Estimating the behind-the-meter solar generation with existing infrastructure

Emre Can Kara,
Michaelangelo Tabone, Sila Kiliccote

AI-at-SLAC

July 11, 2017

SLAC NATIONAL ACCELERATOR LABORATORY

Our team

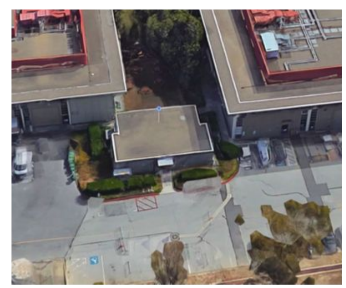
Science Directorate > Applied Energy Program:

- Materials for energy
- Subsurface Technology
- 21st Century Electric Grid

GISMO | Grid Integration, Systems & Mobility

gismo.slac.stanford.edu

- Team of 15 researchers, scientists and engineers.
- Major funding:
  - DOE EERE
  - CEC EPIC
  - ARPA-e



What do we do @GISMo?

Grid Integration, Systems & Mobility

POWER GRID

- Distributed control
- Data-driven planning and operations with DERs
- Pricing mechanisms and market structures
- Business models
- Regulatory frameworks

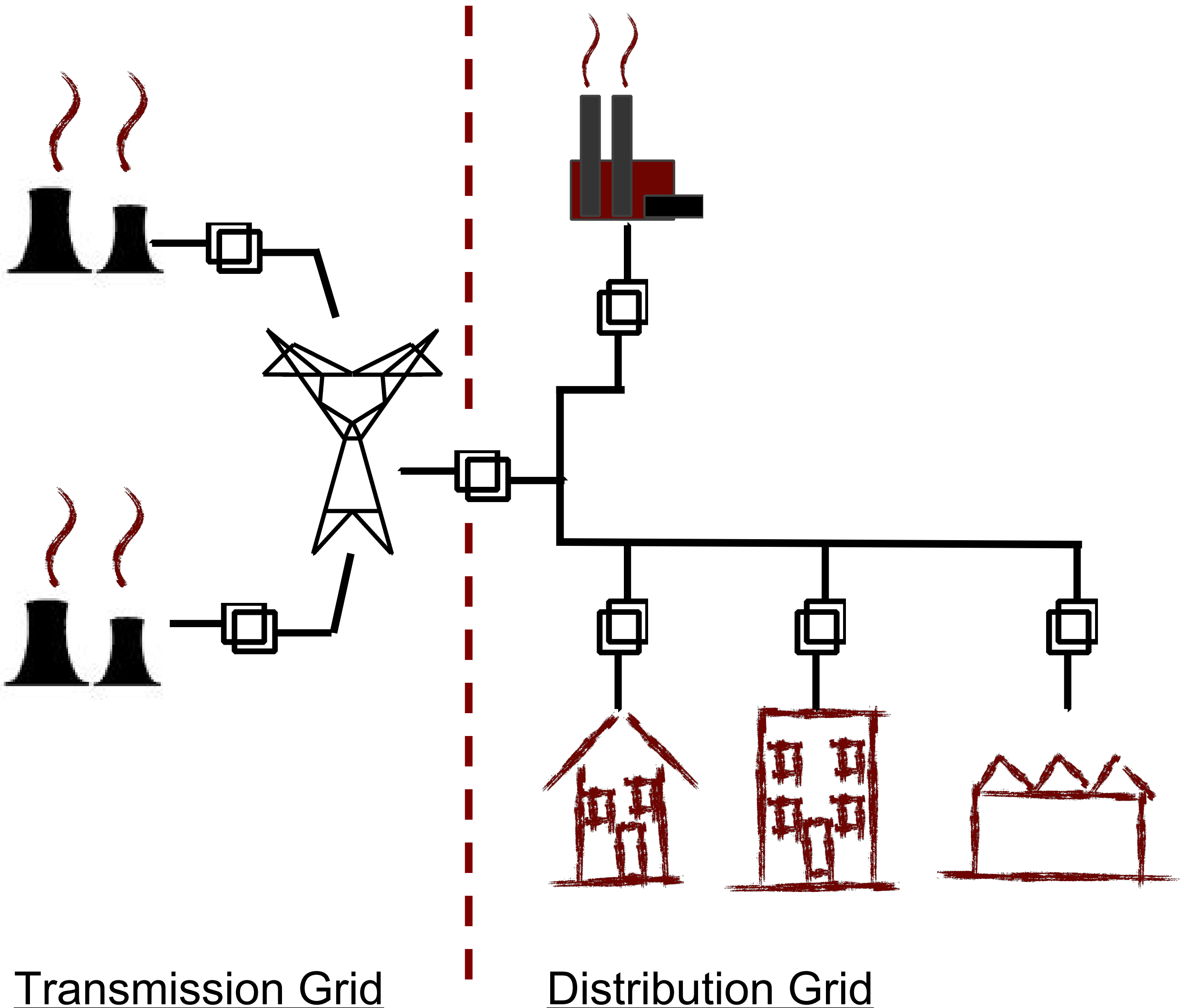
AMBIENT /BUILDING INTELLIGENCE

- Interaction design
- Seamless integration
- Embedded
- Context aware
- Personalized
- Adaptive
- Anticipatory

MOBILITY

- Mobility as a service
- Electrification of transportation
- Data-driven modeling and analysis
- Vehicle-to-building
- Vehicle-to-grid

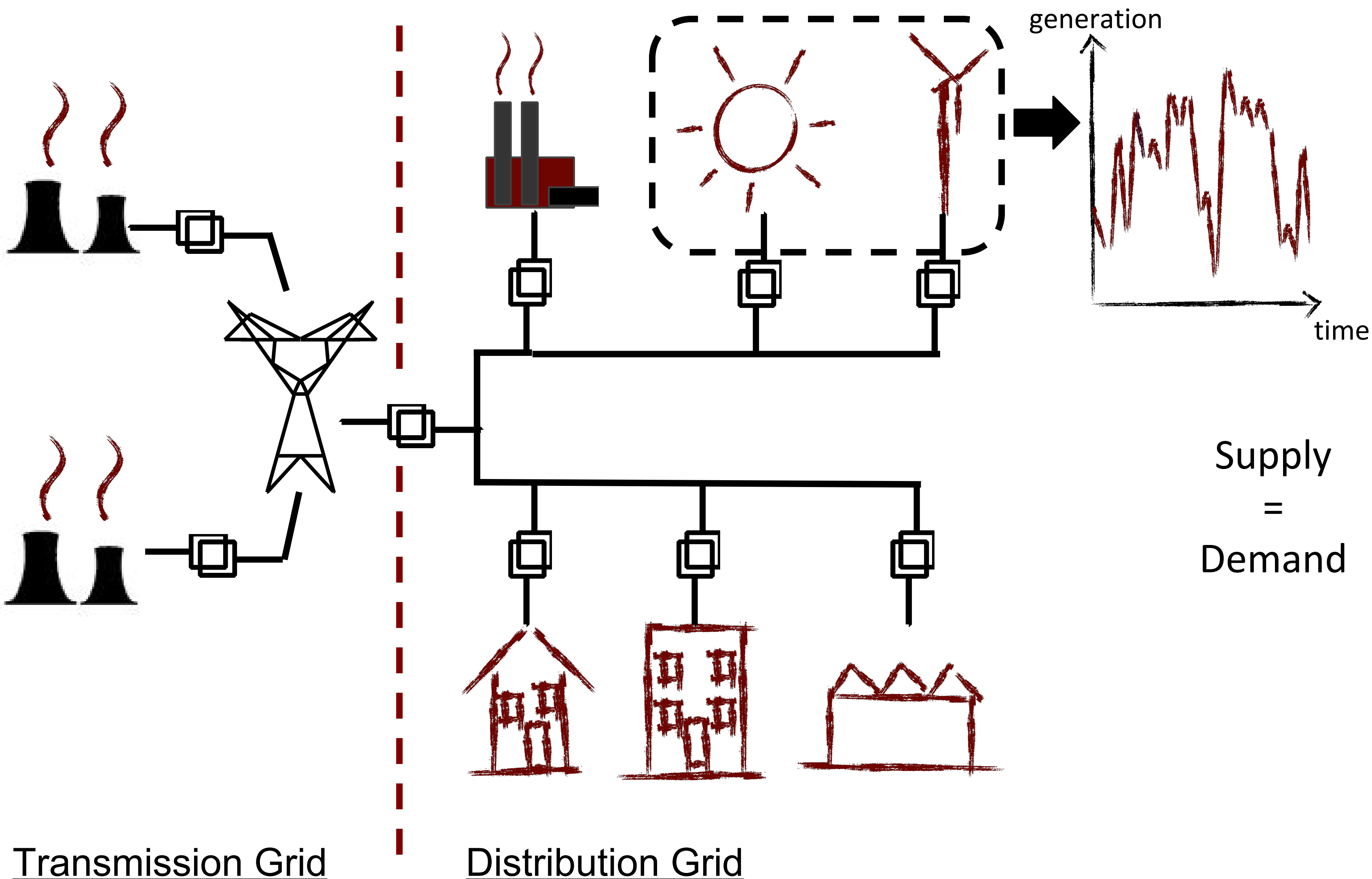
Integrating Renewable Energy



Supply
=
Demand

INFERLab
Intelligent Infrastructure Research Laboratory

Integrating Renewable Energy



Supply
=
Demand

VADER: Visualization and Analytics for Distributed Energy Resources

Enabling high penetration of renewables requires:

- More comprehensive and dynamic situational awareness; and
- Active control of the distribution network.

Traditional approach/technology is inadequate:

- Partial observability
- New devices and models
- Heterogenous data sources and data streams

Plenty of data that is not being used!

Synchrophasors

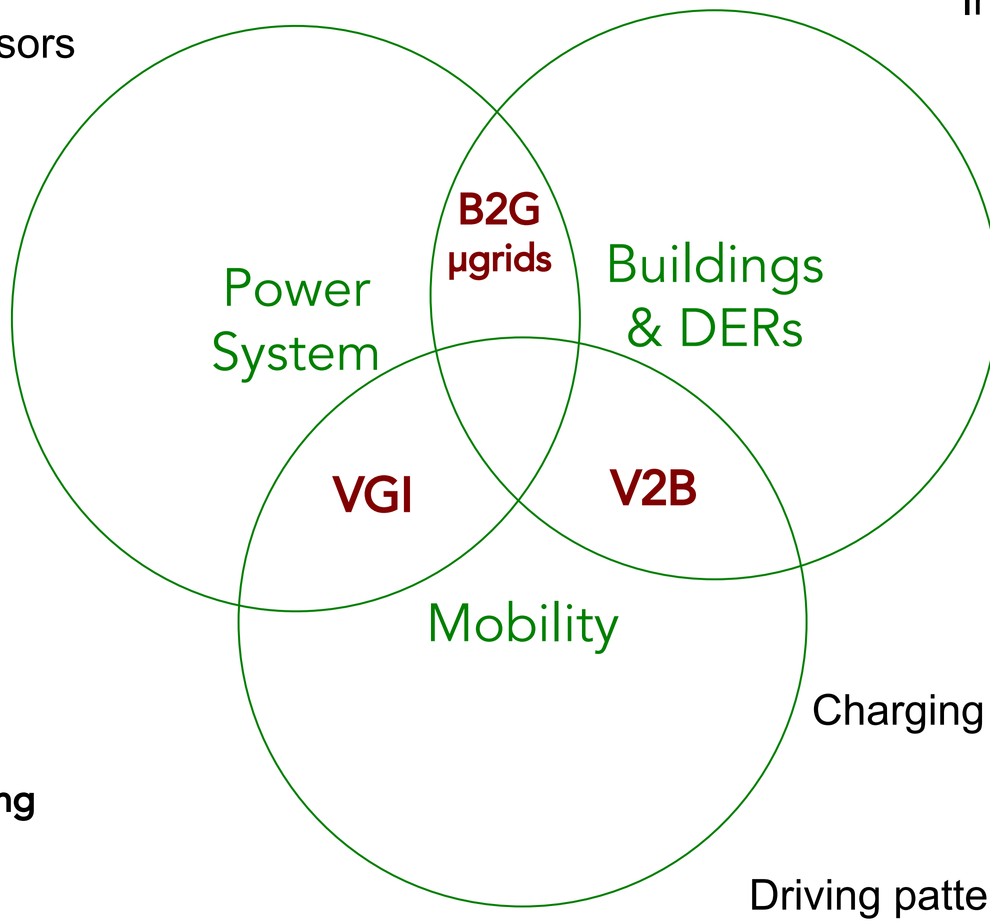
Micro-synchrophasors

Line sensors

Interval meters

Smart meters

Trend logs from building management Systems (BMS)



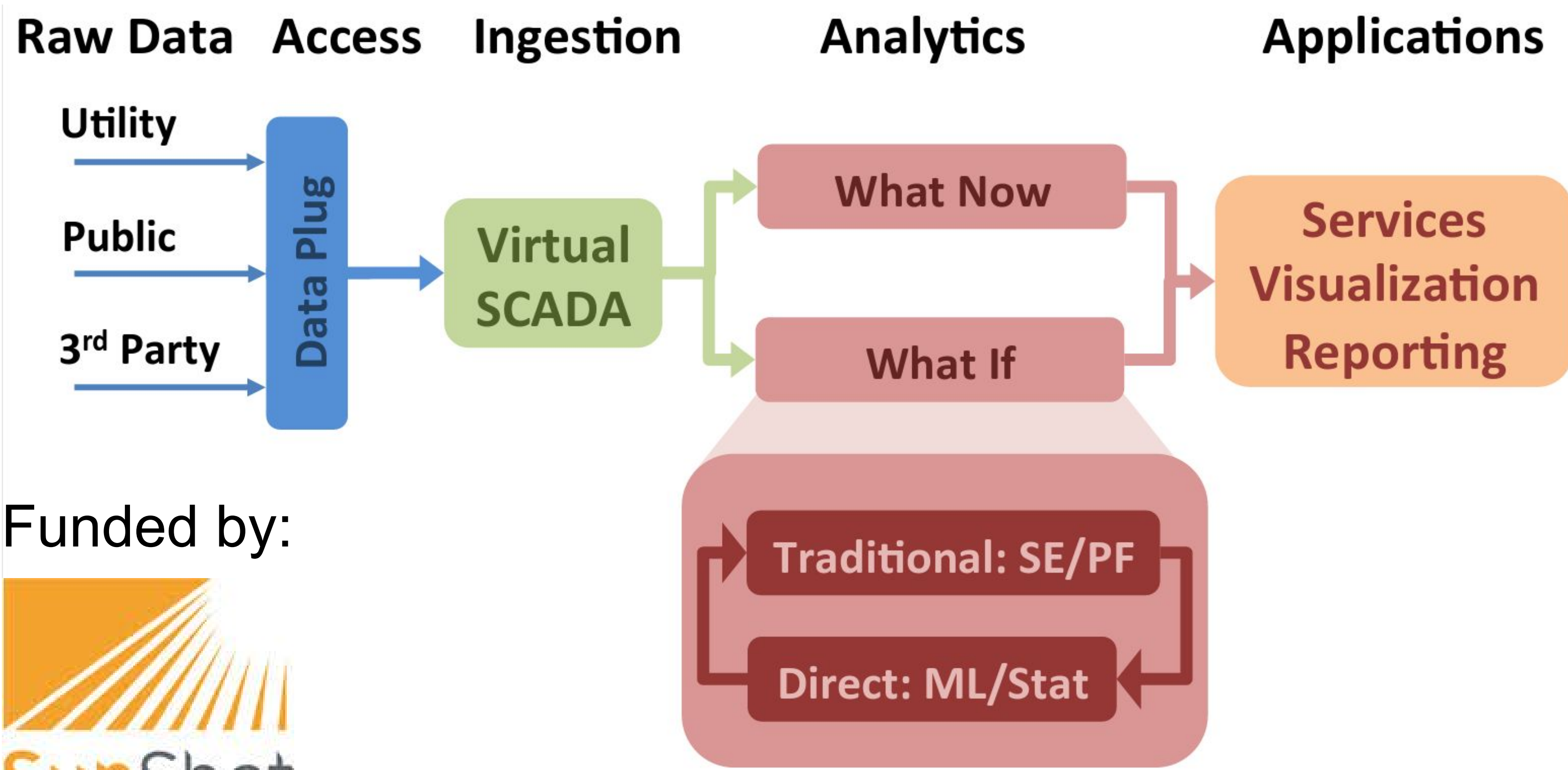
Charging sessions

Driving patterns

B2G – Building to Grid
μgrids – Microgrids
V2B – Vehicle to Building
VGI – Vehicle to Grid

VADER Introduction

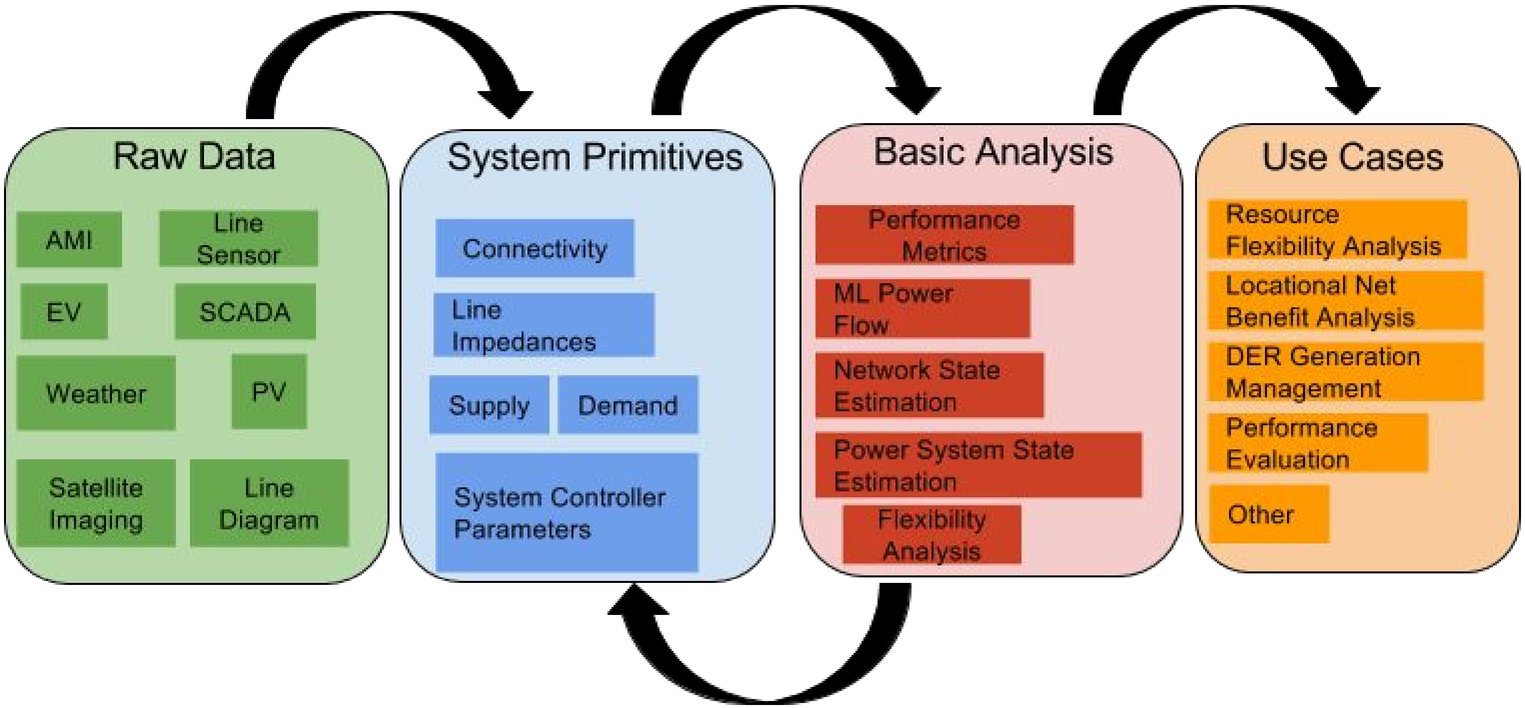
Visualization and Analytics of Distributed Energy Resources



Funded by:

SunShot
U.S. Department of Energy

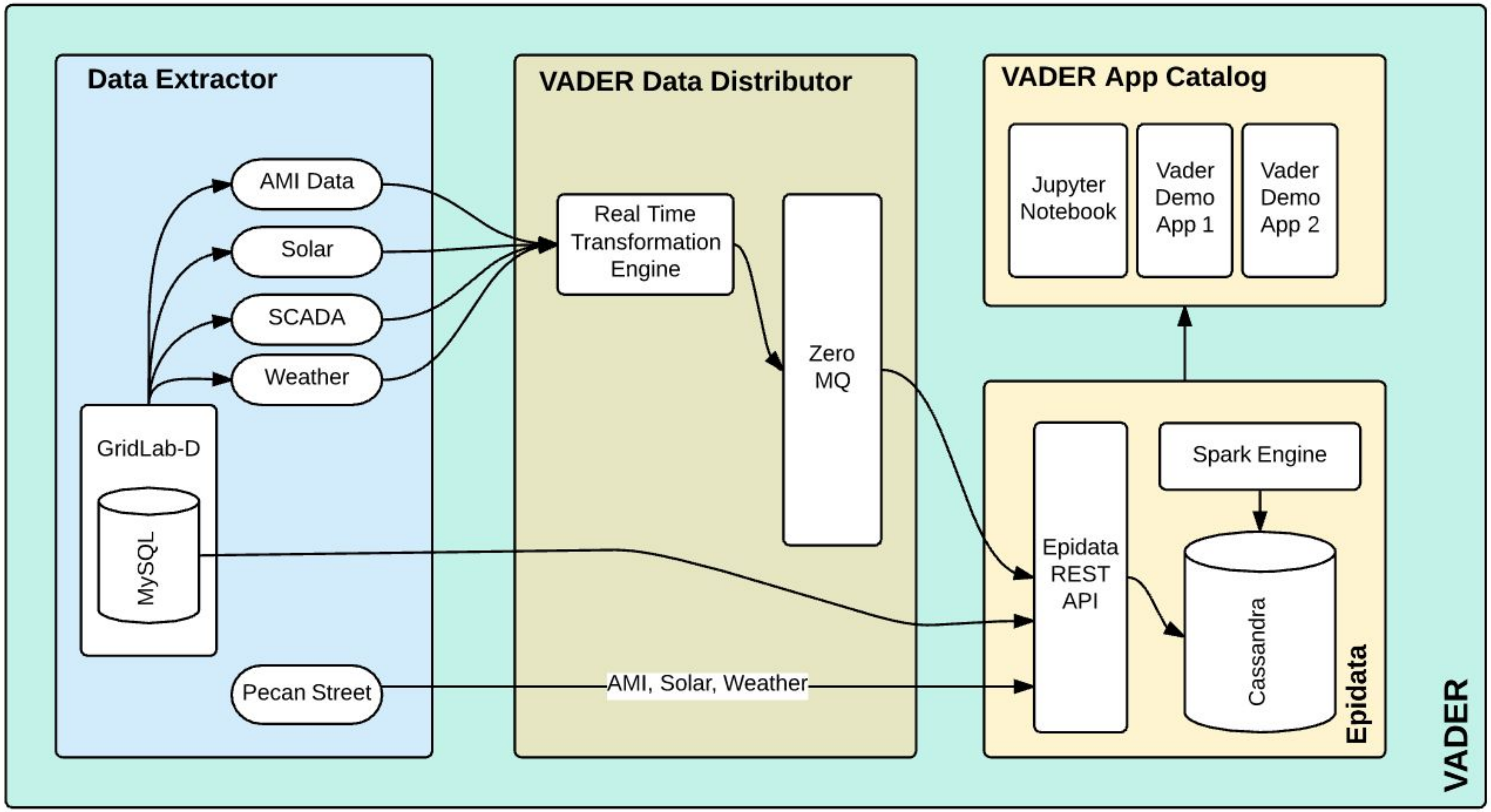
VADER Analytics



VADER key analytics pieces:

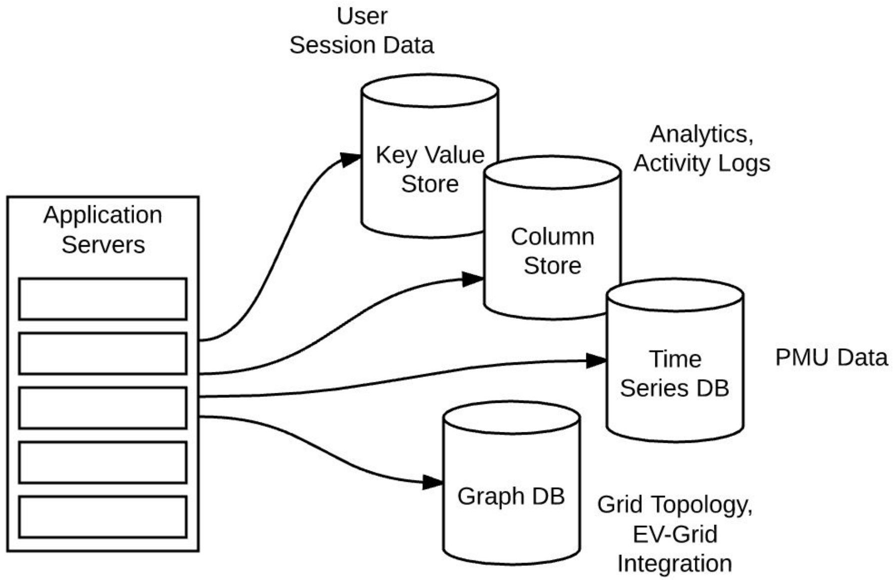
- Solar disaggregation
- Switch configuration
- ML-based power flow analysis
- Network topology detection
- Load forecasting

VADER Infrastructure

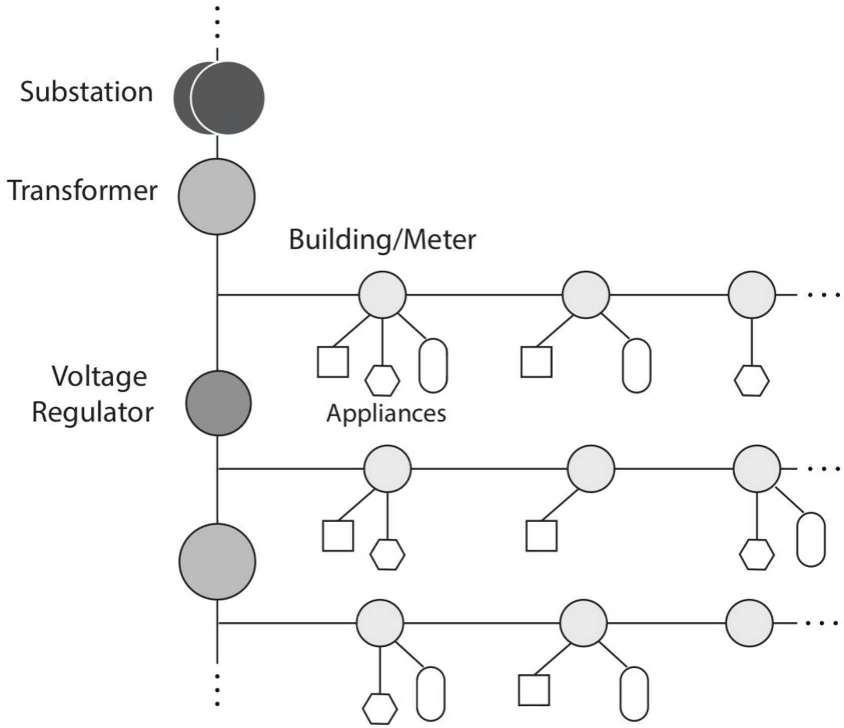


VADER Infrastructure

- Polyglot persistence:
  - Why?



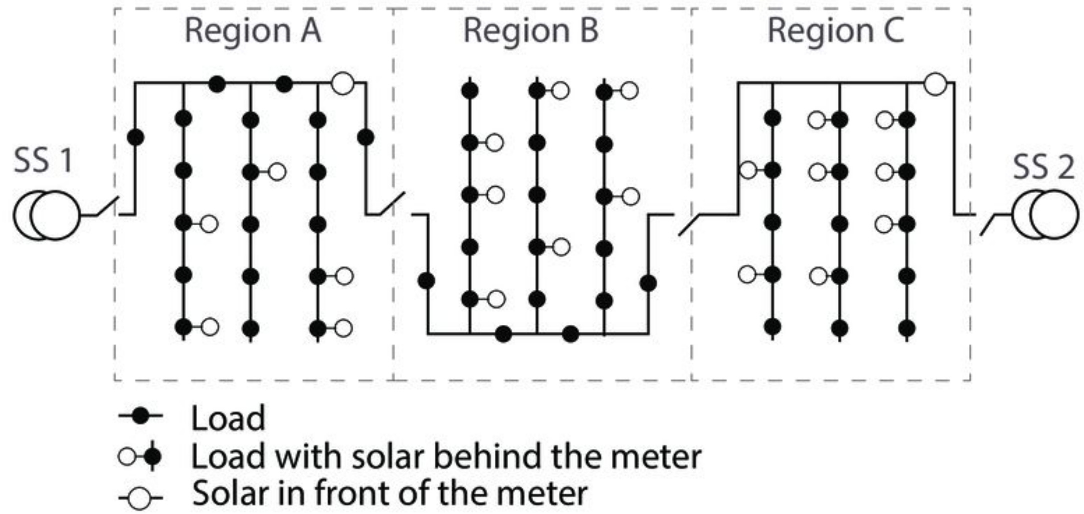
Polyglot Persistence Architecture



GISMO Grid Integration, Systems & Mobility

Introduction

- Increasing solar penetration
  - Behind-the-meter
  - Distribution-level
- Switches maintain a radial structure
- Load is masked
- Visibility into behind-the-meter solar generation is limited



How do we gain more visibility into the load and solar generation?

Introduction

- Traditional approaches include:
  - Physics-based models:
    - Geometry of the array
    - Nameplate capacity
    - Site-specific irradiance measurements (diffuse horizontal, direct normal, and global horizontal irradiance)
      - Costly to obtain DHI and DNI often only GHI is available

Can we use existing measurements that are collected by the utilities to estimate behind-the-meter solar in real-time?
If so, how accurate these estimates would be?

    - Could be costly to do in real-time.

Problem formulation

- Contextually supervised source separation*:
  - Observing an aggregate measurement of signals of interest
  - Each signal can be represented as a linear model
  - Contextual knowledge about signal characteristics
  - Used in Non-intrusive Load Monitoring

*Wytock, Matt, and J. Zico Kolter. "Contextually Supervised Source Separation with Application to Energy Disaggregation." *AAAI*. 2014.

Regularization

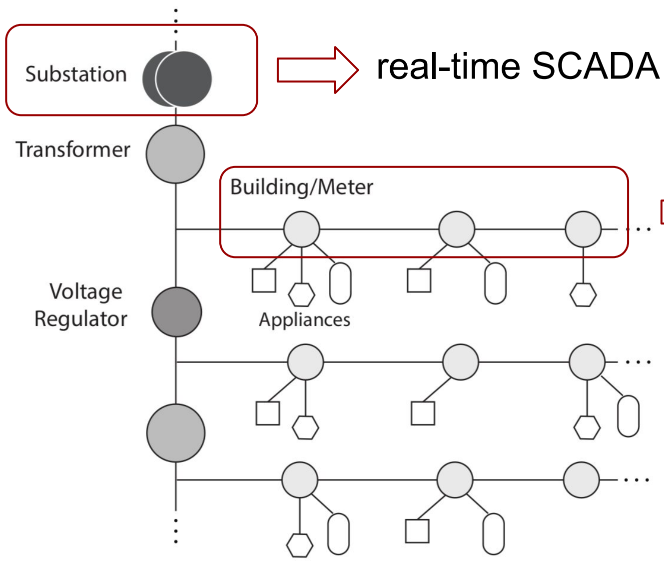
$$\underset{Y_i,\Theta_i}{\text{minimize}} \quad \{\alpha_{\mathbf{i}} \ell_i(\underbrace{(Y_i - X_i\Theta_i)}_{\text{Error between the estimates of the signal and what the linear model would do}}) + \eta_i g_i(Y_i) + \gamma_i h_i(\Theta_i)\}$$

$$\text{subject to} \quad Y_{agg} = \sum_{i=0}^{L} Y_i$$

Contextual knowledge on the shape of individual signals

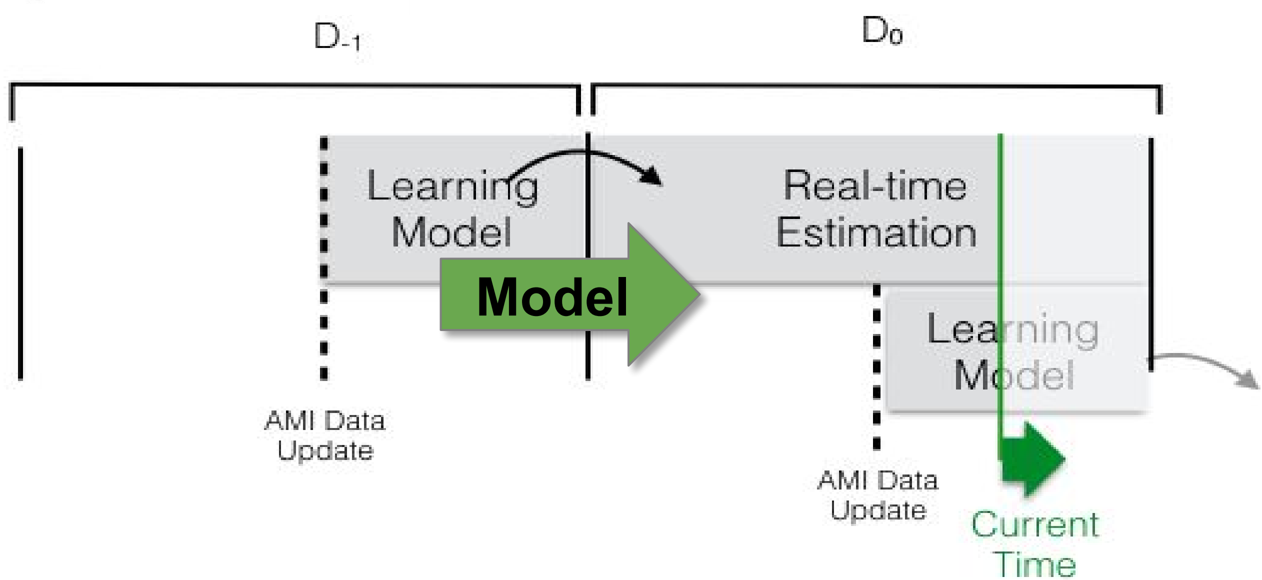
GISMO Grid Integration, Systems & Mobility

Problem formulation



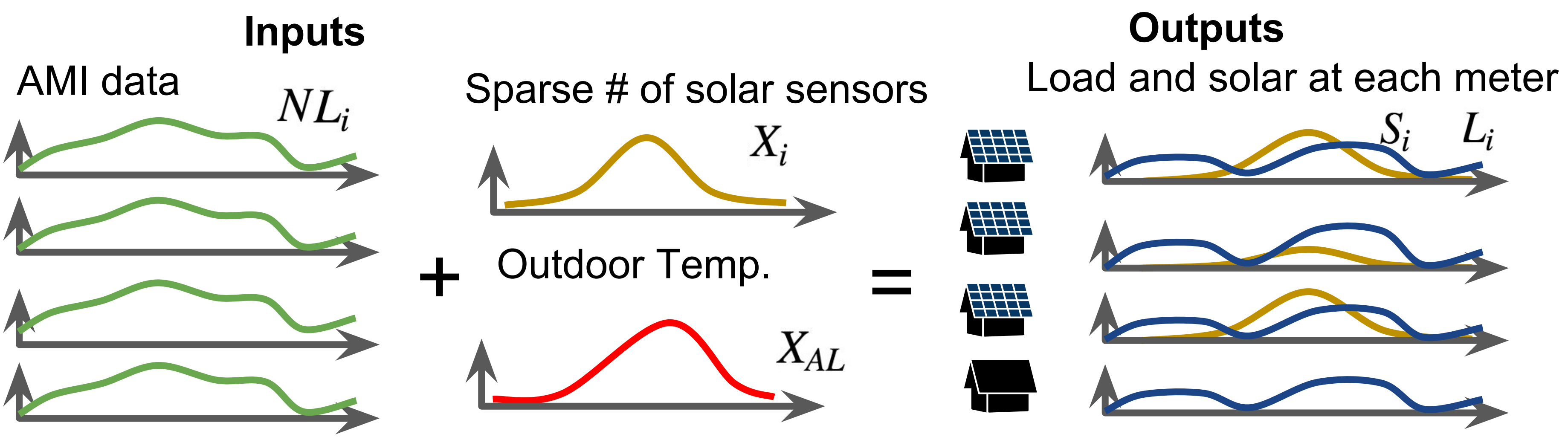
real-time SCADA measurements (typically 4 seconds)

AMI: overnight updates
(1-min to 15-min sampling rate)



Problem formulation

Day-ahead training problem (learning the model):



Problem formulation

Need to be tuned

$$\underset{L, S, \theta}{\text{minimize}} \quad \boxed{\alpha_{AL}} \, ||AL - \theta_{AL} X_{AL}||_2 + \sum_{i=1}^{N} \boxed{\alpha_i} \, ||S_i - \theta_i X_i||_2$$

$$\text{subject to} \quad L_i + S_i = NL_i, \qquad \forall i$$

$$S_i \leq 0 \qquad \forall i,$$

$$L_i \geq 0 \qquad \forall i,$$

$$\boxed{\sum_{i=1}^{N} L_i = AL}$$

Aggregate load is much more predictable

$$\theta_i \geq 0.$$

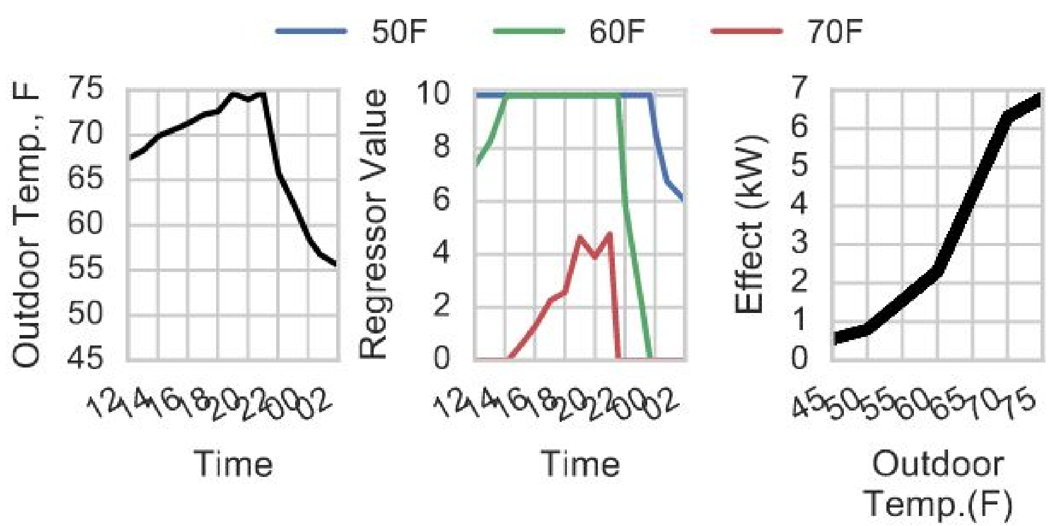
GISMO Grid Integration, Systems & Mobility

Problem formulation

Creation of aggregate load regressors:

$$T_t^{(LB)} = \begin{cases} 0 & T_t < LB \\ 10 & T_t > LB + 10 \\ T_t - LB & LB \le T_t \le LB + 10 \end{cases}$$

	Regressors, X_{AL}			
$Temp_t$	Int	$T_t^{(60)}$	$T_t^{(70)}$	$T_t^{(80)}$
66	1	6	0	0
68	1	8	0	0
70	1	10	0	0
72	1	10	2	0
74	1	10	4	0
80	1	10	10	0
82	1	10	10	2
84	1	10	10	4



Tuning alphas

- Intuitive choice: the inverse of the expected variance of model's errors
- During nighttime there is no solar generation
  - Net load = Aggregate Load
  - Estimate the error for the AL model
- During day time we can estimate the total error.
- Assuming the all errors are independent:

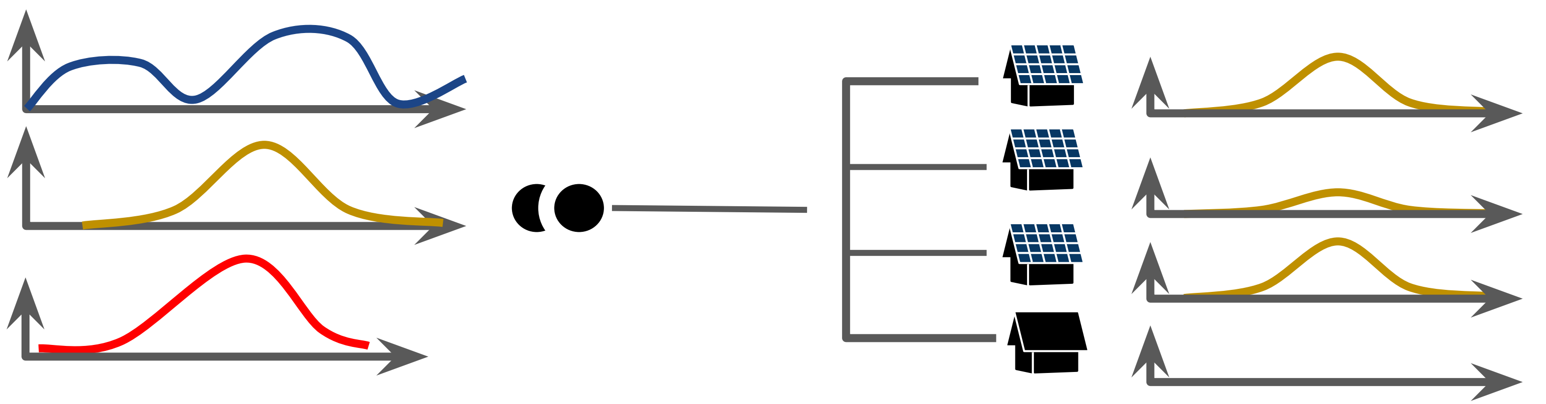
$$Var(\epsilon_{AL}) = Var(\epsilon_{Total}) - \sum_{i=1}^{N} Var(\epsilon_i)$$

Problem formulation

Real-time estimation problem:

Input
Data from load aggregation point (substation)
Solar from sparse # of sensors
Outdoor Temp.

Outputs
Solar at each meter



Problem formulation

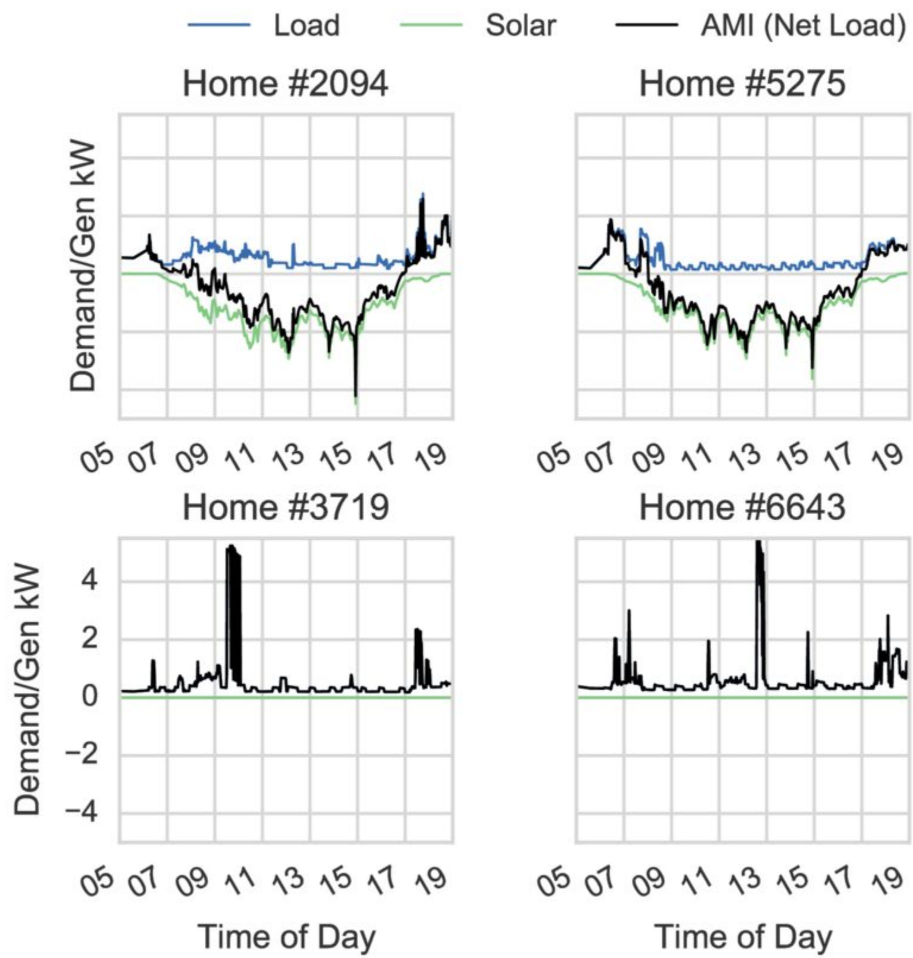
$$
\begin{aligned}
\underset{AL, S_i}{\text{minimize}} \quad & \alpha_{AL} ||AL - \theta_{AL} X_{AL}||_2 + \sum_{i=1}^{N} \alpha_i ||S_i - \theta_i X_i||_2 \\
\text{subject to} \quad & AL + \sum_{i=1}^{N} S_i = \boxed{\sum_{i=1}^{N} NL_i,} \\
& S_i \leq 0 \qquad \forall i.
\end{aligned}
$$

Observed through SCADA

Dataset

Subset of Pecan Street dataset:

- Load and solar sub-metered at 1-minute resolution
- 110 households in TX
- Total duration 7-days
- Divided into:
  - 55 homes with solar
  - 55 homes without solar
  - Varying number of distribution-level solar (used as proxy)



GISMO Grid Integration, Systems & Mobility

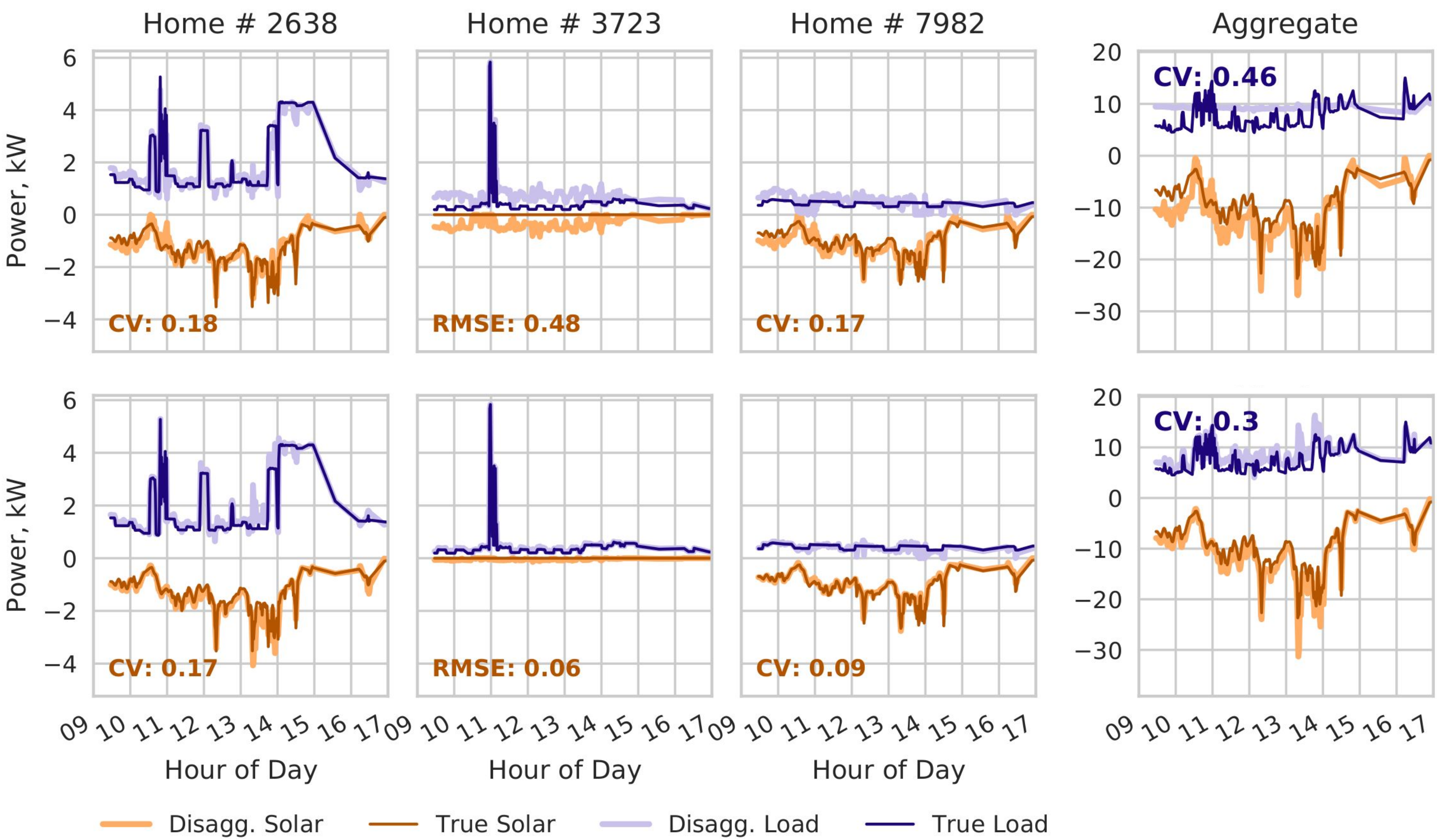
Results

- We use coefficient of variation for homes with solar.
- We report only RMSE for homes without solar.

$$CV = \frac{\sqrt{\sum_{i=1}^{T}(truth_t - estimate_t)^2}}{\frac{1}{T}\sum_{i=1}^{T} truth_t}$$

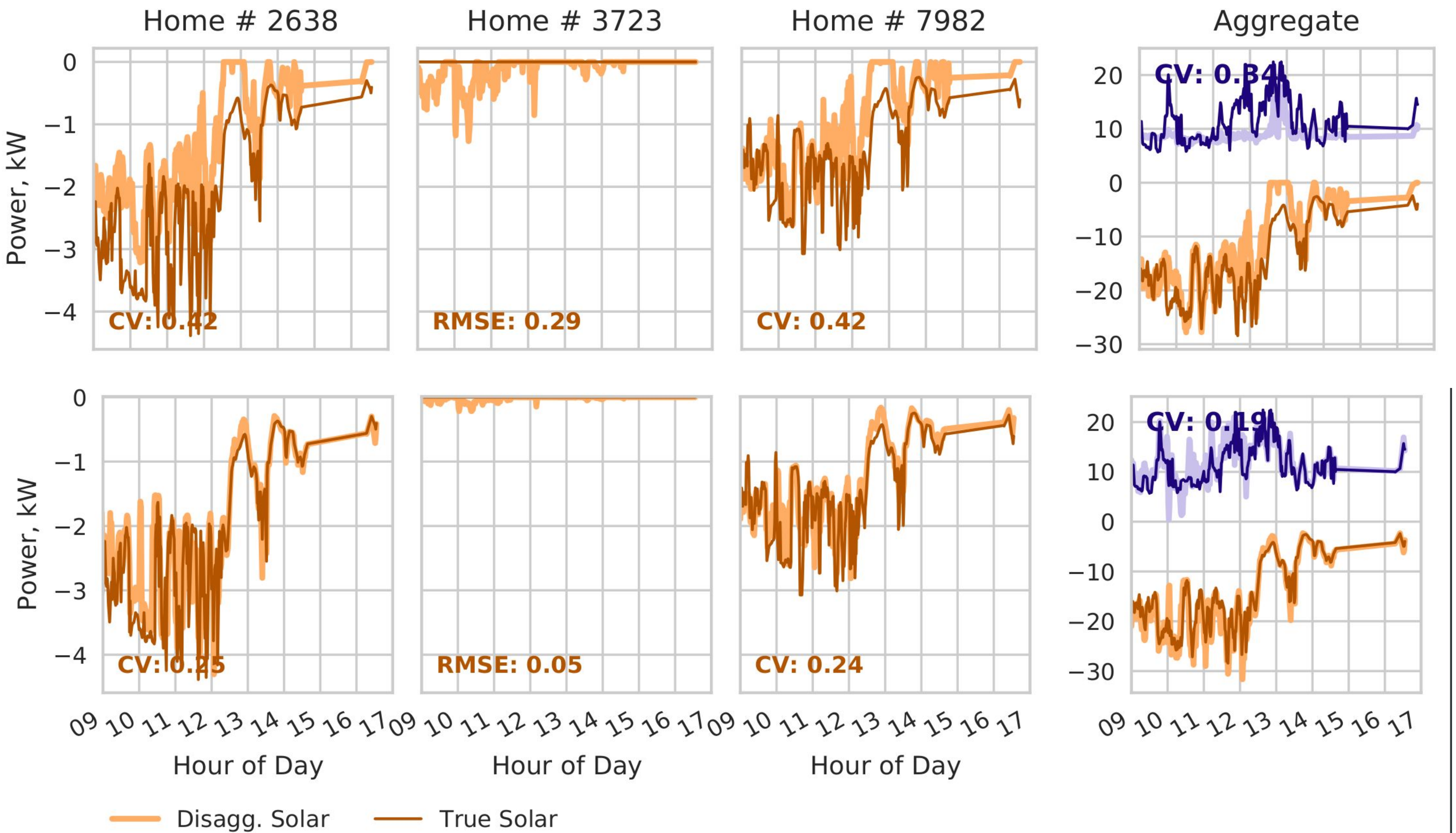
$$= \frac{RMSE}{Mean\ truth}$$

Training Results (Untuned vs. Tuned)



GISMO Grid Integration, Systems & Mobility

Real-time Results (Untuned vs. Tuned)



Conclusions & Future Work

- Useful tool to estimate behind-the-meter solar in real-time.
  - Can also be used to select the optimal number of proxies and their locations for a guaranteed performance of estimation.
- Incorporate losses in the network
- Expanding the disaggregation strategy to incorporate storage
- Running the model on a larger network in SCE’s territory through our partnership in VADER.

Thank you!

Michaelangelo Tabone
Post-doctoral Fellow
Stanford University,
SLAC National Accelerator Laboratory
mtabone@stanford.edu

Emre Can Kara
Associate Staff Scientist
SLAC National Accelerator Laboratory
ekara@stanford.edu

Sila Kiliccote
Staff Scientist
SLAC National Accelerator Laboratory
silak@stanford.edu

Additional Slides