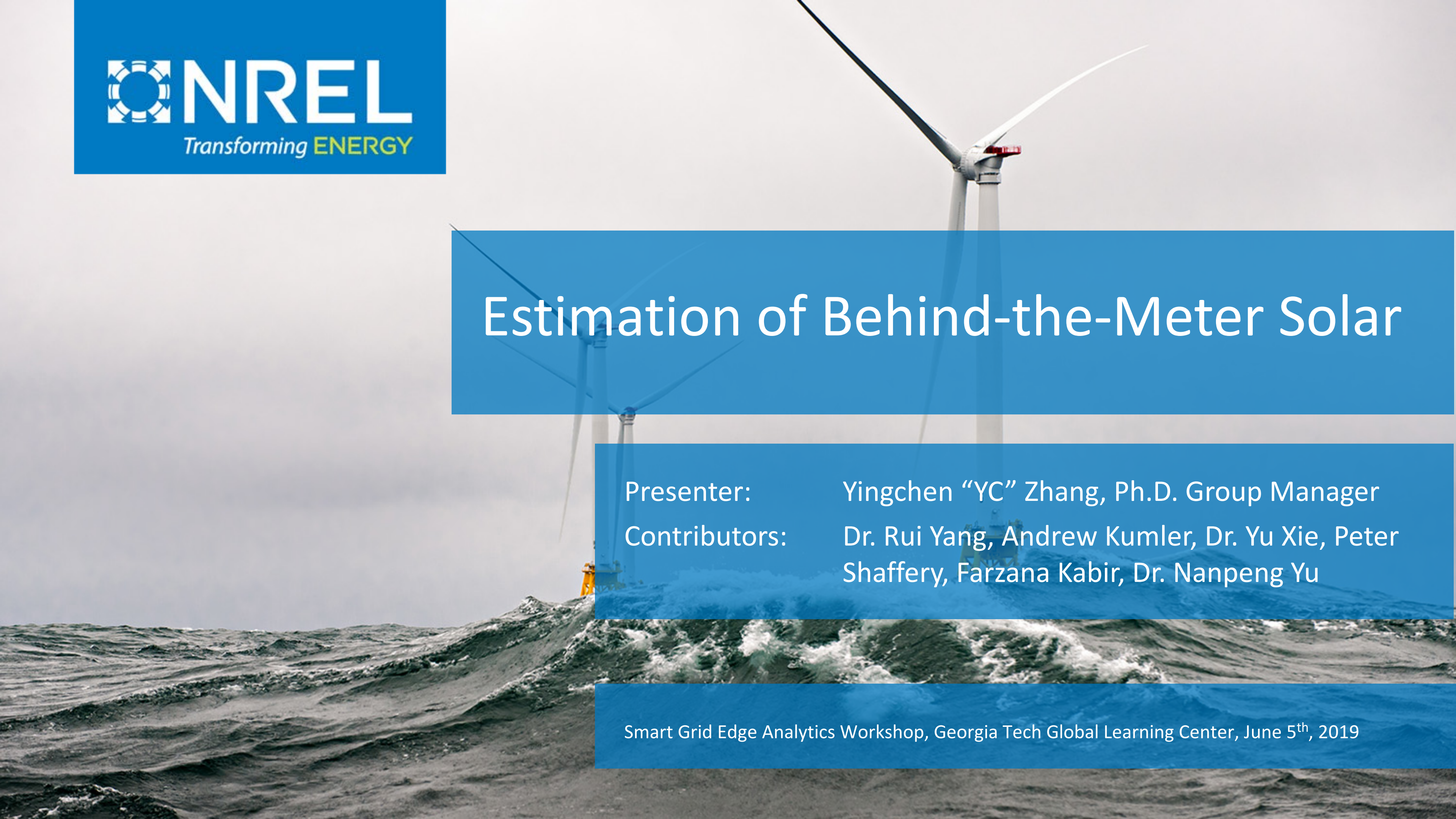


NREL
Transforming ENERGY

Estimation of Behind-the-Meter Solar

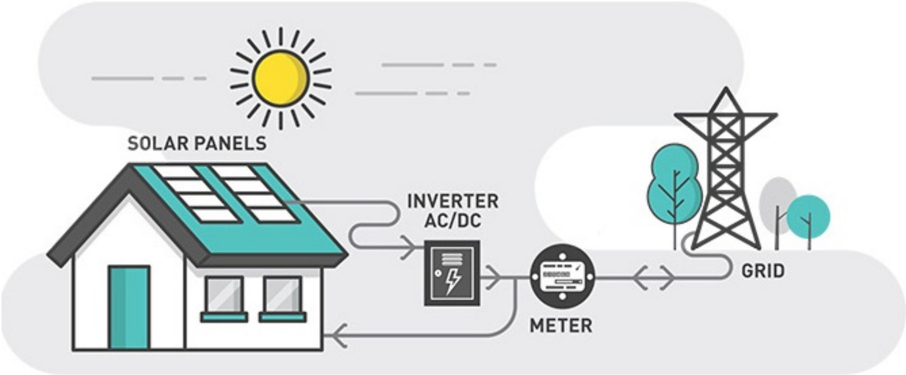
Presenter: Yingchen “YC” Zhang, Ph.D. Group Manager

Contributors: Dr. Rui Yang, Andrew Kumler, Dr. Yu Xie, Peter Shaffery, Farzana Kabir, Dr. Nanpeng Yu

Smart Grid Edge Analytics Workshop, Georgia Tech Global Learning Center, June 5th, 2019

The Problem

- Managing grids with distributed generation (DG) components requires real-time state information
- DG frequently “behind-the-meter”
- Observed net load reflects sum of DG and true, consumption load
- Can we use heterogenous data source (eg. GHI measurements, AMI, SCADA) to estimate behind-meter?



1 A Physics-based Smart Persistence Model

2 Probabilistic Disaggregation at Feeder Level

3 Estimation by Integrating Physical with Statistical Models

1 A Physics-based Smart Persistence Model

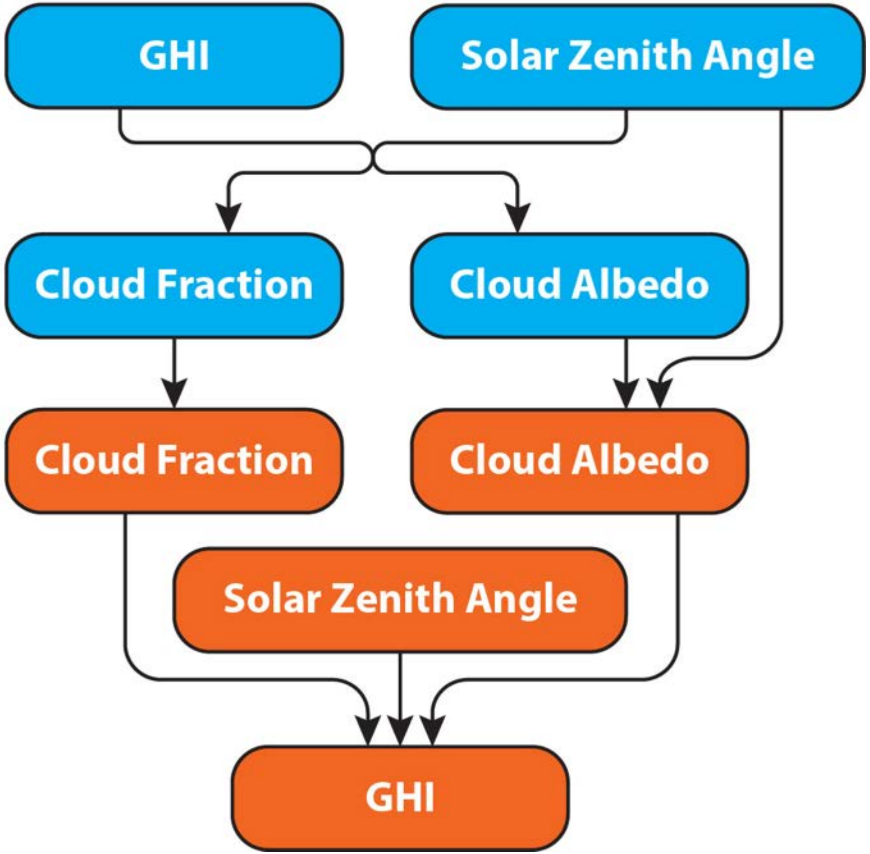
2 Probabilistic Disaggregation at Feeder Level

3 Estimation by Integrating Physical with Statistical Models

Problem(s) formulaiton

- For solar forecasting, clouds are the most difficult problem
  - Type of cloud, duration of cover, etc.
- What type of data can be used
  - Local weather stations
- Computation time
  - Is computation time > forecast horizon?

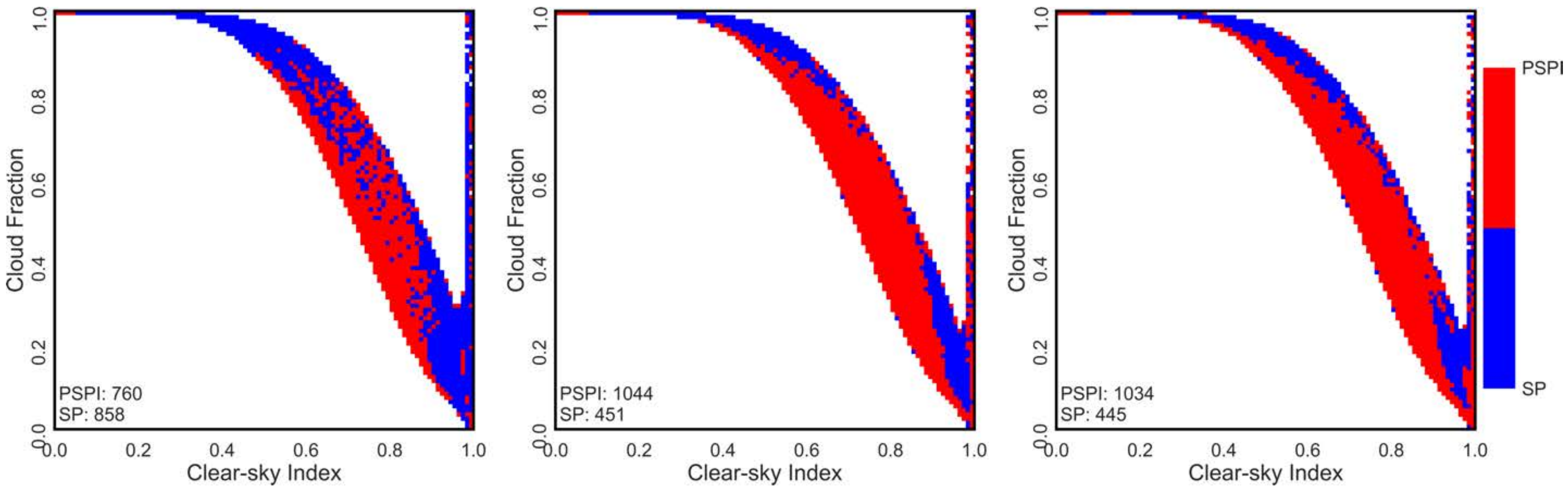
Flowchart for PSPI



Reconstruction of GHI

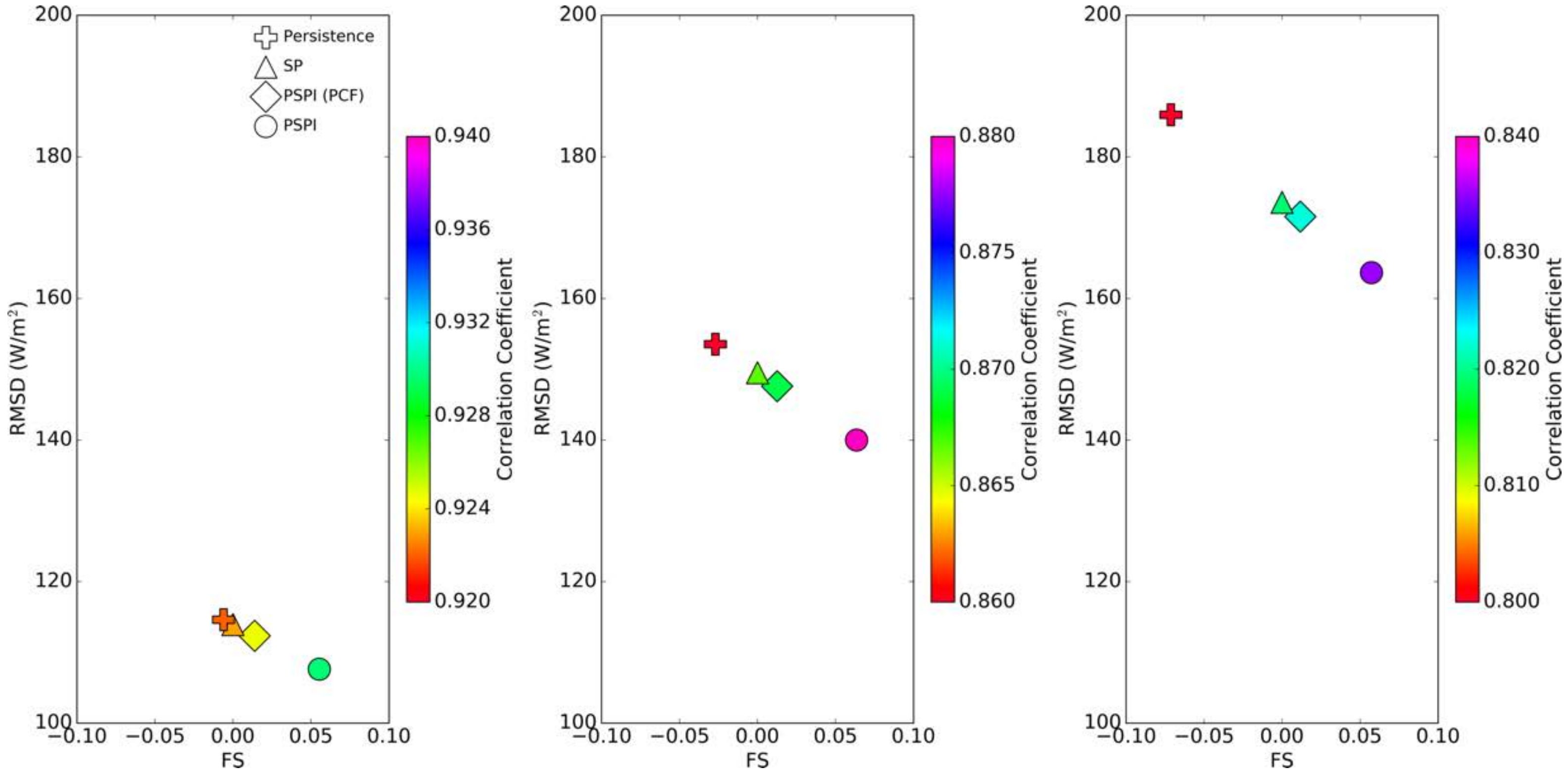
- Before testing the forecasting capabilities of PSPI, PSPI must be able to reconstruct current GHI
- Reconstruction (and thus, forecast) uses simplified atmospheric radiation physics (Xie and Liu (2013))
  - Combines GHI observations, modeled clear-sky variables, and general assumptions about the atmosphere
- Algorithm allows a physics-based representation of GHI without the need to run a entire NWP model

Performance in all-sky conditions



Kumler et al. 2019

Results



Kumler et al. 2019

1 A Physics-based Smart Persistence Model

2 **Probabilistic Disaggregation at Feeder Level**

3 Estimation by Integrating Physical with Statistical Models

Problem(s) formulation

Normally the data available are at the feeder head

Apply “Bayesian Structural Time Series” to disaggregation problem:

- Perform disaggregation probabilistically
- Enables reasoning about uncertainty
- Straightforward, yet flexible, model class

Data

- Pecan Street Austin dataset contains household-level power usage and PV generation data (1-min time resolution, 7 days in both Aug and Jan 2017)
- NSRDB contains GHI and temperature data (30-min time resolution, 1 year total)
- Sum household data to create synthetic feeder data, downsample to 30-min and match to NSRDB

Data

Figure 1: AMI power consumption data for 5 houses in the Pecan Street dataset (January). Global Horizontal Incidence (GHI) overlaid (flipped and scaled) in red.

Data

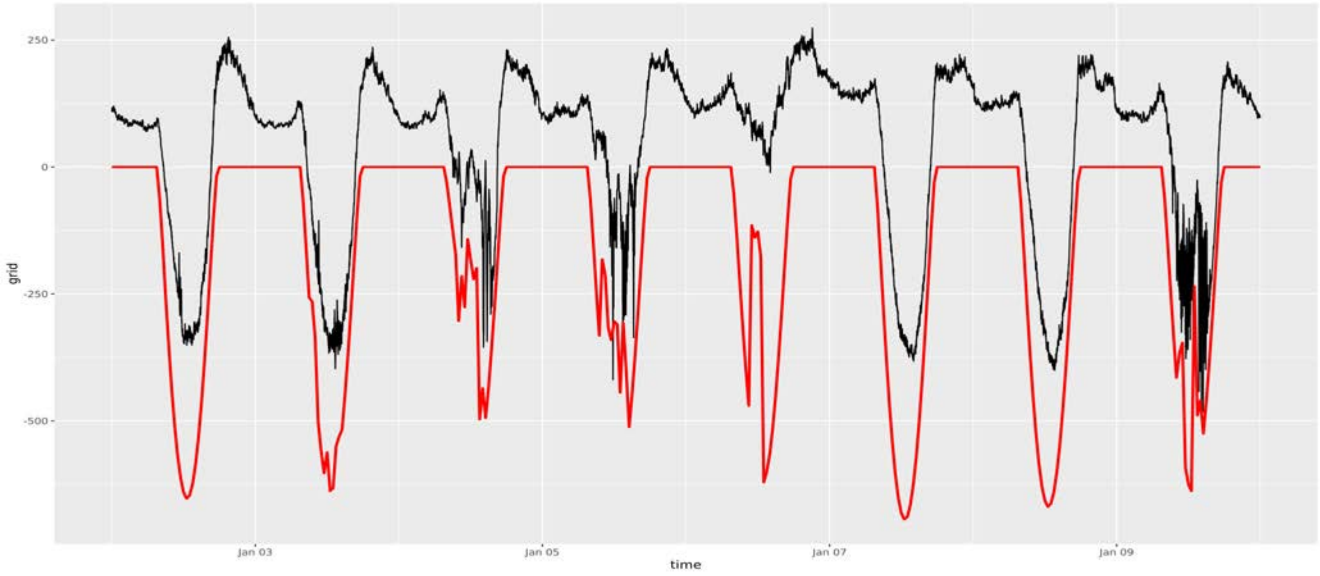


Figure 2: Synthetic feeder measurements of consumption data (summed AMI consumption) for the Pecan Street dataset. GHI again overlaid (flipped and scaled) in red.

Bayesian Structural Time Series

- Formulate a synthetic state space model
- Model structure mimics classic time series model
- Fitting is performed by combining Kalman Filtering and Markov Chain Monte Carlo[1]

1. Scott and Varian. “Predicting the Present with Bayesian Structural Time Series,” June 28, 2013, 21.

Model

Definitions:

Symbol	Def
s_t	(Synthetic) feeder PV gen at time t
l_t	(Synthetic) feeder load at time t
y_t	(Synthetic) feeder measured load
ϕ_t	GHI
X_t	Piecewise-linear temperature covariates

Model

State space model evolves as:

$$s_{t+1} = \beta_{t+1}\phi_{t+1} + \epsilon_{t+1}^{(s)}$$

$$\beta_{t+1} = \beta_t + \epsilon_t^{(\beta)}$$

$$l_{t+1} = X_{t+1}^T\gamma + l_t + \delta_{t+1} + \eta_{t+1}^{(l)}$$

$$\delta_{t+1} = \delta_t + \eta_t^{(\delta)}$$

$$y_{t+1} = s_{t+1} + l_{t+1}$$

$$\text{Where } \epsilon_t^{(\cdot)} \sim N(0, \sigma_\cdot^2) \text{ and } \eta_t^{(\cdot)} \sim T_{\nu_\cdot}(0, \sigma_\cdot^2)$$

Model

Can be brought into conventional Kalman Filtering format by setting:

$$\chi_t = [s_t, \beta_t, 1, l_t, \delta_t]^T$$

Then the state space evolution can be rewritten:

$$\chi_{t+1} = Z_t(\gamma)\chi_t + \omega_t$$
$$y_{t+1} = A\chi_{t+1}$$

Where:

$$A^T = [1, 0, 1, 0, 0] \qquad Z_t(\gamma) = \begin{bmatrix} 0, \phi_t, 0, 0, 0 \\ 0, 1, 0, 0, 0 \\ 0, 0, X_t^T\gamma, 1, 1 \\ 0, 0, 0, 0, 1 \end{bmatrix}$$

BSTS Output

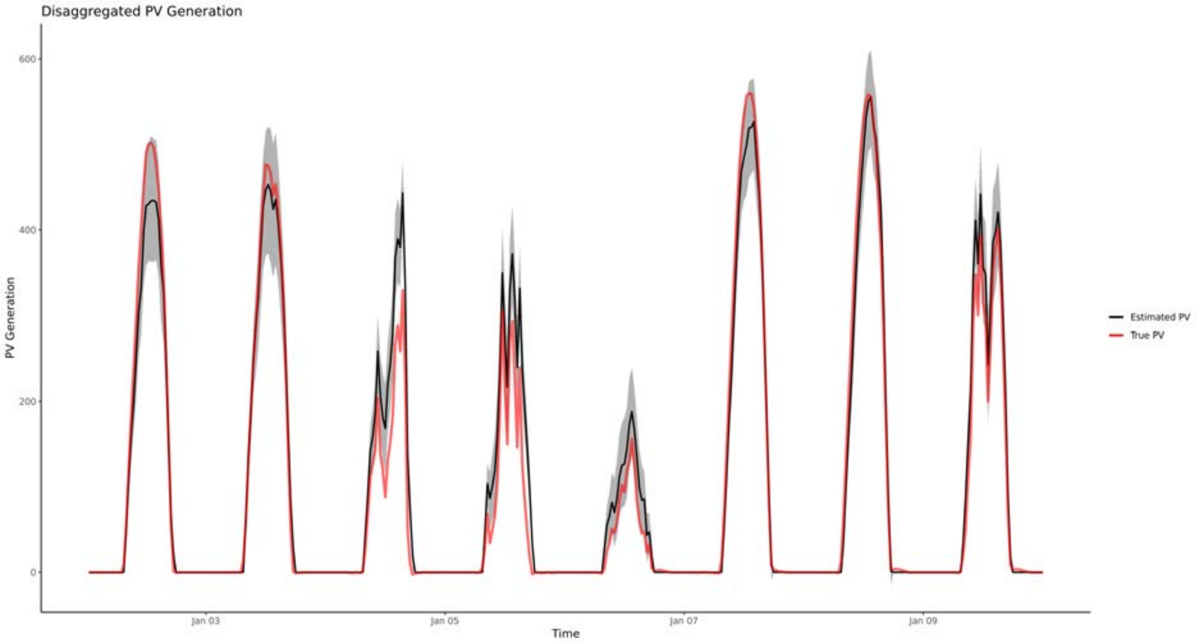


Figure 3: Estimated PV generation occurring over 7 days (black) with 95% credible intervals (gray) against true generation (red)

BSTS Output

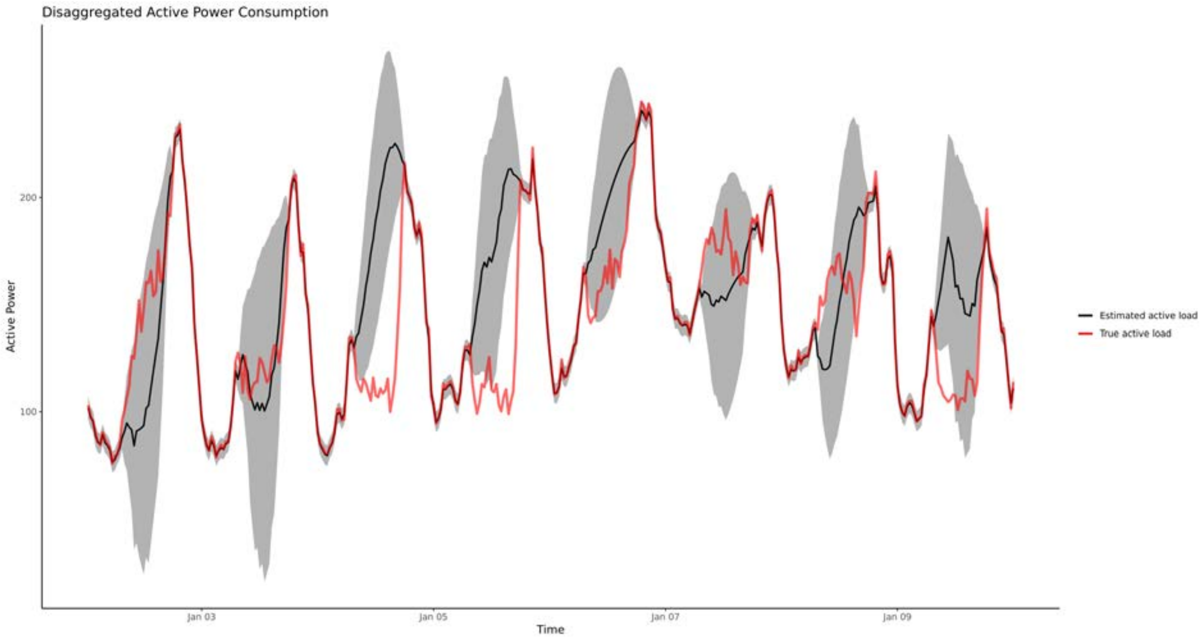


Figure 4: Estimated true load over 7 days (black) with 95% credible intervals (gray) against true generation (red)

1 A Physics-based Smart Persistence Model

2 Probabilistic Disaggregation at Feeder Level

3 **Estimation by Integrating Physical with Statistical Models**

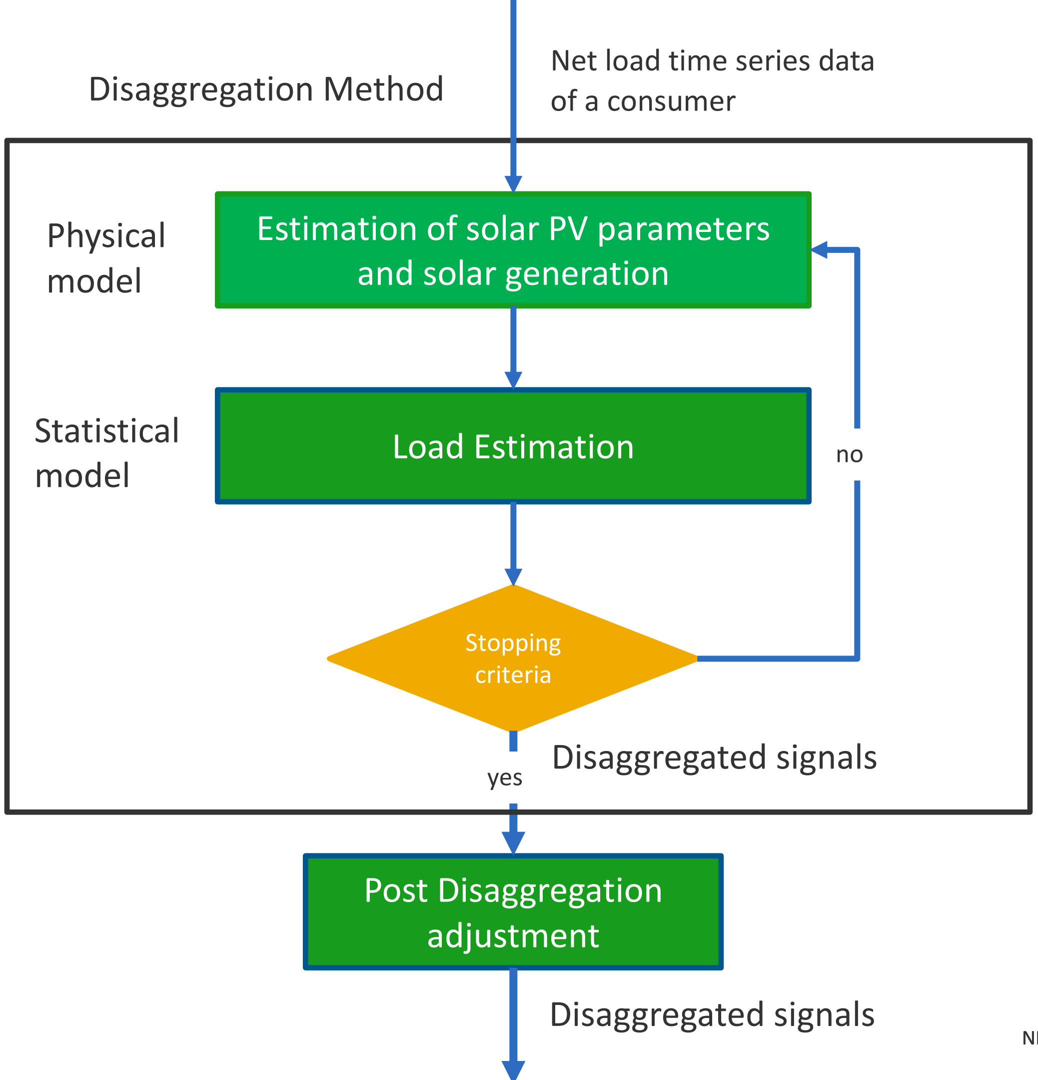
Objective

- From definition of net load

$$NL_t = L_t - S_t; \quad L_t \geq 0, \qquad S_t \geq 0, \forall t$$

- For each residential customer with solar PV installation- disaggregate net load measurement NL_t at each time t into-
  - Load (L_t)
  - Solar generation (S_t)
- Integrate a physical PV system performance model and a statistical load estimation model
- Following information are not available-
  - Historical load and PV generation data
  - Solar panel configuration and parameters (DC size, tilt, azimuth, loss of the PV array and nominal efficiency of the inverter)
  - Exact location of each customer (city's approximate longitude and latitude work as proxy)

Overall Framework



Technical methods

Estimation of solar generation (S)

- Estimation of solar PV parameters θ_S
  - Perform a constrained numerical optimization
  - Solar PV parameters θ_S
    - DC size (P_{dc0})
    - Tilt (θ_t)
    - Azimuth (θ_{az})
    - Loss (l)
    - Nominal efficiency (η_{nom})
- Physical PV system performance model g
  - Estimate solar generation $S = g(\boldsymbol{\theta}_S)$

Estimation of Load (L)

- Statistical hidden Markov model regression

Estimation of Solar PV System Parameters

- Assume the following are available-
  - an estimate of solar generation, S
  - PV system performance model, g
- Solve the constrained nonlinear numerical optimization problem for each customer

$$\underset{\theta_S}{\operatorname{argmin}} \sum_{t=1}^{T} \left(S_t - g_t(\boldsymbol{\theta}_S)\right)^2$$

$$\text{subject to } S_t \geq 0, \boldsymbol{\theta}_{S,min} \leq \boldsymbol{\theta}_S \leq \boldsymbol{\theta}_{S,max}$$

PV System Performance Model

- Computes AC output power P_{ac} (solar generation S) of the PV array when θ_S is known
- Based on PV system performance collaborative (SANDIA)[4] and PVWatts[5](NREL)
- Calculation: $P_{ac} = g(\boldsymbol{\theta}_S) = \eta(\eta_{nom}, P_{dc})P_{dc}$

- $\eta \rightarrow Efficincy\ of\ inverter,\ P_{dc} \rightarrow DC\ output\ power\ of\ the\ PV\ array$

$$P_{dc} = g'(P_{dc0}, \theta_t, \theta_{az}, l) = (1-l)\times\frac{E_{tr}(\theta_t, \theta_{az})}{E_0}P_{dc0}[1+\gamma(T_c(\theta_t, \theta_{az}) - T_0)]$$

- E_0, T_0, γ are known values

[4]Stein, J. S. (2012, June). The photovoltaic performance modeling collaborative (PVPMC). In 2012 38th IEEE Photovoltaic Specialists Conference (pp. 003048-003052). IEEE.
[5]Dobos, A. P. (2014). *PVWatts version 5 manual* (No. NREL/TP-6A20-62641). National Renewable Energy Lab.(NREL), Golden, CO (United States).

PV System Performance Model

- Calculate operating Cell Temperature T_c from Sandia cell temperature model using

$$T_c = T_m + \frac{E_{POA}}{E_0} \Delta T$$

- Sadia module temperature model

$$T_m = E_{POA} \times e^{a+b\times WS} + T_a$$

- Calculate plane of array and transmitted irradiance $(E_{POA}\ and\ E_{tr})$
  - Solar irradiance data $(DNI, DHI\ and\ GHI)$
  - Solar PV installation geometry (PV array tilt and azimuth angle)
  - Solar position data (Solar zenith and azimuth angle → can be calculated from solar position algorithms)

Load Estimation

- No historical load data → Time series models not applicable
- Linear regression models are common
- Explanatory variables:

Variables	Symbols
Temperature (3rd degree polynomial)	c, c^2, c^3
Hour of the day (3rd degree polynomial)	h, h^2, h^3
Weighted moving average of last 24 hours temperature	c_{wmv}
Day of the week	$d = \begin{cases} 1 & if\ weekend \\ 0 & if\ weekday \end{cases}$

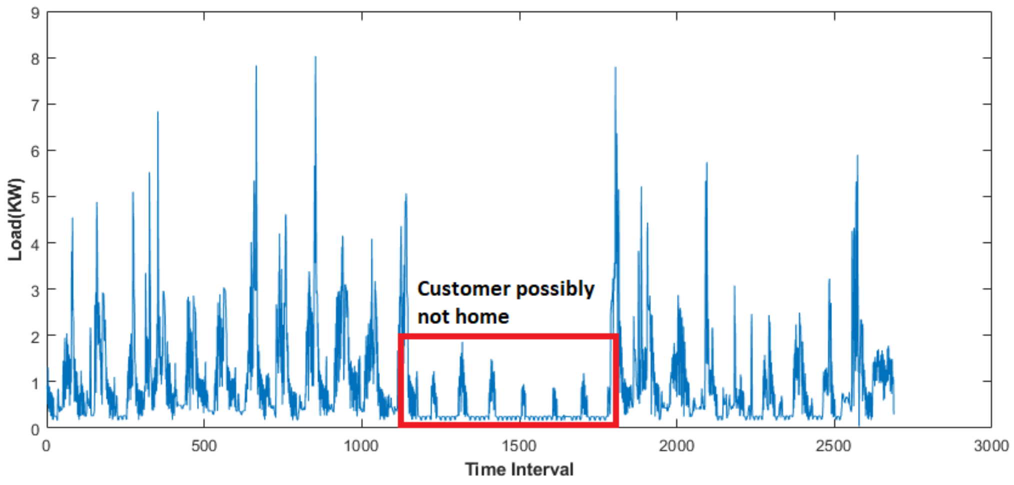


Figure 1: Load time series of a customer

Load Estimation

Load time series can exhibit quite different patterns depending on whether the consumer(s) is present at home or not

Hidden Markov Model Regression

- This change in behavior can be modeled by a hidden Markov model regression[6] (also called Markov switching regression model) given state s_t

$$L_t = \boldsymbol{X}_t^T \boldsymbol{\beta}_{s_t} + \varepsilon_{s_t} \quad \varepsilon_{s_t} \sim N\left(0, \sigma_{s_t}^2\right)$$

- $L_t \rightarrow$ load of a consumer at time t
- $X_t \rightarrow$ explanatory variables at time t
- $s_t = \{s_1, s_2\}$ indicates latent state at time t
- Probability of a change in regime is modeled by a first-order time-invariant two-state Markov chain

$$P\left(s_t = j \mid s_{t-1} = i, s_{t-2} = q, \ldots\right) = P\left(s_t = j \mid s_{t-1} = i\right) = p_{ij}, \qquad \sum_{j=1}^{2} p_{ij} = 1$$

- Can be estimated by maximum likelihood (MS_Regress[7] package of MATLAB used)

[6] Fridman, M. (1994). Hidden markov model regression.
[7] Perlin, M. (2015). MS_Regress-the MATLAB package for Markov regime switching models. *Available at SSRN 1714016*.

Disaggregation Algorithm

Algorithm 1 Algorithm for the disaggregation of net load of each customer and estimation of solar PV parameters

---

Input: Net load of a customer from AMI measurement, **NL**

Output: User consumption $\hat{\mathbf{L}}$, solar generation $\hat{\mathbf{S}}$, and solar PV parameters $\boldsymbol{\theta}_S$

Initialization: Determine M initial solar PV system technical parameters $(\boldsymbol{\theta}_S)_1^{(0)}, \ldots, (\boldsymbol{\theta}_S)_M^{(0)}$

for each starting point $m \in M$ **do**
- Initialize solar generation, $\hat{\mathbf{S}}_m^{(0)} = g\left((\boldsymbol{\theta}_S)_m^{(0)}\right)$
- **for** j=1 to maxiter **do**
  - Estimate user consumption, $\hat{\mathbf{L}}_m^{(j)} = \mathbf{NL} + \hat{\mathbf{S}}_m^{(j-1)}$
  - Fit HMM regression model, denoted by $f(\boldsymbol{X}, \boldsymbol{\theta}_L)$, to $\hat{\mathbf{L}}_m^{(j)}$ and calculate parameters $(\boldsymbol{\theta}_L)_m^{(j)}$
  - Update user consumption, $\hat{\mathbf{L}}_m^{(j)} = f\left(\boldsymbol{X}, (\boldsymbol{\theta}_L)_m^{(j)}\right)$
  - Update solar generation, $\hat{\mathbf{S}}_m^{(j)} = \hat{\mathbf{L}}_m^{(j)} - \mathbf{NL}$
  - Determine $(\boldsymbol{\theta}_S)_m^{(j)}$ from Equation (2) using $(\boldsymbol{\theta}_S)_m^{(j-1)}$ as initial value
  - Update solar generation, $\hat{\mathbf{S}}_m^{(j)} = g\left((\boldsymbol{\theta}_S)_m^{(j)}\right)$
  - Estimate net load, $\hat{\mathbf{NL}}_m^{(j)} = \hat{\mathbf{L}}_m^{(j)} - \hat{\mathbf{S}}_m^{(j)}$
  - Calculate MSE of the net load, $E_m^{(j)}$
  - **if** $\left|(\boldsymbol{\theta}_S)_m^{(j)} - (\boldsymbol{\theta}_S)_m^{(j-1)}\right| \leq \varepsilon$ **then**
    - Break
  - **end if**
- **end for**

end for

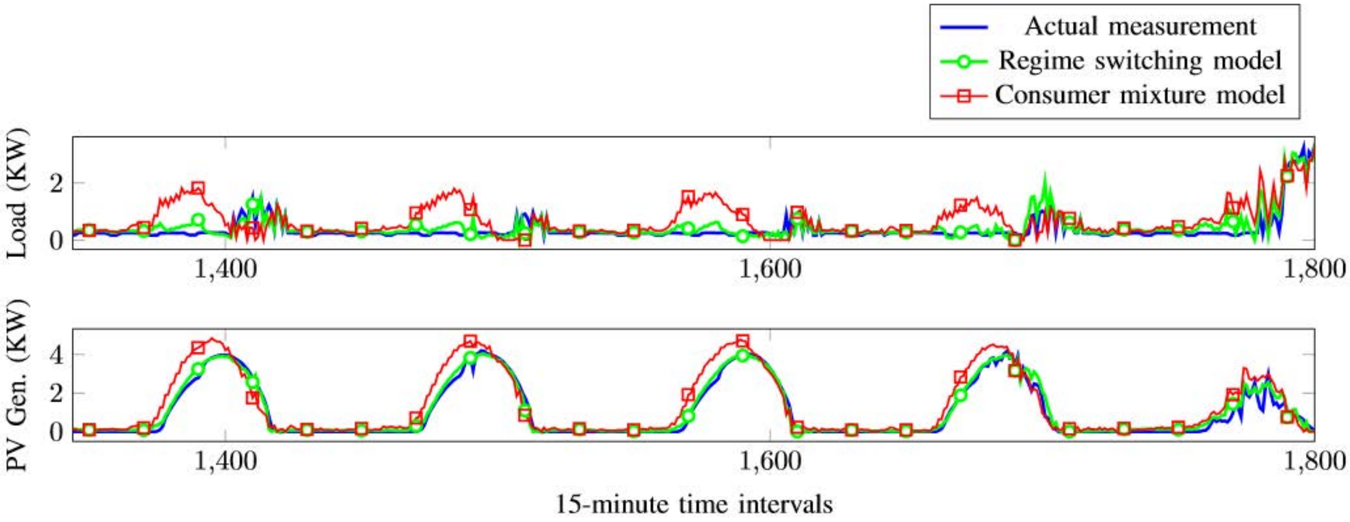
Determine $m^*, j^* = \underset{m,j}{argmin}\, E_m^{(j)}$

return $\hat{\mathbf{L}} = \hat{\mathbf{L}}_{m^*}^{(j^*)}$, $\hat{\mathbf{S}} = \hat{\mathbf{S}}_{m^*}^{(j^*)}$, and $\boldsymbol{\theta}_S = (\boldsymbol{\theta}_S)_{m^*}^{(j^*)}$

Numerical Study

- 15-minute interval data from Pecan Street Dataset[8]
  - Has net load, load, and solar PV generation data
- Customers located in Austin, Texas
- Study period: 10/03/2015-10/30/2015 (28 days)
- 197 consumers with PV installation available
- Solar irradiance and temperature data obtained from National Solar Radiation Database
- Feasible ranges of solar PV system parameters θ_S specified
- 8 initial solar PV system technical sets selected
  - Gradually increase P_{dc0} from 1KW to 8 KW
  - $[\theta_T, \theta_{AZ}, \eta_{nom}, l]$ set at their most common values
- Compared result with consumer mixture model and SunDance model

[8]Holcomb, C. (2012). Pecan street INC.: A test-bed for NILM. In International Workshop on Non-Intrusive Load Monitoring, Pittsburgh, PA, USA.



Result

Comparison of disaggregated load and solar PV generation with actual values for a customer for 5 days from 10/14/2015-10/19/2015

Thank you

www.nrel.gov

A. Kumler, et. al. “A Physics-based Smart Persistent model for Intra-hour forecast of solar radiation (PSPI),” *Solar Energy*, 2018.

F. Kabir, et. al. “Estimation of Behind-the-Meter Solar Generation by Integrating Physical with Statistical Model,” IEEE SmartGridComm, 2019.

P. Shaffery, et. al. “Bayesian Structural Time Series for Distributed Photovoltaic Disaggregation,” in preparation

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy and Office of Electricity. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

NREL Transforming ENERGY