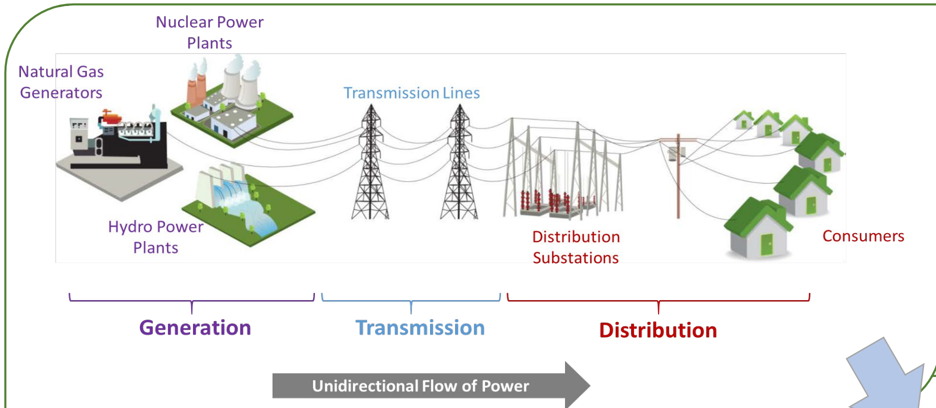
Distributed Energy Resource (DER) Analytics

Santiago Grijalva
Georgia Institute of Technology
sgrijalva@ece.gatech.edu

Acknowledgements

- Advanced Computational Electricity Systems (ACES) Laboratory
- Present and past research contributors
  - Umer Qureshi
  - Sadegh Vejdan
  - Karl Mason
  - Jeremiah Deboever
  - Jouni Pepannen
  - Xiaochen Zhang
  - Matt Reno

Electricity Grid Evolution

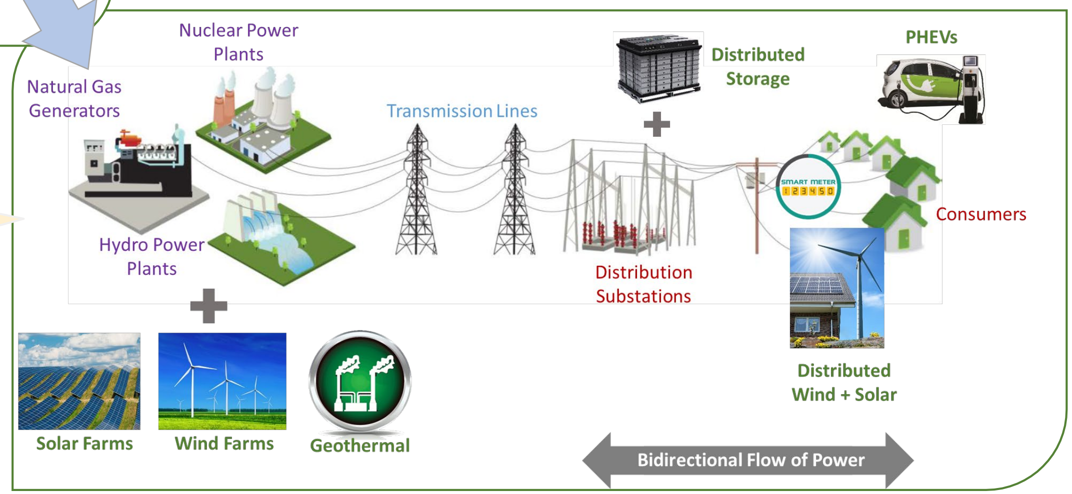


Traditional System:

- Bulk generation
- Central control
- Unidirectional power flow

Future System:

- More renewables and DERs
- Bi-directional power flow
- Decentralized control
- Increased sensing and communication
- Massive new data



Outline

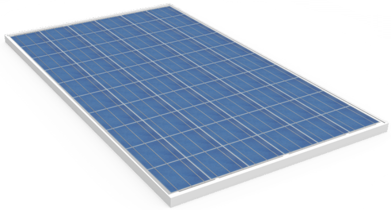
- DERs and emerging DER data
- Applications:
  1. Detecting solar PV installations
     - Change point detection
     - Neural Network
  2. Revenue from energy storage
     - Clustering
     - Optimization
  3. Fast PV hosting capacity
     - Event-driven regression over nonlinear voltage manifolds

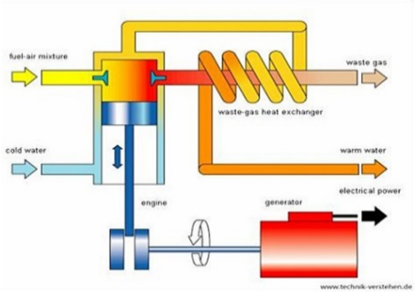
Outline

- DERs and emerging DER data
- Applications:
  1. Detecting solar PV installations
     - Change point detection
     - Neural Network
  2. Revenue from energy storage
     - Clustering
     - Optimization
  3. Fast PV hosting capacity
     - Event-driven regression over nonlinear voltage manifolds

Distributed Energy Resources (DER)

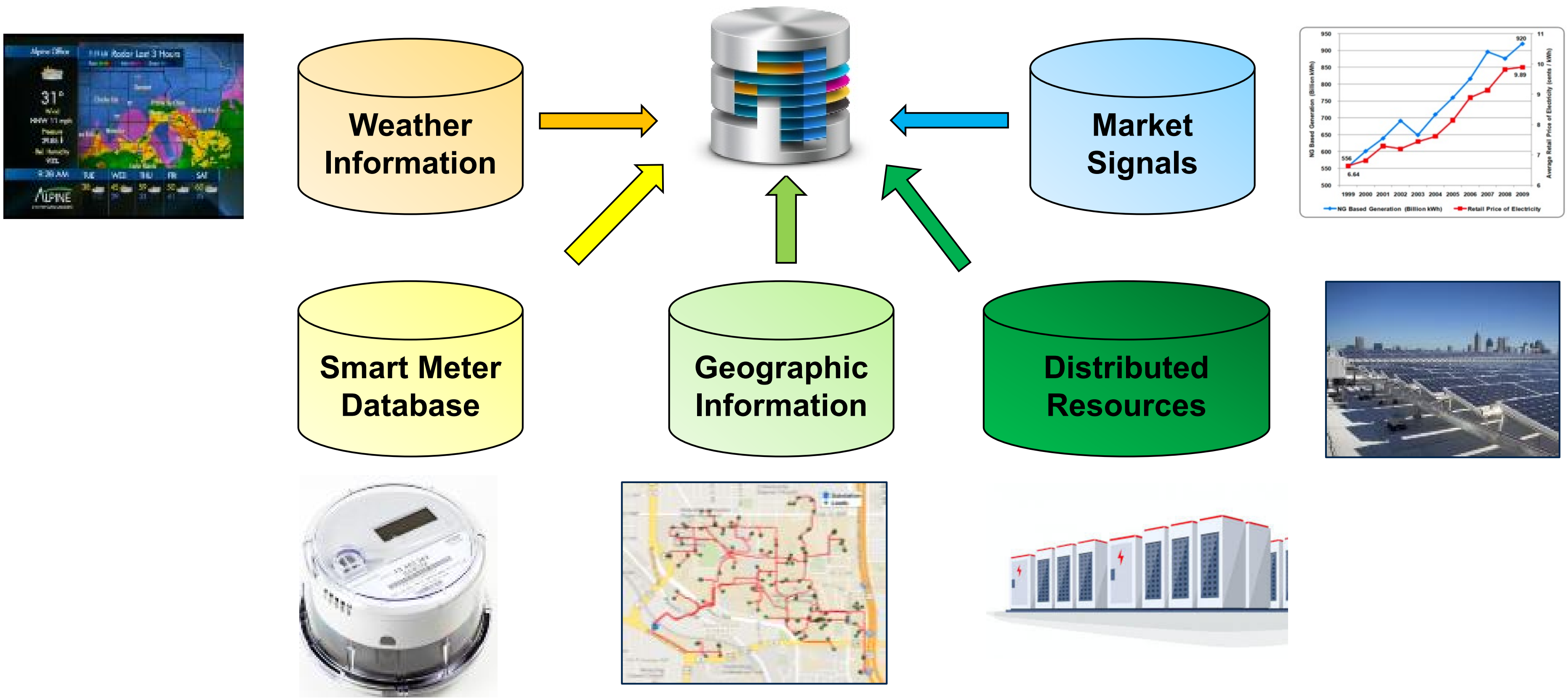
Solar PV
Energy Storage
Distributed Generation
Flexible Loads
Combined Heat and Power (CHP)
Electric Vehicles





DER Related Data Sources

- Data sources for DER analytics:



DER Related Data Sources

- Utilities, DER Providers, Public
- https://openei.org/wiki/Main_Page
- Net load, Appliances: Smart Meter Data Analytics
  - https://smda.github.io/smart-meter-data-portal/

#	Dataset Examples
1	Almanac of Minutely Power Dataset (AMPds)
2	Controlled On/Off Loads Library dataset (COOLL)
3	Dutch Residential Energy Dataset (DRED)
4	Electricity Consumption & Occupancy data set (ECO)
5	GREEND Dataset
6	Indian Dataset for Ambient Water and Energy (iAWE)
7	REFIT Electrical Load Measurements dataset
8	Smart Home Data Set
9	Tracebase
10	UK Domestic Appliance-Level Electricity (UK-DALE)

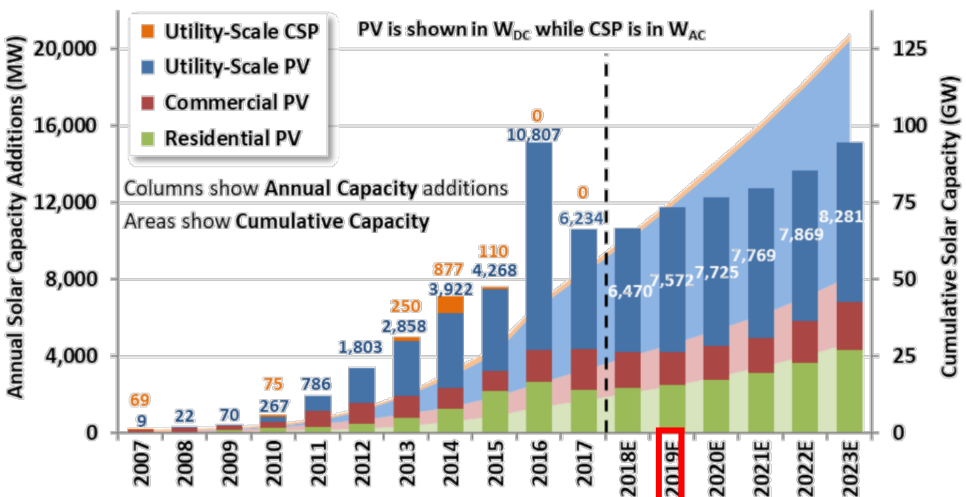
Outline

- DERs and emerging DER data
- Applications:
  1. Detecting solar PV installations
     - Change point detection
     - Neural Network
  2. Revenue from energy storage
     - Clustering
     - Optimization
  3. Fast PV hosting capacity
     - Event-driven regression over nonlinear voltage manifolds

1. PV Detection

Solar Energy Integration

- Solar represented 29% of all U.S. capacity additions in 2018
- U.S. market installed 10.6 GW of solar PV capacity
- Distributed solar PV accounted for 41% of this capacity
- Solar capacity expected to exceed 100 GW by 2021



Sources: *GTM/SEIA Solar Market Insight Reports, Berkeley Lab*

State	PV generation as a % of in-state generation	
	All PV	Utility-Scale PV Only
California	15.2%	10.1%
Hawaii	11.8%	2.0%
Vermont	11.5%	6.2%
Nevada	10.7%	9.7%
Massachusetts	8.1%	3.3%
Utah	6.2%	5.4%
Arizona	5.5%	3.8%
North Carolina	4.4%	4.3%
New Mexico	3.9%	3.3%
New Jersey	3.8%	1.6%
Rest of U.S.	*0.5%*	*0.3%*
TOTAL U.S.	**1.8%**	**1.2%**

Source: EIA's Electric Power Monthly (February 2018)

By 2050, Solar will make up 21% of total installed capacity in the U.S.

PV Detection

Problem:

- PV systems may vary from the interconnection database.
- Keeping PV interconnection databases updated is a major challenge.

Objective:

- Use data driven solutions to detect PV installations.
    - Change Point Detection
    - Convolution Neural Networks

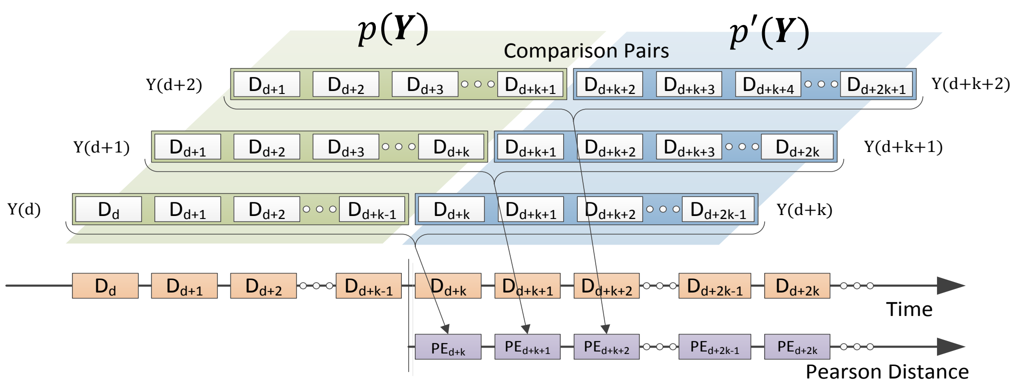
Causes of Discrepancy:

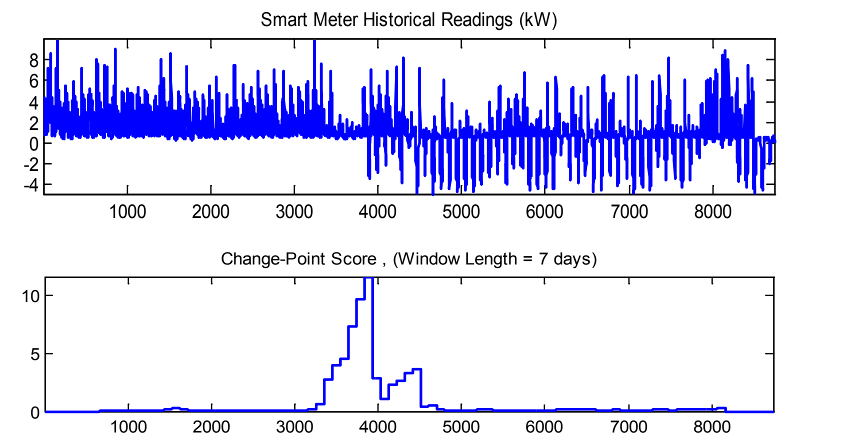
- Not interconnected
- Project delayed
- Changed size
- Module/string failures
- Unauthorized installation

PV Detection using Change Point Method

Change point detection method

- Initially no PV. Then try to detect whether there is a PV



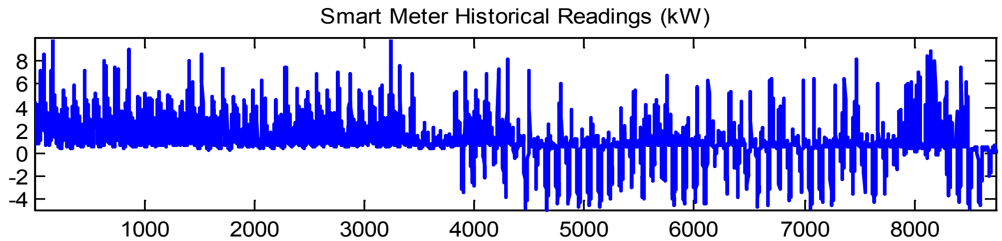


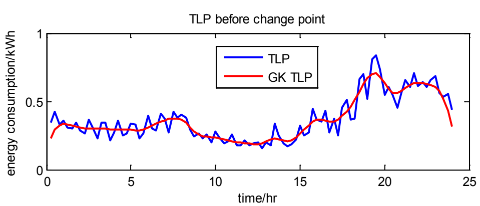
- PE divergence:

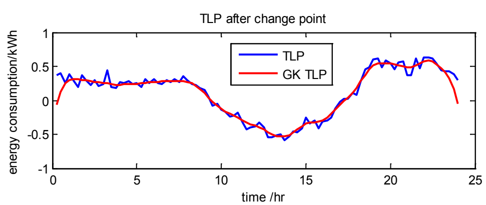
$$PE(P||P') := \frac{1}{2}\int p'(\boldsymbol{Y})\left(\frac{p(\boldsymbol{Y})}{p'(\boldsymbol{Y})} - 1\right)^2 d\boldsymbol{Y}$$

- Measures the difference between two distributions

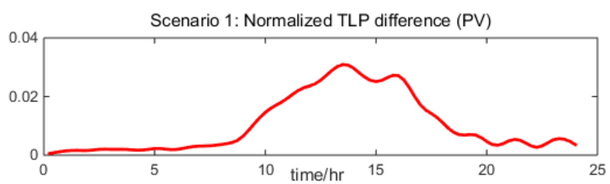
PV Detection using Change Point Method

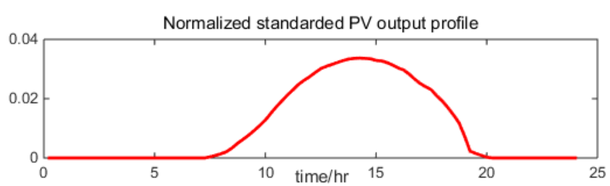






Take the difference





Hypothesis Testing:

$$\begin{cases} H_0: \text{There is NO PV Installed} \\ H_1: \text{There is a PV Installed} \end{cases}$$

Recast:

$$\begin{cases} H_0: \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are not positively correlated} \\ H_1: \boldsymbol{X} \text{ and } \boldsymbol{Y} \text{ are positively correlated} \end{cases}$$

Spearman's rank (r_s) is used instead of Person since $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not normally distributed

Pearson's r	Spearman's rank coefficient	
$\boldsymbol{r}$	$\boldsymbol{r_s}$	p-value
0.9205	0.8351	3.9414e-26

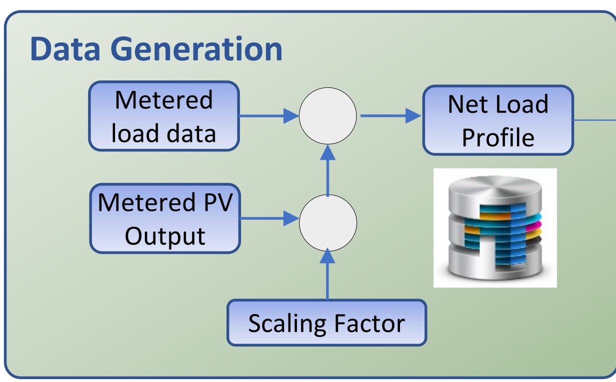
PV Detection using CNN

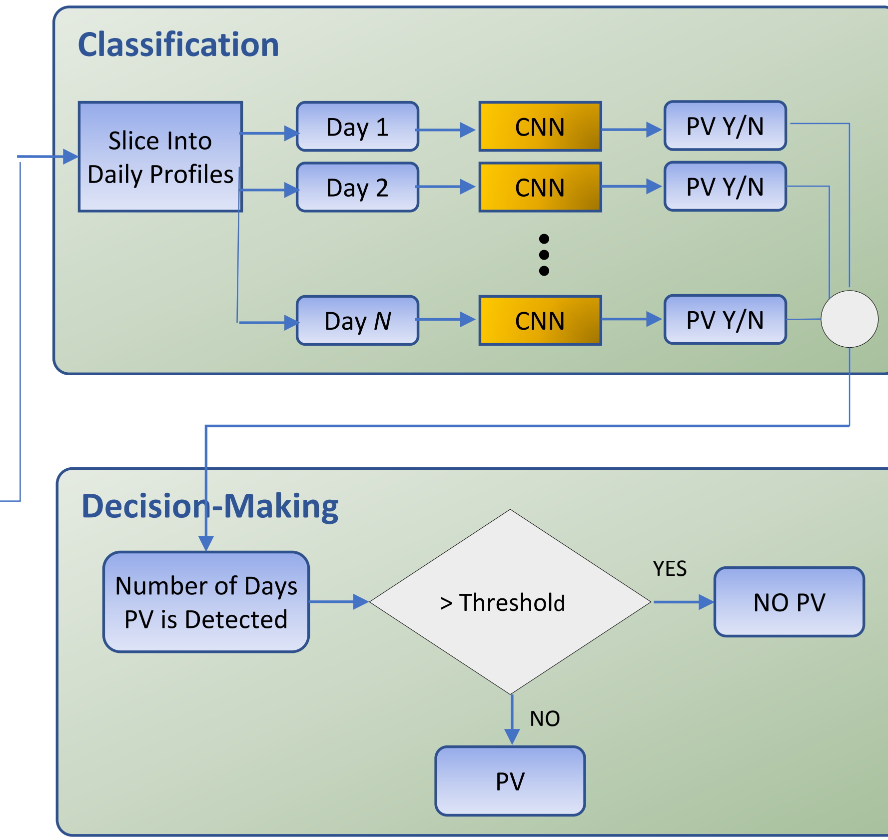
Convolution Neural Network (CNN)

- Data series may not have a change point

Approach

- Synthetic net load **Data Generation**
    - Uses AMI data
- **Classification** of daily net load profiles
- **Decision Making** for each customer





PV Detection using CNN

Generating Synthetic Net Load Data

- Load data sourced from Pecan Street data set.
- PV generation data sourced from the Umass Trace Repository*.
- Synthetic net load data generated by combining each data set using:

$$NetLoad_{n(l,s,p)} = Load_l - ScalingFactor_s \times PV_p$$

- Synthetic data set consisted of 50% customers with PV installed.

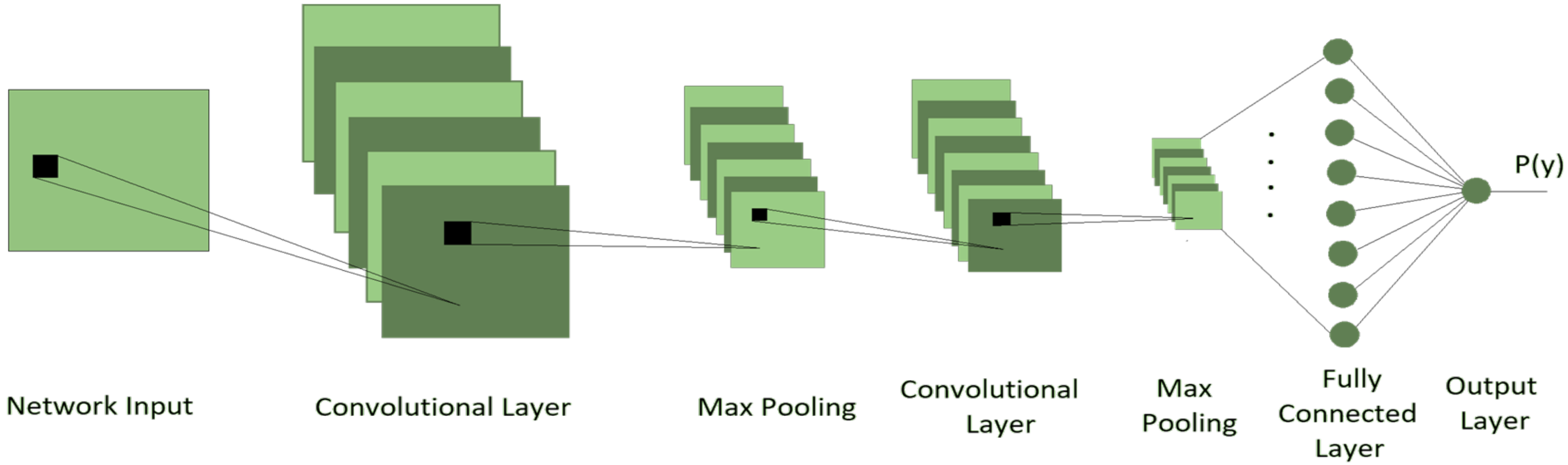
- "On-the-fly" synthetic data generation method to study sensitivities on the various simulation dimensions.

Folds
Epochs
Test Customers
Resolution
Training Customers
Days per Customer
% Mislabeled

[* http://traces.cs.umass.edu/index.php/Smart/Smart](http://traces.cs.umass.edu/index.php/Smart/Smart)

PV Detection using CNN

Classification Using CNN



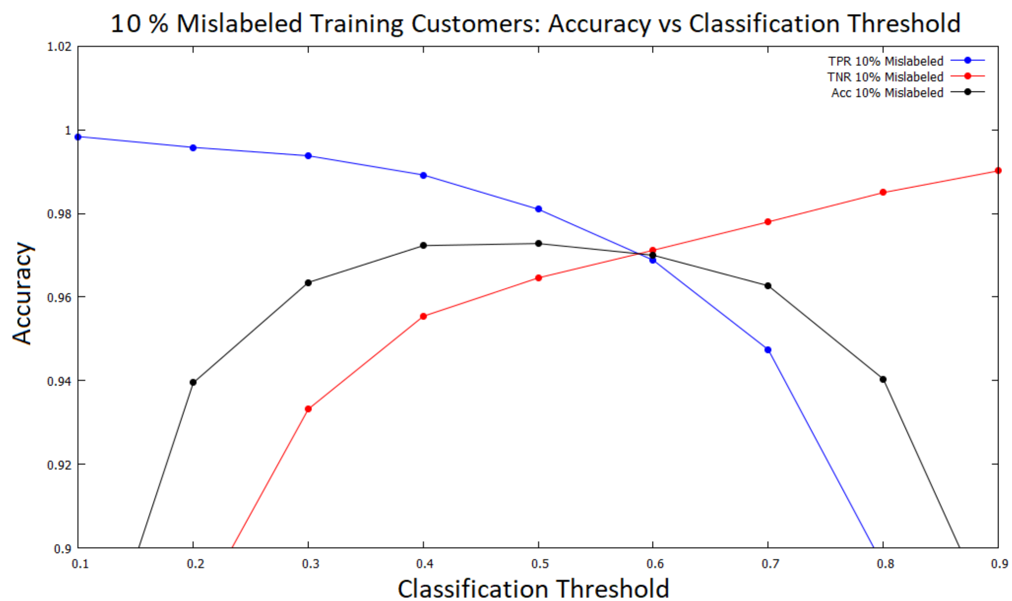
- Input: daily net load profile.
- Output: probability of PV (1 = PV installed, 0 = no PV).
- Architecture: convolutional layer, max pooling, convolutional layer, max pooling, fully connected layer (100), fully connected layer (1).
- CNN trained for 200 epochs using RMSprop optimization.

PV Detection using CNN

Decision Making using Threshold

- Thresholding logic:
- Impact of threshold on TPR, TNR and total customer classification accuracy:

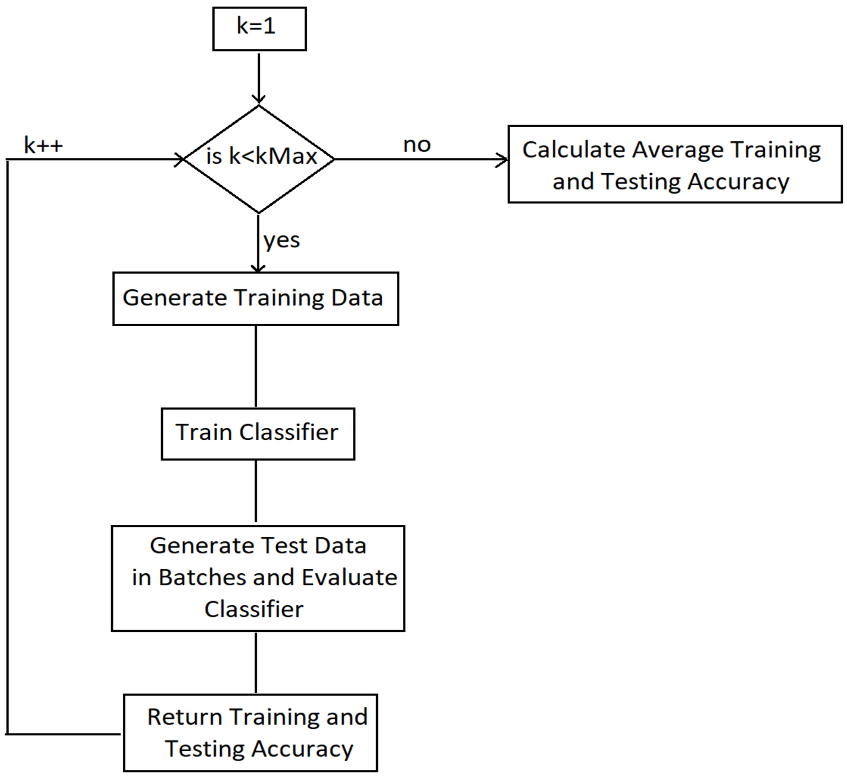
$$if \ \sum_{day} PV\_Detected_{day} > Threshold \times Ndays \Rightarrow Customer \ has \ PV, \ otherwise \ does \ not.$$



PV Detection using CNN

Simulation

- Simulations analyzed the impact of training data parameters on the accuracy of CNN:
  - Number of customers
  - Temporal resolution
  - Level of mislabeled data
- Classifier was trained and tested 10 times for each simulation.
- New training and test data generated at each fold.
- Accuracies and average computational times are recorded.



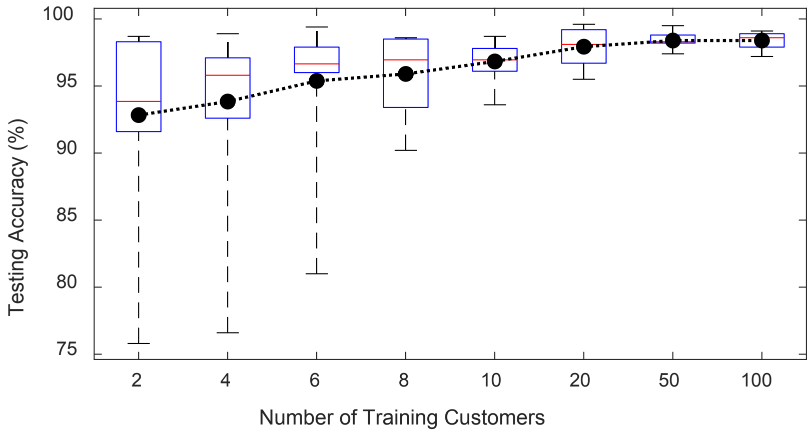
PV Detection using CNN

Results: Number of Training Customers

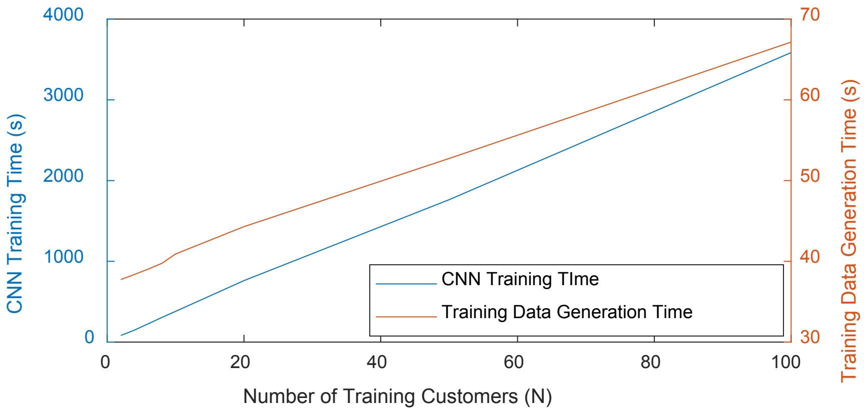
- **Simulation Parameters:**

Folds	10
Epochs	200
Test Customers	1000
Resolution	1 min
Training Customers	**Varying**
Days per Customer	343
% Mislabeled	0.0

Classification Accuracy vs Number of Training Customers



Computational Cost vs Number of Training Customers



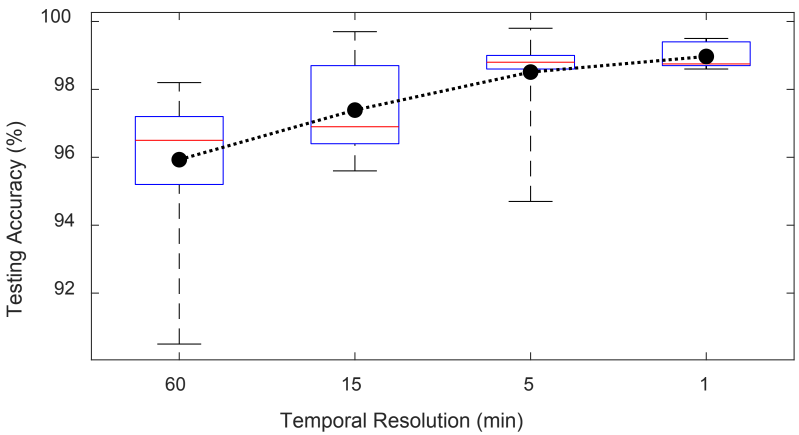
PV Detection using CNN

Results: Temporal Resolution

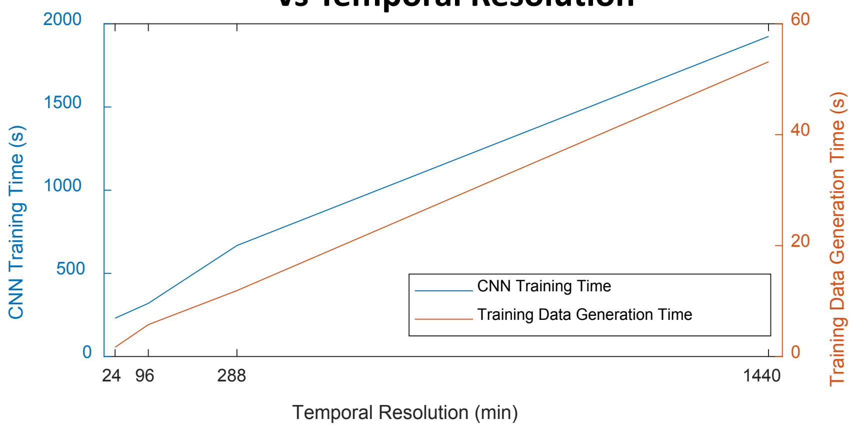
- **Simulation Parameters:**

Folds	10
Epochs	200
Test Customers	1000
Resolution	**Varying**
Training Customers	50
Days per Customer	343
% Mislabeled	0.0

Classification Accuracy vs Temporal Resolution



Computational Cost vs Temporal Resolution



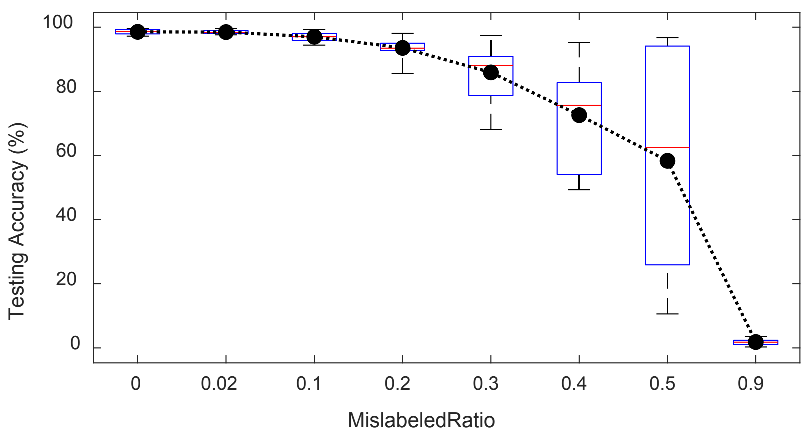
PV Detection using CNN

Results: Mislabeled Training Customers

- **Simulation Parameters**

Folds	10
Epochs	200
Test Customers	1000
Resolution	1 min
Training Customers	50
Days per Customer	343
% Mislabeled	**Varying**

Classification Accuracy vs Mislabeled Ratio



PV Detection using CNN

Summary

- CNN classifier achieves +98% customer classification accuracy. The experiments conducted reveal the following insights about the data and simulation parameter requirements:
  - 50 training customers provides the best performance. More training customer requires more computational time but does not perform significantly better.
  - 1 minute resolution provides the best accuracy. The classifier's performance is robust to lower resolution data however.
  - The classifier maintains a reasonable accuracy even with 10% mislabeled training data (+85%).

Outline

- DERs and emerging DER data
- Applications:
  1. Detecting solar PV installations
     - Change point detection
     - Neural Network
  2. Revenue from energy storage
     - Clustering
     - Optimization
  3. Fast PV hosting capacity
     - Event-driven regression over nonlinear voltage manifolds

2. Energy Storage Revenue Analytics

Introduction

- Context:
  - Energy storage investors and industry stakeholders are interested in the mechanisms for storage services revenue.
  - Find the best measure of “favorable” price volatility to determine the expected revenue using temporal energy arbitrage.
- Two time scales:
  - Day-ahead (DA) market
  - Real-Time (RT) market
- Day-ahead energy market
  - Market clearance of offers and bids of producers and consumers.
  - Power dispatch with the lowest total cost of operation considering network and security constraints.

Day-Ahead Energy Arbitrage Revenue

Optimization Approach:

- Objective: maximizing the revenue
- Decision variables: charging/discharging power, binary variables
- Input data: storage parameters and market prices
- Constraints: storage power and energy limits

$$\max \sum_{t=1}^{T} \pi_t \left(P_t^{dis} - P_t^{chg} \right) \Delta t$$

$$\textbf{s.t.} \quad \forall t \in \mathcal{T}$$

$$0 \le u_t^{dis} + u_t^{chg} \le 1$$

$$P_{\min}^{dis} . u_t^{dis} \le P_t^{dis} \le P_{\max}^{dis} . u_t^{dis}$$

$$P_{\min}^{chg} . u_t^{chg} \le P_t^{chg} \le P_{\max}^{chg} . u_t^{chg}$$

$$E_t = \eta_s E_{t-1} + \left(\eta_{chg} P_t^{chg} - P_t^{dis} / \eta_{dis} \right) \Delta t$$

$$E_{\min} \le E_t \le E_{\max}$$

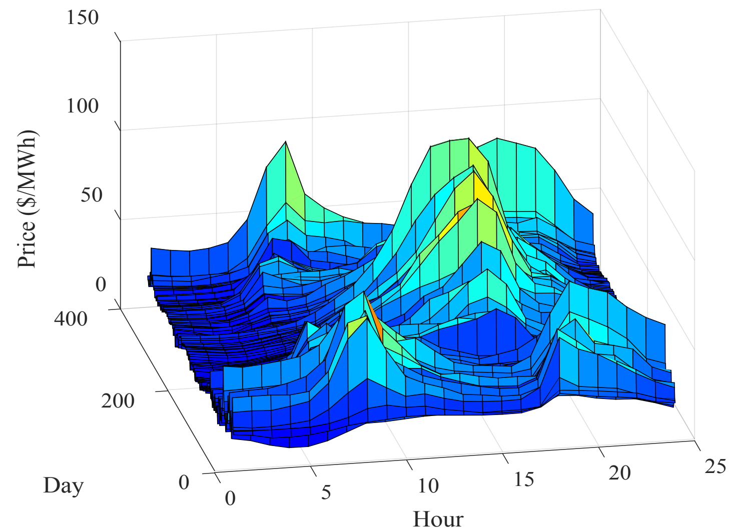
$$E_T = E_0$$

- Assumptions:
  - Price-taker approach: negligible market power
  - Perfect foresight: future prices for the day- ahead horizon are known

Day-Ahead Energy Arbitrage Revenue

Price Patterns

- Given a fixed size of the energy storage system, the arbitrage revenue is dependent on price data patterns and its statistics.
- PJM day-ahead energy market (2016)
  - Seasonal price patterns
  - Summer: one peak in the early evening
  - Winter: two daily peaks, morning and evening
  - Other markets show almost similar pattern
- Revenue quantification based on price data pattern
- Classification criterion:
  - Pearson correlation coefficient

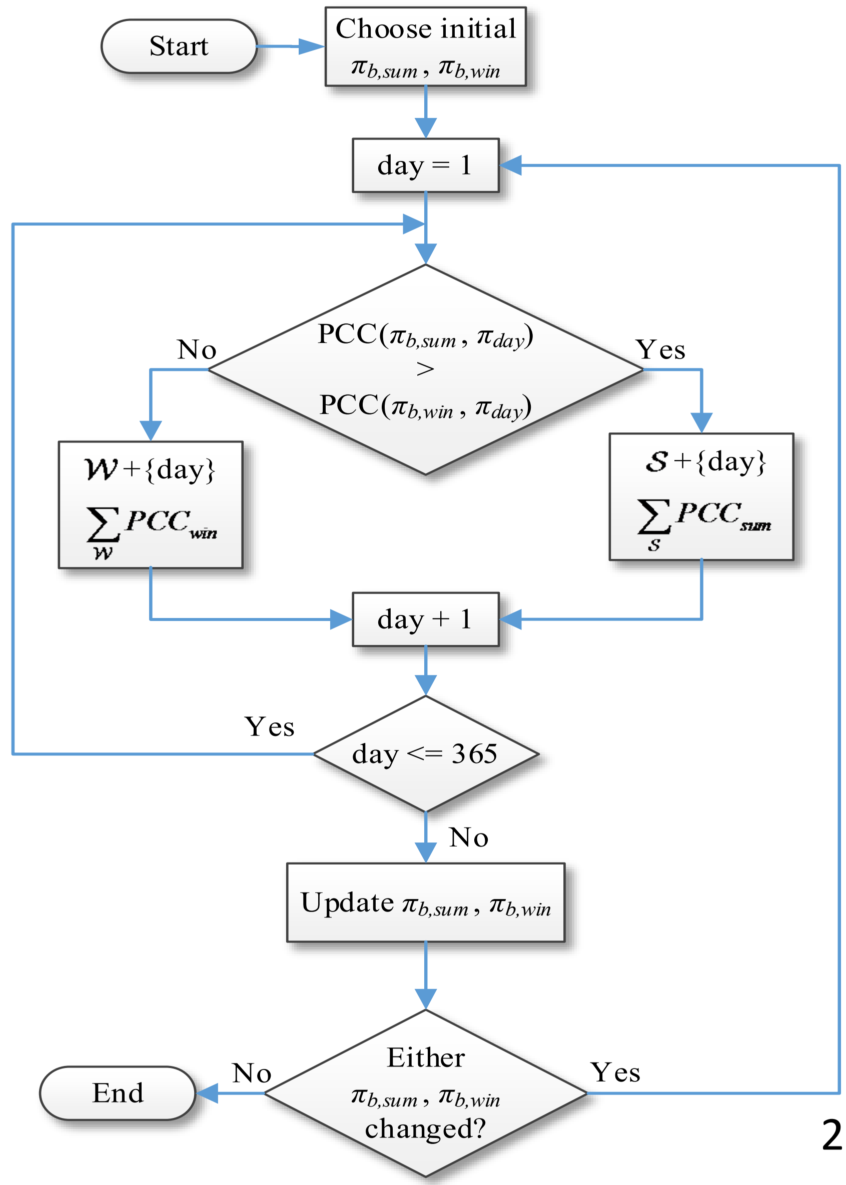


$$PCC = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}}$$

Day-Ahead Energy Arbitrage Revenue

Clustering Algorithm

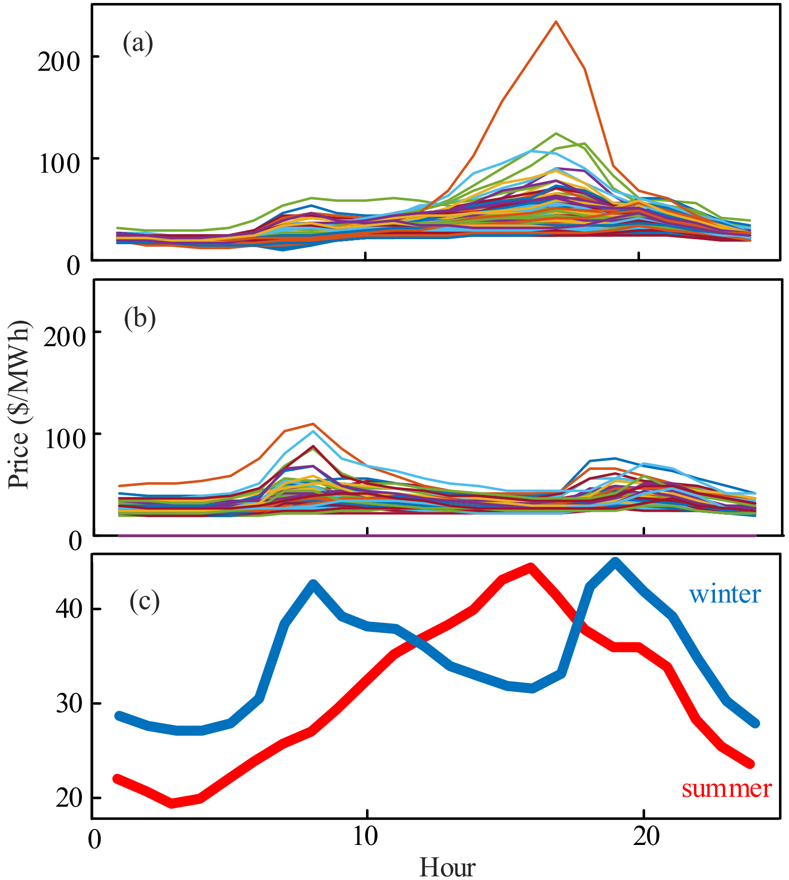
- *K*-means: classifies energy market prices into two groups
- Input: annual price data
- Outputs:
  - Two clusters for summer and winter daily prices
  - Two base prices for each cluster
- Used to find when each season starts (in terms of electricity prices) and how long it lasts.

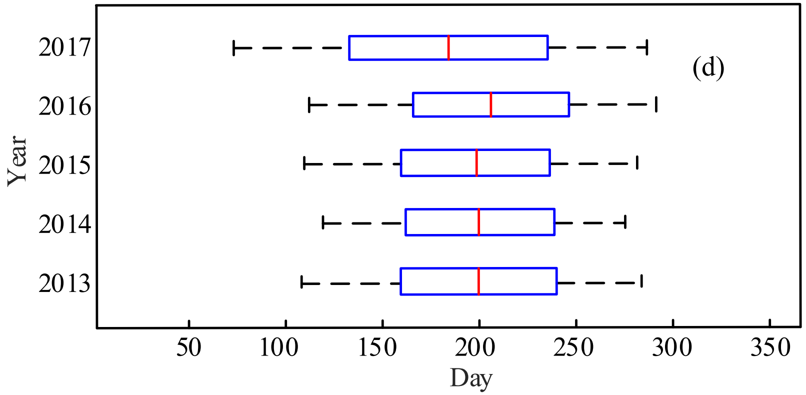


Day-Ahead Energy Arbitrage Revenue

Clustering Results

- Algorithm converges in 2 to 4 iterations.
- Summer and winter prices are clustered: (a), (b)
- The best base prices are found: (c)
- The starting day and duration of each cluster is determined: (d)

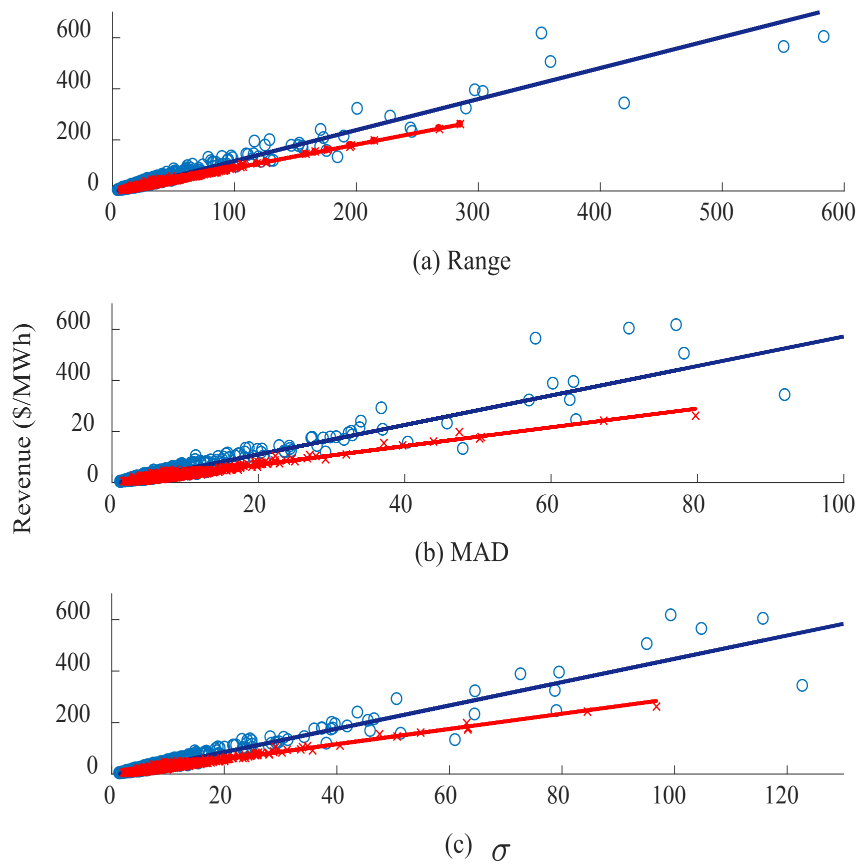




Day-Ahead Energy Arbitrage Revenue

Regression Results

- Good linear fit
- Dispersion statistics: range, mean absolute deviation, standard deviation

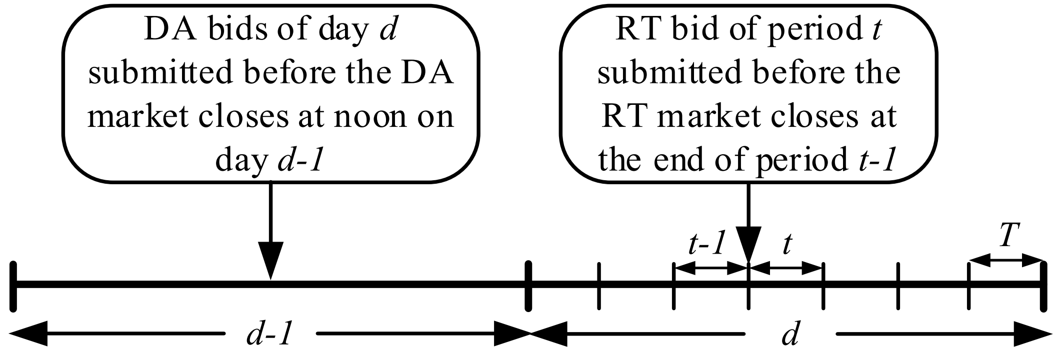


	Season	Range	MAD	σ
β0	summer	-2.55	-2.98	-1.77
	winter	-4.16	-5.01	-5.08
β1	summer	0.92	3.67	2.96
	winter	1.21	5.77	4.53
R^2	summer	**0.9868**	0.9415	0.9619
	winter	**0.9486**	0.9253	0.9484

Real-Time Energy Arbitrage Revenue

Real-Time (RT) Energy Market

- Price updates on a rolling basis (e.g. 5 min) and not known in advance
- Higher variability of RT prices with higher penetration of renewables
- Higher expected arbitrage opportunities in the RT market
- How do RT revenues compare with DA?



Real-Time Energy Arbitrage Revenue

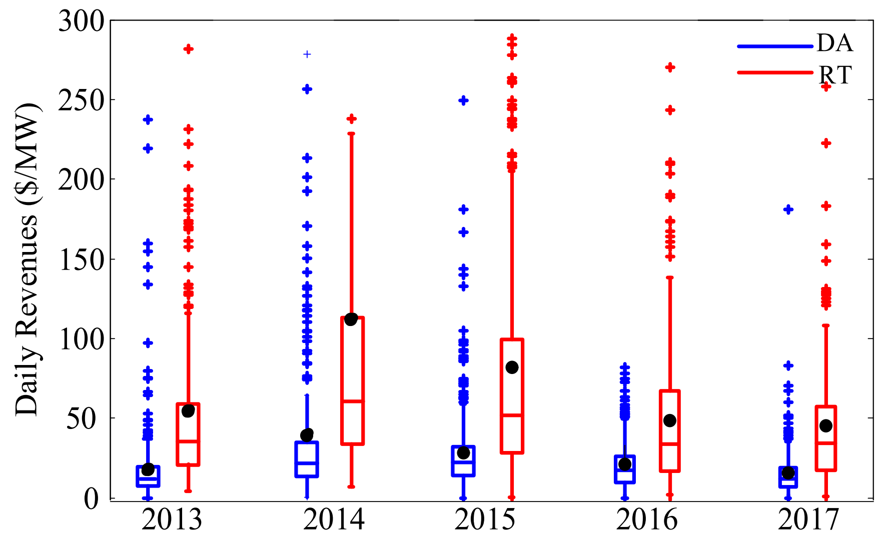
Statistical analysis

- PJM DA and RT price data
- Higher price mean in the day-ahead market
- Higher price variability in the real-time market → Higher expected arbitrage revenues

A. Maximum revenue with perfect foresight

- Similar mixed integer linear optimization model used for the day-ahead market
- Compare maximum RT arbitrage revenues with the day-ahead ones

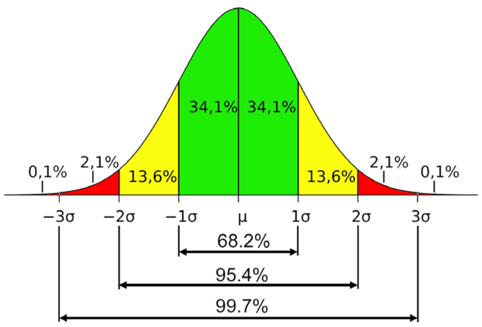
Year	2013	2014	2015	2016	2017
Mean DA	37.15	49.16	35.61	30.01	30.21
Mean RT	36.57	48.40	33.43	27.27	28.97
Median DA	34.62	38.10	30.58	27.48	27.46
Median RT	32.25	34.48	26.62	24.03	25.28
Std DA	15.46	51.87	22.63	11.58	12.02
Std RT	20.69	65.43	27.91	14.64	17.75

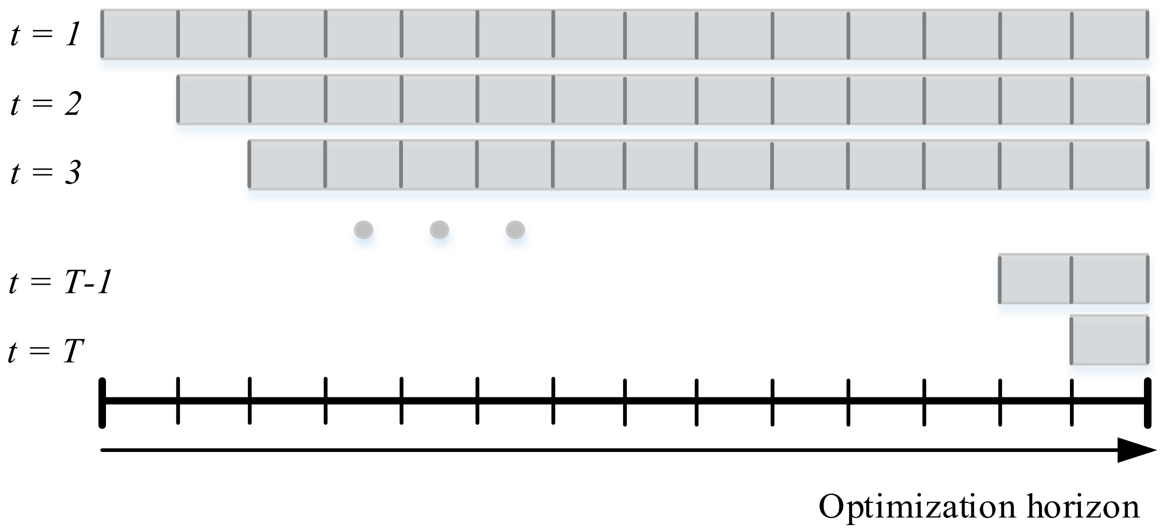


Real-Time Energy Arbitrage Revenue

B. Price with forecasting errors

- Two models: back-casting and random normal errors.
- Proposed optimal dispatch algorithm: optimization based on shrinking horizon dynamic programming.





Real-Time Energy Arbitrage Revenue

Dispatch Algorithm

- Let $\hat{\pi}_i$ be the price forecast for time-period *i*, which can be evaluated by either of the two error models.
- Optimal dispatch decisions are updated at the beginning of each time period *i*.
- State transition times are negligible.

$1: \ t = 0$

$2: \ \text{while } t < T \text{ do}$

$3: \ \text{Solve:}$

$$\max \left[\pi_t (P_t^{dis} - P_t^{chg}) + \sum_{i=t+1}^{T} \hat{\pi}_i (P_i^{dis} - P_i^{chg}) \right] \Delta t$$

$$0 \leq u_t^{dis} + u_t^{chg} \leq 1$$

$$P_{\min}^{dis} . u_t^{dis} \leq P_t^{dis} \leq P_{\max}^{dis} . u_t^{dis}$$

$$P_{\min}^{chg} . u_t^{chg} \leq P_t^{chg} \leq P_{\max}^{chg} . u_t^{chg}$$

$$E_t = \eta_s E_{t-1} + \left(\eta_{chg} P_t^{chg} - P_t^{dis} / \eta_{dis} \right) \Delta t$$

$$E_{\min} \leq E_t \leq E_{\max}$$

$$E_T = E_0$$

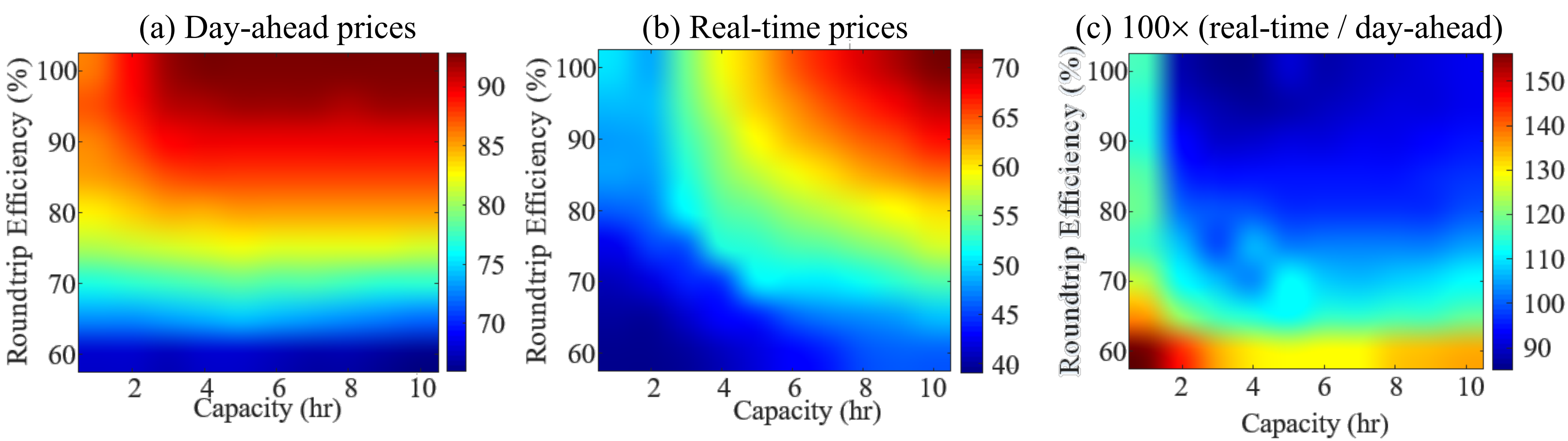
$4: \ t = t + 1$

$5: \ \text{end while}$

Real-Time Energy Arbitrage Revenue

Simulation Results: Back-Casting (BC)

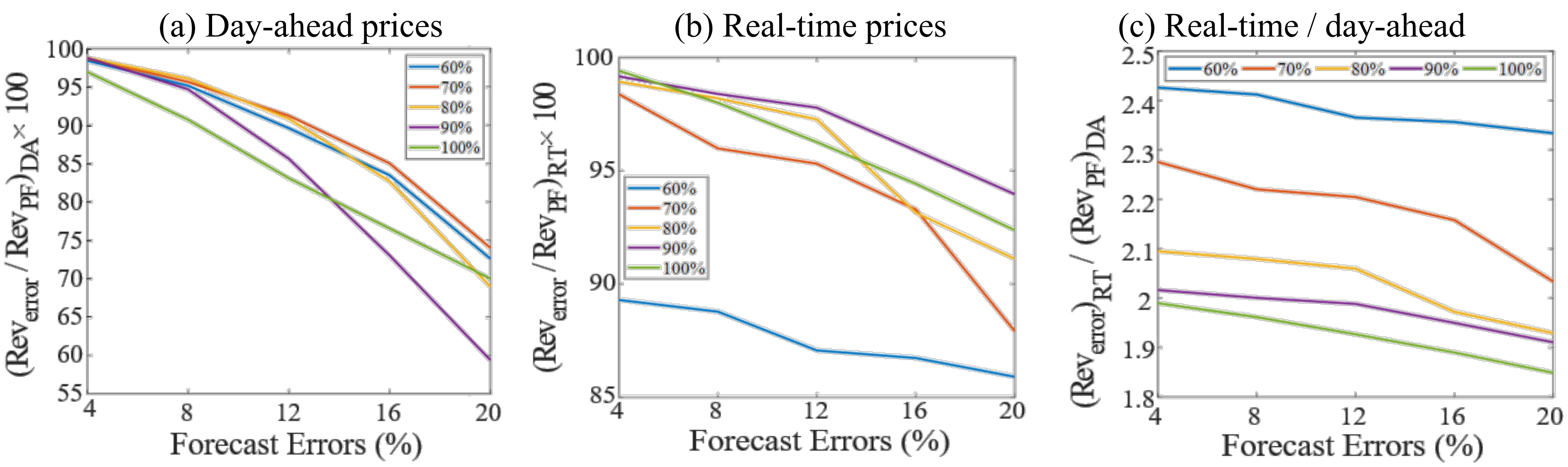
- Revenue (BC) = actual price × dispatch optimized for the day before
- Figure: 100 × [Revenue (BC) / Revenue (perfect foresight)]



Real-Time Energy Arbitrage Revenue

Simulation Results: Normal Errors (NE)

- Revenue (NE) = actual price × dispatch optimized for the simulated price
- Figure: 100 × [Revenue (NE) / Revenue (perfect foresight)]



Energy Storage Revenue Analytics

Conclusions

- Higher value of the real-time energy arbitrage versus day-ahead:
  - Statistical analysis of historical price data
  - Optimization models maximizing the revenue
- Revenue maximization under uncertainty
  - Optimal dispatch based on price forecast error
  - Sensitivities of critical energy storage parameters
- RT arbitrage as an additional revenue stream for energy storage
  - Considerable and reliable if proper optimal dispatch strategies are applied.

Outline

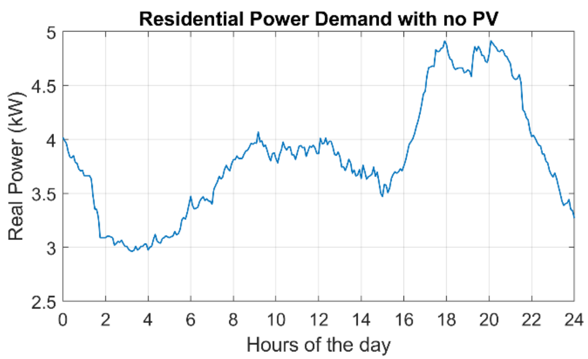
- DERs and emerging DER data
- Applications:
  1. Detecting solar PV installations
     - Change point detection
     - Neural Network
  2. Revenue from energy storage
     - Clustering
     - Optimization
  3. Fast PV hosting capacity
     - Event-driven regression over nonlinear voltage manifolds

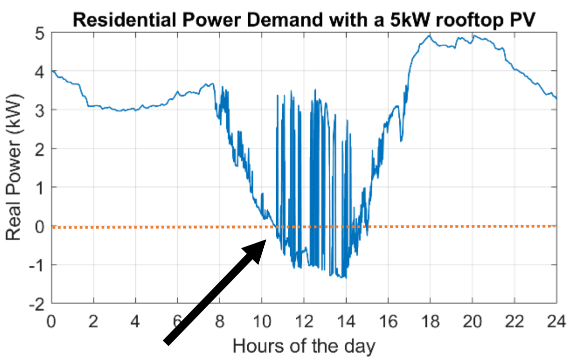
3. Fast PV Hosting Capacity

PV Impacts on Distribution Systems

- Solar PV is an intermittent, non dispatchable resource
- Power output is dependent upon solar irradiance
- Negative impacts include:

- Voltage limit violations
- Thermal overloading
- Increased system losses
- Excessive controller actions
- Power quality
- Protection devices





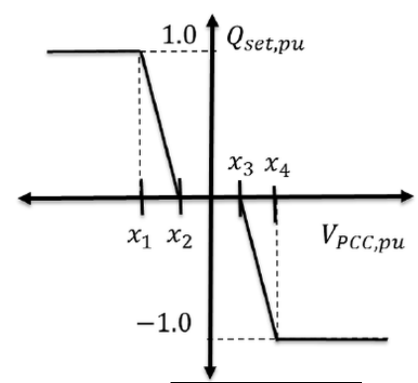
Power exported back to grid

Fast PV Hosting Capacity

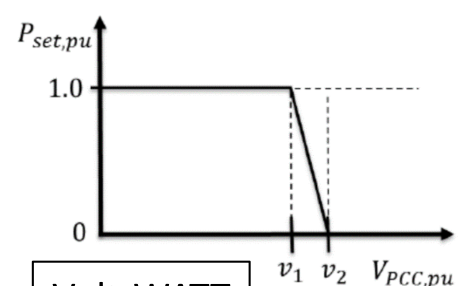
Traditional Interconnection Studies

1) **Static screens:** E.g. PV rated power limited to 15% of peak load
   - Locational impacts ignored
   - Feeder specific conditions not considered
2) **Scenario-based simulation:** Evaluates key scenarios using power flow (e.g. max/min load, max/min PV power)
   - Voltage regulation capability ignored
     - Regulators
     - Capacitors
     - Smart inverters
   - Temporal impacts not captured

Pros: Simple, Fast, Utility Friendly
Cons: Conservative estimates



Volt-VAR



Volt-WATT

Fast PV Hosting Capacity

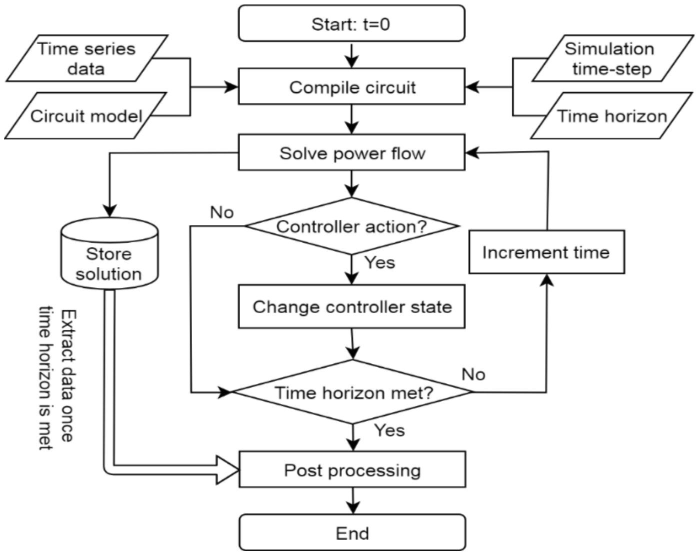
Time Series Simulations

- IEEE P1547.7 recommends:

QUASI-STATIC **TIME SERIES**

Steady state solution
(No numerical integration of
differential equations)

Temporal dependence
between time steps



Brute-force QSTS

- Chronological solution of steady-state power flows
- Discrete controls modeled (tap changers, switches)
- Recommended time step: 1 second to 1 hour
- One-year horizon

- Inputs:
  - Time series data (load and PV)
  - Distribution feeder model
  - Time-step, Time horizon

Temporal Impacts:

1) Regulator tap actions
2) Capacitor bank operations
3) Duration of voltage/thermal limit violation
4) Total line losses
5) Total VAR feed-in
6) Total WATT curtailed

Fast PV Hosting Capacity

Brute-force QSTS Challenges

1) Challenges:
   - Data Requirements: High resolution load (SCADA, AMI) and irradiance data
   - Computational Time: 10-120 hours for a realistic feeder

2) Number of Power Flow (PF) solutions required
   - A single PF flow takes fraction of a second, 31.5 million can take several days.

3) Switched-mode, nonlinear system of equations
   - Voltage regulation equipment $\rightarrow$ *discrete system states*

4) Chronological dependence between time steps
   - Regulators, capacitor banks, switches $\rightarrow$ *hysteresis, deadbands*

5) Multiple valid PF solutions for a given input
   - *Machine learning approaches alone produce large errors*

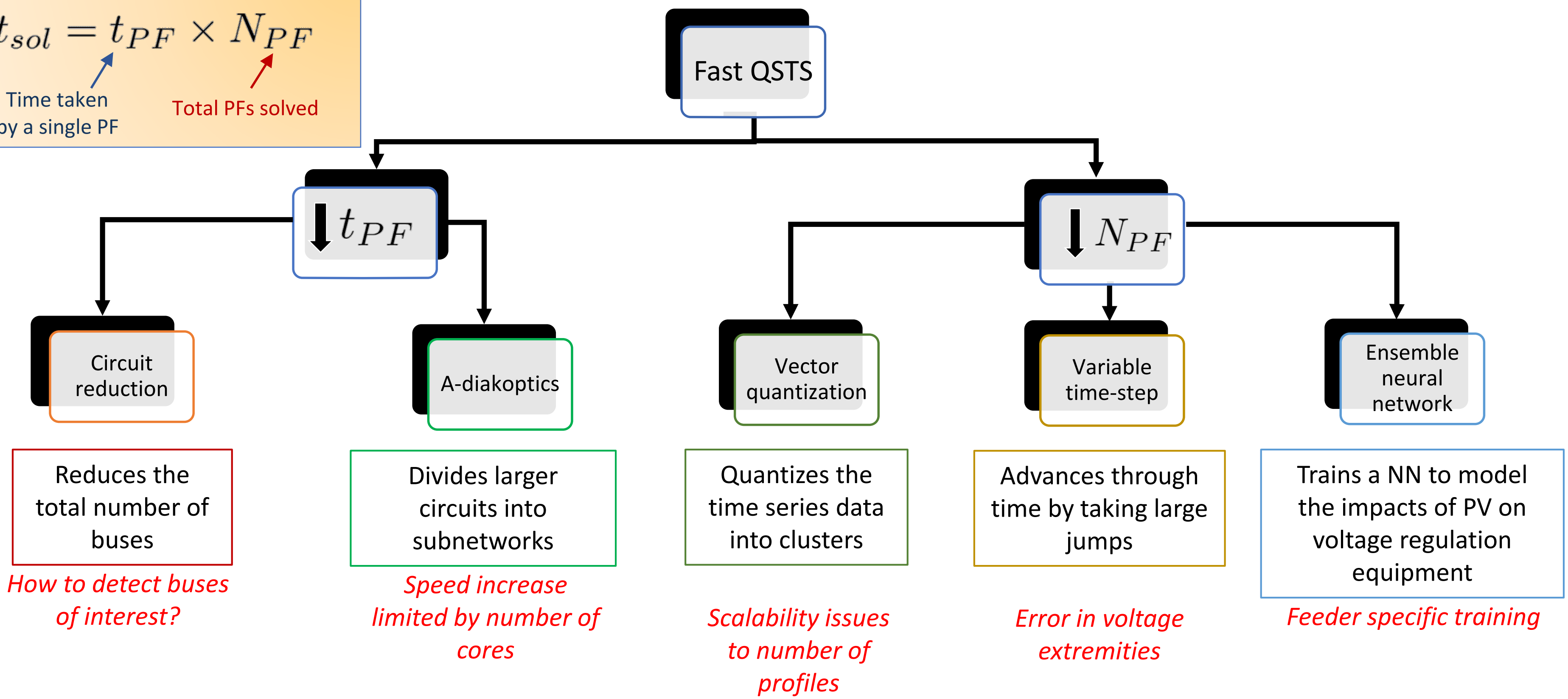
Existing Fast QSTS Methods

Time taken by QSTS:

$$t_{sol} = t_{PF} \times N_{PF}$$

Time taken by a single PF

Total PFs solved



Fast PV Hosting Capacity

Power Flow Voltage Manifold

$$\tilde{V} = \begin{bmatrix} v_1\angle\theta_1 \\ v_2\angle\theta_2 \\ \vdots \\ v_n\angle\theta_n \end{bmatrix} \quad \tilde{S} = \begin{bmatrix} p_1 + jq_1 \\ p_2 + jq_2 \\ \vdots \\ p_n + jq_n \end{bmatrix}$$

- For n-bus network, we have: $\mathrm{diag}(\tilde{V})(Y\tilde{V})^* = \tilde{S}$ ... (1)
- For QSTS, the time-series profiles $(x_l, x_{pv}) \in [0, 1]$ act as inputs,

$$\tilde{S}_i = (p_i + jq_i)x_l \quad \forall i \in \mathcal{L} \longleftarrow$$ Set of all loads assigned x_l

$$\tilde{S}_j = (p_j + jq_j)x_{pv} \quad \forall j \in \mathcal{K} \longleftarrow$$ Set of all PVs assigned x_{pv}

- (x_l, x_{pv}) are '*multipliers*' scaling real and reactive power injections of loads and PV systems
- Let $u = (v, \theta, x_l, x_{pv}) \in \mathbb{R}^{2n+2}$, then we define a power flow manifold $\mathcal{M}$ as,

$$\mathcal{M} := \{u \mid \mathcal{F}(u) = \mathbf{0}_{2n}\} \quad ... \quad (2)$$

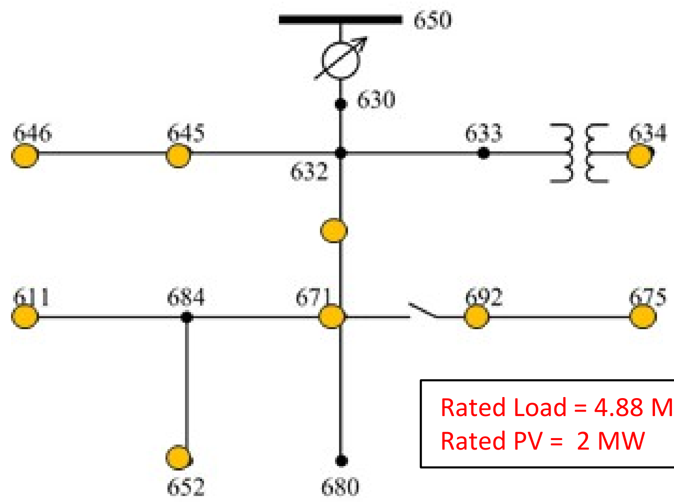
where $\mathcal{F}(u)$ is obtained by rewriting (1) in real and imaginary coordinates

- Without loss of generality, we can extend this notion to any number of time series profiles

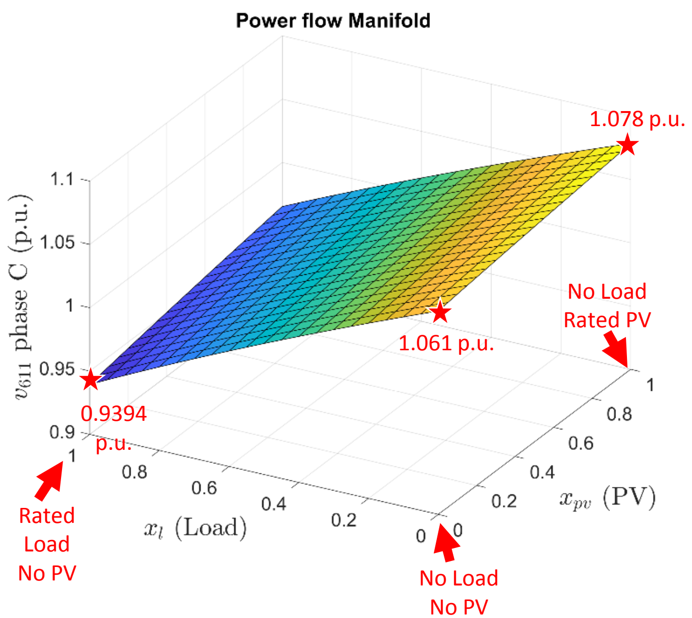
Fast PV Hosting Capacity

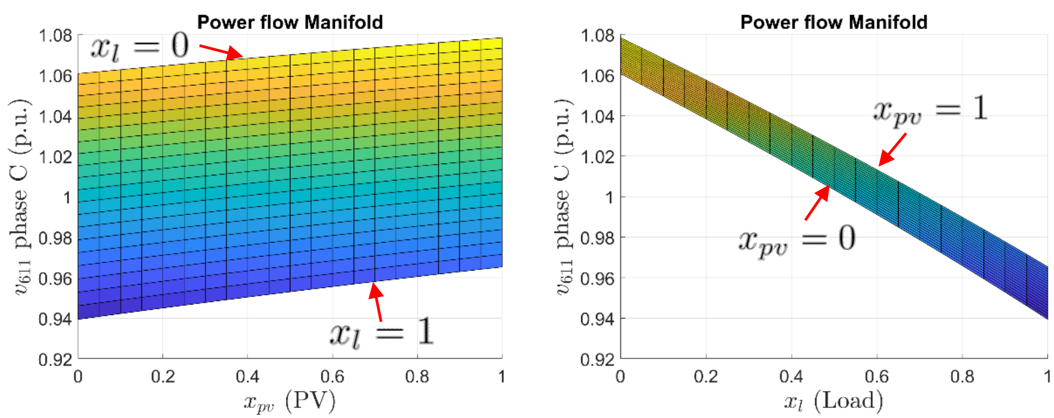
Geometric Interpretation

- Modified IEEE 13-bus test circuit:



Rated Load = 4.88 MVA
Rated PV = 2 MW





Projections of the manifold

Fast PV Hosting Capacity

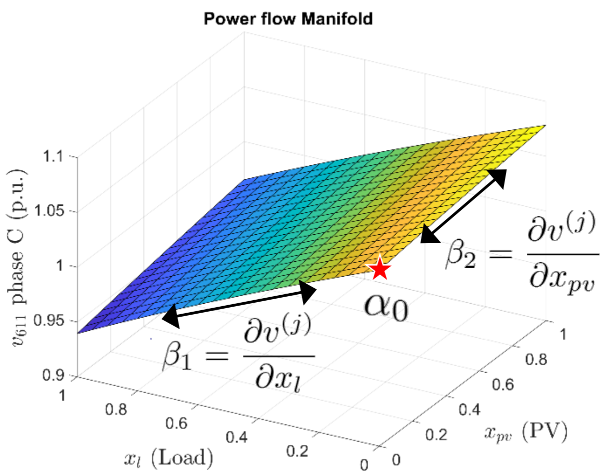
Model Formulation

- The voltage magnitude has a strong correlation with load and PV multipliers
- This correlation can be modeled by a linear approximant of the form:

$$v^{(j)} = \alpha_0 + \underbrace{\beta_1 x_1 + \ldots + \beta_p x_p}_{p\text{ -profiles}}$$

Equation of hyperplane

- β_i is referred to as *voltage sensitivity coefficient*
- Various analytical methods to compute β_i
  - Newton Raphson (inverse of Jacobian)
  - Gauss-Seidel method
  - Adjoint-network technique



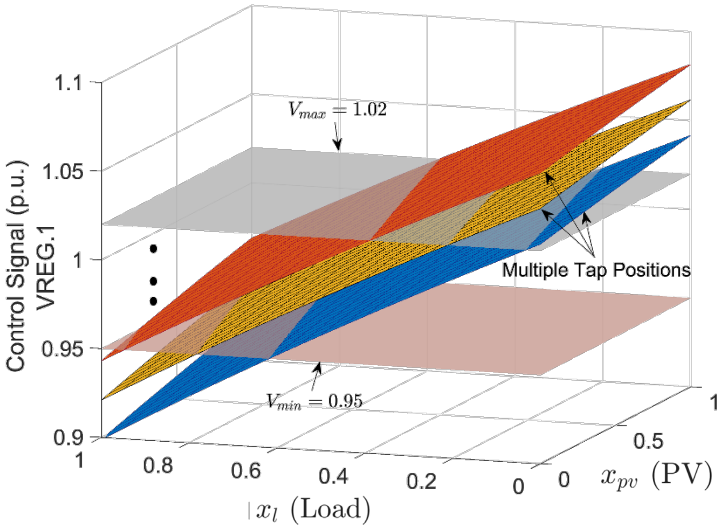
Fast PV Hosting Capacity

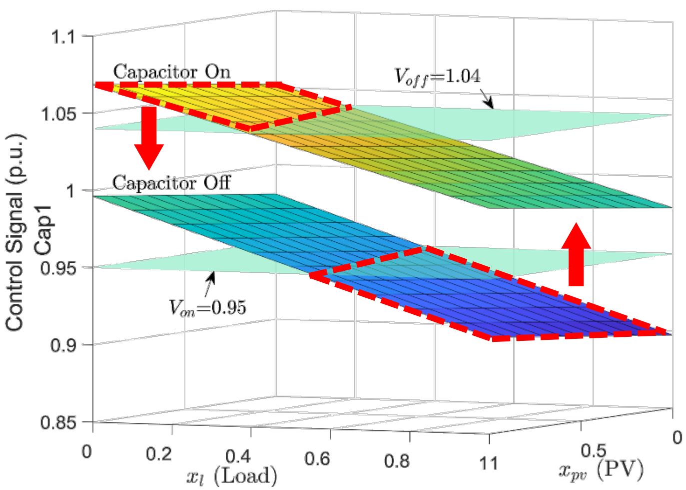
Sensitivity Coefficients

- Perturb-and-observe technique:
  - Introduce small changes in injections
  - Solve the AC power flow problem.
  - Use regression to linearize manifold

Impact of Voltage Regulating Devices

- Maintain voltages +/- 5% of nominal
- Control logic consists of:
  1) A ‘control’ signal (V at secondary winding)
  2) A user-specified voltage set point
  3) Deadband and delays to avoid ‘hunting’
- A change in tap position causes discrete jumps in the power flow manifold
- New sensitivity coefficients determined for each tap position
- Similar impact for capacitor banks.





Fast PV Hosting Capacity

Linear Sensitivity Model

- Let s_t denote the state of system controllers at time t then,
- The nodal voltage is estimated as $s_t = \mathcal{T}(u_1(t), .., u_r(t))$
- The plane coefficients $u_r(t)$ are obtained using,

where,

$\mathcal{T} : \mathbb{Z}^r_{\geq} \to \mathbb{Z}_+$ is a hashing function

For each node

$$\left\{ \begin{aligned} v^{(j)}(t) &= \mathcal{H}^{(j)}_{s_t} \mathbf{x}^\top(t) \\ \mathcal{H}^{(j)}_{s_t} &\triangleq [\alpha_0, \beta_1, .., \beta_p]^{(j)}_{s_t} \end{aligned} \right.$$

$$\mathbf{x}(t) \triangleq [1, x_1(t), .., x_p(t)]$$

Regulator: $u_r(t) \in \{0, 1, .., 32\}$

Cap bank: $u_r(t) \in \{0, 1\}$

$$\mathcal{H}^{(j)}_{s_t} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1} \boldsymbol{X}^\top \boldsymbol{Y}^{(j)}$$

Ordinary Least Squares Estimator (OLS)

Design Matrix
(δ-variations in power injections)

Response Vector
(points on the AC power flow manifold)

Fast PV Hosting Capacity

Design Matrix

- The design matrix $\boldsymbol{X}$ specifies query points on the manifold for the OLS estimator

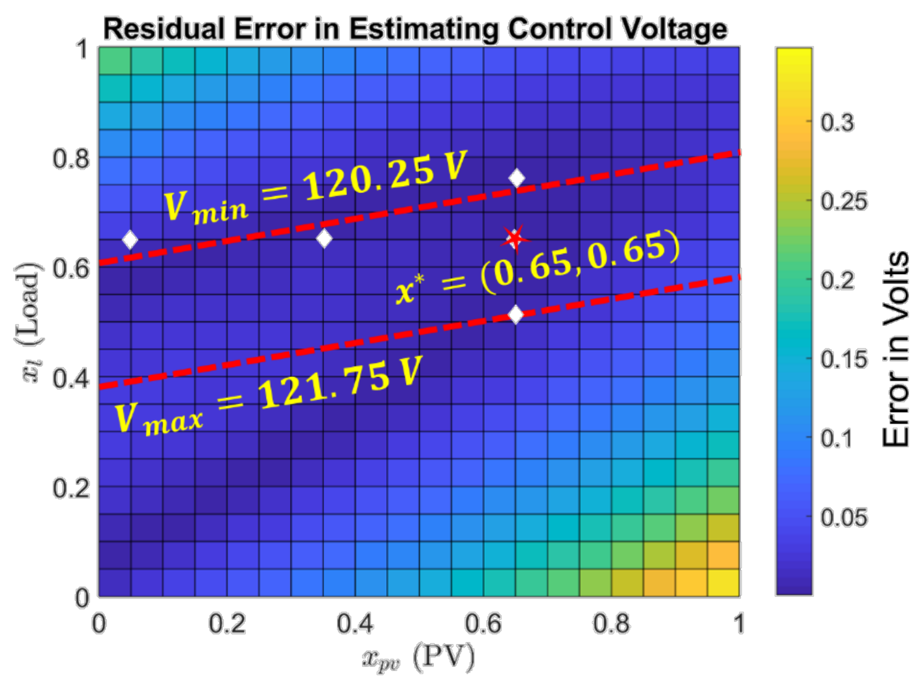
Structure of $\quad$ *for two profiles*

$$\boldsymbol{X} = \begin{bmatrix} 1 & x_l & x_{pv} \\ 1 & x_l + \delta_l & x_{pv} \\ 1 & x_l - \delta_l & x_{pv} \\ 1 & x_l & x_{pv} + 0.5\gamma \\ 1 & x_l & x_{pv} + \gamma \end{bmatrix}$$

$$\gamma = \begin{cases} +\delta_{pv}, & 0 \leq x_{pv} < 0.5 \\ -\delta_{pv}, & 0.5 \leq x_{pv} \leq 1 \end{cases}$$

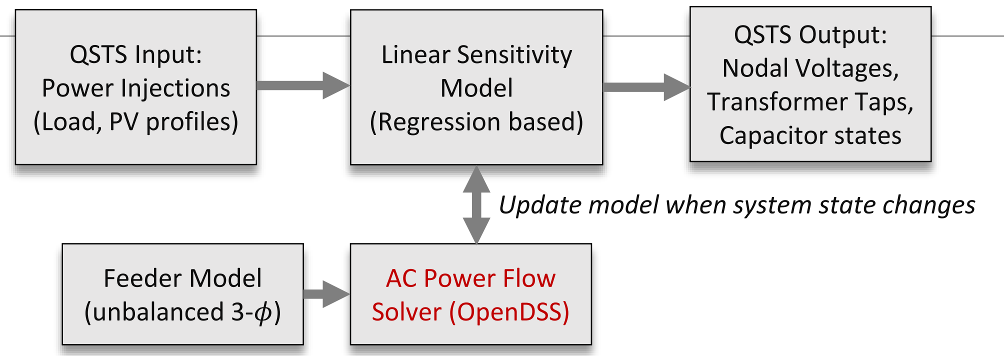
Choose a δ_i that minimizes error in estimating states of system controllers

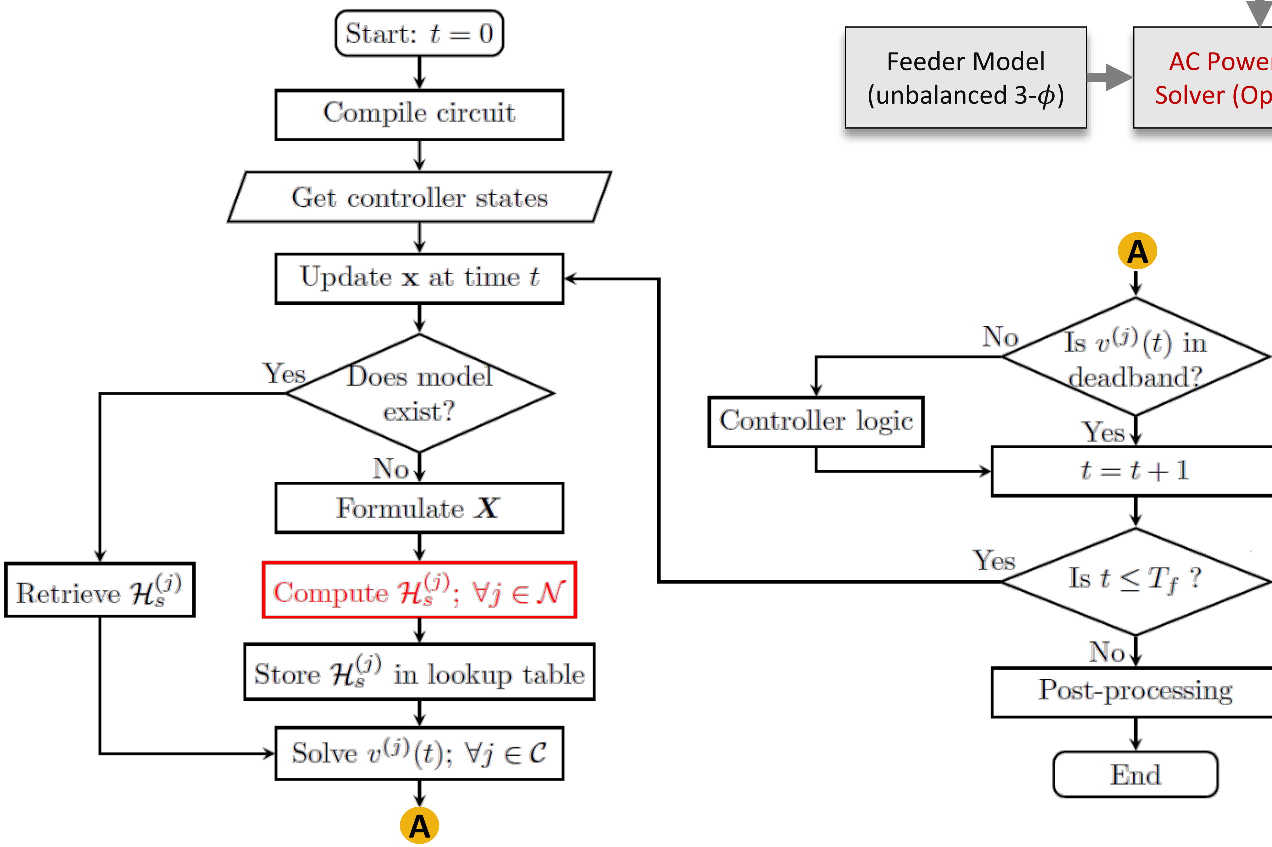
Residual Error Heat Map



Fast PV Hosting Capacity

Fast QSTS Algorithm





where,

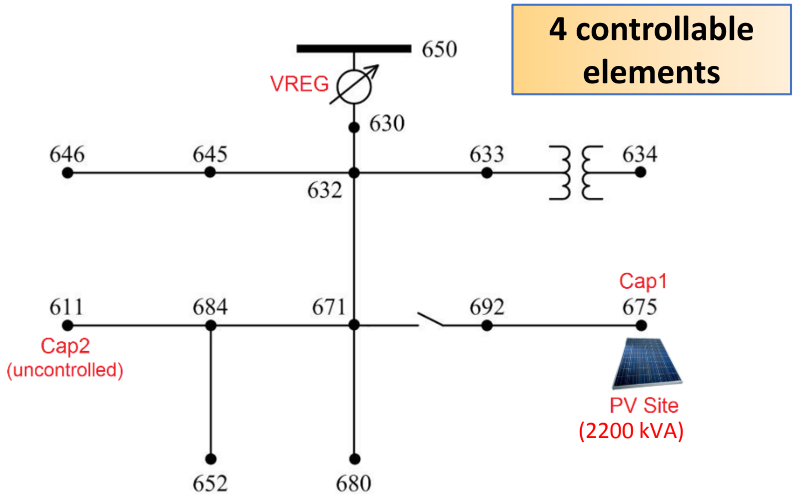
$\mathcal{N}$ is a set of all nodes

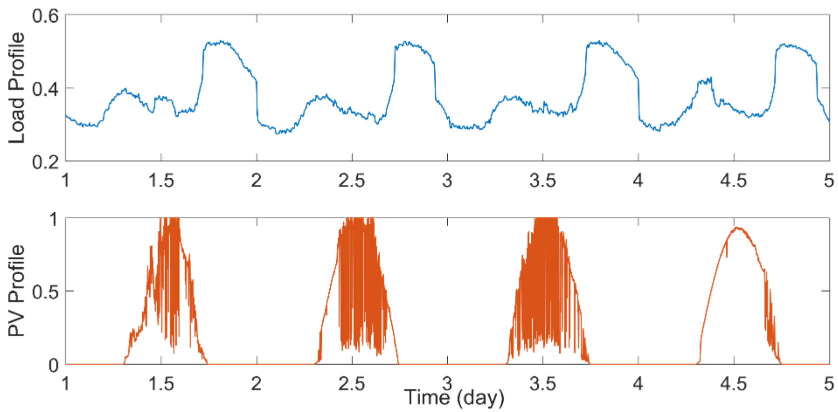
$\mathcal{C}$ is a set of control nodes $\mathcal{C} \subset \mathcal{N}$

T_f is the time horizon

Fast PV Hosting Capacity

Test case 1: IEEE 13-bus





- Controllers:
  - Three 1ϕ voltage regulators (with LDC)
  - One 3ϕ controllable cap bank (600 kVAR)
  - One 1ϕ cap bank fixed (100 kVAR)

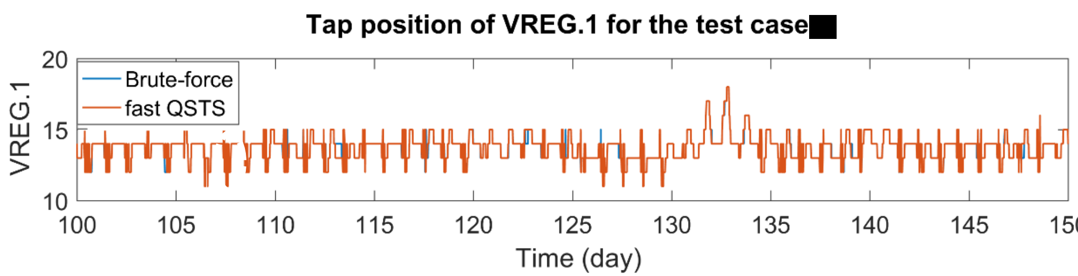
- PV system: 3ϕ 2MW (40% of peak load)
- Time-series Inputs (1 year, 1-sec):
  - 1 Load profile from actual SCADA data
  - 1 PV profile based on irradiance data (Hawaii)

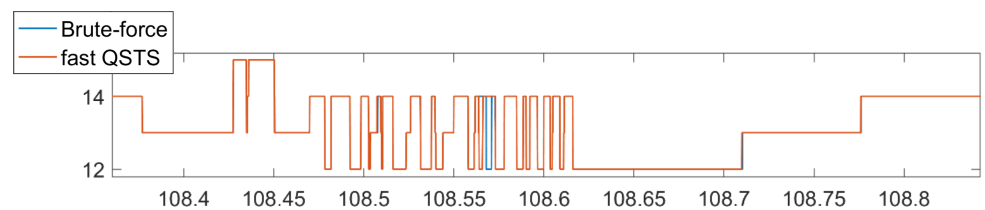
Fast PV Hosting Capacity

Test case 1: IEEE 13-bus, cont.

0.53% RMS Error

QSTS Metric	Brute-force	Fast QSTS (error)
Regulator tap actions:		
VREG.1 ([illegible]	[illegible]	-0.42%
VREG.2 ([illegible]	[illegible]	-0.15%
VREG.3 (C-ϕ)	8449	-0.52%
Capacitor switches: Cap 1 (3-ϕ)	2504	-1.03%
Feeder phase voltage:		
Highest	1.0607 p.u.	<0.0001 p.u.
Lowest	0.9673 p.u.	0.0001 p.u.
Duration of ANSI violations:		
Over voltage	22.13 Hrs	-0.07 Hrs
Under voltage	11.47 Hrs	+0.75 Hrs
Per phase voltage (each bus):	total of 41 nodes	
Highest	0.0003 p.u. (mean error)	
Lowest	<0.0001 p.u. (mean error)	





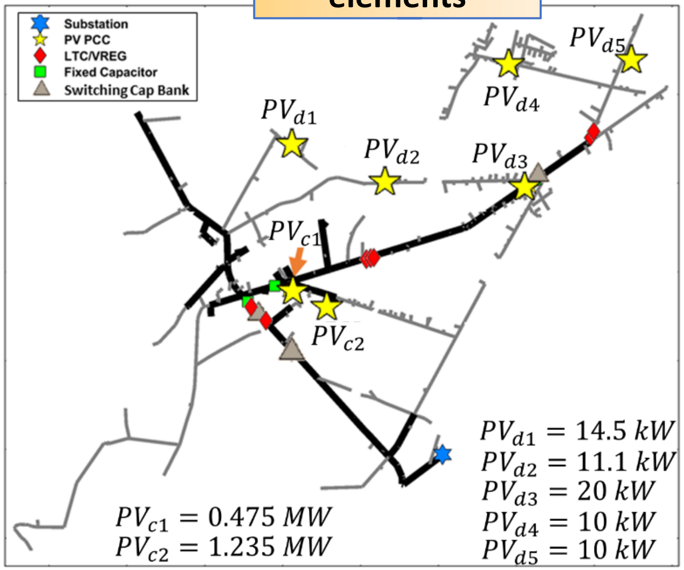
	Brute-force	Fast QSTS	% Reduction
Total time taken	14.25 mins	13.3 secs	98.4%
Power flow solutions	31.5 million	1015	99.9%

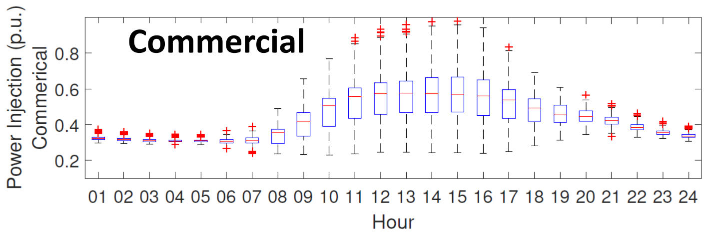
Fast PV Hosting Capacity

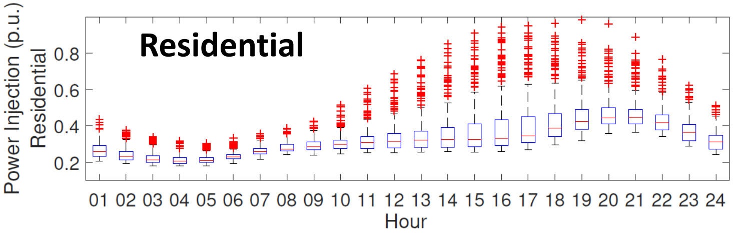
Test case 2: utility feeder j1

- 18.1 km, 12 kV feeder with 4,242 nodes
- 1,300 residential, C&I industrial loads (6.3 MW)
- 12 controllable elements (9-VRs, 3-Cap banks)
- Secondary modeled (wye/delta transformers)
- 7 PV systems installed (centralized, distributed)
- Time-series Inputs (1 year, 1-sec):
    - 3 load profiles (residential, commercial, lumped loads)
    - 7 PV profiles (based on geographic location)

12 controllable elements







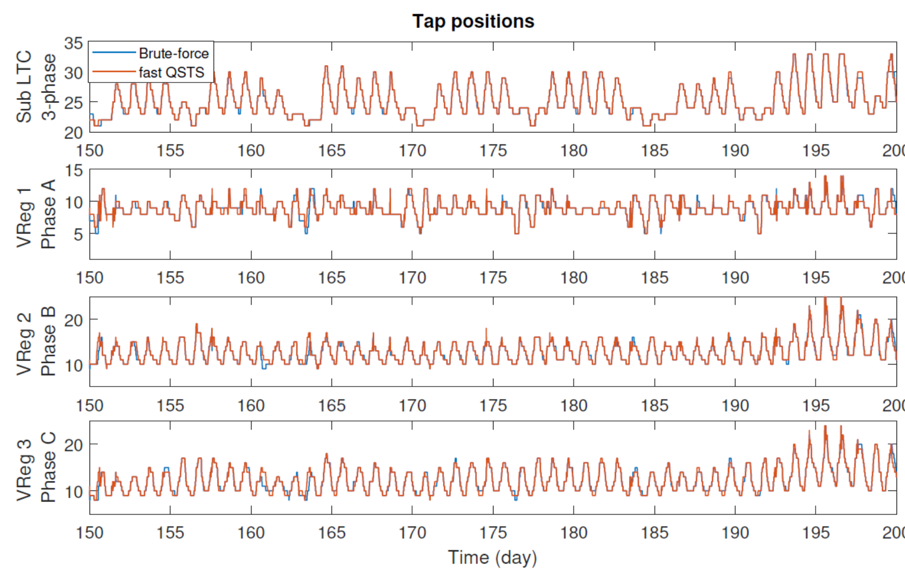
Fast PV Hosting Capacity

Test case 2: utility feeder j1, cont.

QSTS Metric	Brute-force	Fast QSTS (error)
Regulator tap actions:		
Sub LTC (3-ϕ)	433	-0.46%
VReg 1 (A-ϕ)	6194	+0.41%
VReg 2 ([illegible]	[illegible]	0%
VReg 3 ([illegible]	[illegible]	-0.31%
VReg 4 (A-ϕ)	[illegible]	+0.39%
VReg 5 (B-ϕ)	3041	-0.65%
VReg 6 (C-ϕ)	2468	+0.97%
VReg 7 (A-ϕ)	4509	-0.08%
VReg 8 (B-ϕ)	3527	-0.05%
Capacitor switches:		
Cap 1 (3-ϕ)	60	+3.33%
Cap 2 (3-ϕ)	627	0%
Cap 3 (3-ϕ)	11	0%
Feeder phase voltage:		
Highest	1.0883 p.u.	$<$ 0.0001 p.u.
Lowest	0.9365 p.u.	+0.0004 p.u.
Duration of ANSI violations:		
Over voltage	151.34 Hrs	-0.50 Hrs
Under voltage	14.09 Hrs	+1.18 Hrs
Per phase voltage (each bus):	total of 4242 nodes	
Highest	$<$ 0.0001 p.u. (mean error)	
Lowest	0.0001 p.u. (mean error)	

0.55% RMS Error

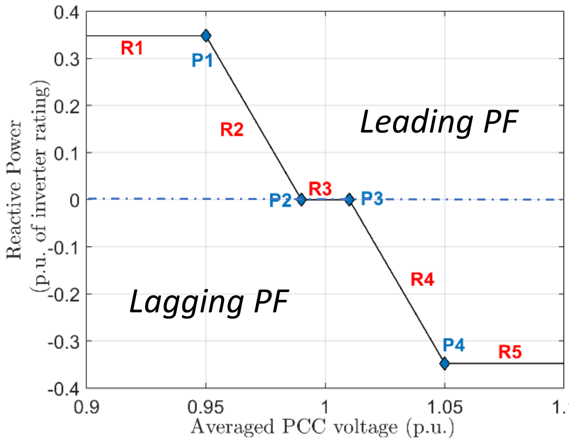
	Brute-force	Fast QSTS	% Reduction
Total time taken	24.3 hours	14.8 minutes	98.98%
Power flow solutions	31.5 million	157,332	99.50%

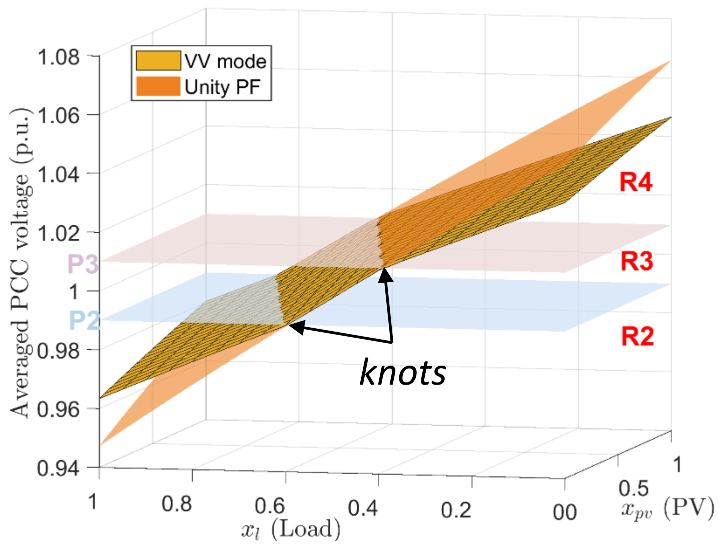


Fast PV Hosting Capacity

Smart Inverter: Volt-VAR (VV) Mode

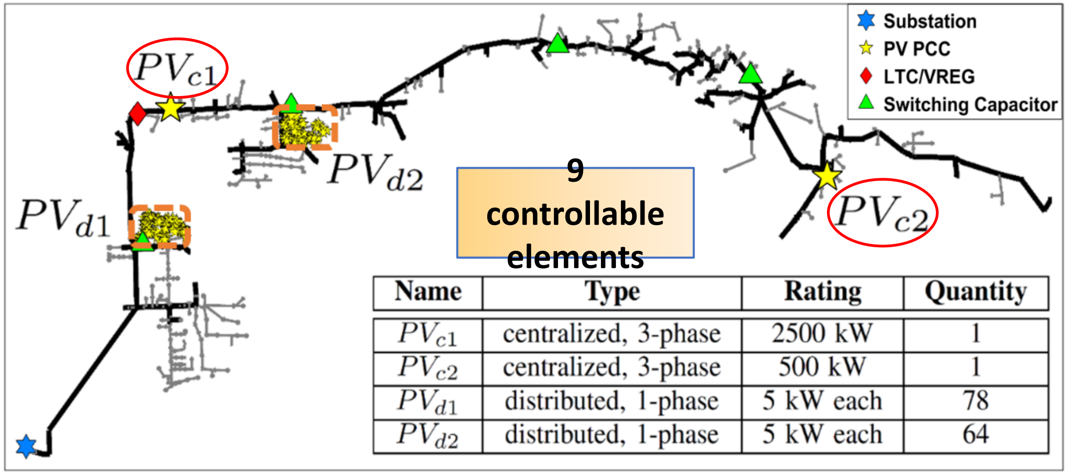
- Inverter varies its reactive power (VAR) feed-in based on the PCC voltage
- Closed loop control
- VV control follows a reference curve:
  - A dead-band (R3)
  - Variable VAR feed-in (R2, R4)
  - Maximum VAR feed-in (R1, R5)
- Each region of the VV curve causes a '*knot*' in the power flow manifold across the entire feeder
- The magnitude of impact is dictated by the size of the inverter and the circuit topology



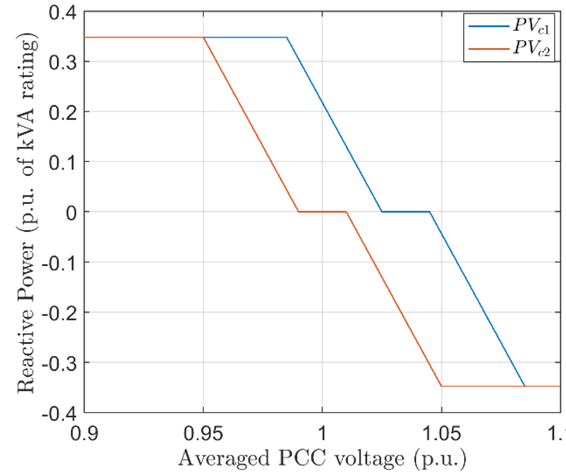


Fast PV Hosting Capacity

Test case 3: Smart Inverter Utility feeder CO1



Name	Type	Rating	Quantity
PV_{c1}	centralized, 3-phase	2500 kW	1
PV_{c2}	centralized, 3-phase	500 kW	1
PV_{d1}	distributed, 1-phase	5 kW each	78
PV_{d2}	distributed, 1-phase	5 kW each	64



- 21.7 km, 12 kV distribution feeder with 5469 nodes
- 9 controllable elements (4-VRs, 5-Cap banks)
- 1,111 single phase loads, 317 three phase loads
- Secondary system modeled (wye and delta transformers)
- 2.71% voltage imbalance

- 144 PV systems (62% penetration)
- PV_{c1}, PV_{c2} in VV mode
- Time-series Inputs (1 year, 1-sec):
  - 2 load profiles
  - 4 PV profiles

Fast PV Hosting Capacity

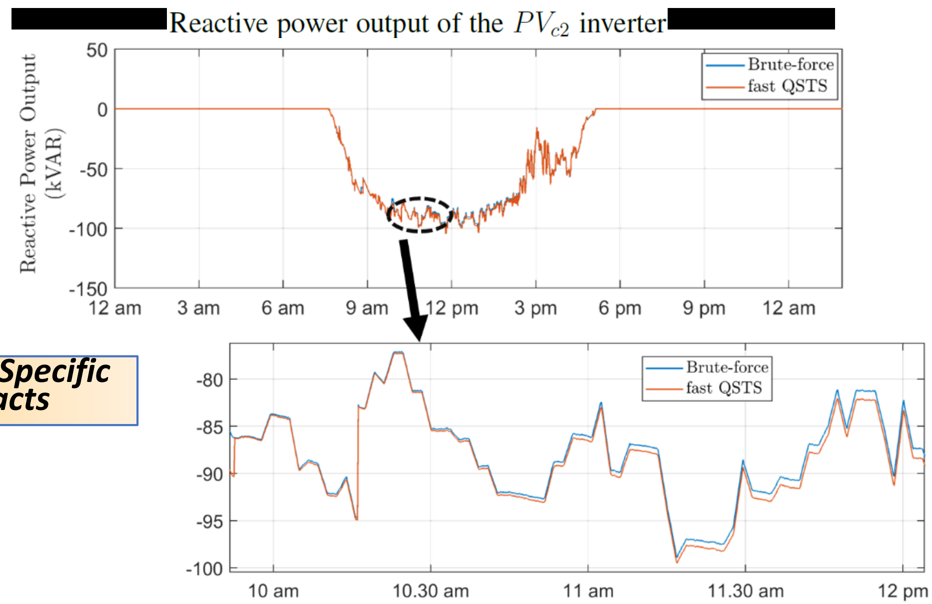
Test case 3: Smart Inverter Utility feeder CO1, cont.

QSTS Metric	Brute-force	Fast QSTS (error)
Regulator tap actions:		
Sub LTC (3-		+0.07%
VReg 1 (A-ϕ)		+1.36%
VReg 2 (B-ϕ)	4822	-0.45%
VReg 3 (C-ϕ)	4704	+1.65%
Capacitor switches:		
Cap 1 (3-ϕ)	360	-0.55%
Cap 2 (3-ϕ)	30	-13.3%
Cap 3 (3-ϕ)	24	-8.33%
Cap 4 (A-ϕ)	526	-0.38%
Cap 5 (B-ϕ)	752	-0.26%
Feeder phase voltage:		
Highest	1.0613 p.u.	-0.0001 p.u.
Lowest	0.9067 p.u.	<0.0001 p.u.
Duration of ANSI violations:		
Over voltage	223.07 Hrs	+11.08 Hrs
Under voltage	129.39 Hrs	-3.56 Hrs
Total kVARh injected:		
PV_{c1} (3-ϕ)	1220.9	-2.69%
PV_{c2} (3-ϕ)	3898.8	-1.46%
Total kVARh absorbed:		
PV_{c1} (3-ϕ)	10.76	-1.46%
PV_{c2} (3-ϕ)	7.923×10^4	+0.58%
Per phase voltage (each bus):	total of 5469 nodes	
Highest	0.0005 p.u. (mean error)	
Lowest	0.0002 p.u. (mean error)	

0.8% RMS Error

Inverter Specific Impacts

	Brute-force	Fast QSTS	% Reduction
Total time taken	67.4 hours	29.8 minutes	99.26%
Power flow solutions	31.5 million	78,884	99.74%



Fast PV Hosting Capacity

Conclusions

- A fast QSTS algorithm is developed
- Leverages the concept of a power flow manifold and dynamic regression.
- On average, 150 times faster than brute-force QSTS
- All voltage and current related PV impacts accurately estimated
- Performance demonstrated on a variety of test cases
- Potential applications:
  - PV interconnection analysis tool
  - Probabilistic hosting capacity
  - Sensitivity-based hosting capacity
  - Optimal smart inverter settings

Summary

- Massive DERs and emerging DER data
- Applications:
  1. Reliable detection of solar PV installations
     - Change point detection
     - Neural Network
  2. Insight into mechanisms for energy storage revenue
  3. Scalable fast PV hosting capacity

Thanks!

References

1. J. Peppanen, M.J. Reno, M. Thakkar, S. Grijalva, R.G. Harley, "Leveraging AMI Data for Distribution System Model Calibration and Situational Awareness", *IEEE Transactions on Smart Grid*, Vol. 6, No. 4, pp. 2050-2059, Jan. 2015.
2. J. Peppanen, M. J. Reno, R. Broderick, and S. Grijalva, "Distribution System Model Calibration with Big Data from AMI and PV Inverters," *IEEE Transactions on Smart Grid*, Vol. 1, No. 99, pp 1-10, March 2016.
3. X. Zhang, S. Grijalva, "A Data Driven Approach for Detection and Estimation of Unauthorized Residential PV Installations", *IEEE Transactions on Smart Grid*, Vol. 1, No. 99, pp 1-10, April 2016.
4. J. Deboever, S. Grijalva, M. Reno, R. Broderick, "Fast Quasi-Static Time-Series (QSTS) for Yearlong PV Impact Studies using Vector Quantization", *Solar Energy Journal*, November, 2017.
5. Y. Seyedi, H. Kulkarni, S. Grijalva, "Irregularity Detection in Output Power of Distributed Energy Resources Using PMU Data Analytics in Smart Grids", *IEEE Transactions on Industrial Informatics*, Vol. 4, pp. 2222-2232, April, 2019, DOI: 10.1109/TII.2018.2865765
6. M.U. Qureshi, S. Grijalva, M.J. Reno, J. Deboever, X. Zhang and R.J. Broderick, "A Fast, Scalable Quasi-Static Time Series Analysis Method for PV Impact Studies using Linear Sensitivity Model", IEEE Transactions on Sustainable Energy, Vol 10, No. 1 Jan, 2019.
7. K. Mason, S. Grijalva, "A Review of Reinforcement Learning for Autonomous Building Energy Management", *Elsevier, Computers and Electrical Engineering*, Submitted, November, 2018.
8. M. J. Reno, K. Coogan, S. Grijalva, R. J. Broderick, and J. E. Quiroz, "PV Interconnection Risk Analysis through Distribution System Impact Signatures and Feeder Zones," *IEEE PES General Meeting*, Washington DC, July 27-31, 2014.
9. X. Zhang, S. Grijalva, and M. J. Reno, "A Time-Variant Load Model Based on Smart Meter Data Mining," *IEEE PES General Meeting*, Washington DC, July 27-31, 2014.
10. X. Zhang, S. Grijalva, "An Advanced Data Driven Model for Residential Plug-in Hybrid Electric Vehicle Charging Demand," *IEEE PES 2015 General Meeting*, Denver, CO, July 26-30, 2015.
11. J. Peppanen, M. J. Reno, R. Broderick, and S. Grijalva, "Secondary Circuit Model Generation Using Limited PV Measurements and Parameter Estimation," *IEEE PES General Meeting*, Boston, MA, USA, July 17-21, 2016 (*).
12. J. Peppanen, X. Zhang, S. Grijalva, M. Reno, "Handling Bad or Missing Smart Meter Data through Advanced Data Imputation", *IEEE Innovative Smart Grid Technologies* (ISGT), Minneapolis, MN, September 6-9, 2016.
13. X. Zhang, S. Grijalva, M. Reno, J. Deboever, R. Broderick, "A Fast, Quasi-Static Time Series (QSTS) Simulation Method for PV Impact Studies Using Voltage Sensitivities of Controllable Elements", *IEEE PV Specialists Conference*, Washington DC, June 25-30, 2017
14. J. Deboever, S. Grijalva, M. Reno, X. Zhang, R. Broderick, "Scalability of the Vector Quantization Approach for Fast QSTS Simulation for PV Impact Studies", *IEEE PV Specialists Conference*, Washington DC, June 25-30, 2017.
15. S. Vejdan, S. Grijalva, "The expected revenue of energy storage from energy arbitrage service based on realistic market data statistics," *Texas Power and Energy Conference*, College Station, TX, February 8-9, 2018. (*)
16. S. Vejdan, S. Grijalva, "The value of real-time energy arbitrage with energy storage systems," *IEEE PES General Meeting*, Portland, OR, August 5-10, 2018.
17. M.U. Qureshi, S. Grijalva, M.J. Reno, "Fast Quasi-Static Time Series Simulation Method for PV Smart Inverters with VAR Control using Linear Sensitivity Model", *IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC)*, Waikoloa, HI, June 10-15, 2018.
18. J. Deboever, S. Grijalva, J. Peppanen, M. Rylander, J. Smith, "Practical Data-Driven Methods to Improve the Accuracy and Detail of Hosting Capacity Analysis", *IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC*), Waikoloa, HI, June 10-15, 2018.
19. S. Vejdan, S. Grijalva, "Maximizing the Revenue of Energy Storage Participants in Day-Ahead and Real-Time Markets", *Clemson University Power System Conference*, Charleston, SC, September 4th-7th, 2018.
20. D. Diaz, A. Kumar, J. Deboever, S. Grijalva, J. Peppanen, M. Rylander, J. Smith, "Scenario Selection for Hosting Capacity Analysis of Distribution Feeders with Voltage Regulation Equipment", *IEEE Innovative Smart Grid Technologies (ISGT)*, Washington, DC, February 17-20. 2019.