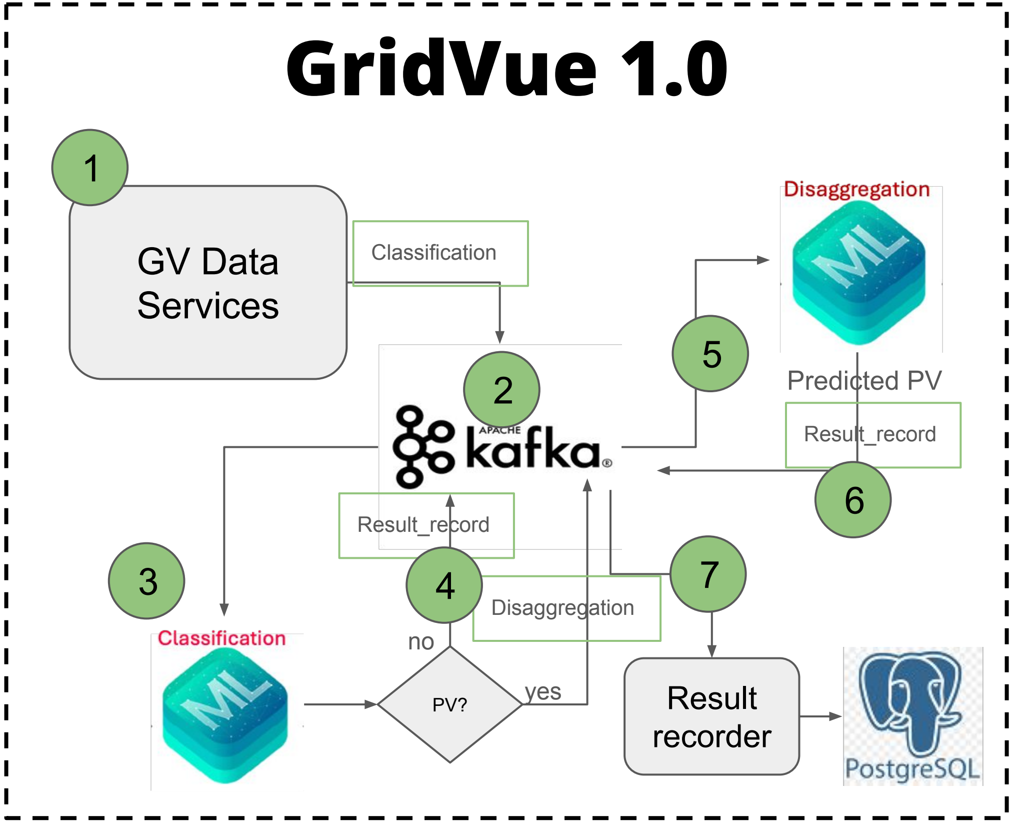
Smart Meter Net-Load Data Feeds



Weather Data Extraction via Solcast API

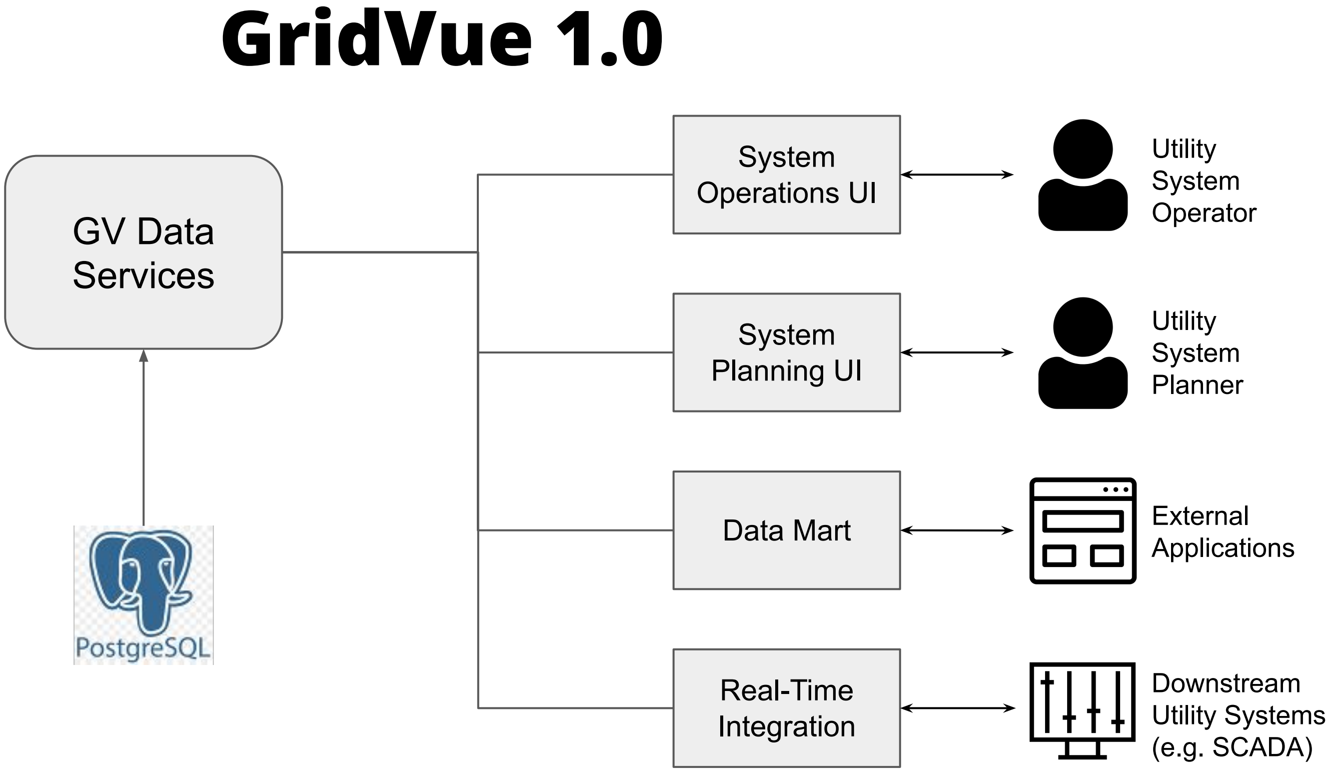


Data Prediction Pipeline



Step	Description
1	GV data services received inputs from smartmeters.
2	GV data services create ‘classification’ request to Kafka.
3	Classification model receives request from Kafka and process classification request.
4	If PV is detected, relay the input data to ‘disaggregation’ queue, otherwise, simply relay the input data to ‘result_record’ queue.
5	Disaggregation model picks up request and generate PV prediction output
6	PV result and other data are being pushed into the ‘result_record’ queue.
7	Result recorder task records the final results to SQL database.

Operational Planning Dashboards



Data Requirements for Each Prediction Model

At time T

Net-load

T-2days 192 frames
(sliding window)

Irradiance

T-2days 192 frames
(sliding window)

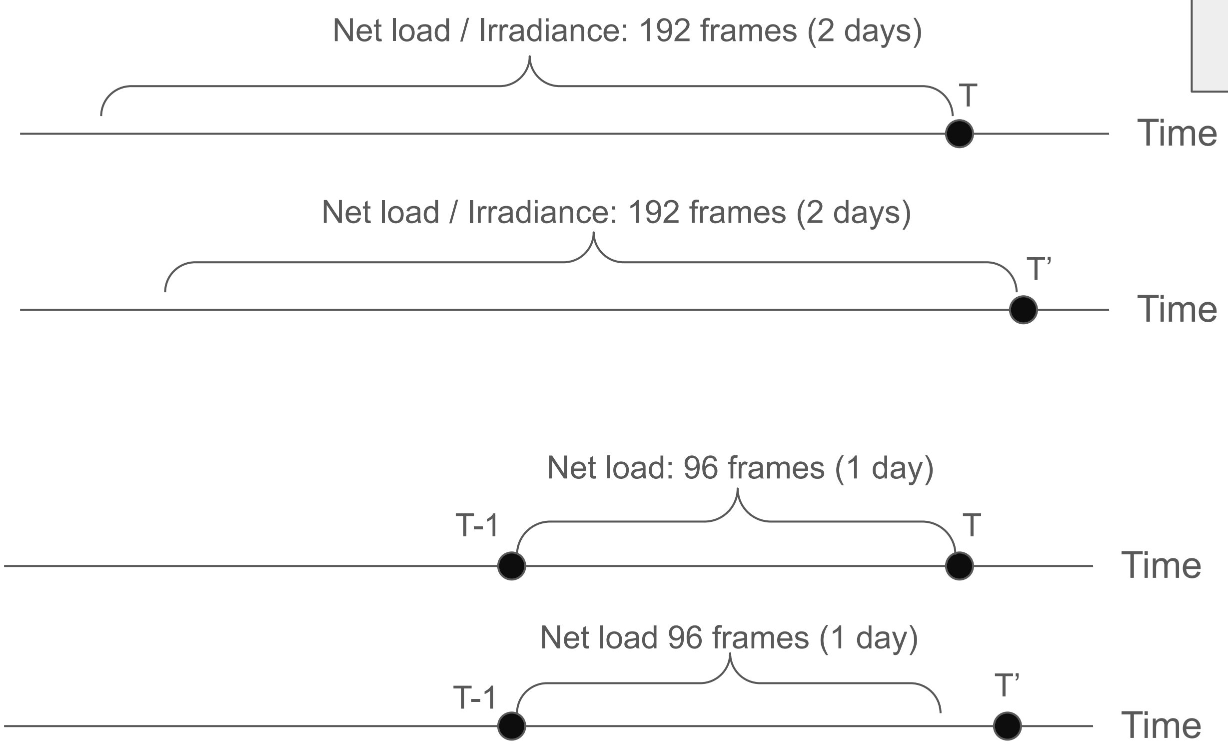
Net-load

T-1day 96 frames

Time Series Prediction
Sliding versus non-sliding windows

Sliding window is expensive!
For classification model, we do not expect customer will switch PV/Non-PV within each day.

For new house prediction, skipped first 2 day of net-load and irradiance data.



Sliding window for Disaggregation model

Non-sliding window for Classification model

Technology Stacks

Data Mining

- AWS S3
- AWS Athena
- AWS Glue
- Spark
- Custom python APIs and tools

Model Building

- AWS SageMaker
- Jupyter notebooks
- Torch model Archiver

AWS Inference Infrastructure

GV Data service, GV Dashboards, prediction pipeline workers

Custom USZip, Irradiance, SQL, Redis connectors and pipeline workers APIs

EC2 - csprod-infra-01

Type: m7i.8xlarge

- 32 vCPU
- 128 GB RAM
- 300GB Storage

GV Model service

TorchServe

EC2 - csprod-mi-dis

Type: inf1.2xlarge

- 8 vCPU
- 16GB RAM
- 100GB Storage

GV Model service

TorchServe

EC2 - csprod-mi-clf

Type: inf1.2xlarge

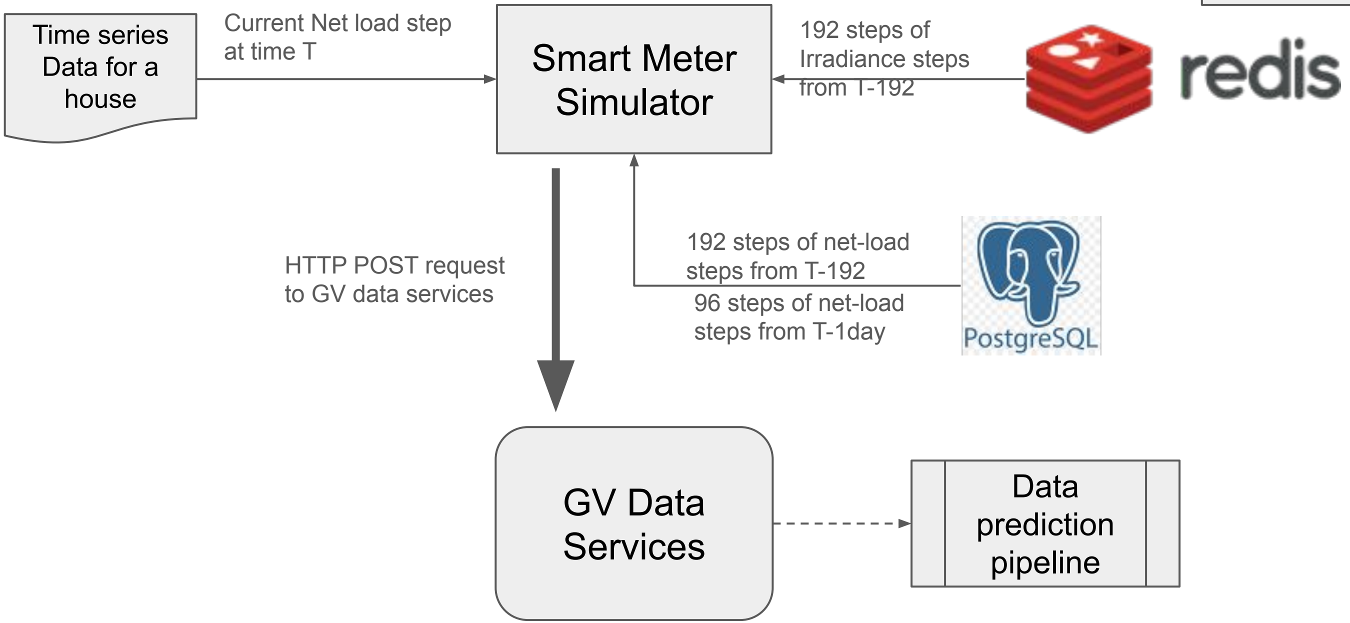
- 8 vCPU
- 16GB RAM
- 100GB Storage

MVP Demo #1 - Real-Time Prediction Pipeline

Smart Meter Simulator (SMS)

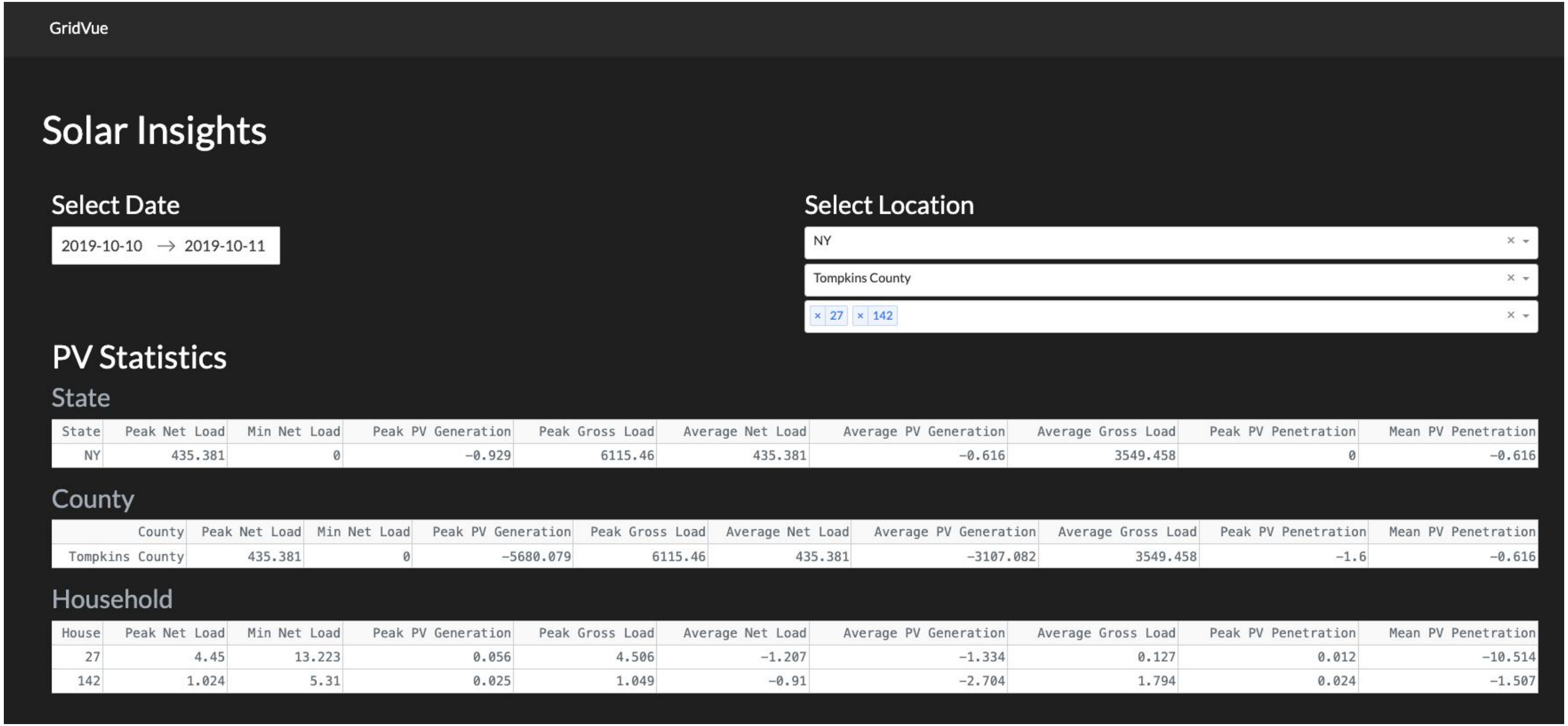
For every step, iterate the step below

The demo house data set includes 1 year of historical data. We pretend that we got each step every 15 minutes from a smartmeter. There is a total of 35K iterations!! Estimated prediction time is around 22 mins. (Each prediction takes around 38ms!!)



MVP Demo #2 - Dashboard UI

Energy Analyst Dashboard



GridVue

Solar Insights

Select Date

2019-10-10 → 2019-10-11

Select Location

NY

Tompkins County

× 27 × 142

PV Statistics

State

State	Peak Net Load	Min Net Load	Peak PV Generation	Peak Gross Load	Average Net Load	Average PV Generation	Average Gross Load	Peak PV Penetration	Mean PV Penetration
NY	435.381	0	-0.929	6115.46	435.381	-0.616	3549.458	0	-0.616

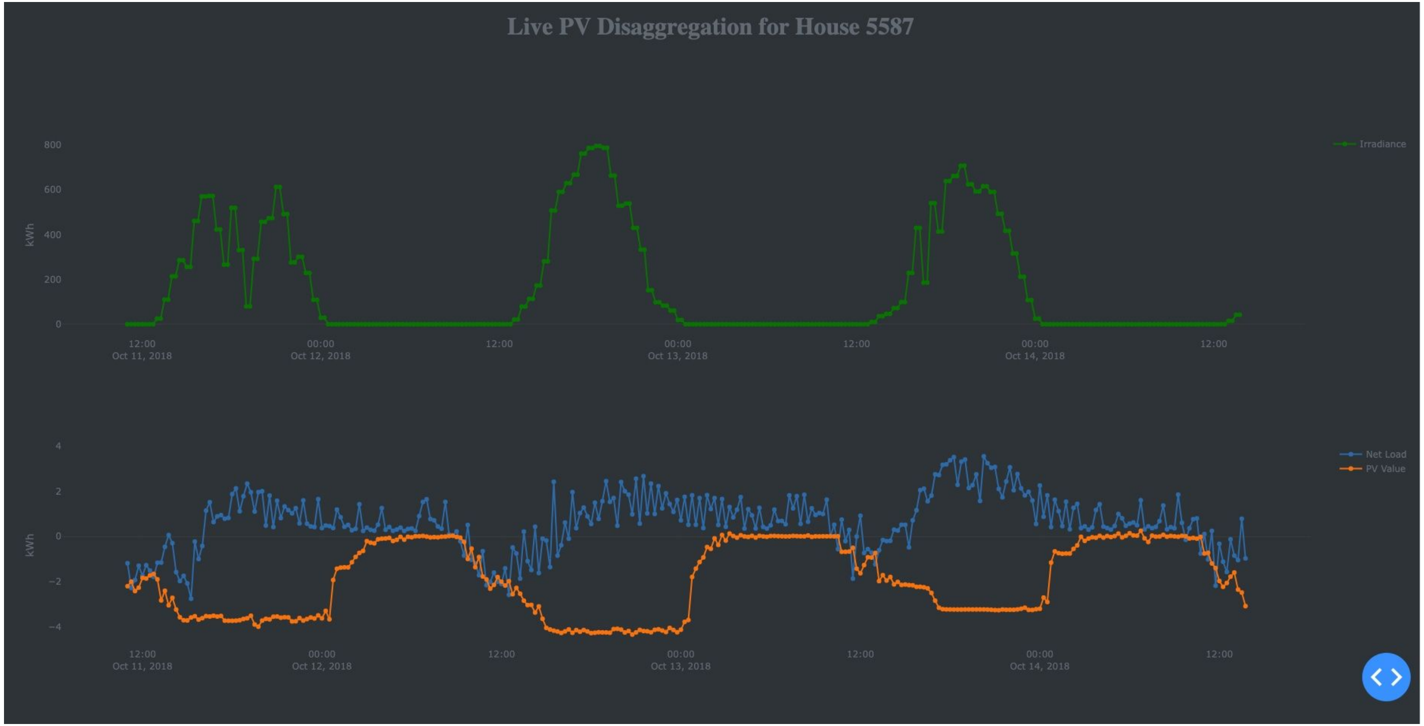
County

County	Peak Net Load	Min Net Load	Peak PV Generation	Peak Gross Load	Average Net Load	Average PV Generation	Average Gross Load	Peak PV Penetration	Mean PV Penetration
Tompkins County	435.381	0	-5680.079	6115.46	435.381	-3107.082	3549.458	-1.6	-0.616

Household

House	Peak Net Load	Min Net Load	Peak PV Generation	Peak Gross Load	Average Net Load	Average PV Generation	Average Gross Load	Peak PV Penetration	Mean PV Penetration
27	4.45	13.223	0.056	4.506	-1.207	-1.334	0.127	0.012	-10.514
142	1.024	5.31	0.025	1.049	-0.91	-2.704	1.794	0.024	-1.507

Operator’s Live dashboard



Learn More at GridVue.org