Bayesian Structural Time Series for Behind-the-Meter Photovoltaic Disaggregation

Preprint

Peter Shaffery, Rui Yang, and Yingchen Zhang

National Renewable Energy Laboratory

Bayesian Structural Time Series for Behind-the-Meter Photovoltaic Disaggregation

Preprint

Peter Shaffery, Rui Yang, and Yingchen Zhang

National Renewable Energy Laboratory

NOTICE

This work was authored in part 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 the National Renewable Energy Laboratory's Laboratory Directed Research & Development Program. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Bayesian Structural Time Series for Behind-the-Meter Photovoltaic Disaggregation

Peter Shaffery

National Renewable Energy Laboratory
Golden, Colorado 80401
Email: peter.shaffery@nrel.gov

Rui Yang

National Renewable Energy Laboratory
Golden, Colorado 80401
Email: rui.yang@nrel.gov

Yingchen Zhang

National Renewable Energy Laboratory
Golden, Colorado 80401
Email: yingchen.zhang@nrel.gov

Abstract—**Distributed photovoltaic (PV) generation often occurs "behind the meter": a grid operator can only observe the net load, which is the sum of the gross load and distributed PV generation. This lack of observability poses a challenge to system operation at both bulk level and distribution level. The lack of real-time or near-future disaggregated estimates of gross load and PV generation will lead to over scheduling of energy production and regulation reserves, reliability constraints violations, wear and tear of controller devices, and potentially cascading failures of a system. In this paper we propose the use of a Bayesian Structural Time Series (BSTS) model with local solar irradiance measurements to disaggregate the summed PV generation and gross load signals at a downstream measurement site. BSTSs are a highly expressive model class that blends classic time series models with the powerful Bayesian state space estimation framework. Disaggregation is done probabilistically, which automatically quantifies the uncertainties of the estimated PV generation and gross load consumption. Depending on the data availability in real-time, it can be used to disaggragate PV and gross load at customer site, or can be used at the feeder level. In this paper, we focus on solving the problem at feeder level. We compare the performance of a BSTS model as well as a handful of state-of-the-art methods on a Pecan Street AMI dataset, using the National Solar Radiation Database (NSRDB) to estimate local irradiance.**

I. INTRODUCTION

In the last decade distributed photovoltaic (PV) generation has become a significant source of renewable energy, and its adoption is projected to continue growing for the next several decades [1]. Increased penetration of distributed PV poses new challenges to the operation and control of distribution power grids, which are frequently compounded by the fact that these resources are typically unobserved [2]–[4]. Most of distributed PV systems are located behind-the-meter on customer premises, where utilities do not have direct access to their output but only net load measurements. Real-time estimation of the behind-the-meter PV generation is

crucial to improving the situational awareness and reducing the uncertainties of distributed PV generation in distribution systems. By disaggregating PV generation from net-metered measurements, distribution system operators can improve the forecasting accuracy of net load, reduce the reserve deployment, and optimize the voltage profile.

To this end, a number of methods have been proposed to disaggregate net load measurements into PV generation and gross load components at several scales. These methods can be divided into supervised disaggregation and unsupervised approached. Supervised methods use labelled, pre-disaggregated data to perform disaggregation [5]–[7], while unsupervised methods use only the aggregate signal, structural knowledge of the system (eg. panel configurations), and environmental measurements (eg. solar irradiance or temperature) [8]–[10]. More recently, [11] proposed using a "synthetic" PV generator model combined with a Hidden Markov regression model for gross load to perform disaggregation at the household-level. Similarly, [12] also used a synthetic PV generator model as a component of a probabilistic, load forecasting pipeline for disaggregation at the regional level. Notably, while the forecast is probabilistic, the disaggregation itself is not.

In this paper, we introduce a probabilistic model whose structure mimics well-understood time series models, and for which high-performing fitting algorithms already exist: Bayesian Structural Time Series (BSTS) [13]. In this approach we construct a synthetic state space model, which is designed so that the model has a structural similarity to classic time series models (eg. the AR/I/MA class of models). This methodology fundamentally differs from the aforementioned unsupervised approaches by providing probabilistic estimation of PV generation and gross load consumption. Instead of producing a deterministic or "point" estimate of the disaggregated signals, the BSTS model defines a joint probability density function (PDF) over the gross load and PV generation at each time point, and the dispersion and covariance structure of the joint PDF automatically quantifies the uncertainties associated with the disaggregated PV and load signals.

In addition to quantifying uncertainty, the BSTS model has the advantage of being robust to noise in the input data. The underlying methodology of BSTS disaggregation is to consider all possible disaggregated signal shapes, and weight them according to their plausibility with the model covariates

and net load data. The BSTS approach will therefore only make significant changes to its estimates of PV and gross load when the input data exhibit a change large enough to alter the plausibility of a substantial range of potential disaggregations, reflecting real changes in the system.

These features of the BSTS approach can improve the efficiency of distribution system operations. The joint PDF produced by the BSTS model provides real-time bounds for gross load and PV generation, and also accounts for correlation in the errors between the two quantities, allowing the operator to more precisely determine necessary reserve deployment. Furthermore, robustness in the disaggregation estimate helps prevent the operator from taking unnecessary control actions due to noise in the data. These unique advantages contribute to the novelty of our proposed method.

II. DATA

AMI data from Pecan Street Inc. Dataport was used to assess the performance of our disaggregation model. This data includes household-level power usage and PV generation data at a 1-minute time resolution for 123 houses in the Mueller neighborhood of Austin, TX. We test our proposed method on 7 days in January and August, 2017. As our model relies on local measurements of solar irradiance and ambient temperature to disaggregate PV generation (quantities highly correlated with PV and gross load), we supplemented the Pecan Street data with data from the National Solar Radiation Database (NSRDB). The NSRDB provides irradiance and climate data (temperature, humidity, etc.) at a 30-minute time resolution and a 4 km^2 spatial resolution. To create a synthetic distribution feeder, we summed the household-level Pecan Street time series data, and then downsampled the sum to 30-minute resolution, matching the NSRDB data. We did not include line loss or other grid effects in this feeder model.

III. DISAGGREGATION MODEL

A. Structural Time Series Model

Let P_t denote the observed, feeder-level net power consumption at time t, S_t denote the downstream, distributed PV generation, and L_t denote the gross load consumption. In addition, we denote the solar irradiance metric as ϕ_t, and the $1 \times m$ gross load covariate vector as $\boldsymbol{X}_t$ (see III-B below for an explanation of the temperature covariates). We model these quantities as evolving as follows:

$$
\begin{aligned}
S_t &= \beta_t \phi_t + \epsilon_t^{(s)} \\
\beta_t &= \beta_{t-1} + \epsilon_t^{(\beta)} \\
L_t &= \boldsymbol{X}_t \gamma + L_{t-1} + \epsilon_t^{(l)} \\
P_t &= S_t + L_t
\end{aligned}
\tag{1}
$$

where the random terms $\epsilon_t^{(\cdot)} \sim N(0, \sigma_\cdot^2)$, which in the case of the terms $\epsilon_t^{(s)}$ and $\epsilon_t^{(l)}$ represents evolution in the gross load and PV generation not explained by their respective covariates.

Since the right-hand side of Eq. (1) is linear in the state variables, we can define a "synthetic" state variable $\boldsymbol{y}_t = [S_t, \beta_t, L_t, 1]^T$ and then rewrite (1) into a state space model:

$$
\begin{aligned}
\boldsymbol{y}_t &= \boldsymbol{Z}_t(\gamma)\boldsymbol{y}_{t-1} + \omega_t \\
P_t &= \boldsymbol{A}\boldsymbol{y}_t
\end{aligned}
\tag{2}
$$

where:

$$
A^T = [1, 0, 1, 0]
$$

$$
Z_t(\gamma) = \begin{bmatrix} 0, & \phi_t, & 0, & 0 \\ 0, & 1, & 0, & 0 \\ 0, & 0, & 1, & \boldsymbol{X}_t\gamma \\ 0, & 0, & 0, & 1 \end{bmatrix}
\tag{3}
$$

$$
\omega_t = [\epsilon_t^{(s)}, \epsilon_t^{(\beta)}, \epsilon_t^{(l)}, 0]^T
$$

In addition to specifying the model structure as in (1) or (2), the Bayesian approach also requires setting prior distributions over all unknown parameters. We use the following priors:

$$
\begin{aligned}
S_1 &\sim N(\mu_S, \Sigma_s^2) \\
\beta_1 &\sim N(0, \Sigma_\beta^2) \\
L_1 &\sim N(\mu_L, \Sigma_l) \\
\gamma &\sim N(0, \Sigma_\gamma)
\end{aligned}
\tag{4}
$$

There are several options to handle the unknown $\Sigma_\cdot^2$ in (4), as well as the $\sigma_\cdot^2$ and from (1). The package used here, `bsts` for the R statistical computing language, assigns them empirically-derived inverse-gamma priors and then estimates them simultaneously with the other unknown parameters [14]. However, for simplicity of presentation we will treat them as simply known.

B. Load Model

We model the relationship between gross load and temperature using a lagged temperature model, which reflects that load commitments are often not due to instantaneous ambient temperature. Furthermore, model performance using lagged temperature demonstrated improvements over other temperature models used in previous model iterations, which for brevity are not discussed here [15]. To determine the appropriate temperature lags we randomly designated one day in both August and January as training data, and the remaining six days as testing data. All performance metrics were averaged over all such partitions of training and test data.

Given training temperature data $\tilde{T}_t$ and gross load $\tilde{L}_t$, temperature lags were determined using the cross-correlation function:

$$
\text{CC}(k) = \mathcal{F}^{-1}\left[\mathcal{F}[\tilde{\boldsymbol{T}}]^*(\omega)\mathcal{F}[\tilde{\boldsymbol{L}}](\omega)\right](k)
\tag{5}
$$

Where $\mathcal{F}$ is the Fourier transform. All timepoints $k_1, ..., k_m$ corresponding to local maxima of $|\text{CC}(k)|$ were selected as lags for our model. Define the k-lag operator by:

$$
L_k(T_t) = \begin{cases} 0, & t < k \\ T_{t-k}, & t \geq k \end{cases}
$$

Then the covariate vector $\boldsymbol{X}_t$ was defined as:

$$
\boldsymbol{X}_t = [L_{k_1}(T_t), ..., L_{k_m}(T_t)]
$$

2

C. Model Fitting

Let $\boldsymbol{y} = [y_1, ..., y_t]^T$ and $\boldsymbol{P} = [P_1, ..., P_t]^T$ (with similar vectors defined for all other times series). Given the priors in (4), we can now write the joint posterior distribution over the state space:

$$\Pr[\boldsymbol{y}, \gamma | \boldsymbol{P}] \propto \mathcal{L}(\boldsymbol{P}; \boldsymbol{y}, \gamma) \pi[y, \gamma]$$

$$\Pr[\boldsymbol{S}, \boldsymbol{L}, \boldsymbol{\beta}, \boldsymbol{\delta}, \gamma | \boldsymbol{P}] \propto N(\mu_S, \Sigma_s^2) N(0, \Sigma_\beta^2) N(\mu_L, \Sigma_l^2) \times$$

$$N(0, \Sigma_\gamma^2) \times \qquad (6)$$

$$\prod_{t=1}^{n} [N(\beta_t \phi_t, \sigma_S^2) N(\beta_{t-1}, \sigma_\beta^2) \times$$

$$N(\boldsymbol{X}_t^T \gamma + L_{t-1}, \sigma_L^2)]$$

To estimate the disaggregated $\boldsymbol{S}$ and $\boldsymbol{L}$ we marginalize over the state variables $\boldsymbol{\beta}$ and γ:

$$[\boldsymbol{S}, \boldsymbol{L} | \boldsymbol{P}] = \int [\boldsymbol{S}, \boldsymbol{L}, \boldsymbol{\beta}, \gamma | \boldsymbol{P}] d\boldsymbol{\beta} d\boldsymbol{\delta} d\gamma$$

A point estimate of the disaggregated PV generation and true load can then be obtained by taking the expectation of the joint posterior over $\boldsymbol{S}$ and $\boldsymbol{L}$:

$$(\hat{\boldsymbol{S}}, \hat{\boldsymbol{L}}) = E_{[\boldsymbol{S}, \boldsymbol{L} | \boldsymbol{P}]}[S, L]$$

Uncertainty in this point estimate is represented by the dispersion of $[\boldsymbol{S}, \boldsymbol{L} | \boldsymbol{P}]$ about the expectation. While there are a number of measures of dispersion, the most straightforward choice for this purpose (and the one we will use) is the $p\%$ probability region centered at the mode of $[\boldsymbol{S}, \boldsymbol{L}, \boldsymbol{\beta}, \boldsymbol{\delta}, \gamma | \boldsymbol{P}]$ (ie. the $p\%$ quantile region).

In practice, the functional form of Eq. (6) is too complex to handle analytically. We therefore use Markov Chain Monte Carlo (MCMC) to sample from it, and use the empirical statistics of the samples (mean, quantiles, etc.) as approximations of the true quantities. Given the model (2), we can perform MCMC using a generalization of the Kalman smoother [13], [16].

IV. RESULTS

A. Model Comparisons

We compared the performance of our model to five other state-of-the-art methods for disaggregation which use optimization approaches. We tested the Contextually Supervised Generation Estimation (CSGE) model [8], as well as the four models proposed in [9], which we will refer to as Models A-D (reflecting their labels in the original source). Multiple versions of CSGE have been proposed, however due to data availability constraints we implemented an older iteration found in [17]. Full descriptions of these models can be found in their corresponding citations.

To fit the BSTS model we iterated the sampling algorithm for $N = 10^5$ iterations, using the `bsts` package for the R statistical computing language. Burn-in length was estimated automatically by `bsts`. Fitting the comparison models was performed with the `CVXOPT` and the SciPy `Optimize`

packages for Python. All fitting was performed on a laptop computer (Intel Core i5-4300U with 16GB of RAM). Posterior sampling for the BSTS model took approximately 6 minutes and 38 seconds, while fitting the CSGE model took .7 seconds, and fitting Models A-D took .08 to 82.1 seconds.

We compare the performance of the six models under consideration using the relative Root Mean Square Error (rRMSE). For an estimate $\hat{x}_t$ being compared to true values x_t, this is defined:

$$\text{rRMSE} = \frac{\sqrt{\frac{1}{N} \sum_{t=1}^{N} (\hat{x}_t - x_t)^2}}{\max_t(|x_t|)} \qquad (7)$$

Each model's performance estimating PV generation and gross load is listed in Table I. For our proposed method, rRMSE has been averaged over all training/testing partitions. We see that the BSTS method achieves significantly lower rRMSE than other state-of-the-art methods, and that therefore on average it estimates PV and gross load more closely than the optimization methods. While in January the BSTS model performs only slightly better than CSGE at estimating true PV, it performs substantially better at estimating gross load than any other method.

In addition to these metrics, which are valid for all considered models, we also assess the performance of the BSTS estimates using metrics which are only defined for probabilistic estimates. Specifically, we look at the Prediction Interval Coverage Probability (PICP) [18], which is defined as:

$$\text{PICP} = \frac{1}{N} \sum_{t=1}^{T} \mathbf{1}_t \qquad (8)$$

where $\mathbf{1}_t$ is an indicator variable which is 1 when the target variable is within the bounds of the prediction interval (in this case the 95% quantile of the posterior density), and 0 otherwise. Since our model assumes that PV generation is strongly determined by GHI, when GHI=0 the prediction interval collapses and so PICP for these times will not accurately reflect model performance. We therefore calculated PICP using all time-points, as well as using only time-points where GHI

TABLE I
DISAGGREGATION PERFORMANCE FOR JANUARY AND AUGUST DATASETS

Model	S_t rRMSE (Jan)	S_t rRMSE (Aug)
BSTS	**5.37%**	**9.74%**
CSGE	6.75%	22.36%
Model A	11.54%	15.93%
Model B	10.41%	15.80%
Model C	10.52%	18.54%
Model D	12.07%	15.70%
	L_t rRMSE (Jan)	L_t rRMSE (Aug)
BSTS	**11.77%**	**9.40%**
CSGE	18.43%	21.53%
Model A	24.88%	15.25%
Model B	24.48%	15.17%
Model C	24.01%	30.07%
Model D	26.02%	15.04%

3

TABLE II
BSTS CALIBRATION AND PRECISION BY TARGET VARIABLE

Metric	S_t (Jan)	L_t (Jan)	S_t (Aug)	L_t (Aug)
PICP	24.41%	78.08%	44.96%	85.50%
PICP, daytime	53.41%	52.84%	77.30%	76.40%
PINAW	5.89%	21.15%	15.29%	20.74%

> 0 (ie. the daytime), which gives a better picture of model accuracy.

While the PICP is determined by prediction accuracy, it also depends on the choice of prediction interval width. By choosing an especially wide interval, PICP can be increased, without a concurrent increase in model performance. We therefore pair it with the Prediction Interval Normalized Average Width (PINAW), which measures the interval width of our estimates. It is defined:

$$\text{PINAW} = \frac{1}{NW} \sum_{t=1}^{T} (U_t - L_t) \tag{9}$$

where W is the range of the target variable, and U_t and L_t are the prediction upper and lower bounds at time t. When the PICP is near 1 (high accuracy) and the PINAW is near 0 (high precision) it implies that the probabilistic algorithm is performing very well.

The performance of the BSTS model in terms of these probabilistic metrics is given in Table II. The daytime PICP indicates that our 95% prediction intervals contain the true value only about 50-75% of the time, despite being wide relative to the scale of the data. This suggests that our model is overconfident, which is a common problem with Kalman Filters when the underlying model doesn't fully capture the true dynamics of the data. Indeed we see that this problem is worse for the January data, where the load model seems structurally unable to capture an unusual mid-day reduction in gross load (as shown in Fig. 1). We expect that further developing our model to more faithfully track load dynamics, for example by leveraging reactive power measurements, could yield improvement in this area.

V. DISCUSSION

The BSTS model generally performed quite well relative to the benchmark comparisons (Table I), however there were some cases where no model performed well. Figs. 1 and 2 show disaggregated signal estimates from the BSTS and CSGE models for a day where both models struggled to accurately estimate true load. This appears to be due to a sudden decrease in load which occurred mid-day several times in the January dataset that was not captured by the covariates used in either model. This unexpected drop leads to different behaviors in both the BSTS and CSGE models, which highlight some of the core differences between the BSTS approach and the other optimization approaches.

In Fig. 2 we see that the estimates produced by CSGE exhibit large, unfounded swings during the middle of the day. This is likely because the surface of the loss function used by
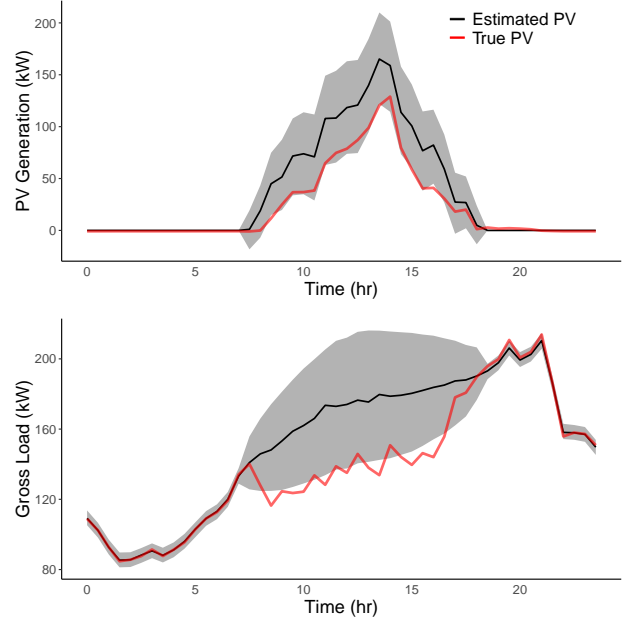

Fig. 1. BSTS estimated PV generation and gross load for one day in the January dataset. Gray ribbon indicates 95% credible interval.
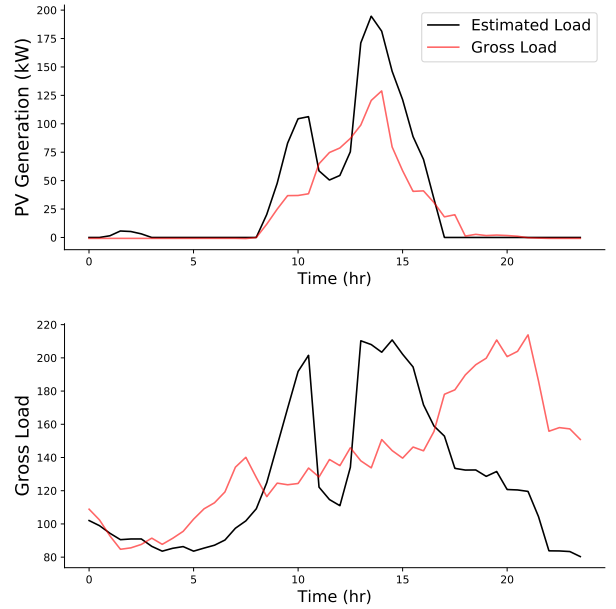

Fig. 2. CSGE estimated PV generation and gross load for one day in the January datase

CSGE is relatively flat near its minimum and so its argmin is not a good estimate, as there are many solutions near it which perform almost identically; the CSGE model is "indifferent" between a range of possible disaggregations. The flatness of the loss function is likely exacerbated by the drop in gross load being inconsistent with the model (although similar problems occurred on days where the drop was not present). On the other hand, the BSTS estimate does not exhibit this erratic behavior. Instead, the model represents its indifference with

4

increasing uncertainty (visible in the gray confidence band in Fig. 1), and since the posterior expectaion averages over this uncertainty, the resulting $\hat{L}$ in the BSTS model is better-fitting and less volatile.

Another feature of the BSTS model which contributes to its performance is that it allows the β_t term to evolve in time in a systematic way. This is a unique feature of the BSTS approach, and it allows our model to overcome deficits in the GHI data, which does not always correlate well with PV generation due to the discrepancy in spatiotemporal resolution between the NSRDB and the Pecan Street dataset, as well as the absence of PV generation site configuration data (the PV cell zenith and azimuth at each generating site). While we employed a fairly simple model of the coefficient's evolution, more sophisticated models are possible accounting for seasonalities in the solar site generating capacity. Similar models can be constructed for other covariates, depending on the particular needs of the dataset, for example allowing the power factor to have time-dependence in a model which includes reactive power.

While our developed BSTS method performs well, relative to other methods, it also faces some challenges. In Table II we see that our model is somewhat overconfident, suggesting the our model is unable to capture some dynamics in the PV generation and gross load. Furthermore, the MCMC algorithm used to fit the BSTS model had a higher computational cost than any of the models based on optimization. Although our model was fit within the time resolution of the data, shorter timescales may prove challenging. Fitting time can be reduced by running multiple sampling chains over several nodes. Furthermore, the comparison in fitting time is not perfect as the BSTS model also produces uncertainty estimates "for free". Estimating uncertainty in an optimization approach requires, such as bootstrapping, which adds computational cost not reflected in their implementation here.

VI. CONCLUSION

In this paper we have analyzed the efficacy of Bayesian Structural Time Series for probabilistically disaggregating consumer load and PV generation at the feeder-level. Using Pecan Street Inc. AMI data we demonstrated that our proposed method performs better than other state-of-the art PV disaggregation models, particularly in the August dataset. In addition to accurately estimating the component signals, the novelty in our approach is that it also quantifies the uncertainty in the estimates. We also show that it has more favorable failure modes, behaving less erratically when the problem is weakly-constrained. The work presented here suggests a number of ways to develop the BSTS model further. Estimation accuracy of disaggregated PV generation and gross load can be achieved by employing more sophisticated load models, and fitting time can be reduced through a more sophisticated implementation of the sampling algorithms. This framework can also be extended to estimate generation and load at the household level, although this would likely require higher resolution irradiance and environmental data. Finally, given forecasts of the relevant covariates, the BSTS approach can be extended

to provide probabilistic forecasts, both of the net load as well as of the disaggregated, behind-the-meter PV generation.

REFERENCES

[1] "Annual Energy Outlook 2019 with projections to 2050," Jan. 2019.
[2] Y. Liu, J. Bebic, B. Kroposki, J. d. Bedout, and W. Ren, "Distribution System Voltage Performance Analysis for High-Penetration PV," in *2008 IEEE Energy 2030 Conference*, pp. 1–8, Nov. 2008.
[3] M. E. Baran, H. Hooshyar, Z. Shen, and A. Huang, "Accommodating High PV Penetration on Distribution Feeders," *IEEE Transactions on Smart Grid*, vol. 3, pp. 1039–1046, June 2012.
[4] "Final Report for Assessment of Visibility and Control Options for Distributed Energy Resources," tech. rep., California ISO, June 2012.
[5] H. Shaker, H. Zareipour, and D. Wood, "Estimating Power Generation of Invisible Solar Sites Using Publicly Available Data," *IEEE Transactions on Smart Grid*, vol. 7, pp. 2456–2465, Sept. 2016.
[6] Y.-M. Saint-Drenan, L. Wald, T. Ranchin, L. Dubus, and A. Troccoli, "An approach for the estimation of the aggregated photovoltaic power generated in several European countries from meteorological data," in *Advances in Science and Research*, vol. 15, pp. 51–62, Copernicus GmbH, May 2018.
[7] C. Ming Cheung, W. Zhong, C. Xiong, A. Srivastava, R. Kannan, and V. K. Prasanna, "Behind-the-Meter Solar Generation Disaggregation using Consumer Mixture Models," pp. 1–6, Oct. 2018.
[8] E. C. Kara, C. M. Roberts, M. Tabone, L. Alvarez, D. S. Callaway, and E. M. Stewart, "Disaggregating solar generation from feeder-level measurements," *Sustainable Energy, Grids and Networks*, vol. 13, pp. 112–121, Mar. 2018.
[9] F. Sossan, L. Nespoli, V. Medici, and M. Paolone, "Unsupervised Disaggregation of Photovoltaic Production From Composite Power Flow Measurements of Heterogeneous Prosumers," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 3904–3913, Sept. 2018.
[10] T. Landelius, S. Andersson, and R. Abrahamsson, "Modelling and forecasting PV production in the absence of behind-the-meter measurements," *Progress in Photovoltaics: Research and Applications*, vol. 0, no. 0.
[11] F. Kabir, N. Yu, W. Yao, R. Yang, and Y. Zhang, "Estimation of Behind-the-Meter Solar Generation by Integrating Physical with Statistical Models," p. 6.
[12] Y. Wang, N. Zhang, Q. Chen, D. S. Kirschen, P. Li, and Q. Xia, "Data-Driven Probabilistic Net Load Forecasting With High Penetration of Behind-the-Meter PV," *IEEE Transactions on Power Systems*, vol. 33, pp. 3255–3264, May 2018.
[13] S. L. Scott and H. Varian, "Predicting the Present with Bayesian Structural Time Series," p. 21, June 2013.
[14] S. L. Scott, "bsts: Bayesian Structural Time Series," June 2019.
[15] J. L. Mathieu, P. N. Price, S. Kiliccote, and M. A. Piette, "Quantifying Changes in Building Electricity Use, with Application to Demand Response," *IEEE Transactions on Smart Grid*, Nov. 2010.
[16] J. Durbin and S. J. Koopman, "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, vol. 89, pp. 603–616, Aug. 2002.
[17] E. C. Kara, M. Tabone, and C. Roberts, "Poster Abstract: Estimating Behind-the-meter Solar Generation with Existing Measurement Infrastructure," p. 2, Nov. 2016.
[18] D. W. van der Meer, M. Shepero, A. Svensson, J. Widn, and J. Munkhammar, "Probabilistic forecasting of electricity consumption, photovoltaic power generation and net demand of an individual building using Gaussian Processes," *Applied Energy*, vol. 213, pp. 195–207, Mar. 2018.

5