AI and ML Applications for PV Reliability & System Performance

K. Perry[1], C. Deline[1], D. Jordan[1], M. Deceglie[1], M. Muller[1], M. Springer[1], R. White[1], B. Meyers[2], D. Ragsdale[2], S. Miskovich[2], G. Ogut[3]

(1) National Renewable Energy Laboratory, Golden, CO 80401, USA

(2) SLAC National Acceleratory Laboratory, Menlo Park, CA 94025, USA

(3) Stanford University, Palo Alto, CA 94305, US

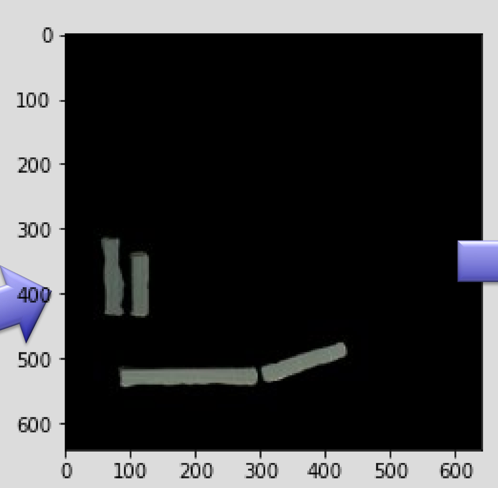
PV Fleets: Automated Data QA and Metadata Verification

Automated Metadata Satellite Analysis using Deep Learning

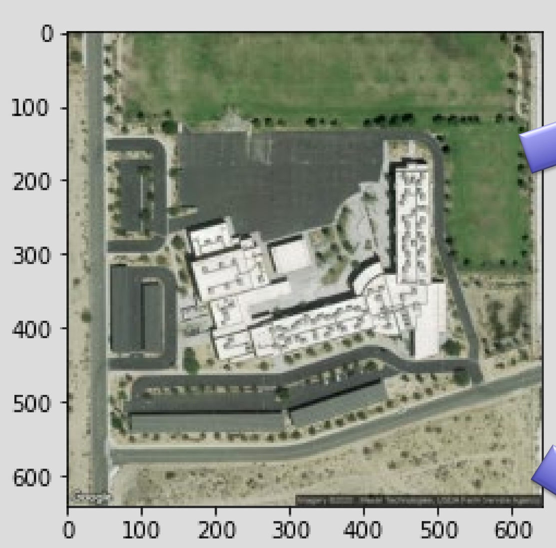
- NREL Panel-Segmentation Package
- Uses deep learning models to automatically do the following:
  - Locate solar installation in Google Maps satellite image
  - Extract solar azimuth
  - Determine the mounting configuration (rooftop, carport, ground; fixed or tracking)
- Imagery analysis great use case for deep learning
- Useful for analyzing fleets where metadata is unknown or incorrect

Panel-Segmentation package:

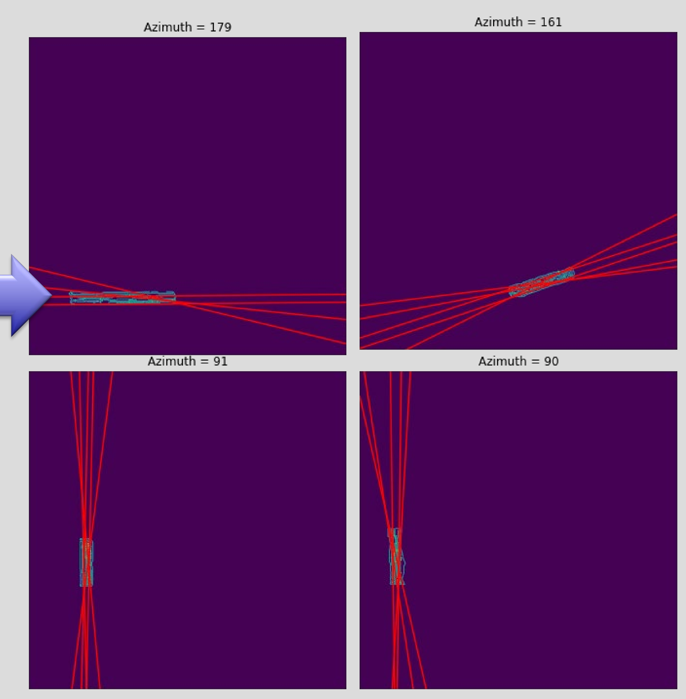
Semantic segmentation: VGG16 model. Find pixel mask of arrays



Google Maps Image



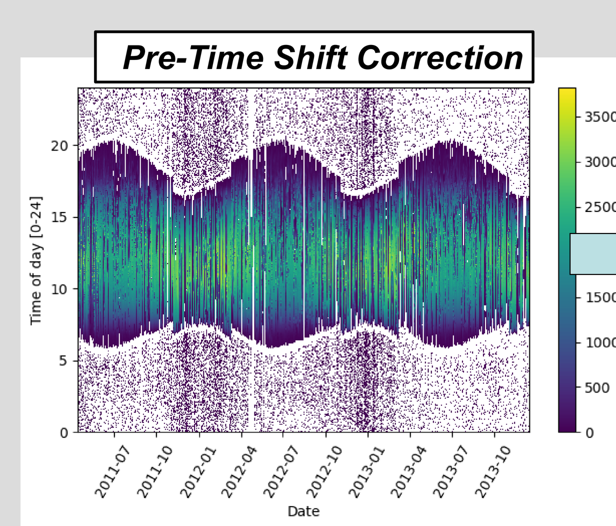
Mounting configuration detection: Resnet-50 model. All arrays determined to be carport-fixed tilt

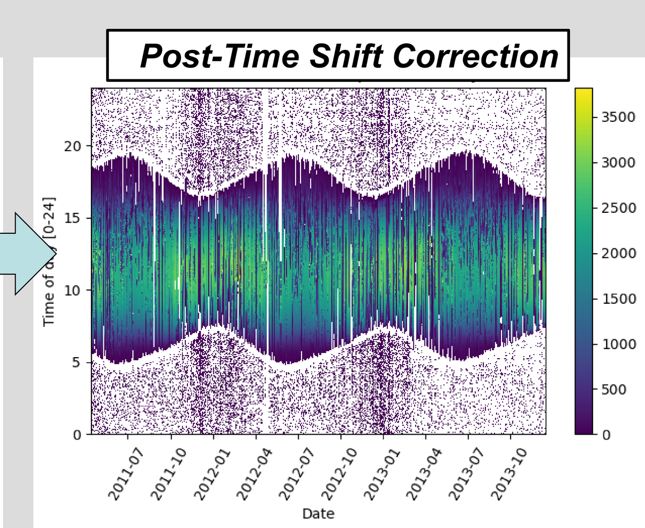


Edge Detection and Hough Line Transform: get the azimuth associated with each array

Automated QA: Clipping Detection, Time Shift Estimation, Capacity Shift Detection

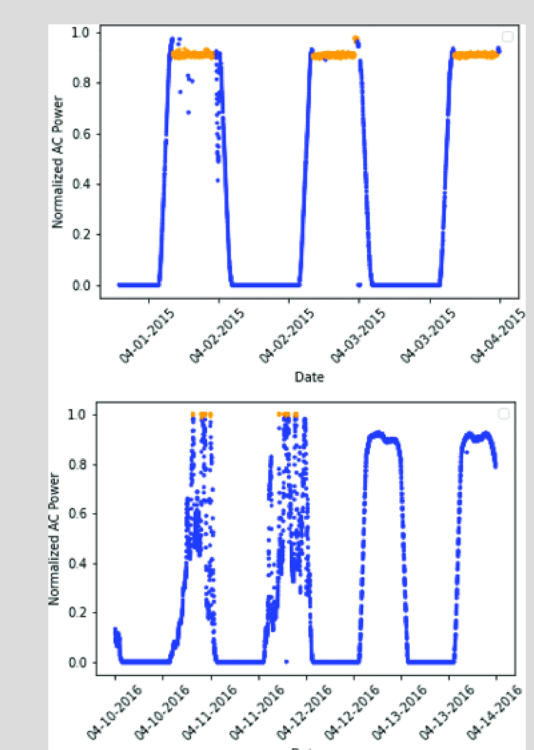
- Developed supervised and unsupervised ML algorithms for finding issues/features in measured PV data
  - **Clipping/curtailment detection:** Logic-based AI method and supervised ML method (XGBoost). Creates mask of clipped/non-clipped periods
  - **Time shift detection:** Unsupervised changepoint detection (CPD) to identify time shifts between modeled and measured solar noon
  - **Capacity shift detection:** Unsupervised CPD to detect abrupt capacity shifts in measured PV data
- All functions validated with "ground-truth" labeled data and results published
- Functions publicly available in Python PVAnalytics package and Rdtools package (clipping only)



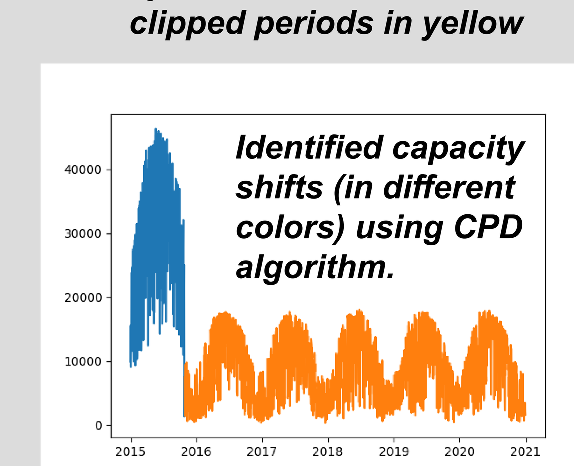


Daily heatmap of AC power values before automated time shift correction and after, respectively, for a data stream with daylight savings time (DST).

PVAnalytics package:

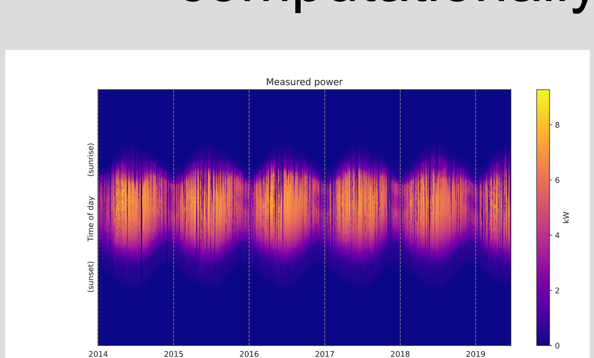


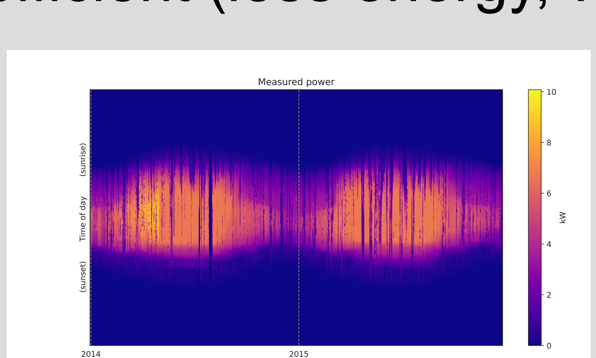
Curtailed and clipped systems, with AI-identified clipped periods in yellow

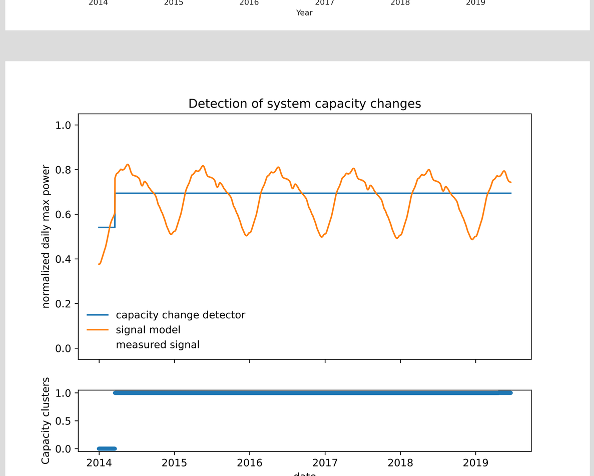


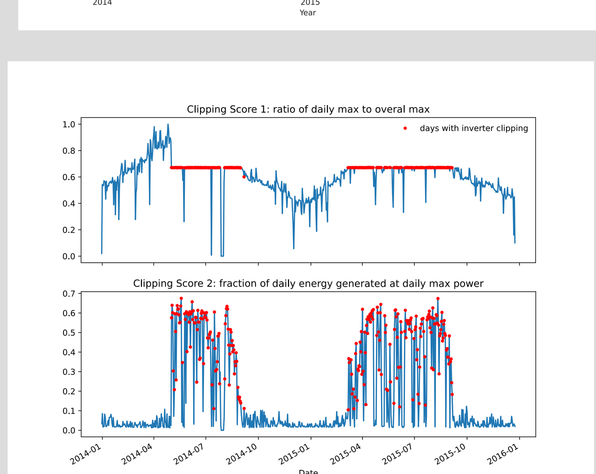
Statistical Learning in PVInsight

- Developing *white-box* machine learning models based on statistical signal processing, convex optimization, and domain expertise
- Deep neural networks are not part of our toolkit!
- **Methods:** we have a *monograph*[1] and a no-math, no-code *tutorial*[2]
- **Applications:** check out this *report*[3] and this *dissertation*[4]
- As opposed to neural networks, this flavor of machine learning is
  - interpretable (good for science *and* troubleshooting!)
  - highly data efficient (good models with 75% data loss!)
  - computationally efficient (less energy, water, cost, …)





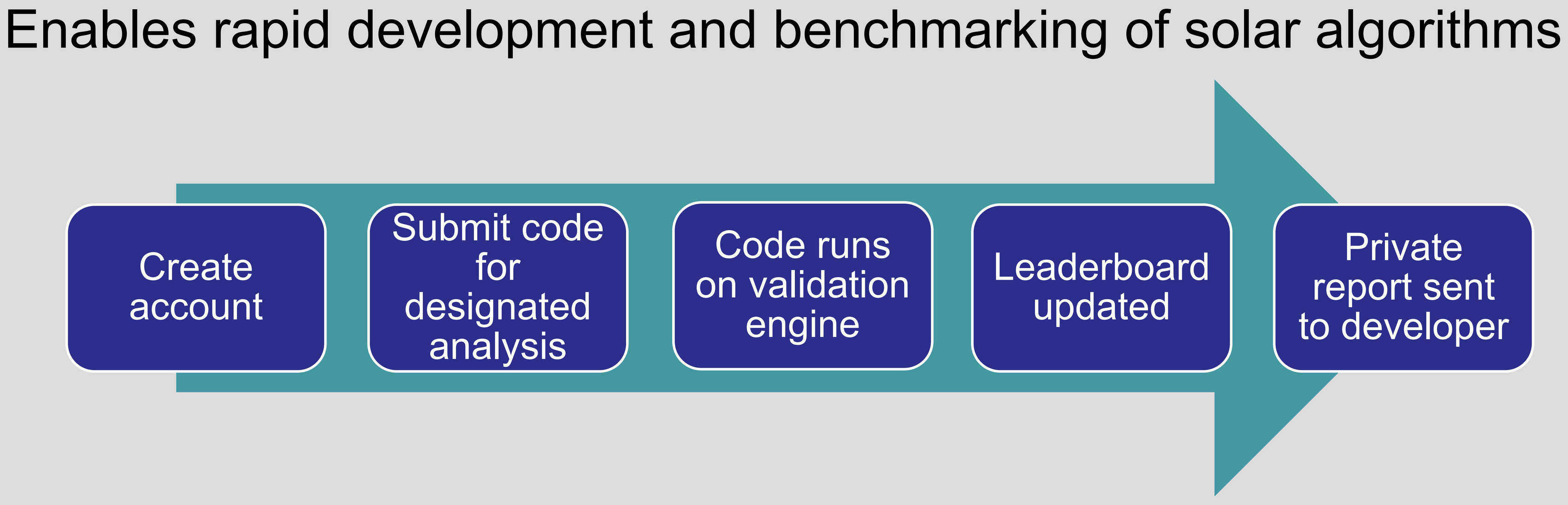


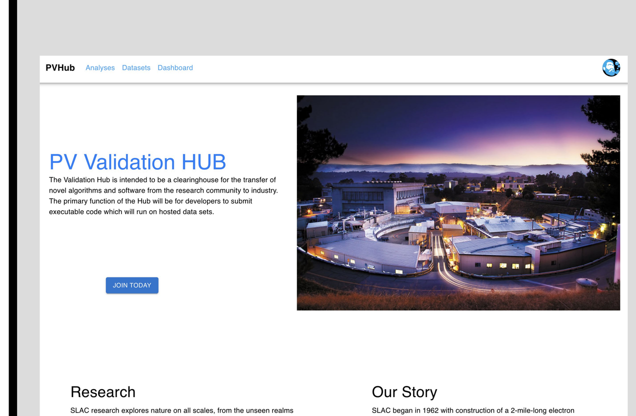


1 2 3 4

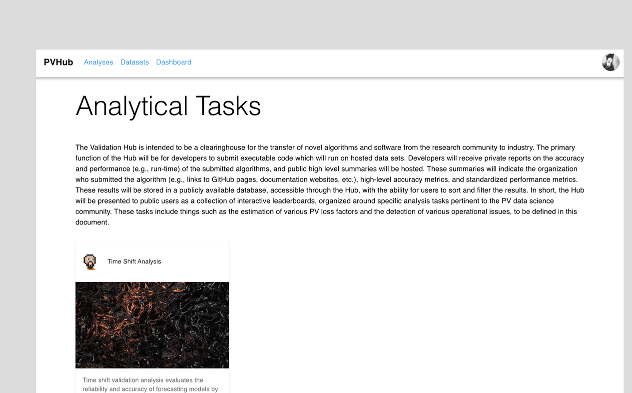
PV Validation Hub

- Allow developers to submit PV analytics algorithms for validation.
  - *Degradation, soiling, tilt/azimuth estimation, etc.*
- Well-curated validation data sets and procedures
- Consistent labeled data sets allow for side-by-side comparison of different algorithms
- Public leaderboards and documentation facilitate tech transfer
- Enables rapid development and benchmarking of solar algorithms

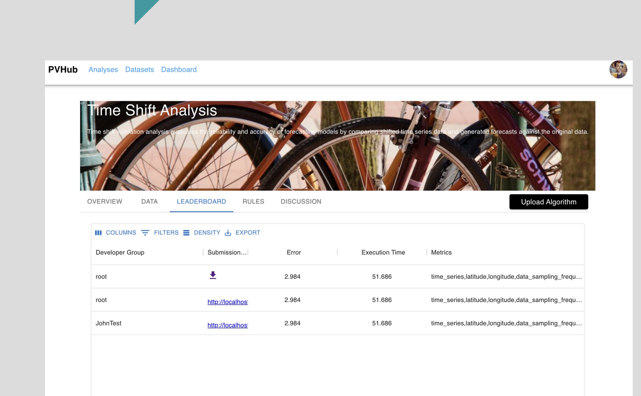




Main page of the site



List of tasks



Example leaderboard

This work was supported by the U.S. Department of Energy under SETO Contract Numbers 38258 and 38529 with the National Renewable Energy Laboratory and SLAC National Accelerator Laboratory.

SETO Workshop on Solar Applications in AI and ML, Alexandria VA, October 31st 2023

Funding was provided by the U.S. Department of Energy
Solar Energy Technologies Office
Building Technologies Office

AI-based Optimal Design and Controls Can Greatly Reduce Carbon Emissions and Enhance Resilience in Residential Communities in Cold Climates

AI-Driven Smart Community Control for Accelerating PV Adoption and Enhancing Grid Resilience

Principal Investigator: Xin Jin
PI Email: Xin.Jin@nrel.gov

Additional Project Contributors:

NREL: Fei Ding, Rajendra Adhikari, Sathya Balamurugan, Michael Blonsky, Jianli Chen, Lieko Earle, Rawad El Kontar, Utkarsh Kumar, Jeff Maguire, Prateek Munankarmi, Paul Norton, Harsha Padullaparti, Bethany Sparn, and Jing Wang

Partners:

- Habitat for Humanity Roaring Fork Valley
- Holy Cross Energy
- Copper Labs
- Conservation Labs
- Thrive Home Builders
- City of Fort Collins
- A.O. Smith

Introduction

- Net-zero energy residential communities are crucial for achieving decarbonization goals, but the high-penetration PV in those communities is posing challenges to the distribution grid.
- Traditional design and operation of net-zero communities rely on rule-of-thumb methods and may not work in complex scenarios.
- AI/ML methods can optimally size PV for net-zero energy, identify user preferences and usage patterns, and fully unlock the potential of DERs to address distribution grid issues.

How is AI/ML used in this project

- **PV sizing:** ML-based automated workflow identifies the optimal placement of rooftop PV in a residential community to maximize solar production and operational cost savings.
- **Data-driven learning:** Various data-driven methods were used to identify building models, user preferences, and user behavior to inform decision-making.
- **Control:** Optimization-based control of BTM resources in a residential community to improve grid reliability and resilience.

Challenges and best practices

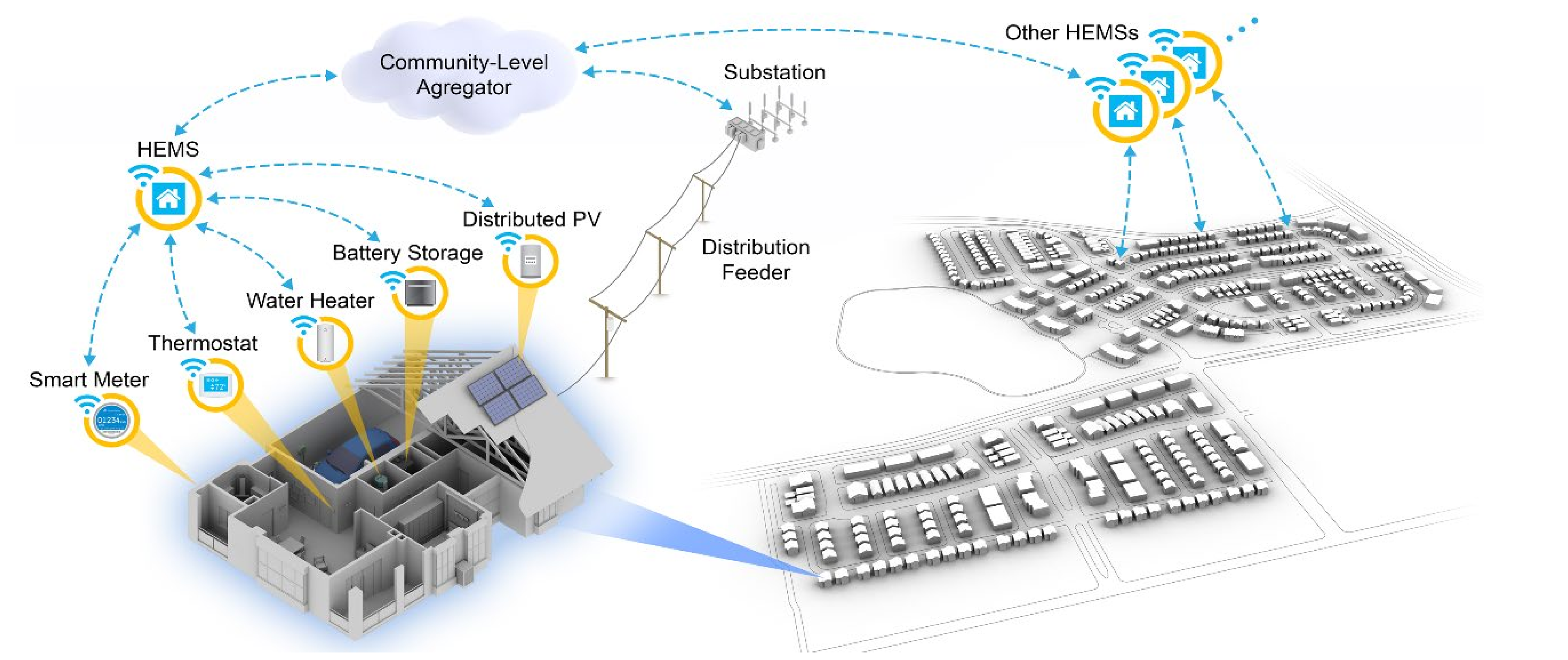
- User preferences and behavior are uncertain. *Solution*: Retrain ML models periodically with a mix of new data and old data and focus on predicting behavior that has a higher impact on control.
- Optimization-based control of a large population of BTM resources is computationally challenging. *Solution*: Formulate the complex problem in a hierarchical manner to make it scalable.

Key takeaways and future work

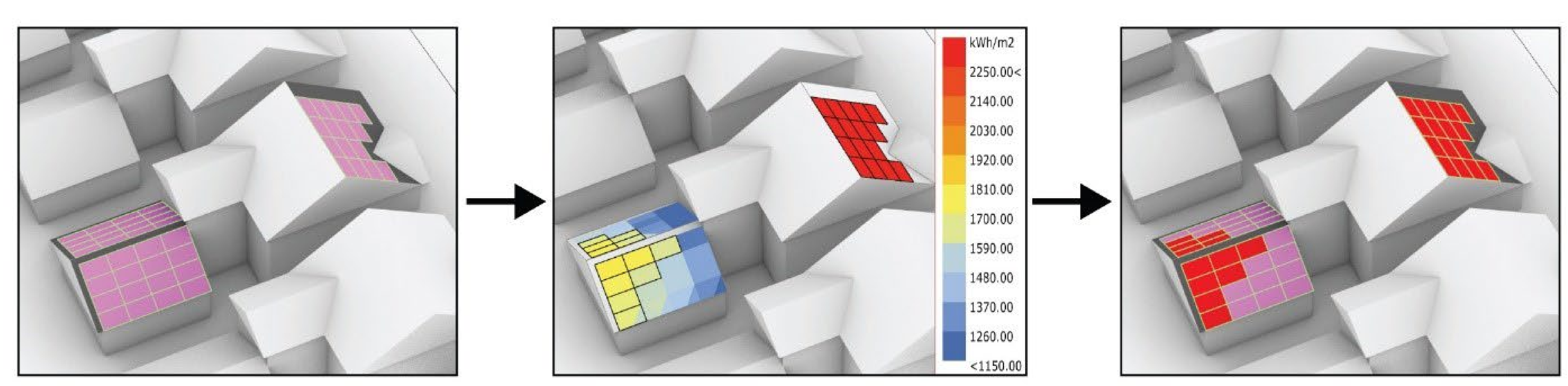
- **Key takeaways**: AI/ML can help residential communities meet the net-zero energy design goal without over sizing the PV and improve grid reliability and resilience through advanced controls.
- **Future work**: Large-scale demonstration in real-world environment under various operational scenarios.

Research Highlights

AI-driven Hierarchical Control for Scalable Management of DERs: HEMS manages each home's behind-the-meter DERs, and community-level aggregators coordinate the HEMS and the grid.

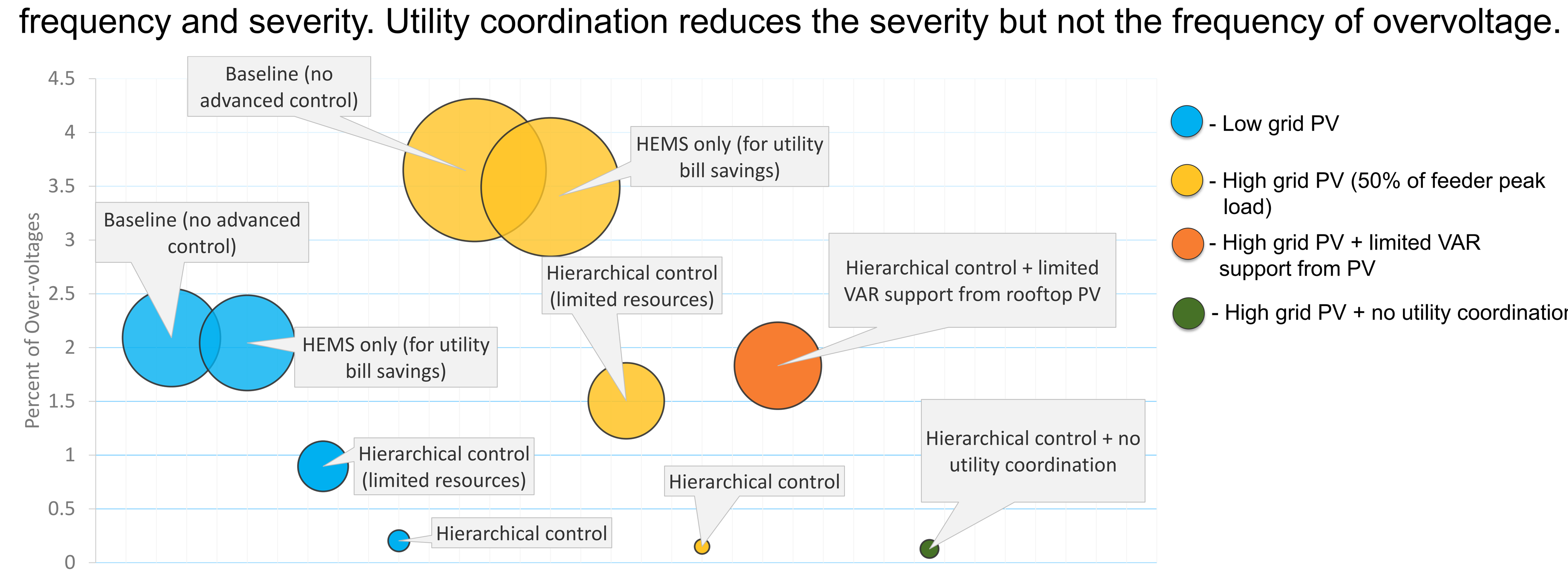


ML-based Optimal Rooftop PV Placement: An automated, ML-based workflow was implemented in architectural design software to optimally place rooftop PV in a residential community to achieve the net-zero energy goal considering roof geometry, orientation, shading, irradiance, etc.



(a) Automated PV panel layout (b) Solar radiation study (c) Panels selection

Addressing overvoltage issues caused by high-penetration PV: Unlike HEMS that focuses on utility bill savings, community aggregators and VAR support effectively reduce the overvoltage frequency and severity. Utility coordination reduces the severity but not the frequency of overvoltage.



Field Demonstration in an Affordable Housing Community: Performed field demonstration in four homes at the Basalt Vista community in Colorado. Achieved 3.1 kW average load reduction and 4.5 kW peak demand reduction during a 5-hour peak period in field experiments.

2023 SETO Workshop on the Solar Applications of Artificial Intelligence and Machine Learning
Alexandria, VA
10/31/2023 – 11/1/2023
NREL/PO-5500-87892

AI-Based Protection Schemes for DER

Award # DE-EE0036533: "Adaptive Protection and Control for High Penetration PV and Grid Resilience"

Principal Investigator: Matthew Reno

Introduction:

- Distribution system protection is becoming more complicated with DER producing reverse fault current, decreasing short circuit capabilities, and inverters' unpredictable response for current injection characteristics and angles
- Embedding machine learning in relays can improve the protection system reliability, speed, and accuracy
- AI algorithms in relays can provide:
  - Backup under resilience scenarios when communication is lost
  - Faster response time than some conventional protection algorithms

Challenges:

- There are very few faults in the field that can be used as training data
  - Therefore simulation data generally has to be used, which requires extensive simulation time and is reliant on the model accuracy
- Each relay is unique with different surrounding system topologies, types of protection (overcurrent vs. distance), and experiences different fault currents
- Protection operates in milliseconds, so the AI/ML algorithm has to be able to run in real-time very quickly
- Even with 99.9% success rate of an AI/ML protection scheme, dozens of daily misoperations or nonoperations would result if the technology is widely deployed across a major grid
- Adoption is challenging for black-box AI methods that are not explainable (training) or verifiable (certification testing)

Best Practices:

- Working with IEEE PSRC Standards to develop best practices for "Applications of Artificial Intelligence and Machine Learning in Power System Protection and Control"
  - This includes discussion on best practices for data types, data structures, hardware, software, implementation, redundancy in case of failure, testing and validation, and user training
- While algorithms are trained based on simulations of the actual system, in order to evaluate the accuracy of the algorithm, it must be tested on real fault data
- In order to test the speed of the algorithm, it must be implemented in actual hardware (not just simulation) to evaluate the real-time speed using hardware-in-the-loop (HIL) testing or field testing

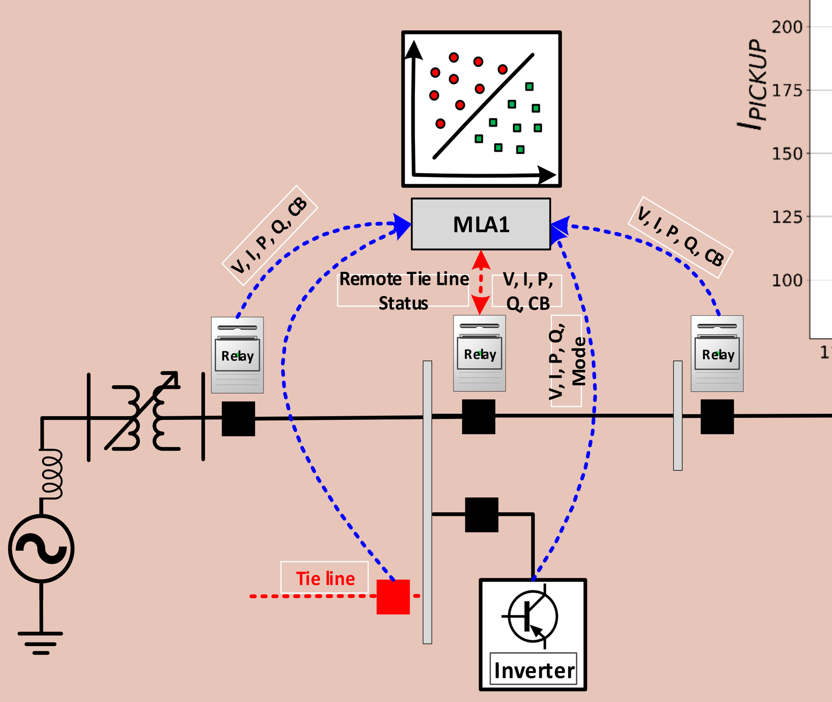
Key Takeaways:

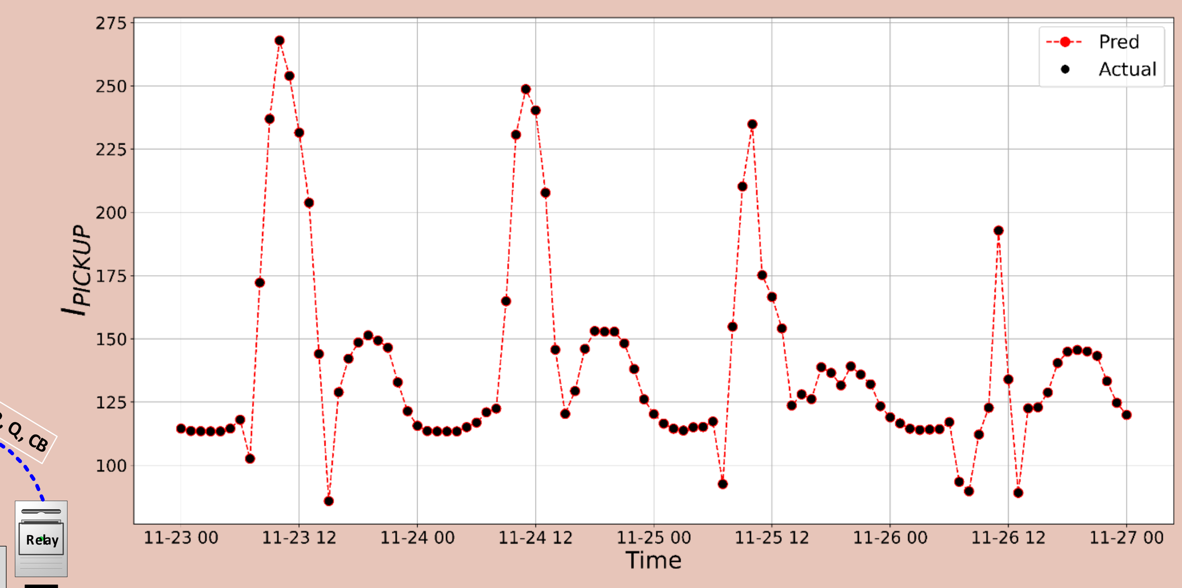
- If AI uses the same features/patterns used by relays, why would it be able to create a more dependable and secure classifier?
- If AI uses more complex and abstract features/patterns that have no transparent relation to the underlying physics, and thresholds (separation planes) are created simply by learning through data, why would it work better?
- AI/ML should be applied to protection problems where physics-based models are not as applicable or well-understood

How is AI/ML used in this project?

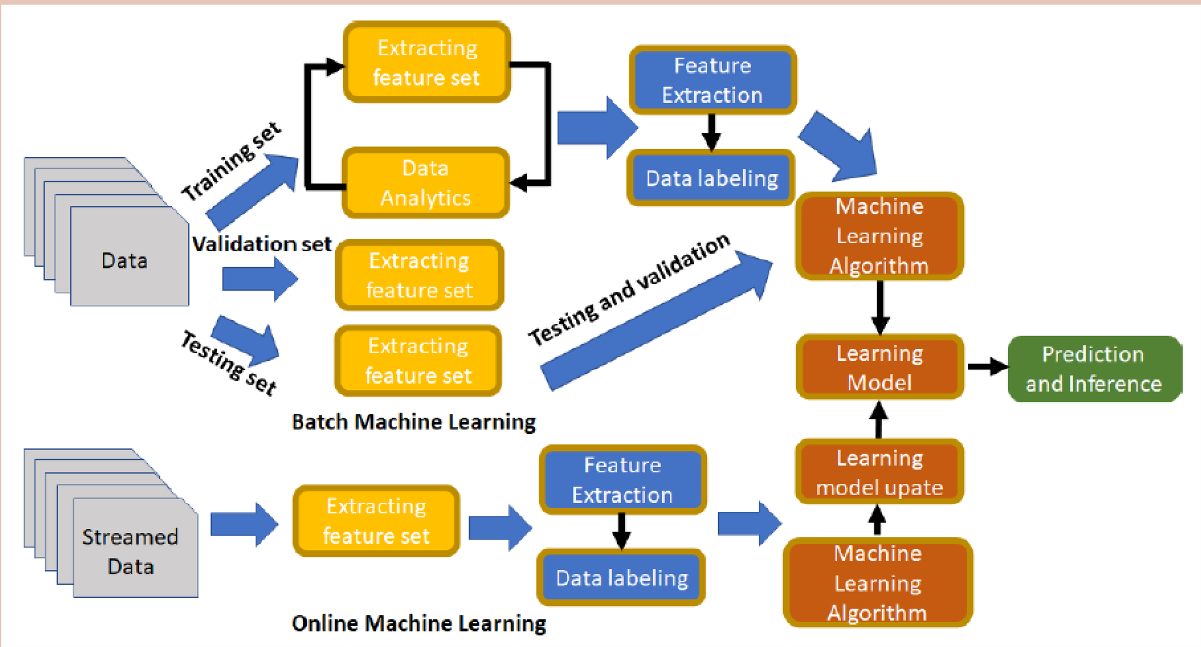
Multiple Use Cases for AI/ML in Protective Relays Considered:

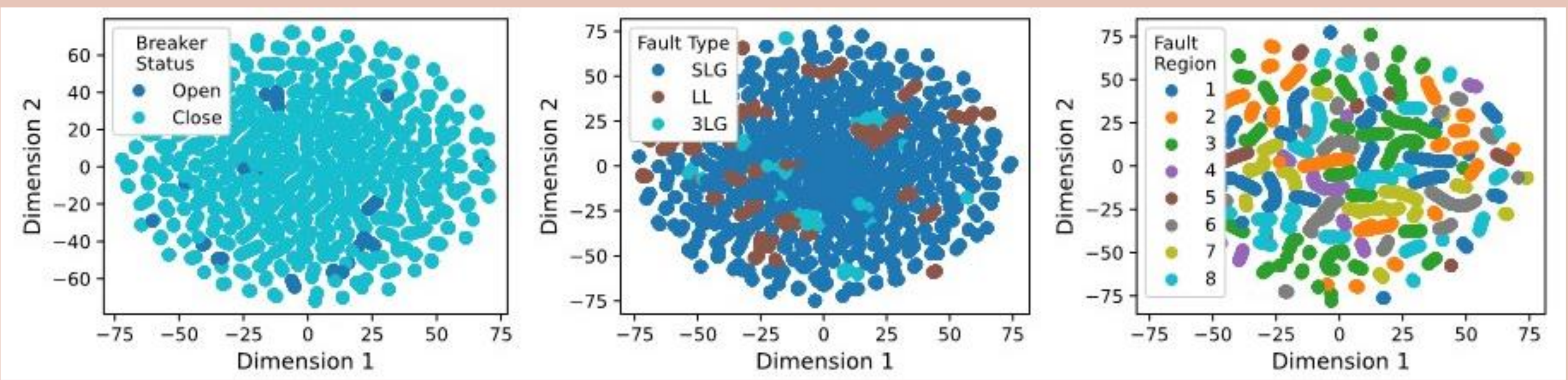
1) For adaptive protection, ensures reliable communication-free operation by local learning at each relay of expected communication from other devices for the grid state – learning settings/trends [1]



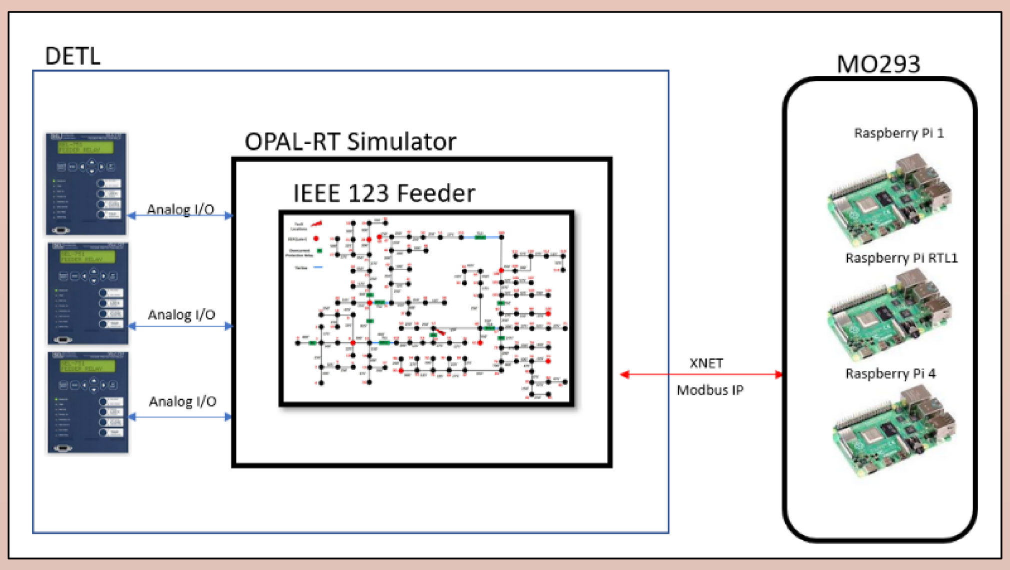


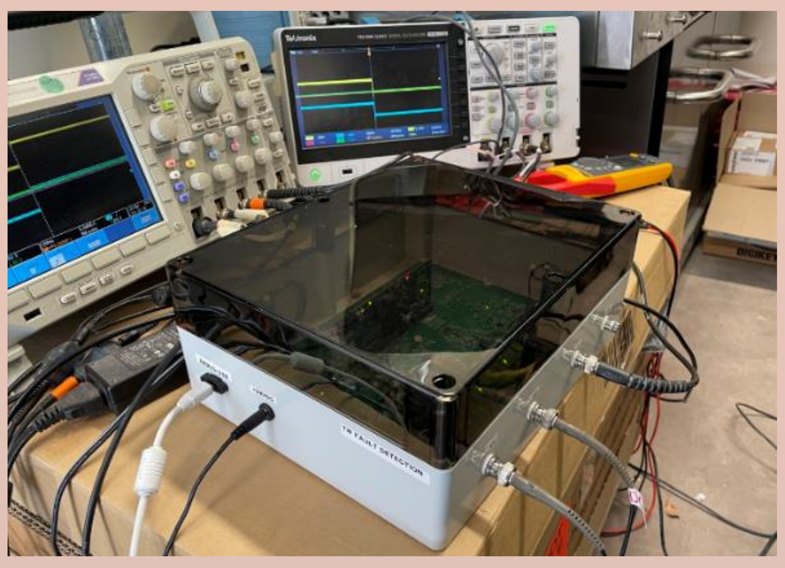
2) AI/ML based fault detection, classification, and location using NB/NN [2], CNN [3][4], transfer learning [5], Graph NN [6], SVM [7-9], and Random Forest [10]





3) Developed custom AI-based relays to test the algorithms in Opal-RT HIL side-by-side with the SEL relays to validate the performance [11] and finally deployed for field testing [12]





[1] A. Summers, T. Patel, R. Matthews, and M. J. Reno, "Prediction of Relay Settings in an Adaptive Protection System", IEEE Innovative Smart Grid Technologies (ISGT), 2022.
[2] S. T. Ojetola, M. J. Reno, J. Flicker, D. Bauer, and D. Stoltzfuz, "Testing Machine Learned Fault Detection and Classification on a DC Microgrid", IEEE Innovative Smart Grid Technologies (ISGT), 2022.
[3] A. R. Kurup, A. Summers, A. Bidram, M. J. Reno, and M. Martinez-Ramon, "Ensemble models for circuit topology estimation, fault detection, and classification in distribution systems", Sustainable energy, grids and networks, 2023.
[4] S. Paul, S. Grijalva, M. Jiménez-Aparicio, and M. J. Reno, "Knowledge-based Fault Diagnosis for a Distribution System with High PV Penetration", IEEE Innovative Smart Grid Technologies (ISGT), 2022.
[5] S. Paruthiyil, A. Bidram, and M. J. Reno, "A Physics-informed Learning Technique for Fault Location of DC Microgrids Using Traveling Waves," IET Generation, Transmission, and Distribution, 2022.
[6] M. Jimenez-Aparicio, J. Hernandez-Alvidrez, A. Y. Montoya, M. J. Reno, "Embedded Real-Time and Distributed Traveling Wave Fault Location Method using Graph Convolutional Neural Networks" Energies, 2022
[7] C. B. Jones, A. Summers, M. J. Reno, "Machine Learning Embedded in Distribution Network Relays to Classify and Locate Faults," IEEE Innovative Smart Grid Technologies (ISGT), 2021.
[8] B. P. Poudel, A. Bidram, M. J. Reno, and A. Summers, "Zonal Machine Learning-based Protection for Distribution Systems", IEEE Access, 2022.
[9] R. Montoya, B. Poudel, A. Bidram, M. J. Reno, "DC Microgrid Fault Detection Using Multiresolution Analysis of Traveling Waves," International Journal of Electric Power & Energy Systems, 2022.
[10] F. Wilches-Bernal, M. Jiménez-Aparicio, and M. J. Reno, "Algorithm for Fast Fault Location and Classification Based on Mathematical Morphology and Machine Learning", IEEE Innovative Smart Grid Technologies (ISGT), 2022.
[11] M. Jimenez-Aparicio, J. Hernandez-Alvidrez, A. Y. Montoya, and M. J. Reno, "Micro Random Forest: A Local, High-Speed Implementation of a Machine-Learning Fault Location Method for Distribution Power Systems", IEEE PES General Meeting, 2023.
[12] S. Paruthiyil, A. Bidram, M. Jimenez Aparicio, J. Hernandez, and M. J. Reno, "Hardware Implementation of a Traveling Wave Protection Device for DC Microgrids" IEEE Kansas Power & Energy Conference, 2023.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. ____________

Artificial Intelligence-based PV Power Forecast and Energy Management Systems of Power Plants and Utility-Scaled Hybrid PV+ BESS

SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

FAMU-FSU College of Engineering

Hector Akuta[1], Yuan Li[2], Caleb Crow[3], Uthandi Selvarasu[4], Shumeng Wang[5], Mahshid Amirabadi[6], Brad Lehman[7]

[1,2]Florida State University, [3]City of Tallahassee Electric Utility, [4,5,6,7] Northeastern University

This material is partially based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number DE-EE0009340, Unified Universal Control and Coordination of Inverter-Based Resources, and Validation for a PV + Battery Hybrid Plant, PI: Fang Zheng Peng.

Introduction

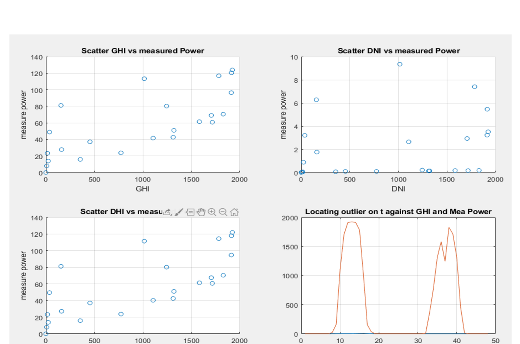
- Research team received field data from City of Tallahassee Electric & Gas Utility (TAL), Florida, of ~62 MW hourly power generation in two years (2020-2021), including two PV power plants (20 MW and 42 MW), and CC/IC generators to perform PV power forecast and battery energy storage system (BESS) management.
- Energy storage is assessed in terms of economic benefits, battery sizing, reserve reduction for power grid.
- **Problem statement**
  - PV Forecast error rises energy storage size, power plant/grid operation cost.
  - BESS is applied to replace conventionally IC-based reserve capacity, however, high capital investment of BESS, makes it crucial to utilize it profitably.
  - PV generation is weather-sensitive → Uncertainty in battery power → State-of-charge cycles → Cyclic degradation → Cost increasing.
- **Tasks performed**: Energy management controller design and development
  - 24 hr.-ahead (HA) PV power forecast implementation, evaluation and assessment
  - Battery life-cycle and economic costs modeling and simulation
  - Energy scheduling control algorithms development and implementation

How is AI/ML used in this project

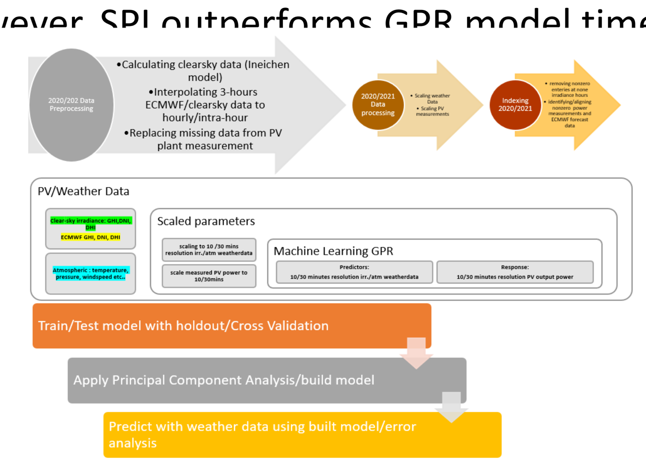
- Multiple machine learning approaches are evaluated and assessed for prediction to increase PV power forecast accuracy over the baseline (persistence method).
- A dynamic programming (DP) based battery management system (BMS) algorithm is developed to substitute the IC units for reserves to improve economic benefit.
- A real-time degradation estimation approach for batteries in hybrid power plant is developed.

PV Power Forecast

✓ Gaussian Process Regression with different kernels and outlier detection, which scores > 30% error improvement on 24 –HA nRMSE index vs. persistence and solar power index (SPI) , SVM, boost and bagged tree. However, SPI outperforms GPR model timescales < 30mins.



Residual vs. Predicted value plots



ML-based solar power forecast process

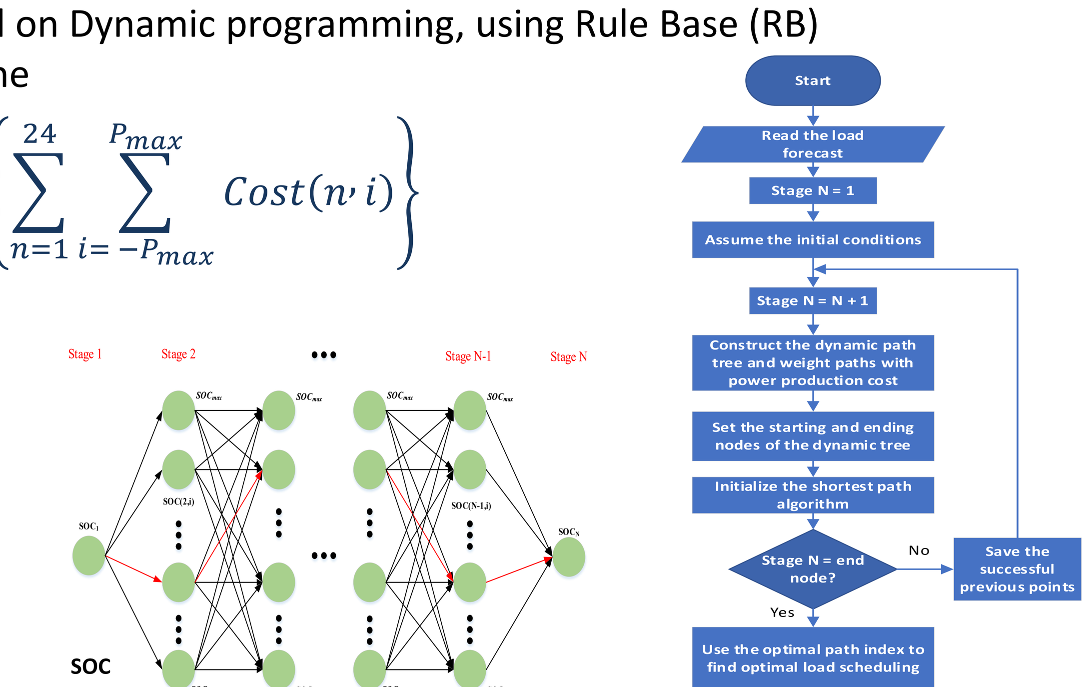
Battery Management System

✓ BESS management based on Dynamic programming, using Rule Base (RB) management as a baseline

$$Total\ cost = min\left\{\sum_{n=1}^{24}\sum_{i=-P_{max}}^{P_{max}} Cost(n, i)\right\}$$

Li iron battery degradation cost

DOD (%)	Li-ion battery degradation cost(cents/kWh)
20	0.84
40	2.30
60	5.06
80	7.92
100	10.13



DP Algorithm and flow chart

Battery Degradation Estimation

Limitations: significant burden for battery real-time degradation estimation system → timely battery health information monitoring is impractical & cost tracking for capacity fade during battery operation is impractical.

Cost Policy Proposed: Real-Time Method

Recurrent structure with memory $\mathbf{S}_t$

$$deg_{inc,t\to t+1} = f\left(\mathbb{O}(\tilde{\mathbf{S}}_{t+1})\right) - f(\mathbb{O}(\mathbf{S}_t))$$

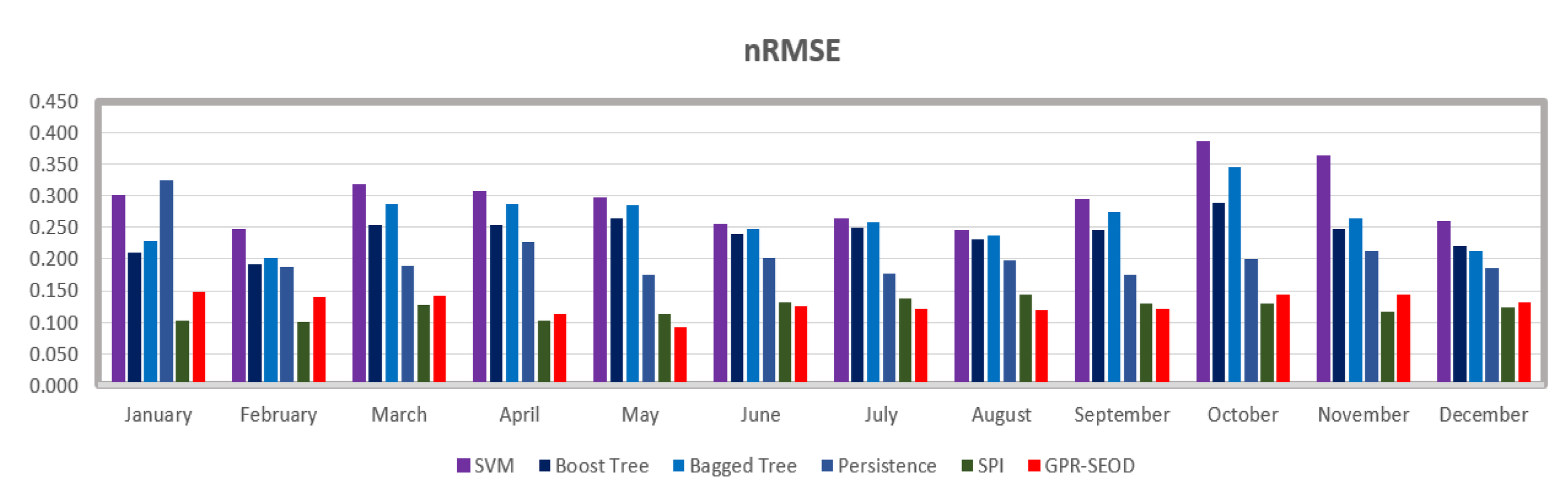
- $\mathbf{S}_t := [\mathbb{I}(SOC_0)SOC_0, \mathbb{I}(SOC_1)SOC_1, \cdots, \mathbb{I}(SOC_t)SOC_t]$
- $\mathbb{I}(SOC_0) \in \{0,1\}$ is the **indicator function**.
- $\tilde{\mathbf{S}}_{t+1} := [\mathbf{S}_t|SOC_{t+1}]$ is the appended memory

Memory management $\mathbb{I}(SOC_i)$

- In each recurrence, the **"Drop"** block evaluates the memory by setting indicator functions, drops noncritical samples.
- **"Drop"** evaluates 4 cases sequentially: Half-cycle Extension → Full-cycle → Unextendible half-cycle → Extendible half-cycle.
- $\mathbf{S}_{t+1}$ is the outcome of "Drop" and passed to the next recurrence (incremental degradation of next SOC sample).
- The size of actively-updated memory $\mathbf{S}_t$ is independent from the length of evaluation period. Computation complexity of $\mathbf{D}_{0\to t}$ with n samples: $\mathcal{O}(n)$

Best Practices

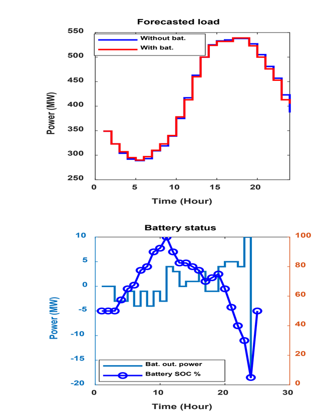
Result: PV Power Forecast



Average monthly prediction error of GPR vs. persistent, SPI, SVM, boosted, and bagged tree methods

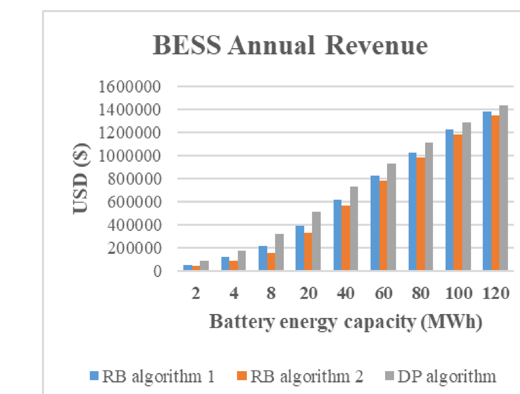
- ✓ 24h-ahead prediction with nRMSE performance index
- ✓ GPR improved accuracy by > 30% over persistent method.
- ✓ SPI approach outperforms other approaches at timescales < 30 mins, consistent with NREL report (E. Ibanez, et al).

Result: Battery Management System



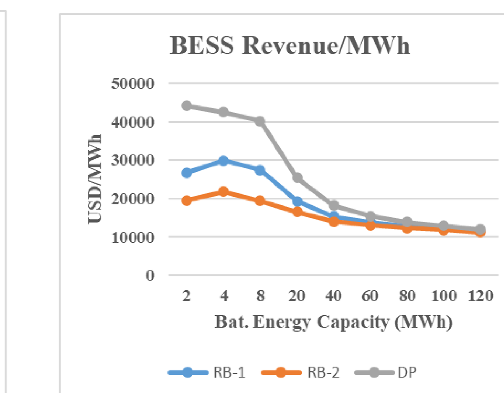
BMS by DP algorithm

- BMS is optimized using the DP algorithm, particularly with the battery capacities which can substitute the smallest IC unit for less than an hour.



Battery annual revenue

- RB and DP algorithms are simulated for each day of the annual load for various battery capacities (2MW-120MW).
- The DP algorithm generates more revenue than RB algorithms for all the battery capacities, which reduce the payback period of the BESS.



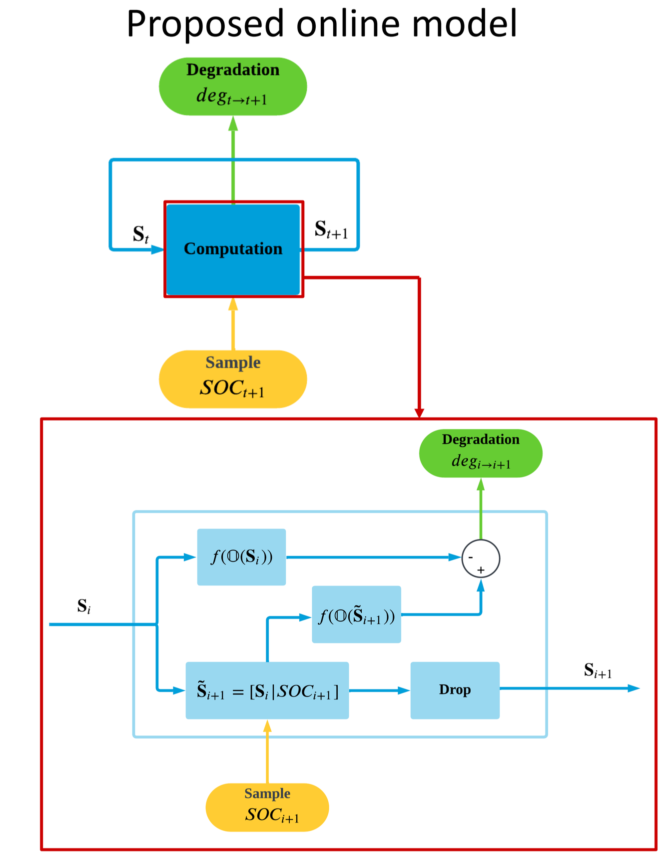
DP vs. RB

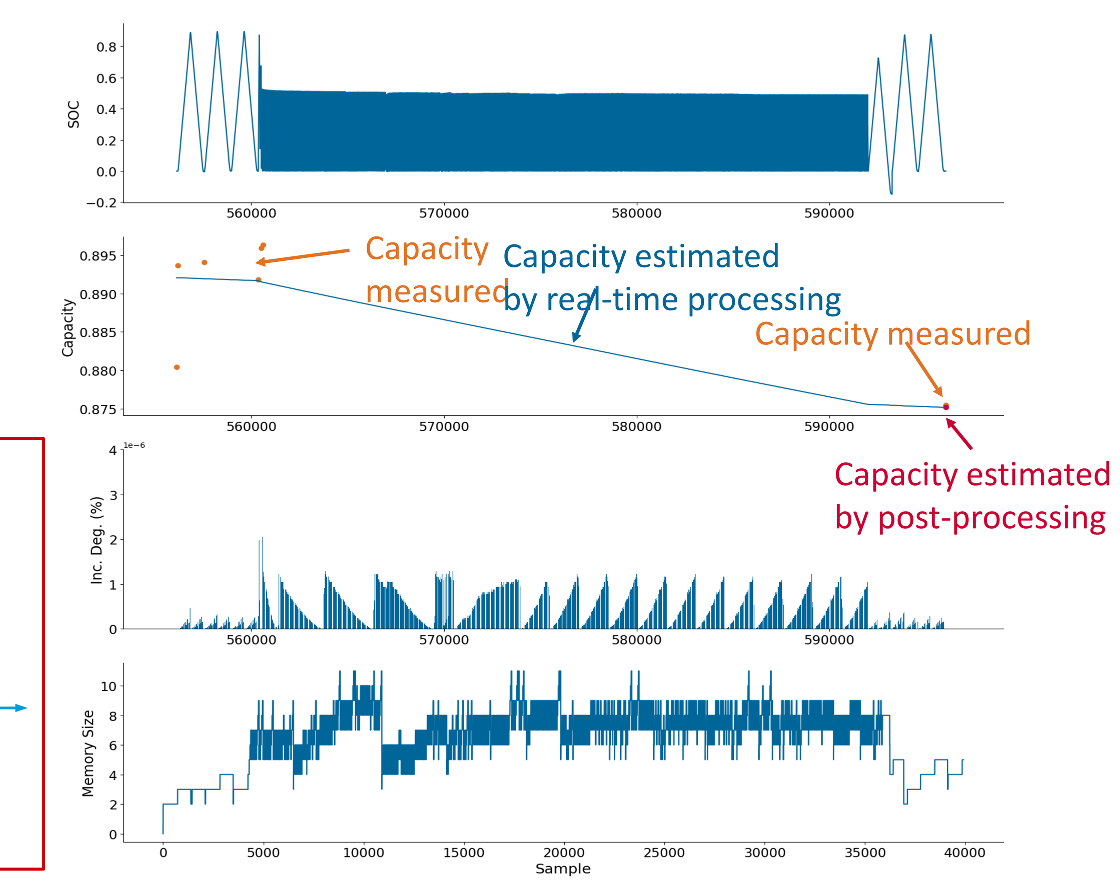
- The revenue/MWh is declining rapidly and reaches the saturation after 40 MWh.
- It shows that a BESS with a capacity smaller than the smallest IC backup unit has more revenue/MWh rate than larger batteries.

Result: Battery Degradation Estimation

Using the subset of 39907 consecutive samples of LFP battery testing from *Sandia National Laboratories.*

Proposed online model





Contribution: Critical SOC history is accurately stored, computation complexity is greatly reduced, enabling real-time degradation estimation for batteries.

Key Takeaways

- ✓ GRP algorithm compares favorably with naive machine learning algorithm, and persistent forecast. The forecast error improves on day-ahead forecast horizon, which is have more than 30% nRMSE errors improvement compared to persistence methods used by utilities.
- ✓ Improves forecast errors achieved better battery sizing, and cost saving on reserves.
- ✓ DP-based optimization of BESS to substitute low-efficiency IC units is more advantageous than traditional RB scheduling methods and yields economic benefits.
- ✓ The proposed method successfully predicts the actual measurement capacity fade with a low error. Additionally, the memory list size during operation demonstrates lower computational complexity compared to the conventional evaluation methods for real-time operation.

Publications

[1] S. Wang, M. Mahshid, B. Lehman, "A Real-Time Degradation Estimation Approach for Batteries in PV and Battery Hybrid Plant Operation," IEEE Energy Conversion Conference Expo, Nashville, Nov. 2023.

[2] U. Selvarasu, M. Mahshid, Y. Li, C. Crow, B. Lehman, "DP-Based Optimization of BESS to Substitute RICE Reserves for Improved Economic Benefits," IEEE Energy Conversion Conference Expo, Nashville, Nov. 2023.

Bayesian GAN-based False Data Injection Attack Detection in Active Distribution Grids with DERs

Jian Xie, Airin Rahman, and Dr. Wei Sun, ECE Dept., Univ. of Central Florida

Secure and Resilient Operations Using Open-Source Distributed Systems Platform (OpenDSP), DE-EE0009339

Background and Motivation

- This work is part of the project, Secure and Resilient Operations Using Open-Source Distributed Systems Platform (OpenDSP).
- The goal is to develop a **multi-layer multi-channel cyber-physical defense and survival mechanism** for operating distribution networks with high penetration of solar, IBR, and DER.
- **Challenges** include, 1) new cyber-physical vulnerabilities from grid-edge DERs; 2) renewable energy sources with uncertainty and variability; 3) imbalanced data between normal operations and compromised or attacked states; 4) traditional machine learning and data-driven approaches are difficult to train, leading to long computation time, and less suitable for large-scale power systems.

Key Contributions

This work proposes a Bayesian Generative Adversarial Network (BGAN)-based approach for cyber attack detection.

- A data-driven cyber-attack detection approach that, while building upon existing Bayesian GANs, introduces a customized model architecture and modified training procedure to enhance both speed and sensitivity in performance.
- A solution to the imbalanced data distribution problem, which is commonly encountered in practical applications.
- The method accounts for the uncertainty of renewable energy sources and generates scenarios that align with the historical data distribution, leading to more precise detection results.

New Security Challenges

- Uncertainty modeling with renewable: take PV as an example, as their distributed nature represents different modes of uncertainty.
- Targeted cyber attacks on renewable DERs: coordinated attacks, like **multiwave attacks** where the attackers sequentially target PVs by manipulating their respective inverter reference voltages. In Fig. 1, even if the attacker targets only a small subset of DERs, any altered measurements can disrupt system functionalities.
- Unbalanced data sampling: the historical training data of the secure class is much more than that of the attacked class.
- Unobservable cyber attacks: the attacker can manipulate the measurements to bypass the residual-based detection methods.

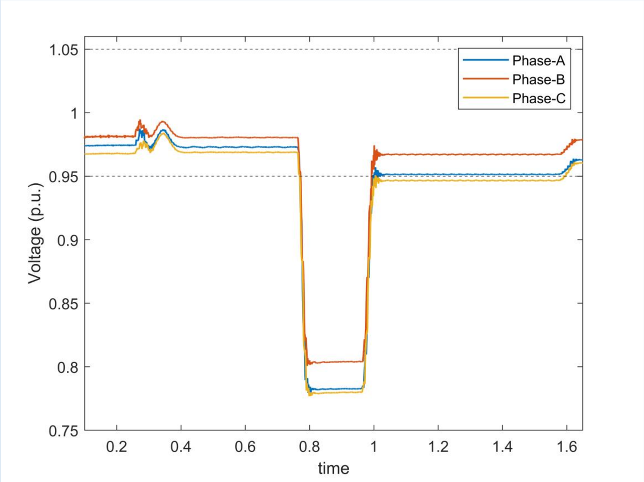


Fig. 1. Voltage disturbance due to multiwave cyber attack on DERs

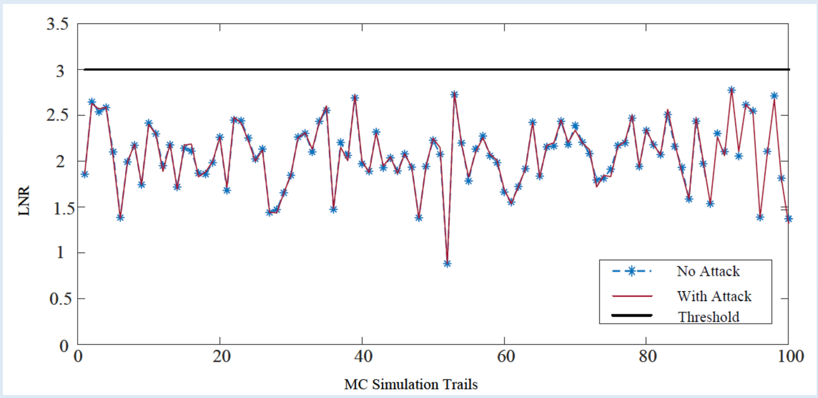


Fig. 2. Residual-based detection with unobservable attack

Bayesian GAN for FDIA

- Addressing the imbalanced data: the proposed BGAN-based approach combines Bayesian probability and GAN for fully probabilistic inference. Accurate estimation can be obtained by posteriors of θ^G and θ^D after training with different samples.

$$\boldsymbol{logp}\{\boldsymbol{\theta^D}|\boldsymbol{\theta^G}\} = \boldsymbol{E_X}[\boldsymbol{D}(\boldsymbol{X},\boldsymbol{\theta^D})] - \boldsymbol{E_Z}[\boldsymbol{D}(\boldsymbol{G}(\boldsymbol{Z},\boldsymbol{\theta^G}),\boldsymbol{\theta^D})] + \boldsymbol{logp}\ \{\boldsymbol{\theta^D}|\boldsymbol{\gamma^D}\}$$

- The Bayesian approach is employed to update the distributions through parameters with adversarial feedback. Inference is performed by iteratively sampling from the following conditional posteriors:

$$p(\boldsymbol{\theta^G}|\boldsymbol{\theta^D},z) \propto \prod_{i=1}^{n} D(G(z_i;\boldsymbol{\theta^G});\boldsymbol{\theta^D})p(\boldsymbol{\theta^G}|\boldsymbol{\gamma^G})$$

$$p(\boldsymbol{\theta^D}|\boldsymbol{\theta^G},\boldsymbol{z},\boldsymbol{X}) \propto \prod_{i=1}^{n} \boldsymbol{D}(\boldsymbol{x_i};\boldsymbol{\theta^D}) \prod_{i=1}^{n}\left(\boldsymbol{1}-\boldsymbol{D}(\boldsymbol{G}(\boldsymbol{z_i};\boldsymbol{\theta^G});\boldsymbol{\theta^D})\right)\boldsymbol{p}(\boldsymbol{\theta^D}|\boldsymbol{\gamma^D})$$

- Accounting for PV uncertainty: extend the Bayesian approach to capture the uncertainty of PV output by teaching the model to generate data that not only resembles the tube PV output but also accounts for its inherent uncertainty.

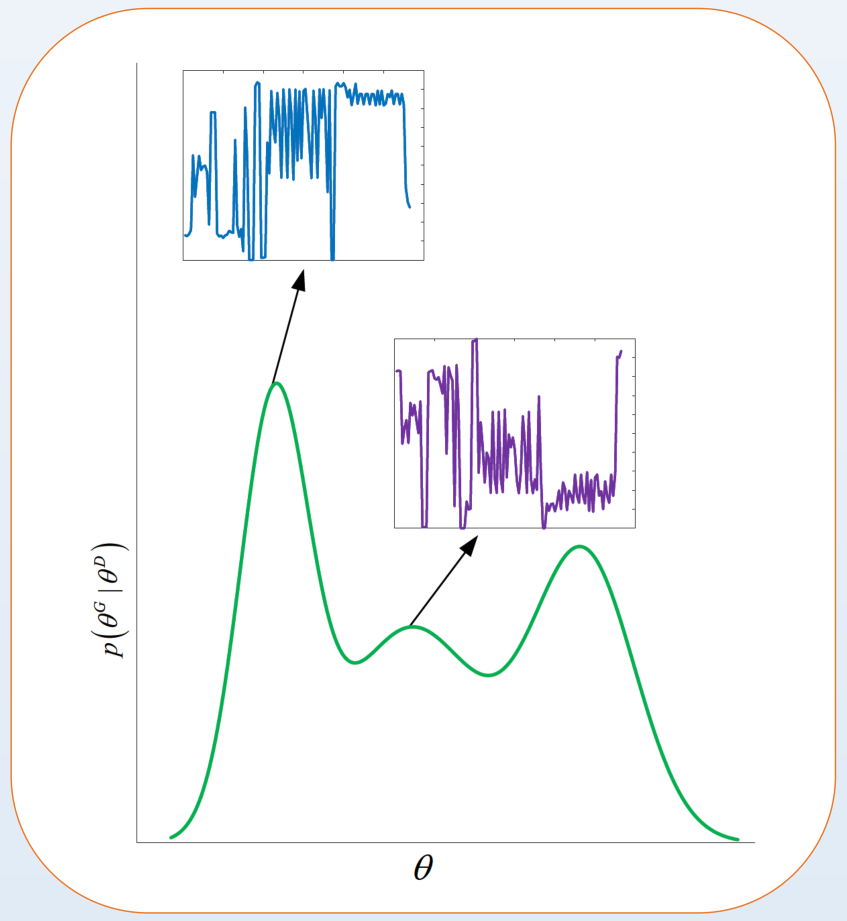


Fig. 3. An example that different generators capture unique behaviors

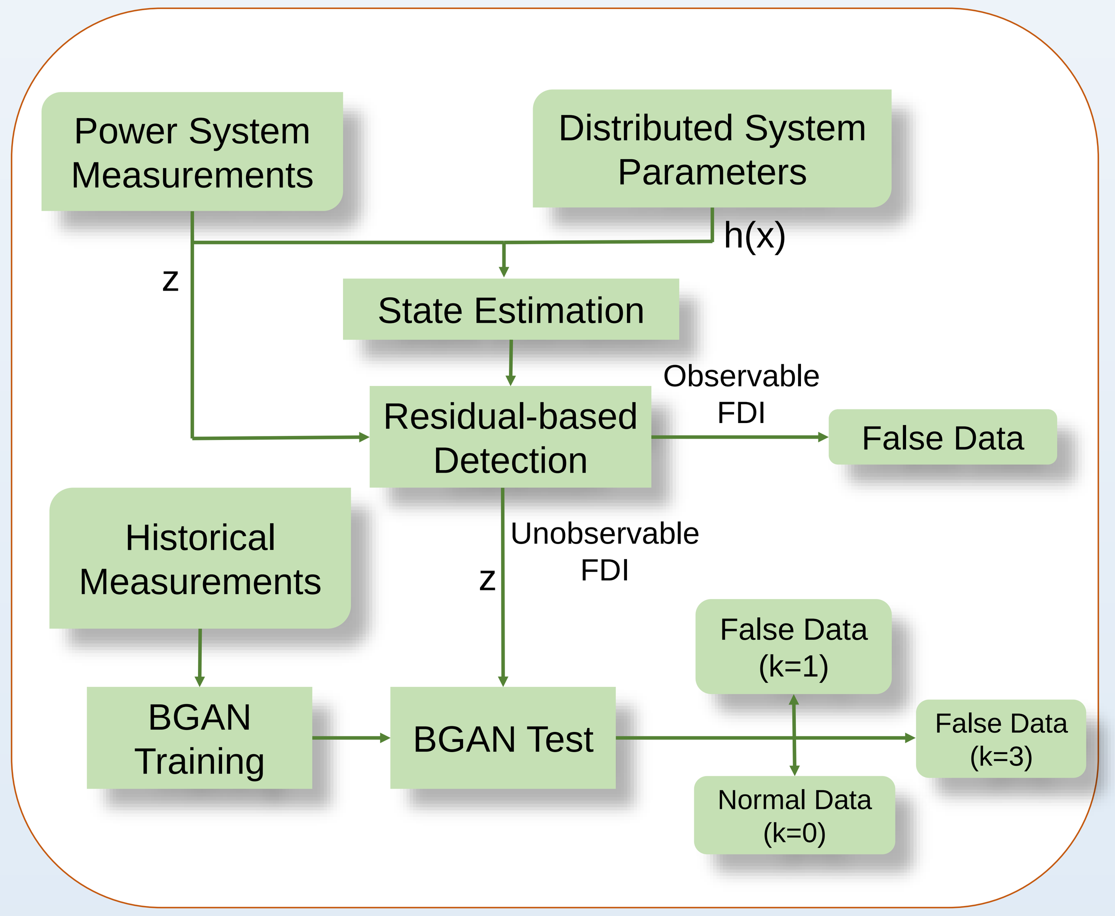


Fig. 4. The flowchart of the proposed scheme

Attributes of BGAN-based FDIA

The BGAN-based attack detection method offers the following features:

- Robust performance against uncertainties related to renewable energy sources.
- High performance and accurate detection of erroneous data, even with imbalanced training data and measurement noise.
- Maintain high accuracy under varying and extreme load conditions.
- Efficient for practical applications in real-world active distribution systems.

Case Studies

Table I. Detection performance on IEEE 13-node and 123-node systems

	Modified IEEE 13-node system		Modified IEEE 123-node system	
Data	Accuracy%	Detection time/ms	Accuracy%	Detection time/ms
Train-Labeled	100	-	100	-
Train-Unlabeled	97.83	-	98.58	-
Test	97.52	8.97	98.13	16.77

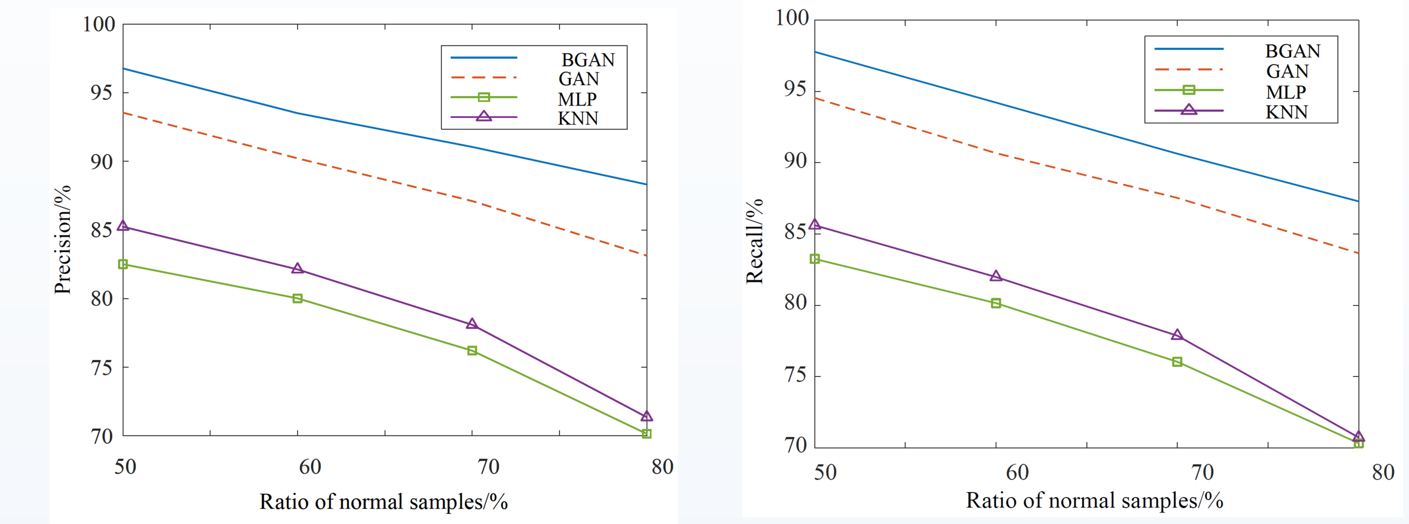


Fig. 5. Comparison of precision and recall scores from different methods under different imbalanced scenarios.

Table II. Performance comparison

Algorithm	Accuracy IEEE 13-node system	Accuracy IEEE 123-node system
BGAN	97.52	98.13
GAN	94.27	95.01
MLP	83.21	84.06
KNN	85.35	86.22

Table III. Detection Performance with different load

Load Level	Precision	Recall
10%	0.9026	0.9031
90%	0.9114	0.9095
100%	0.9152	0.9131
110%	0.9167	0.9123

Table IV. Detection Performance with different number of generator

Num of G	Precision	Recall
7	0.8943	0.8915
8	0.9024	0.9061
9	0.9107	0.9084
10	0.9152	0.9131
11	0.9115	0.9120

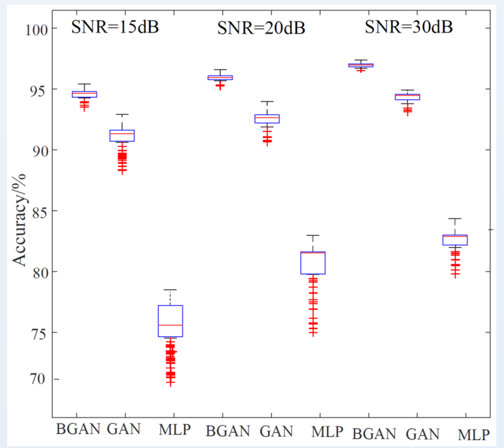


Fig. 6. Accuracy comparison in different SNR level conditions.

Conclusions

This work proposes a Bayesian GAN-based approach for detecting FDIA in active distribution systems. The suggested method utilizes a novel Bayesian GAN to achieve accurate FDIA detection by learning data features with a small amount of imbalanced training data. BGAN is completely data-driven and does not require any system model or typology parameters. Simulation results and comparisons demonstrate that the method exhibits higher accuracy and robustness.

Acknowledgement

This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technology Office (SETO) Award Number DE-EE0009339.

Coordinated High-Speed Voltage Control in Real-Time Unobservable Active Distribution Systems

Shiva Moshtagh, Dhaval Dalal, and Anamitra Pal
School of Electrical, Computer, and Energy Engineering (ECEE) at Arizona State University (ASU)

I. INTRODUCTION

- State estimation (SE) in distribution systems can be done at high-speeds if only micro-phasor measurement unit (μPMU) data is used
- Leveraging the high-speed of a μPMU-only state estimator, a coordinated inverter control algorithm can be implemented to achieve high photovoltaic (PV) hosting capacity (HC) in active distribution systems
- **Challenges:**
  - Linear state estimation (LSE) needs the distribution system to be fully observed by μPMUs
  - Weighted least squares, which is most often used for μPMU-based LSE, is not robust against non-Gaussian measurement noise
  - The distribution system often undergoes topology changes, which, if not accounted for, can degrade SE performance
  - Traditional inverter control solutions do not: (a) use full state information, (b) have a sufficiently accurate sensitivity matrix (SM), and (c) take cross-phase sensitivity into account
- **Solution:**
  - Deep neural network (DNN)-based topology identification (TI) is done first to track topology in real-time from sparsely placed μPMUs
  - A DNN-based state estimator is developed next to estimate states in a fast, time-synchronized manner
  - Transfer learning is employed to account for the effects of topology changes on DNN-based SE
  - A control algorithm is proposed that iteratively refines the voltage-reactive power SM to mitigate voltage violations at high PV penetrations
  - The proposed control recognizes the diminishing effects of cross-phase sensitivity and incorporates it in the optimization
  - The proposed control avoids active power curtailment (APC) as well as changes in capacitor bank (CB) and step voltage regulator (SVR) settings

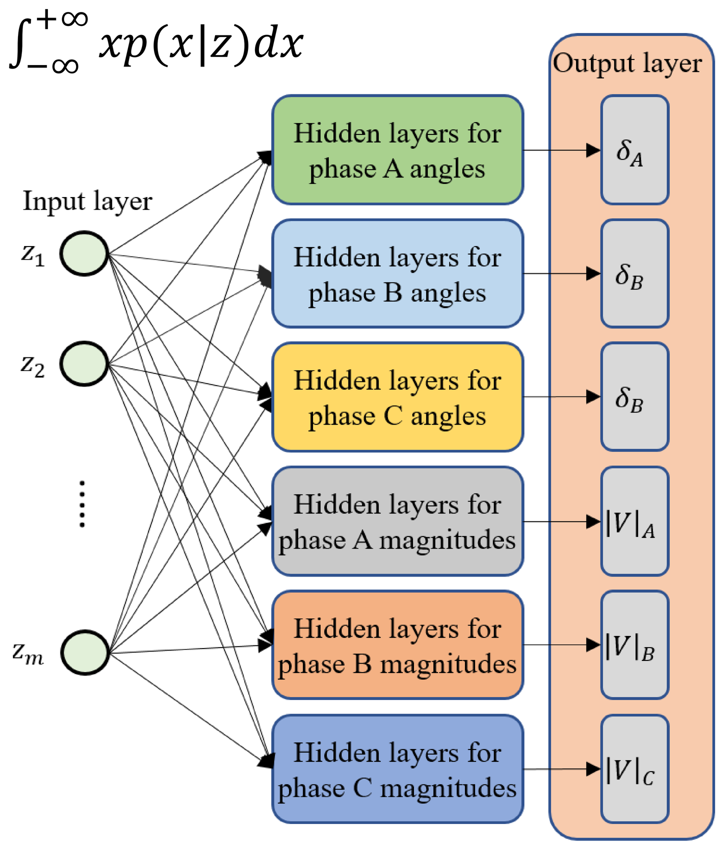
II. DEEP NEURAL NETWORK-BASED SE (DNN-SE)

- Consider measurements, z, and states, x, as random variables
- Create a minimum mean-squared error (MMSE) estimator to minimize the estimation error:

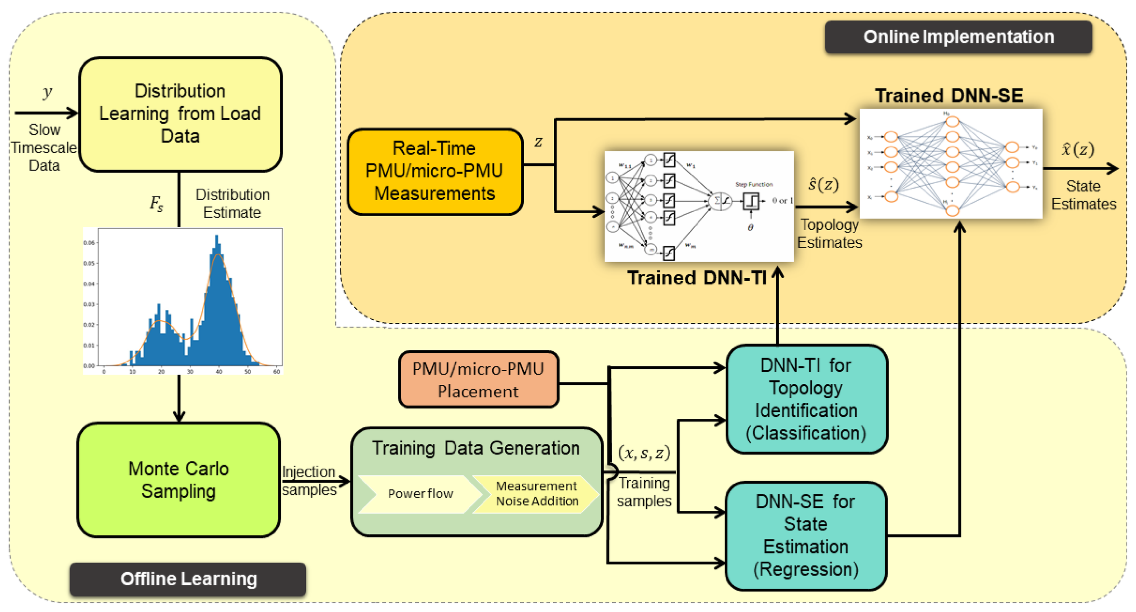
$$\min_{\hat{x}(\cdot)} \mathbb{E}(\| x - \hat{x}(z) \|^2) \Longrightarrow \hat{x}^*(z) = \mathbb{E}(x|z)$$

$$\text{where, } \mathbb{E}(x|z) = \int_{-\infty}^{+\infty} xp(x|z)dx$$

- **Challenges:**
  - Requires the knowledge of $p(x|z)$
  - Closed form solution for $\int_{-\infty}^{+\infty} xp(x|z)dx$ can be difficult in presence of unobservability
- **Solution**:
  - A DNN is used to find the mapping function that relates x and z, and thereby, approximate the MMSE state estimator

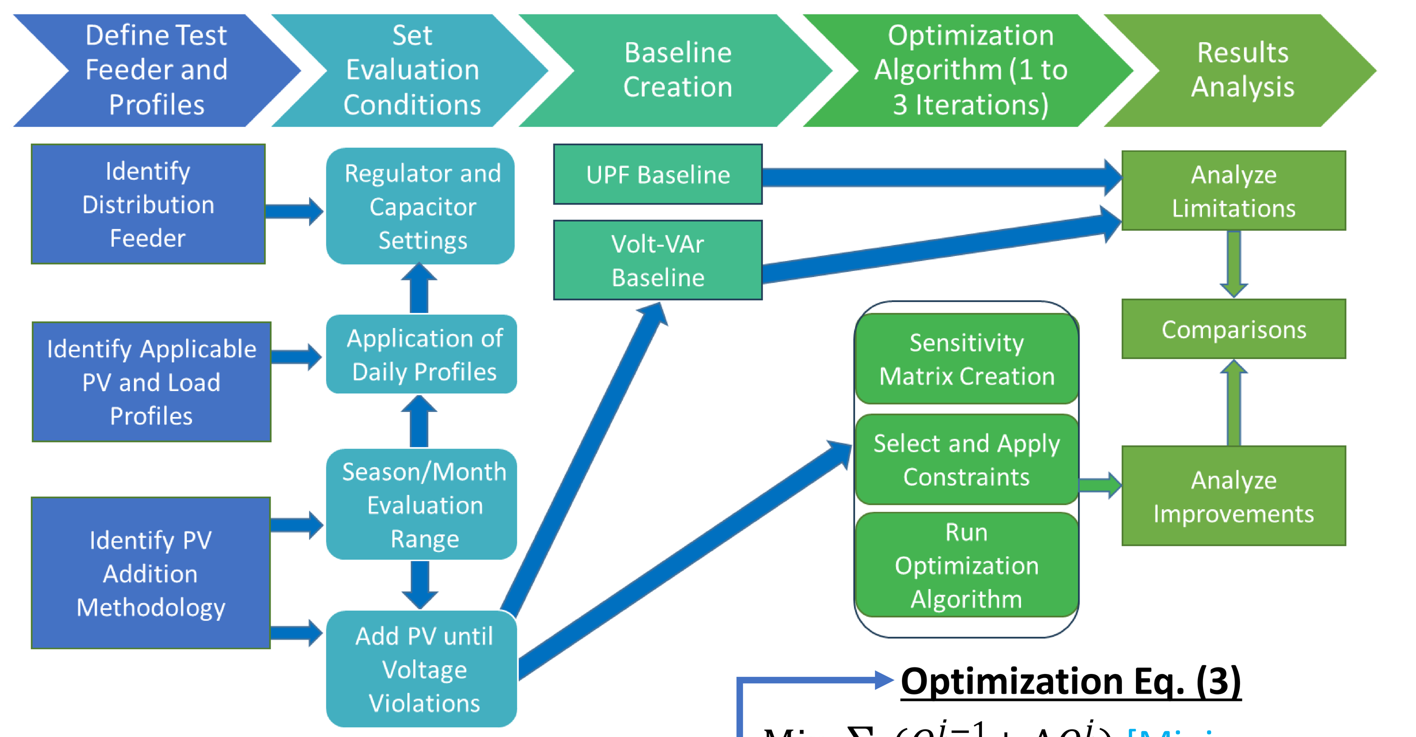


- **Challenges:**
  - The DNN is trained for the nominal topology
  - What if the topology changes?
- **Solution**:
  - Implement framework for joint TI and SE
  - Perform transfer learning to update DNN parameters in real-time when topology changes



- Sequential feature selection is used to find the location of μPMUs for TI
- Spearman's correlation coefficient followed by hierarchical clustering is used to find the location of μPMUs for SE

III. COORDINATED INVERTER CONTROL



Algorithm 1 Iterative Optimization Algorithm

1: Load network information into a distribution system solver
2: Set SVR, OLTC, and/or capacitor bank settings
3: Add PV systems
4: Apply appropriate PV and load profiles
5: Select evaluation condition (time instance)
6: Determine base Q values:

$$\mathbf{Q^0} = [Q_0, Q_1, \ldots Q_j, \ldots, Q_M]$$

7: Run power flow with base conditions and calculate base voltage magnitudes: $\mathbf{V^0} = [V_0, V_1, \ldots V_i, \ldots, V_N]$
8: $\mathbf{V} \leftarrow \mathbf{V^0}$
9: $\mathbf{Q} \leftarrow \mathbf{Q^0}$
10: **while** $(\max(\mathbf{V}) > V_{thmax})$ or $(\min(\mathbf{V}) < V_{thmin})$ **do**
11: Calculate Sensitivity Matrix, SM, with $\mathbf{Q}$ values
12: Solve (3) to determine $\mathbf{Q^n} = \mathbf{Q} + \Delta\mathbf{Q}$
13: Apply $\mathbf{Q^n}$ values to PV and run power flow to get $\mathbf{V^n}$
14: $\mathbf{V} \leftarrow \mathbf{V^n}$
15: $\mathbf{Q} \leftarrow \mathbf{Q^n}$
16: **end while**

Optimization Eq. (3)

Min $\sum_j (Q_j^{i-1} + \Delta Q_j^i)$ [Minimum control intervention]

s.t. ($\forall k$ nodes, $\forall j$ PV sites, & i^{th} iter.):

- Max Q constraint: [No APC]

$$-Q_{jmax} \leq (Q_j^{i-1} + \Delta Q_j^i) \leq Q_{jmax}$$

- Regulation Constraint:

$$V_{pumin} \leq \left(V_k^{i-1} + \Delta V_k^i\right) \leq V_{pumax}$$

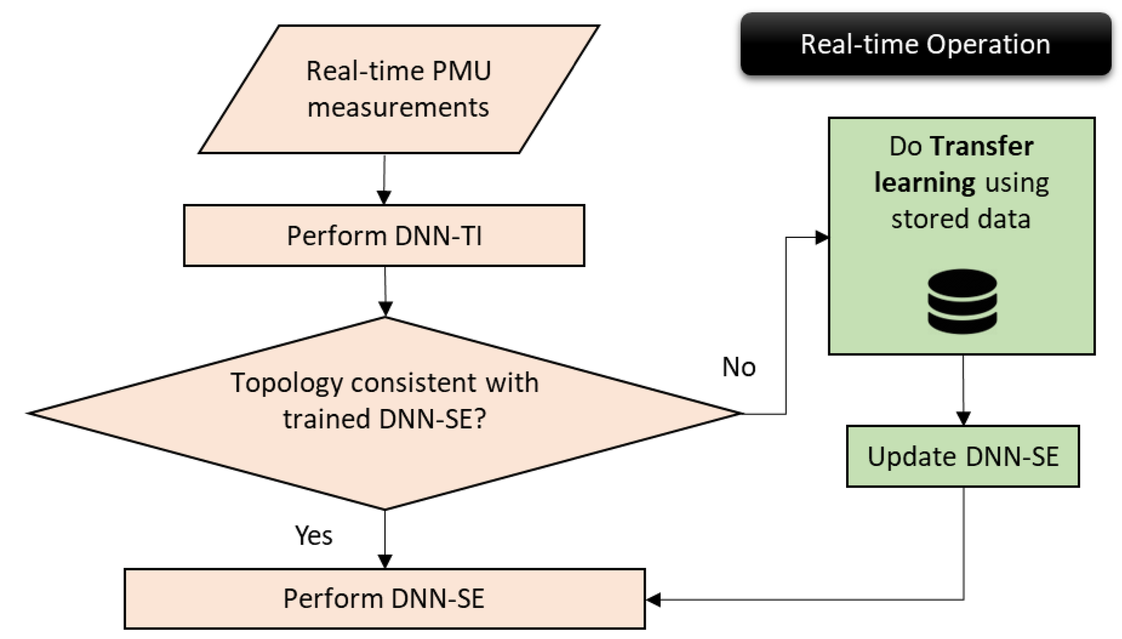
- Sensitivity Constraint:

$$\Delta V_k^i = \sum_j sm_{k,j}^i * \Delta Q_j$$

$$SM = \begin{bmatrix} sm_{1,1} & sm_{1,2} & \cdots & sm_{1,M} \\ sm_{2,1} & sm_{2,2} & \cdots & sm_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ sm_{N,1} & sm_{N,2} & \cdots & sm_{N,M} \end{bmatrix}$$

$$Q_{jmax} = \sqrt{{S_j}^2 - P_{jmax}^2} \quad sm_{i,j} = \frac{v_{i,j}^Q - v_i^0}{\Delta Q_j}$$

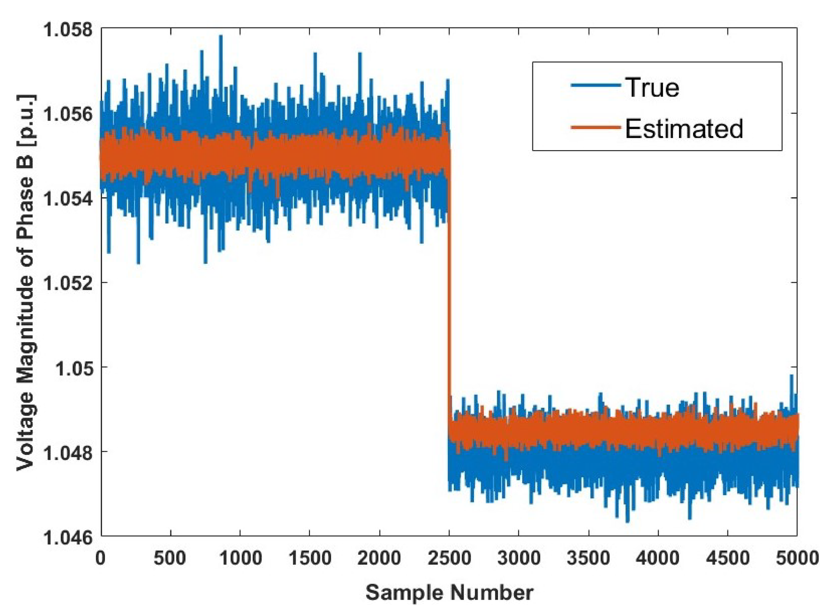
IV. DNN-SE RESULTS WITH TRANSFER LEARNING



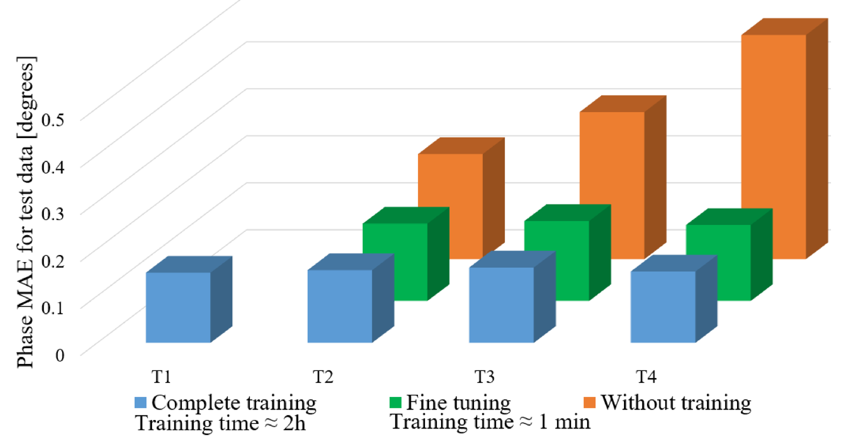
Compare estimation error between LSE and DNN-SE for 240-node system

Method	Noise model	Phase Angle MAE (degrees)	Magnitude MAPE (%)	#μPMU
LSE	1% Gaussian TVE	0.14	0.25	113
DNN-SE	1% Gaussian TVE	0.02	0.04	6
	1% non-Gaussian	0.02	0.05	6

- Actual and estimated voltage before and after application of Volt-VAr (VV) control
- Estimation performed at milli-second timescale using one μPMU at feeder-head
- Control applied at sample number 2500



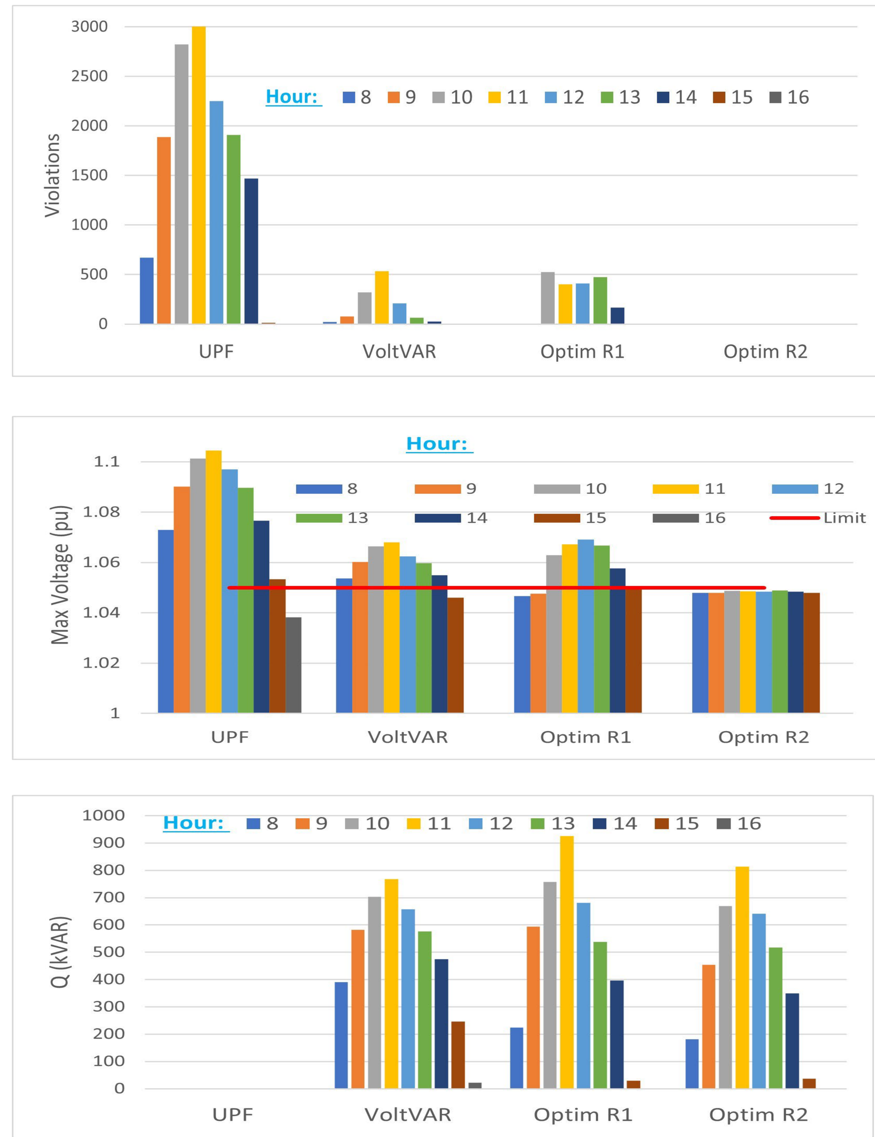
- By using Transfer Learning, DNN-SE can be done quickly and accurately despite changes in network topology



V. HOSTING CAPACITY (HC) IMPROVEMENTS

HC=>	UPF	VV	Optimized	Improvement (Δ to VV)
kW	2113.6	2253.6	6413.6	4160
% peak load	19.3%	20.6%	58.6%	38.0%
% 11AM load	26.5%	28.3%	80.5%	52.2%

- VV control reduces the violations to ≈500
- VV also reduces max voltage from 1.1045 p.u. to 1.0681 p.u.
- Optim R2 removes all violations (from >3000 at 11 AM under UPF)
- Optim R2 restricts max voltage to 1.0485 p.u.
- Total Q values are similar for VV and Optim R2, but much better results with the latter
- Q values reduce from Optim R1 to Optim R2 (while improving results)!!!!



VI. CONCLUSIONS AND FUTURE WORK

- The proposed Bayesian approach for high-speed time-synchronized SE in real-time unobservable distribution systems via DNNs facilitates creation of a coordinated inverter control strategy that achieves high PV HC in comparison to the state-of-the-art

Key Takeaways

- The proposed DNN-based TI and DSSE framework:
  1. Does not require complete observability by μPMUs
  2. Does not rely on slow timescale data during online operation
  3. Outperforms LSE with a significantly smaller number of PMUs
  4. Can quickly detect network topologies in reconfigurable systems
  5. Ensures reliable SE for different topologies
  6. Is robust against non-Gaussian measurement noise, non-parametric load variations, and renewable energy fluctuations
- The proposed coordinated control algorithm demonstrated tripling of HC compared to unity power factor (UPF) and VV control. These results were achieved with:
  1. No APC (no revenue loss)
  2. No participation of SVRs and capacitor banks (low maintenance costs)
  3. Minimum reactive power intervention (lower control burden and lower stress on inverters)
  4. Worst-case conditions (equivalent to cold-start)
  5. Consideration of wide range of use cases
- These results are possible due to the following reasons:
  1. Availability of high-resolution, synchronized, accurate system-wide voltage data
  2. Iterative refinement of sensitivity matrix
  3. Recognition of the cross-phase sensitivity effects and their incorporation in the optimization algorithm
- **Ongoing Work:**
  - We are currently developing strategies to provide robustness and trustworthiness guarantees to our DNN-SE performance
  - We are trying to reduce the runtime of the optimization algorithm by using a subset of the most critical network voltages
  - We are exploring use of a trained DNN to by-pass the optimization during real-time operation

Papers

Day-Ahead Probabilistic Forecasting of Net-Load and Demand Response Potentials with High Penetration of Behind-the-Meter Solar-plus-Storage (Award # DE-EE0009357)

Presenters: Saumil Shah, John Yu, PI: Dr. Wenyuan Tang

Introduction

The rapid integration of behind-the-meter (BTM) solar-plus-storage in the grid has made point load forecasting less accurate.

To overcome the variability of renewable resources, it is advantageous to leverage the recent advances in predictive data analytics and big data. Moreover, for improved decision-making and risk assessment, it is beneficial to use probabilistic forecasting instead of point forecasting.

Load Forecasting Methods

- White Box (Statistical First Principle)
- Data-Driven (ML)
- Hybrid (Physics-Informed ML)

Two Machine Learning Models

Thrust I : Net-Load Probabilistic Forecasting using Fuzzy Decision Tree and NGBoost

- Suitable for short-term forecasts where only small amounts of training data is available.
- Predicts the probabilities of all net-load outcomes for each forecasted time point.
- Fast training and real-time prediction.

Thrust II : Net-Load Probabilistic Forecasting using Transformer-Based Architecture

- Suitable for long-term forecasts with large complex datasets.
- Lower Mean Absolute Percentage Error (MAPE) than the model in Thrust I.
- Multi-objective function coherently optimizes all the quantile predictions together.

Challenges and Best practices

Challenges

- Inferring and aligning daylight savings from missing or duplicated hours in March and November.
- Identification and repair of incorrect, missing, or out-of-range net-load and temperature samples.
- Interpretation of ML-learned internal parameters.
- Obtaining theoretical lower bound for accuracy.

Best Practices

- Verify the time zone alignment and daylight savings when combining data sources.
- Test the ML model extensively using a variety of data sources to obtain trustworthy performance metrics.

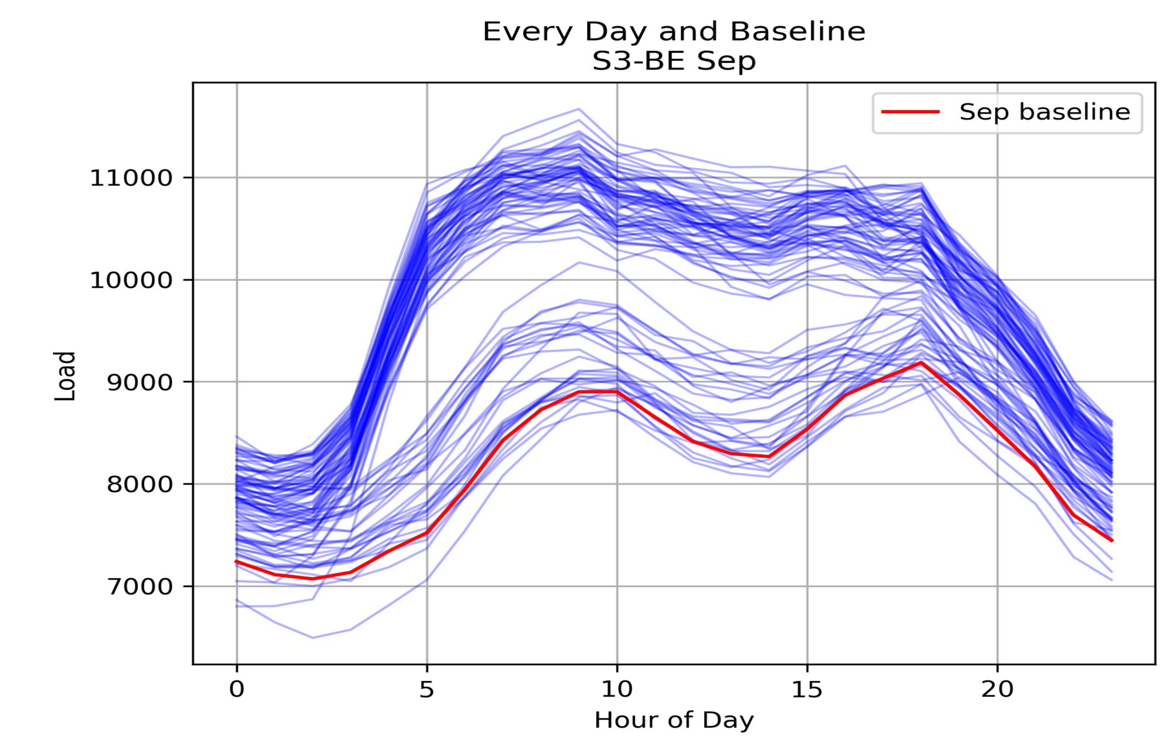
Key takeaways and Future work

Key Takeaways

- Machine-learning is suitable for prediction tasks where little to no information is available about the data-generating process.

Future Work

- Combining load forecasts with optimization algorithms to perform energy resource planning.
- Forecasting the real-time demand response potential to assist in the decision-making in demand response events.

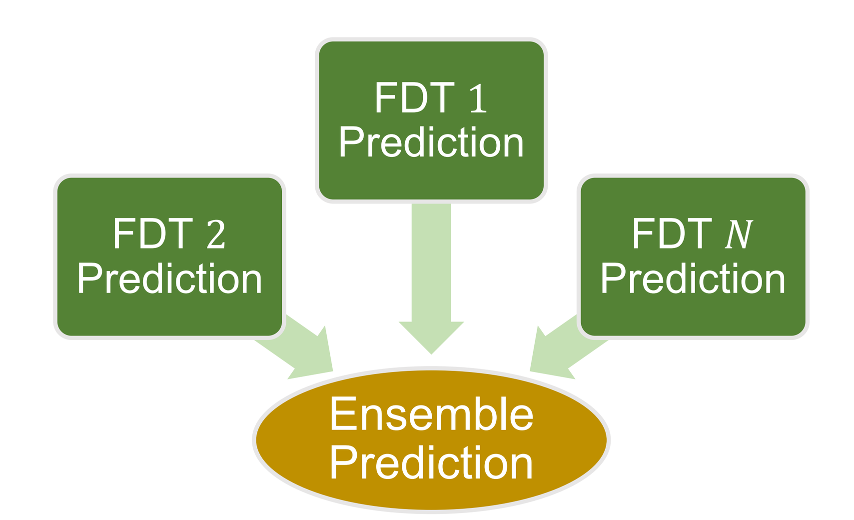


- Investigating combination of Channeled Aligned Dual Transformer (CARD) with the Thrust II model.

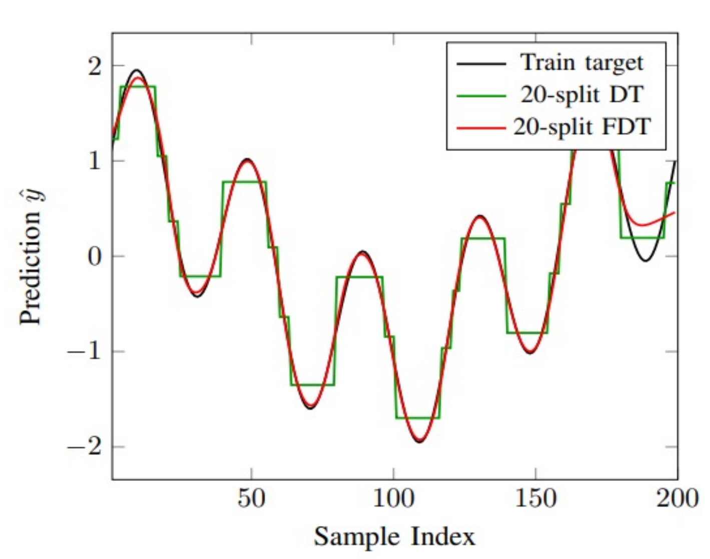
Methodology

Thrust I : Net-Load Probabilistic Forecasting using Fuzzy Decision Tree and NGBoost

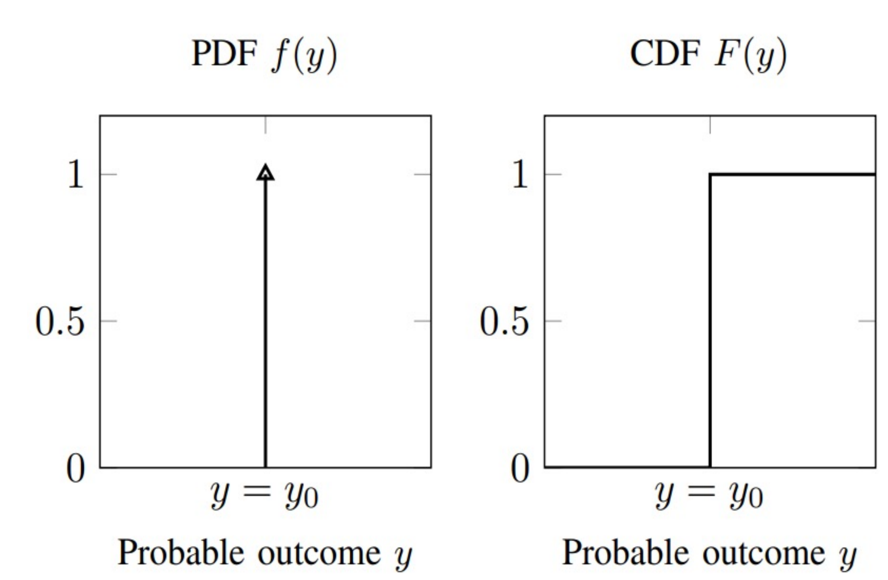
Boosting - Intelligence of the Crowd



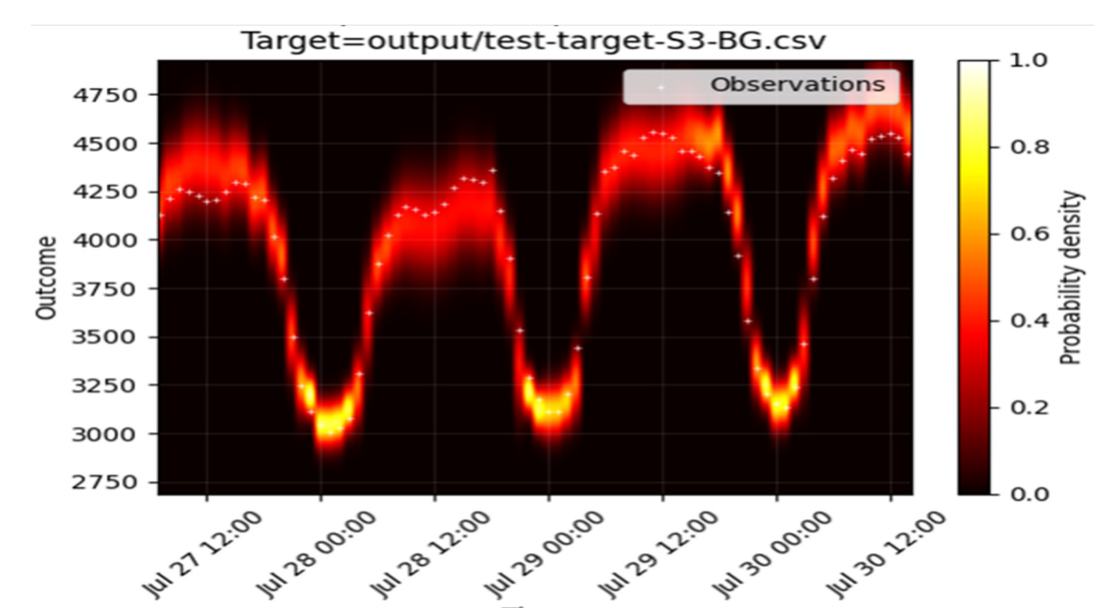
FDT – Continuous Output



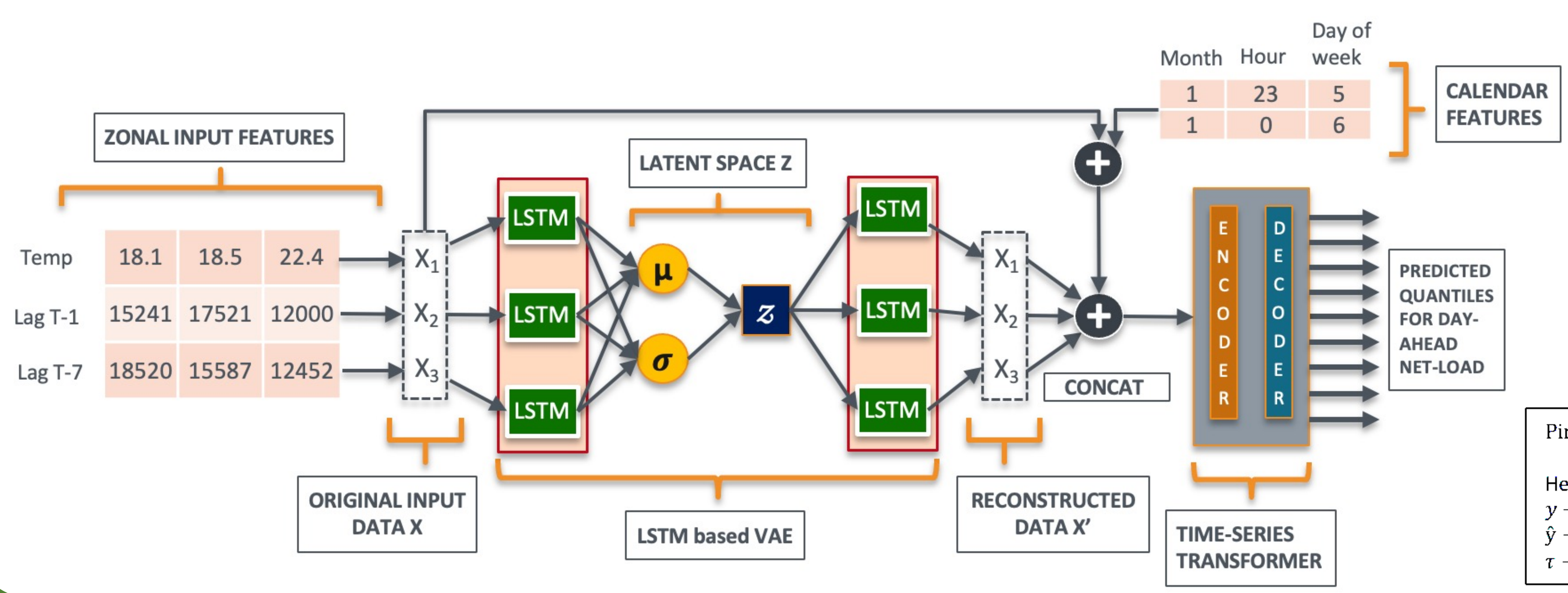
The Ideal Probabilistic Forecast

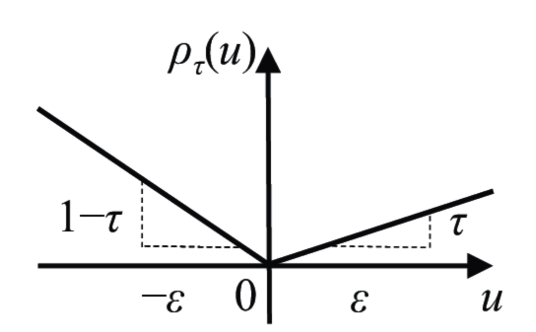


Test Set – Prediction vs Actual Observations



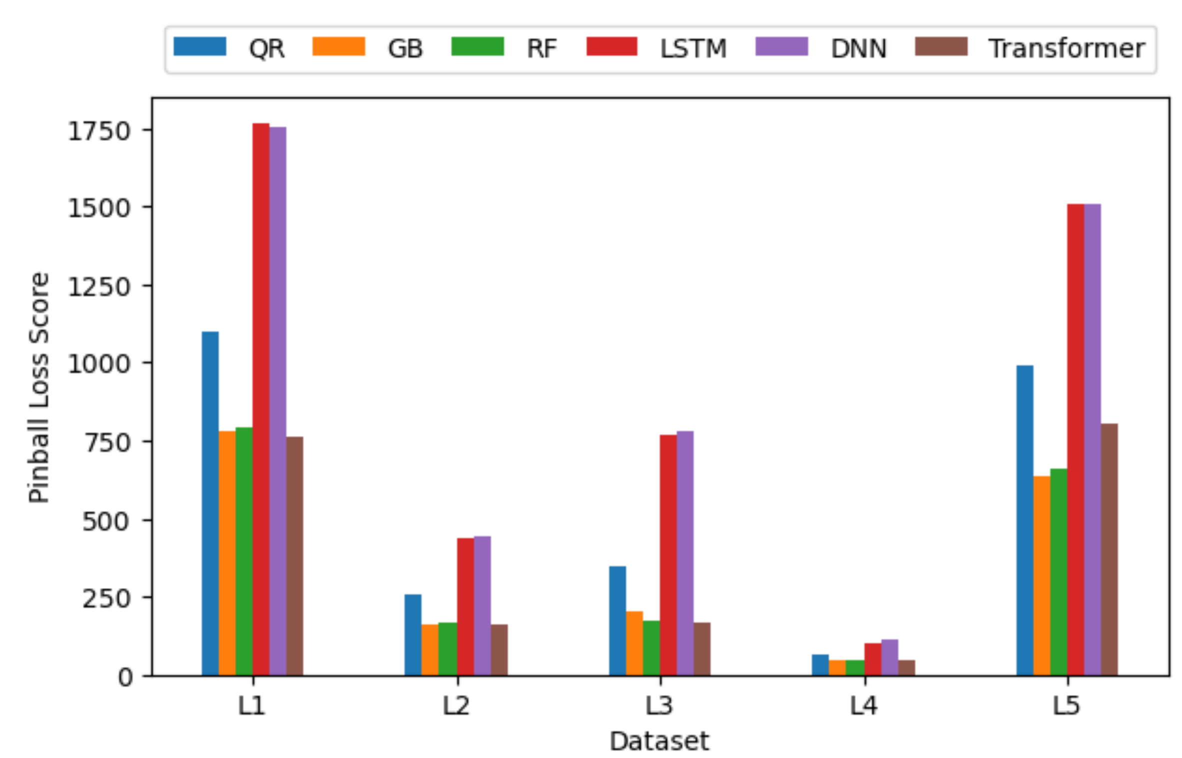
Thrust II : Net-Load Probabilistic Forecasting using Transformer-Based Architecture





Pinball loss function $L_\tau = (y - \hat{y}) \cdot \tau$ if $y \geq \hat{y}$
$= (\hat{y} - y) \cdot (1 - \tau)$ if $\hat{y} > y$

Here,
$y \rightarrow$ True data
$\hat{y} \rightarrow$ Predicted data
$\tau \rightarrow$ Quantile value



Partners

Dominion Energy
SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy
PURDUE UNIVERSITY
NC Electric Cooperatives
Your Touchstone Energy Cooperatives

Deep Learning-Computer Vision Framework for Monitoring Avian Interactions with PV Solar Facilities (#36473)

Adam Szymanski[1], Paul Tarpey[1], Xijun Wang[2], Yuki Hamada[1], Andrew Ayers[1], Nicola Ferrier[1], Leroy Walston[1], Heidi Hartmann[1]

[1] Argonne National Laboratory, [2] Northwestern University

SOLAR ENERGY TECHNOLOGIES OFFICE
U.S. Department Of Energy

Argonne NATIONAL LABORATORY

INTRODUCTION

How do birds respond to photovoltaic (PV) solar facilities? How many birds collide with solar panels? Our understanding remains incomplete without observations of bird activities, including collisions, and their outcomes. Since March 2020, Argonne researchers have been developing an edge-computing, machine-vision system to collect data on such opportunistic events at PV solar facilities. The technology aims to continually monitor bird activities around PV facilities, more specifically fly-over, fly-through, perching, landing, and collisions, during daytime to answer critical questions on avian-solar interactions.

CAMERA SYSTEM

The Sighthound DNNCam system (Fig.1), a true-color or visible-spectrum camera, is built with an NVIDIA Xavier edge computing processor capable of running computer vision (CV) and artificial intelligence/machine learning (AI/ML) algorithms on the camera itself. The system uses a docker based framework for running custom applications and allows for direct graphical processing unit (GPU) access. We utilize this framework to build our software and optimize our code for GPU execution.

Fig.1 DNNCam (https://www. sighthound.com/products/hardware).

CHALLENGES & BEST PRACTICES

CHALLENGE	SOLUTION / BEST PRACTICE
Creating ML models that would execute fast enough to be near-real time on the DNNCam edge computing processor. We needed to execute the combination of moving object detection and tracking; object classification; and collision detection algorithms fast enough for near real time notifications.	(1) Create models complex enough to achieve high accuracy but not so complex as to take a long time to compute and (2) Utilize CV and ML libraries optimized for fast execution on GPUs. We used the TensorRT framework to convert our ML models to execute quickly on the camera's NVIDIA Xavier processor, and we used several GPU-enabled OpenCV libraries. This required re-writes of several pieces of code and model optimization and porting.
Limited computational power of current camera technology for executing all models simultaneously.	Prioritize daytime computation for detecting and tracking moving objects, classifying objects, and detecting collisions. Activity classification is performed at nighttime (Fig.2).
Not being able to collect video of bird collisions with PV because such events are very rare.	Simulate bird collisions using decoys. We threw the decoys from the rooftop of a 2-story building to mimic realistic collisions. Trajectory, speed, and shape of descending birds could not be exactly replicated.

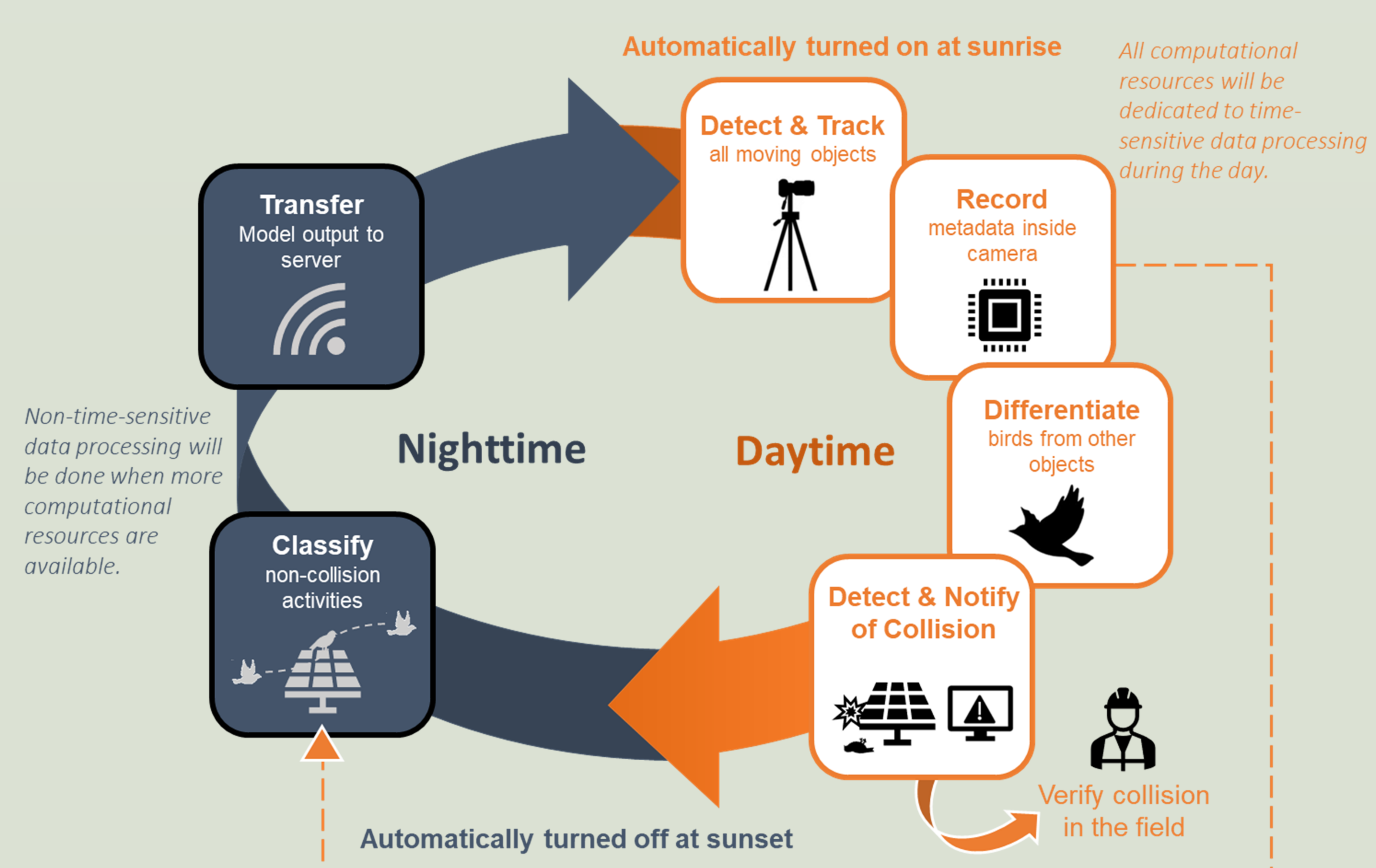
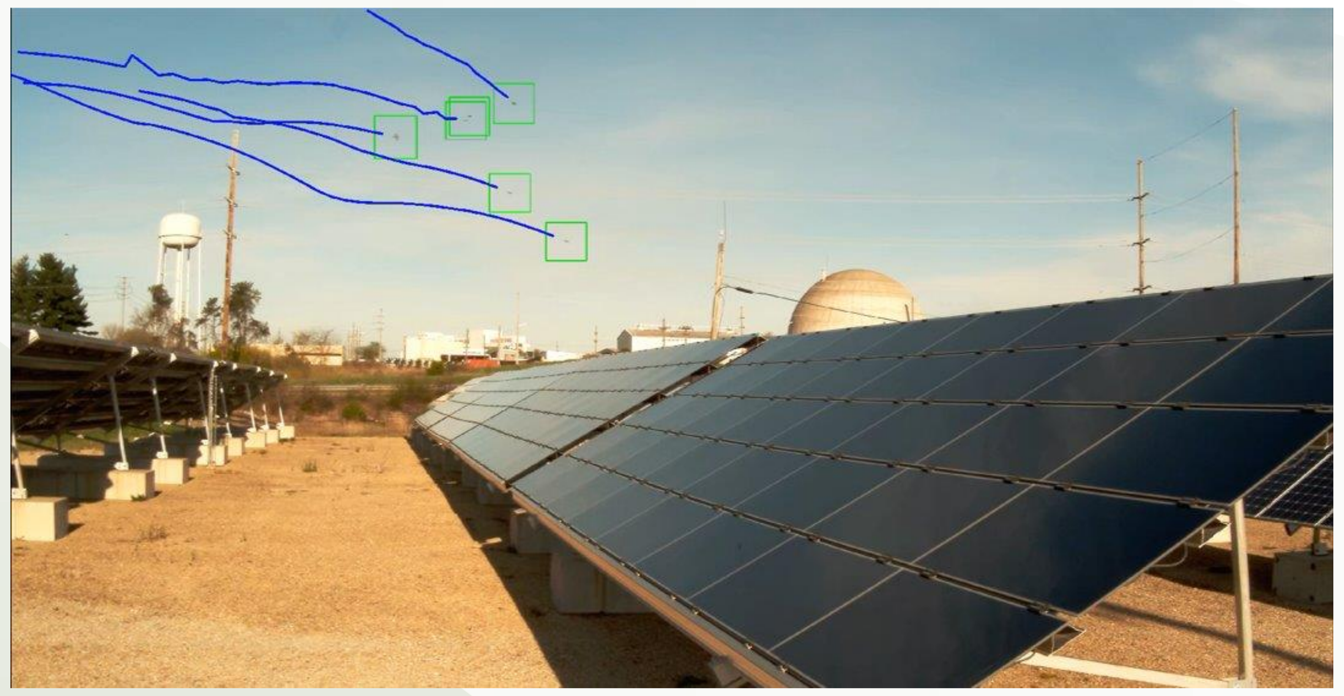


Fig.2 Timeline of video monitoring, model execution, and output transfer.

HOW IS AI USED IN THIS PROJECT?

To maximize the distance from which the camera was able to detect birds, we needed an AI/ML algorithm that could classify relatively small and indistinct objects. Therefore, we utilize an architecture capable of analyzing an entire sequence of images of the same object through time to perform the classification. We first employ a moving object detection algorithm in conjunction with a tracking algorithm to obtain a sequence of images for a **"track" (Fig.3)**. Then we use this "track" data as input into a series of different AI/ML models to complete the moving object identification and bird activity classification software.



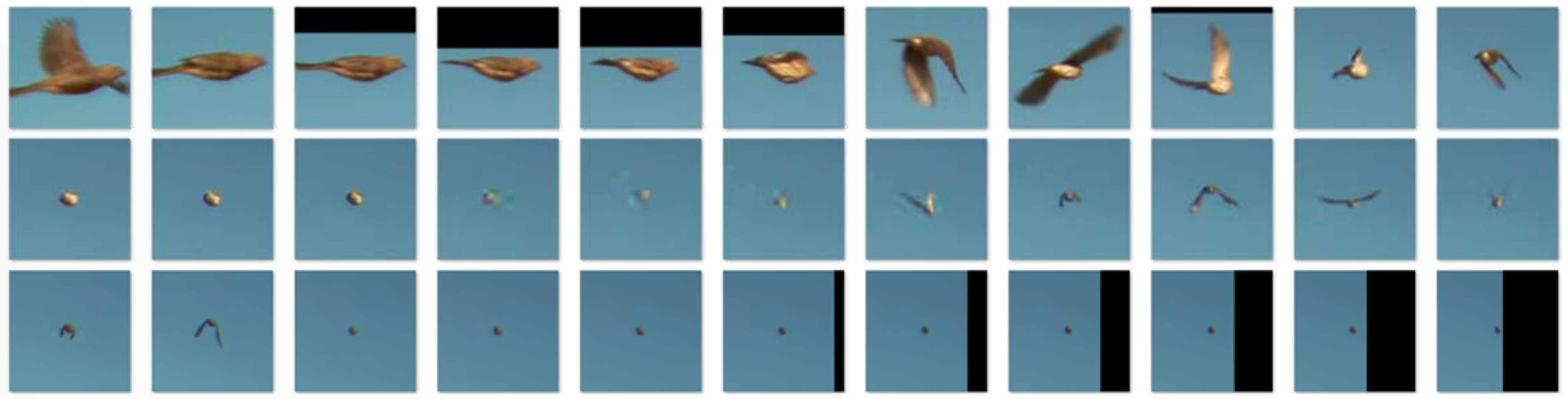
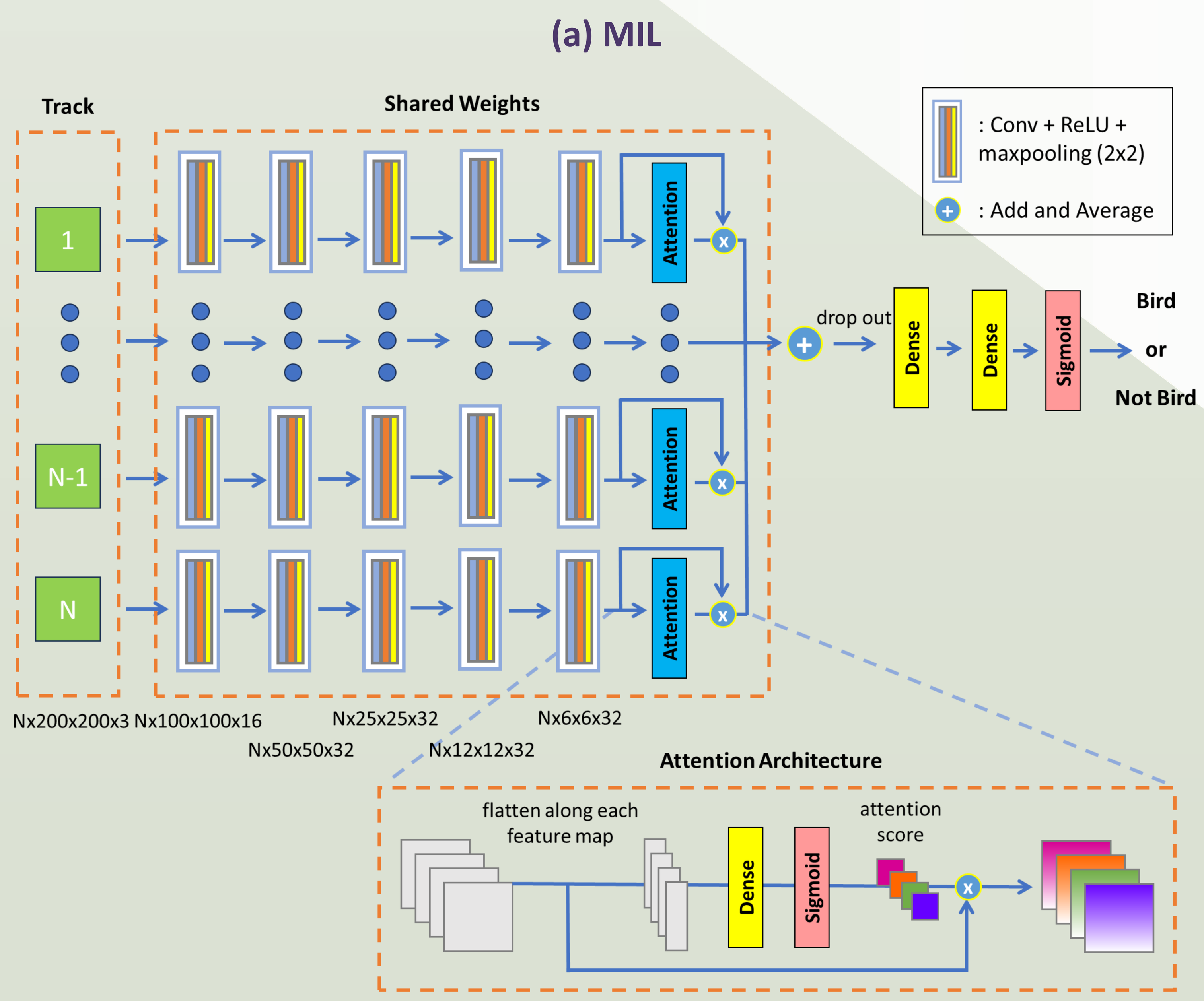


Fig.3 Birds detected and tracked at a PV site (Left) and a "track" of images, in sequence, generated by the moving object detection and tracking model (Right). The system executes the model continually during daytime, triggering the collision detection model when an object is identified as a bird, while storing data on all bird activities. It executes the activity classification model during nighttime using the stored data and transfers the output to the server. The stored data is discarded each night to clear space for the next day.

The moving object detection component is the first area where ML is used. We utilize a background subtraction algorithm known as a Gaussian Mixture Model (MOG2) that learns what the background of a sequence of video frames looks like for removal. We then use a ML model to classify the object as bird or not-bird based on the track meta-data. The entire sequence of frames along with information about the x/y location, speed, and area are used in a **multiple instance learning (MIL) ML model (Fig.4a)**. When an object is classified as a bird, a **hybrid ML model (Fig. 4b)** is used to classify collision vs. non-collision events, by utilizing Bidirectional Long Short-Term Memory (BiLSTM) layers to generate feature vectors as input to a Support Vector Machine (SVM). Finally, a **Fusion-BiLSTM model (Fig.4c)** consisting of a combination of several convolutional layers feeding into a 1D-input LSTM, using the x/y location and a 2D-input LSTM, using the x/y location as well as cropped images of the bird from the track, classifies the birds' activity.



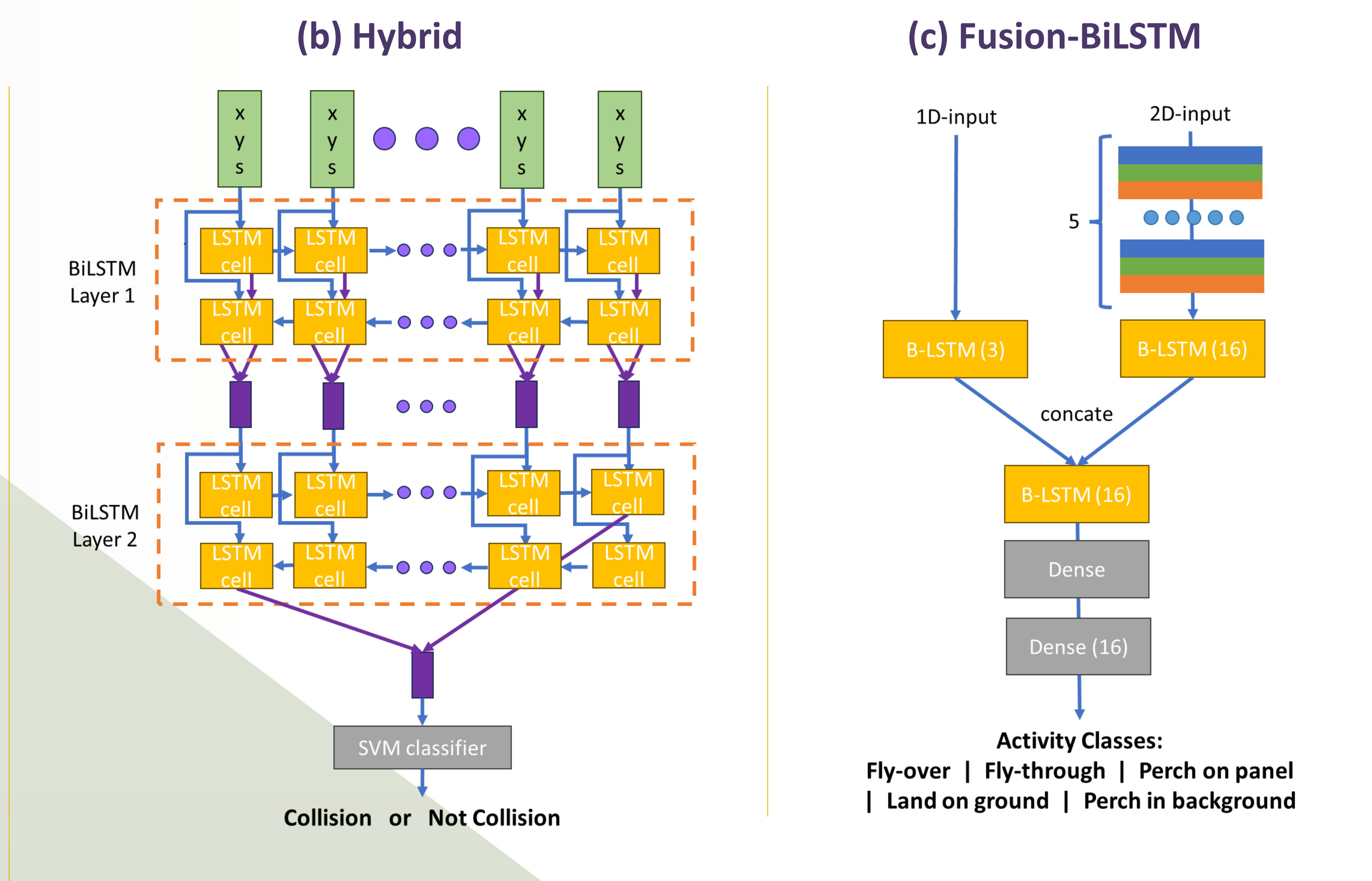


Fig.4 Architectures of machine learning models. (a) Multiple instance learning (MIL) model for differentiating birds from other moving objects, (b) Hybrid model with support vector machine (SVM) for classifying collision vs. non-collision events, and (c) Fusion-Bidirectional Long Short-Term Memory (BiLSTM) model for classifying bird activities into 5 classes (i.e., fly-over, fly-through, perch on panel, land on ground, and perch in background).

KEY TAKEAWAYS & FUTURE WORK

Building a portable machine-vision based, edge-AI system that maximizes accuracy at great distances requires trade-offs in model architectures used and software execution pipelines. We achieved the desired result by breaking the object detection, classification, and activity determinization into separate model components executed at different times while creating novel architectures for track-based data input. The challenge of having limited computational resources for deploying complex AI models also required a focus on model optimization and use of certain GPU-enabled algorithms to maximize accuracy along with efficiency and speed.

COMING SOON! Starting in Spring 2024, we will deploy our system at operational PV solar facilities at multiple U.S. regions. By incorporating conventional bird carcass surveys, we will answer outstanding questions on avian-solar interactions and clarify misconceptions on the topic.

THANK YOU! Aggelos Katsaggelos and Yuri Balasanov for expert guidance in AI model development; University of Chicago Master of Science in Analytics students for model architecture research; Sighthound for technical support integrating models into the DNNCam; partner solar facilities for providing study sites; and the advisory committee for guidance with keeping our technology relevant for solar energy development.

DeepSolar: Machine-Learning-Based Mapping and Modeling of Solar Energy with Ultra-High Spatiotemporal Granularity

Stanford University

PI: Ram Rajagopal **Co-PI :** Arun Majumdar

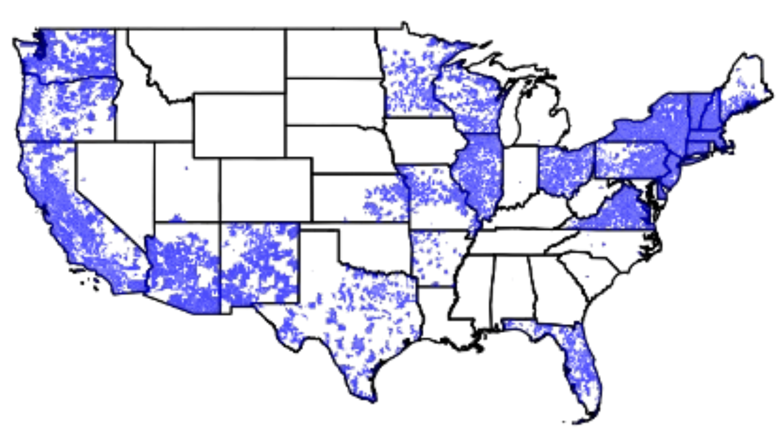
Key Personnel: Zhecheng Wang, Chad Zanocco, Rajanie Prabha, Moritz Wussow, June Flora, Chin-Woo Tan

Funded by:

SOLAR ENERGY TECHNOLOGIES OFFICE U.S. Department Of Energy

Introduction

- Granular data of solar PV installations is essential for tracking the progress of decarbonization, designing energy policies, identifying energy injustice issues, and integrating renewable energy into the grid.
- Existing datasets of PV installations are either not comprehensive in geographical scope, not sufficiently granular—from both spatial and temporal perspectives—or not publicly accessible to enable spatiotemporal analysis of PV adoption at the nationwide scale.
- Conventional data collection approaches relying on data reporting, surveying, or crowdsourcing are incapable of constructing and maintaining a complete solar PV dataset covering the entire country, especially states that lag behind in PV adoption.

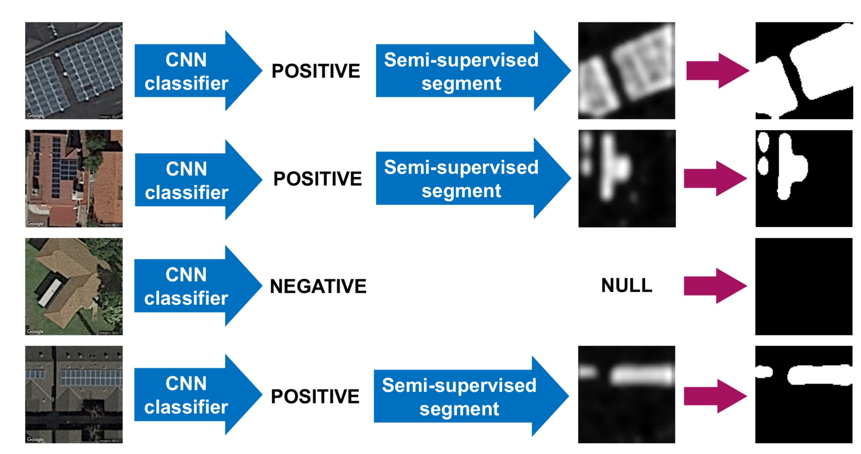


The geographic coverage of the "Tracking the Sun" dataset

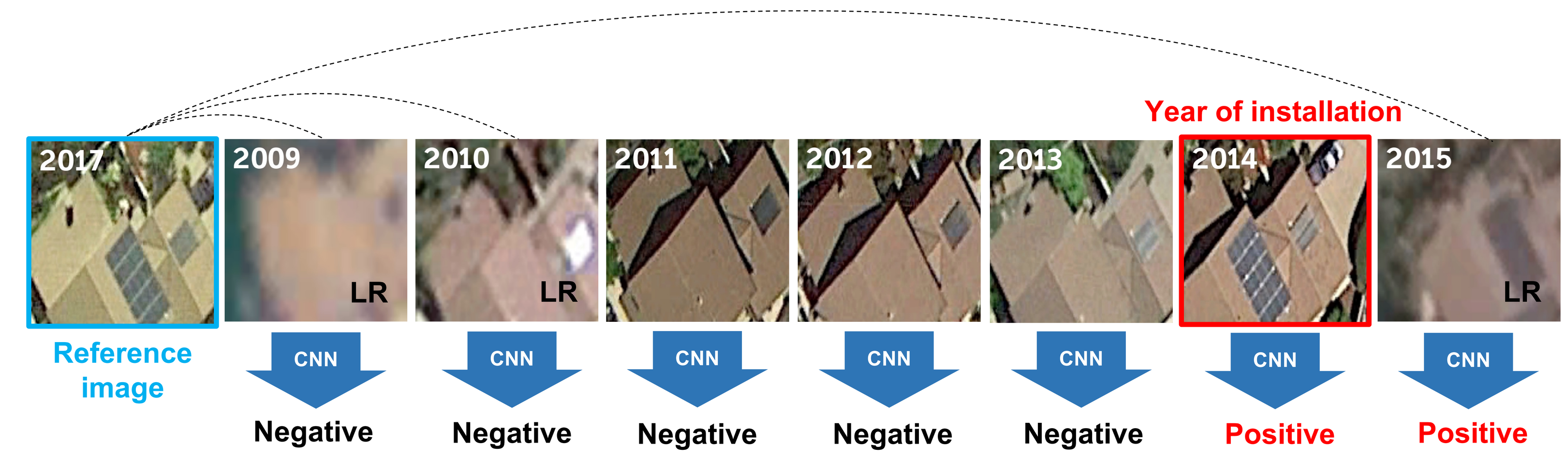
Our contributions:

- Leveraged AI to construct a U.S.-wide spatiotemporal solar PV installation database.
- Identified heterogeneity in the dynamics of PV adoption.
- Identified certain types of incentives which can mitigate the heterogeneity.

AI for mapping solar PVs

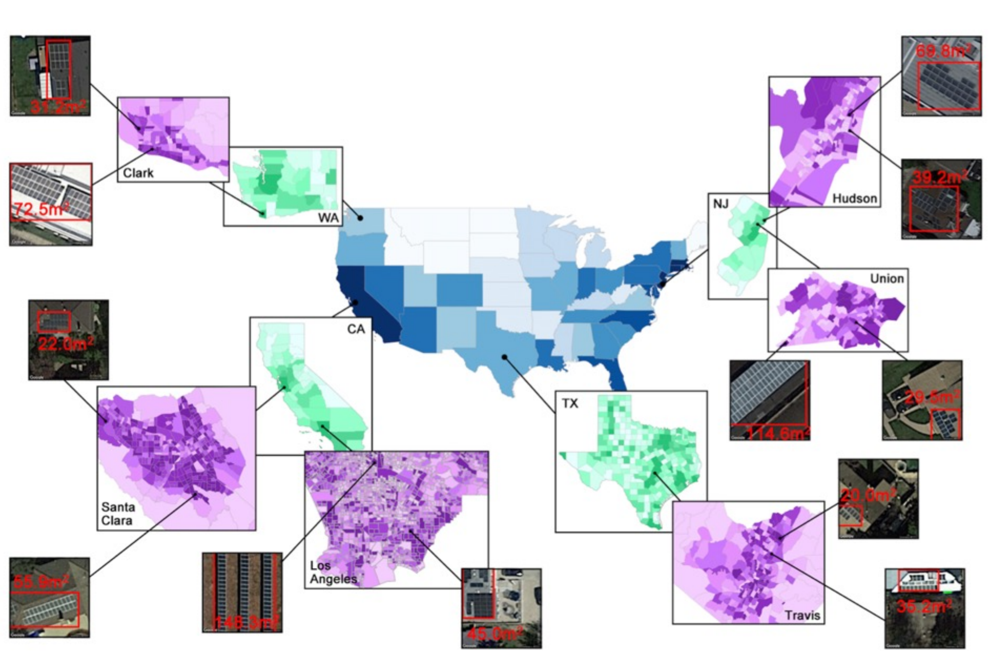


- **Input data:** Satellite and aerial images
- **Geospatial mapping:** Fully-supervised image-level binary classification → semi-supervised size estimation → subtype classification
- **Timelapse uncovering:** Deep Siamese Network for identifying the year of installation (86% accuracy)
- **Deployment:** 1 billion image tiles

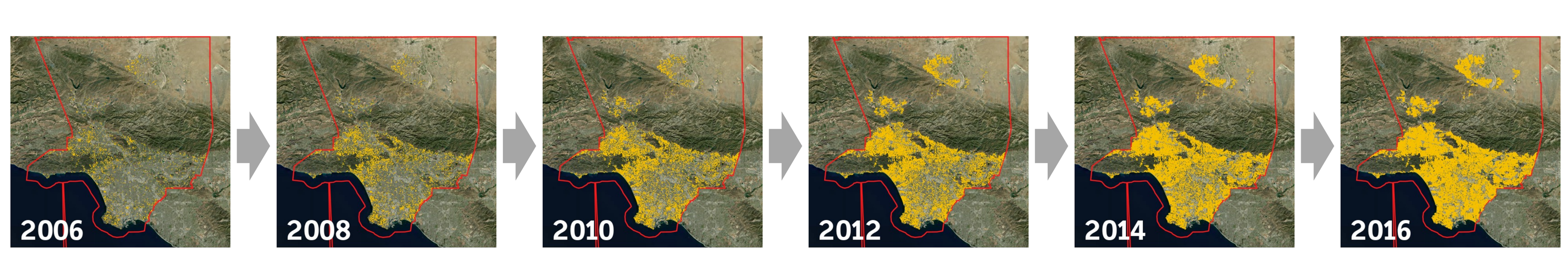


DeepSolar dataset

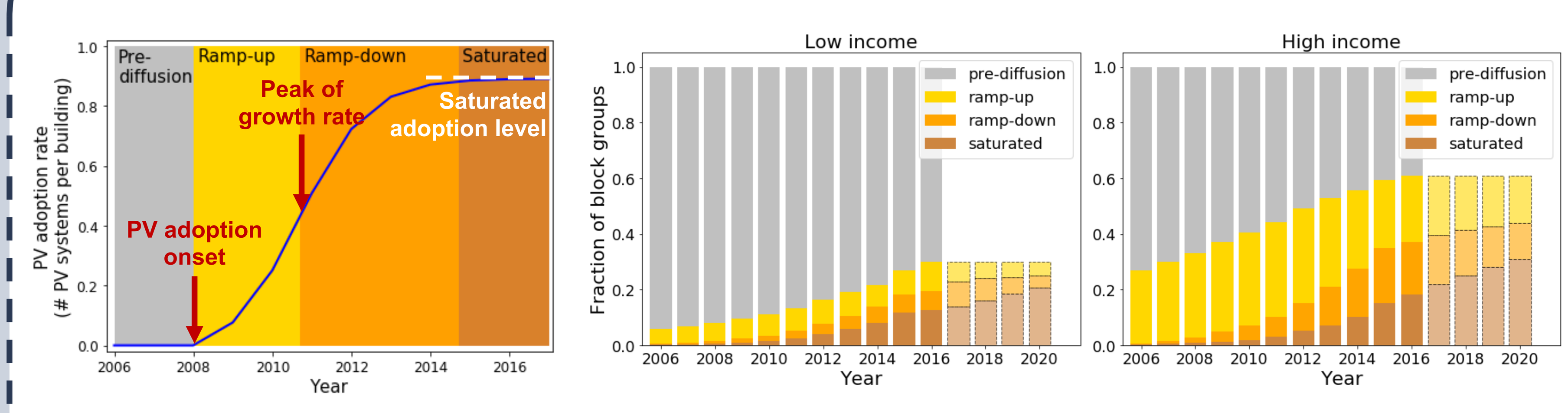
- The geospatial dataset constructed in 2017 contains 1.5 million solar installations in the contiguous U.S. with geo-coordinate, size, and subtype information.
- Its 2023 updated version contains 3 million solar installations in the whole U.S. including Alaska and Hawaii, which will be made publicly available soon.



- DeepSolar++ dataset further included the year of installation information for every residential and commercial PVs.
- Datasets can be downloaded at https://deepsolar.web.app
- **Usages:** socioeconomic, policy, grid integration, resilience analyses, etc.



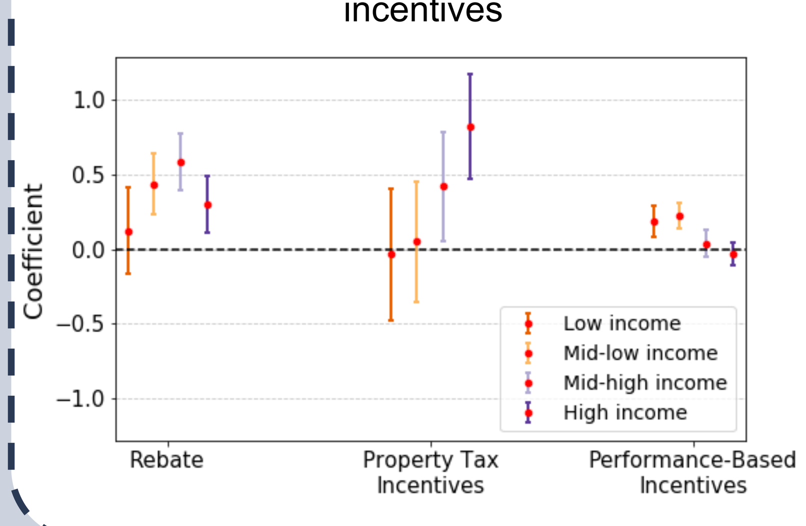
Solar adoption dynamics



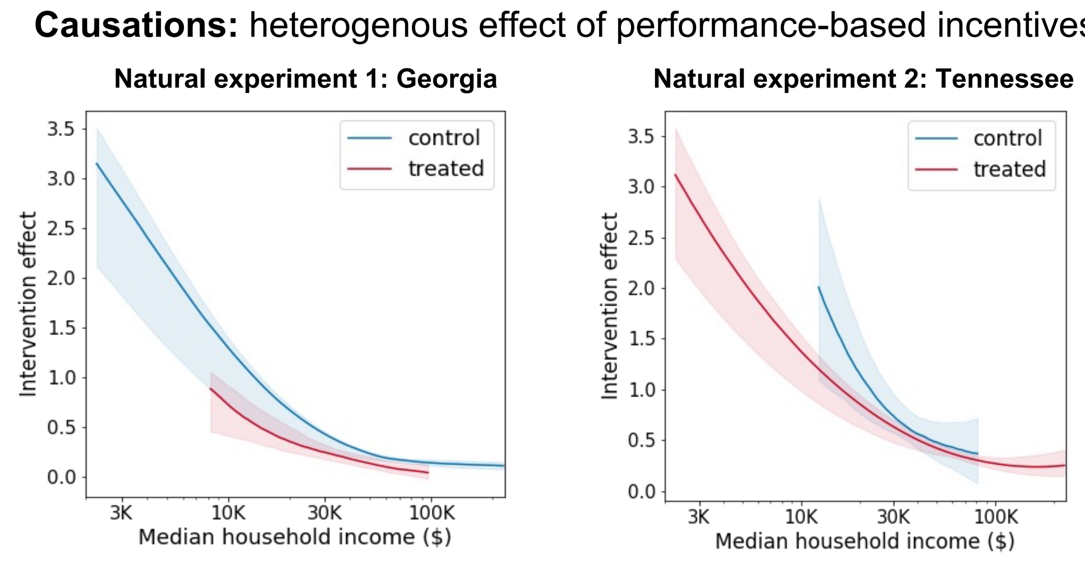
- We analyze the non-linear dynamics of residential PV adoption using technology diffusion model (Bass model). Specifically, we use Bass model to estimate the time of adoption onset and saturated adoption level in each census block group.
- We find that, low-income communities are not only delayed in their adoption onset, but also more likely to saturate at lower levels.
- The median saturated adoption level is 28 systems per 1000 buildings in low-income communities and 58 systems per 1000 buildings in high-income communities.
- Wealthier communities started adoption at lower levels of PV benefit (with rebate/ grant), yet the PV benefit (with rebate/grant) they experienced is higher. This suggests a potential for re-distribution of existing upfront subsidies to lower income communities to make PV adoption more equitable.

Policy insights

Correlations: saturated adoption level vs. incentives



Causations: heterogenous effect of performance-based incentives



Correlational analysis:

- Although rebate and property tax incentives are positively correlated with saturated adoption level for high and mid-high income communities, there is no statistically significant correlation in low-income communities.
- By contrast, performance-based incentives are associated with higher saturation in low-income communities but not in high-income ones.

Causal analysis:

- By utilizing causal forest model to estimate the heterogeneous intervention effect, we find that **performance-based incentive is more effective in low-income communities,** corroborating the correlational evidence.

Publications

- Zhecheng Wang, Marie-Louise Arlt, Chad Zanocco, Arun Majumdar, and Ram Rajagopal (2022). **DeepSolar++: Understanding Residential Solar Adoption Trajectories with Computer Vision and Technology Diffusion Models.** *Joule*.
- Zhecheng Wang, Michael Wara, Arun Majumdar, and Ram Rajagopal (2023). **Local and Utility-Wide Cost Allocations for a More Equitable Wildfire-Resilient Distribution Grid.** *Nature Energy*.
- Kevin Mayer, Benjamin Rausch, Marie-Louise Arlt, Gunther Gust, Zhecheng Wang, Dirk Neumann, and Ram Rajagopal (2022). **3D-PV-Locator: Large-Scale Detection of Rooftop-Mounted Photovoltaic Systems in 3D**. *Applied Energy*.
- Kevin Mayer, Zhecheng Wang, Marie-Louise Arlt, Dirk Neumann, and Ram Rajagopal (2020). **DeepSolar for Germany: A Deep Learning Framework for PV System Mapping from Aerial Imagery**. *International Conference on Smart Energy Systems and Technologies (SEST).*
- Zhengcheng Wang*, Zhecheng Wang*, Arun Majumdar, and Ram Rajagopal (2019). **Identify Solar Panels in Low Resolution Satellite Imagery with Siamese Architecture and Cross-Correlation**. *NeurIPS Tackling Climate Change with Machine Learning Workshop*.
- Jiafan Yu*, Zhecheng Wang*, Arun Majumdar, and Ram Rajagopal (2018). **DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States**. *Joule*.

Challenges and Future Work

- **Privacy challenge.** Geo-coordinate information of solar installations is not made publicly available due to privacy concern. Future work can explore advanced methods for sharing the data in its highest granularity while preserving privacy.
- **Fairness challenge.** Machine learning models may not perform equally well across different regions and communities, which could bias the PV adoption analysis. This deserves further investigation.
- **Continuous update.** We aim to build a data processing pipeline to update the dataset on a yearly basis.
- **Community solar.** We aim to identify community solar and analyze its adoption.

Acknowledgement

This project is funded by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office (SETO) Fiscal Year 2020 Funding Program (award number DE-EE0009359).

End-of-Line Solar Cell Binning via Machine Learning

Lena Bruno, Adrienne L. Blum, Harrison Wilterdink, Ronald A. Sinton
Sinton Instruments, Boulder, CO, USA

ABSTRACT: Sinton Instruments is developing an in-line current-voltage (I-V) and line-scan photoluminescence (PL) tool to characterize solar cells. The tool will operate at line-speed to provide near-contactless binning of solar cells and defect recognition. We are in the process of creating a Machine Learning model that classifies cells based on their line-scan PL and I-V characteristics. We anticipate this will have a wide range of applications with the solar industry.

Line-Scan PL System

Line-scan PL [1] is a powerful imaging technique that can be performed on-the-fly at line speed. It shows many of the same features as conventional electroluminescence (EL) without the need for probe bars that shade the cell during imaging.

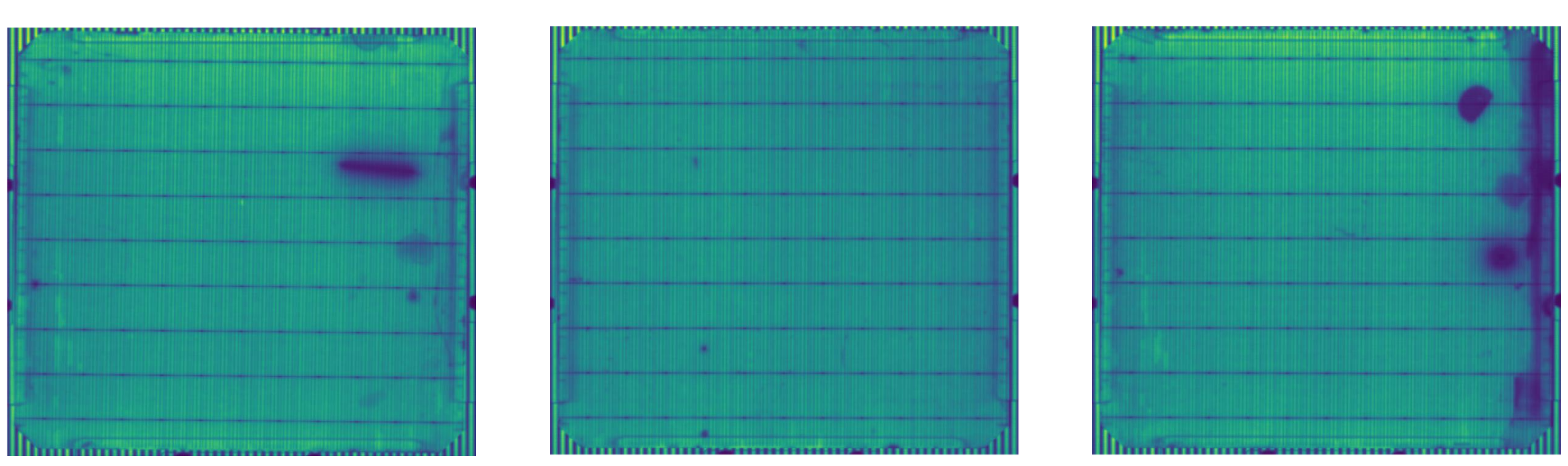
Binning

End-of-line characterization and binning of solar cells is necessary to both sort out defective cells and group similarly performing cells to avoid power mismatch in PV modules. We use a Convolutional Neural Network (CNN) to automate this process, sorting cells based on their line-scan PL image and I-V data.

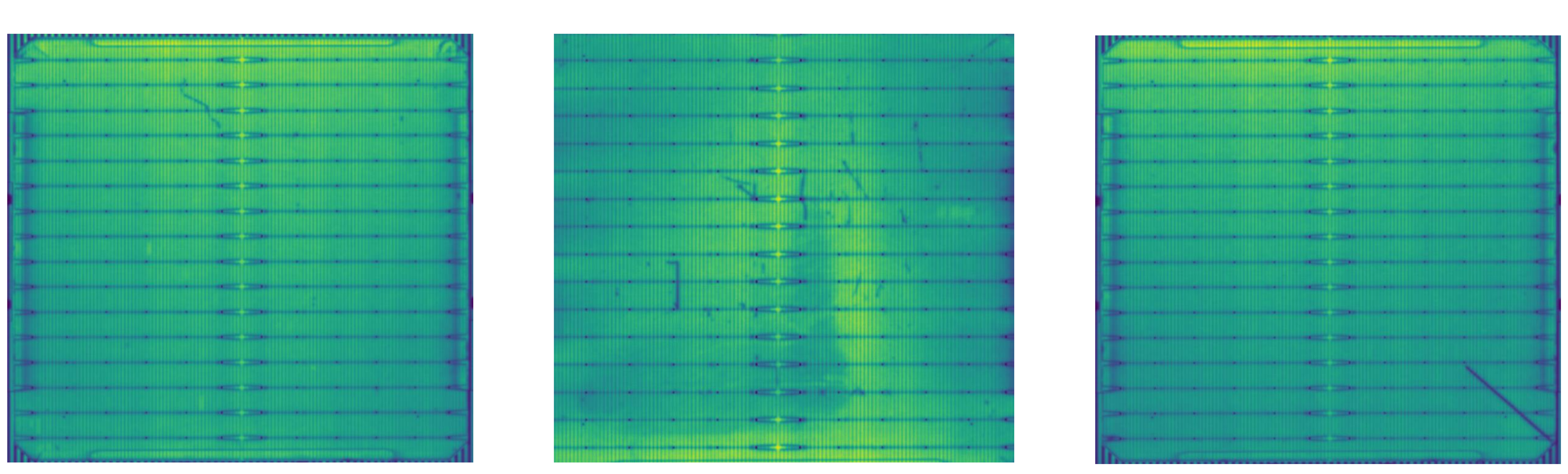
Cells with Defects:

Defects can manifest in a number of visually detectable ways as seen below. False color is applied to enhance visibility of some defects.

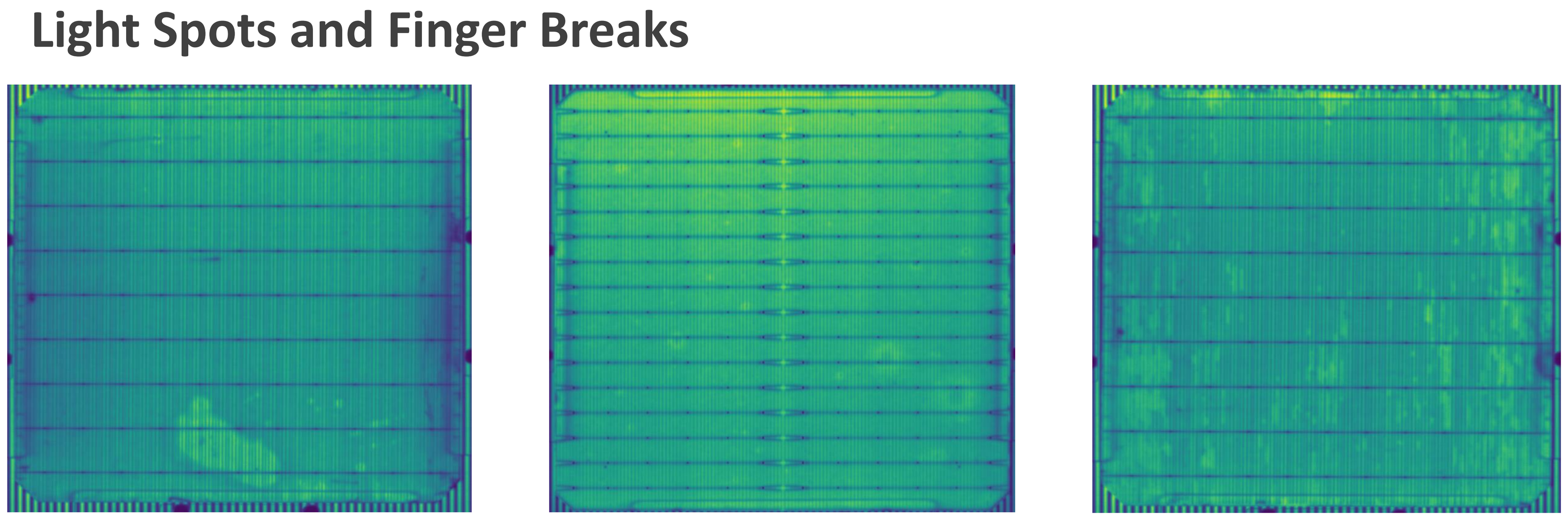
Dark Spots



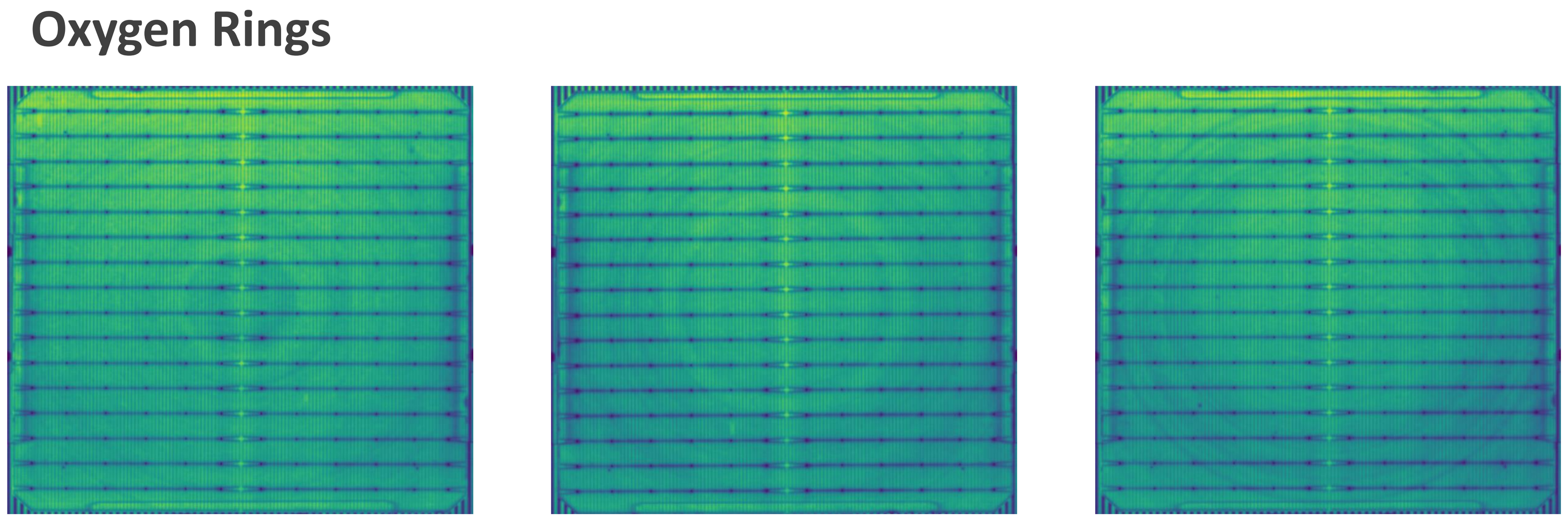
Cracks and Scratches



Light Spots and Finger Breaks

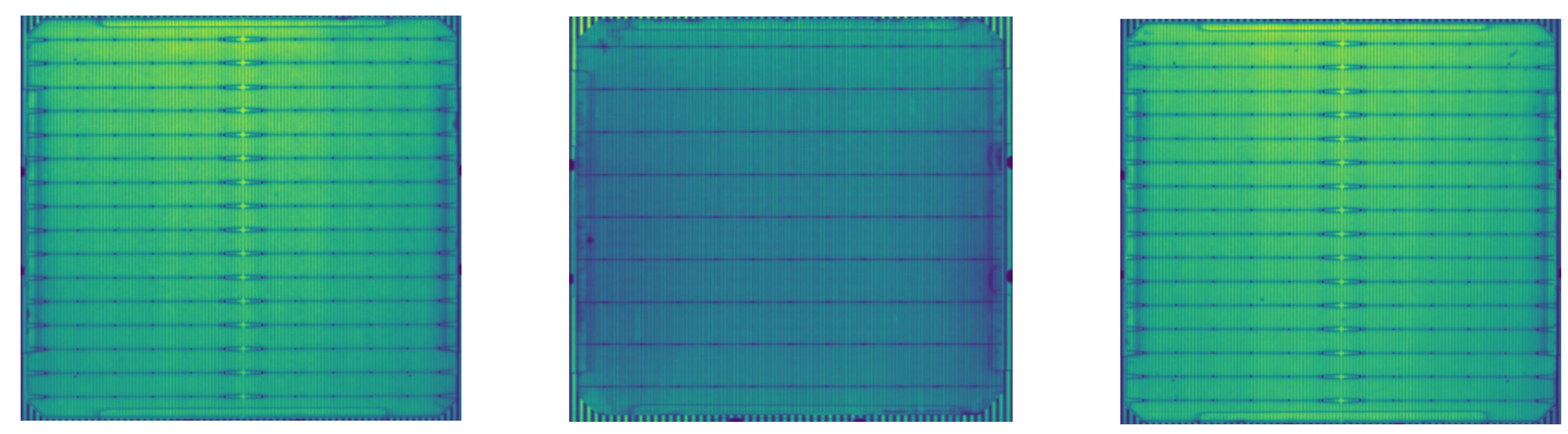


Oxygen Rings



Cells without Defects:

Acceptable cells can span a range of brightness while still being functional.



Machine Learning Code

Currently, the code performs binary sorting: classifying cells as either good or bad. This utilizes a number of CNNs built into Pytorch.

Based off a sample set of 1000 line-scan PL images, the binary sorting code classifies them with the following efficiencies:

97.1 % True Positive

56.6 % True Negative

28.4 % False Positive

1.0 % False Negative

		Predicted Class	
		Good	Bad
Actual Class	Good	True Positive	False Negative
	Bad	False Positive	True Negative

Binary sort is inadequate for end-of-line cell characterization, but provides a proof-of-concept for using CNNs in this application. This iteration of the code also only uses the line-scan PL data, and does not yet take into account the I-V data.

Code Goals and Open Questions

- **Sorting by Defect**: Sort cells by a number of defects
  - How to accurately address concurrent defects?
- **Utilization of I-V Data**: Incorporate I-V data from cells into program
  - Can we quantify power lost to individual defects?
- **Binning Thresholds**: Allow users to fine-tune reject criteria
  - What thresholds are appropriate to make adjustable?
- **Defect Localization**: Have the code output the location of defects detected on the cell.
  - Can we do so using UNSW's trained CNN (LumiNet [2])?
- **Data Aggregation**: Ability to save defect binning/localization outputs so that customers can use it for production improvements.
  - What data is the most valuable/reliable for this purpose?

References:

[1] I. Zafirovska, M. Juhl, J. Weber, O. Kunz, and T. Trupke, "Module inspection using line scanning photoluminescence imaging", in *32nd EU PVSEC*, Munich, 2016
[2] Y. Buratti, A. Sowmya, R. Evans, T. Trupke, and Z. Hameiri, "Half and full solar cell efficiency binning by deep learning on electroluminescence images", in Progress in Photovoltaics: Research and Applications, Vol. **30**, Issue 3, 2021
[3] J. Wong, P. Teena, and D. Inns, "Griddler AI: new paradigm in luminescence image analysis using automated finite element methods", in *44th IEEE PVSC*, 2017

Harnessing Sensor Data for Degradation Analytics and Operations & Maintenance Optimization in PV Systems: A Prognostic Approach

Principal Investigator: Feng Qiu (Argonne National Laboratory)

Other Contributors: Sandia National Laboratories, Wayne State University, Iowa State University
Ideal Energy Inc., CB Solar Inc., Prime Partners Engineering, LLC, Pacific Gas and Electric Company, SunEnergy1

Funded by: SOLAR ENERGY TECHNOLOGIES OFFICE U.S. Department Of Energy

Acknowledgment: This material is based upon work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Solar Energy Technologies Office Award Number 38458

Project Overview

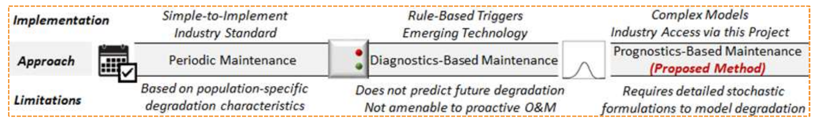
OBJECTIVE:

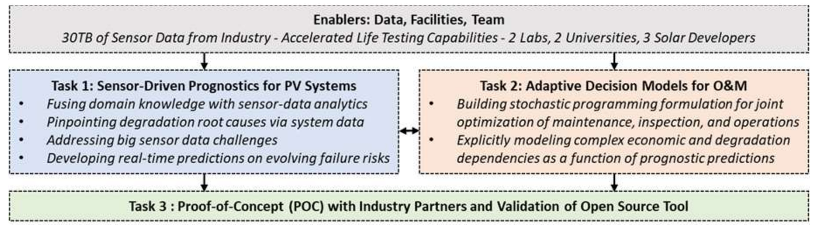
To provide a new generation of prognostics, operations and maintenance (O&M) approaches for PV inverters in an effort to:

- ✓ ***Reduce maintenance costs and LCOE*** (through accurate determination of inverters in need of service, optimizing opportunistic maintenance, reducing operational implications due to outages)
- ✓ ***Enhance inverter service life*** (by replacing them efficiently before onset of failure, rather than relying on a set schedule)
- ✓ ***Reduce inverter unavailability*** (by minimizing failures and modeling economic dependencies)

APPROACH:

- ✓ Develop prognostic models for inverters by analyzing a large-scale sensor data repository from industry and laboratory data
- ✓ Develop efficient reformulations and modeling approaches to embed the developed RLD predictions and prognostic models within stochastic programming
- ✓ Build large-scale O&M optimization models and associated solution methodologies
- ✓ Co-develop open-source tools for prognostics and O&M schedules, in cooperation with the industry partners





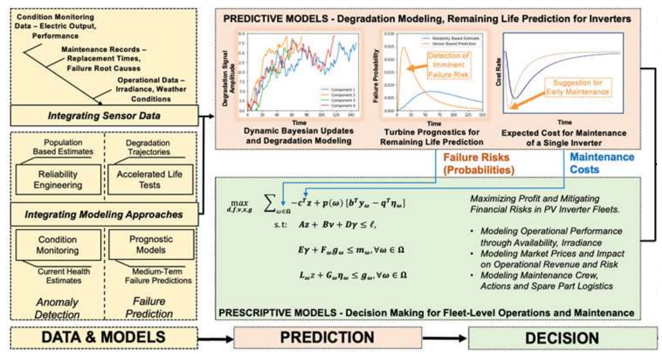
Project Approach – *from Streaming Sensor Data to Experiments*

DATA:

- ✓ **Accelerated life testing (ALT) for solar inverter and corresponding database**
  - Experiment types: (1) High Temperature Operating Life Testing (HTLT); (2) Powered Thermal Cycling (PTC) Test; (3) Damp heat test; (4) Lightning surge
  - Temporal coverage of over 2 years, temporal resolution up to every 1 second
- ✓ **Industrial inverter data and maintenance records**
  - 751 inverters in 443 sites, with 4 inverter manufacturers, string inverters and micro-inverters
  - Inverters locates in 6 states across U.S. (IA, MN, MI, TX, NC, VA)
  - Diversified terrains and applications (agricultural and urban)
  - Temporal coverage of over 10 years, temporal resolution up to every 5 minutes
  - Over 2300 alert records related to maintenance and failure history

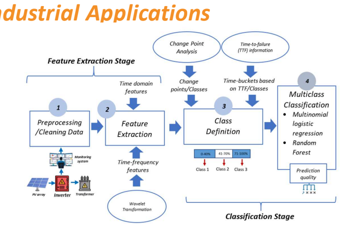
EXPERIMENTS:

- ✓ **Predictive Model #1: Inverter diagnostics and prognostics using ALT experiment data**
- ✓ **Predictive Model #2 : Inverter diagnostics using industrial data**
  - (1) Health Indicator-Based Degradation; (2) Anomaly Events-Based Degradation
- ✓ **Prescriptive Model: Prognostics-driven operations and maintenance**



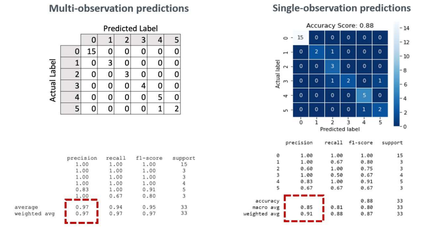
Prediction of Inverter Life – *from Laboratory Experiments to Industrial Applications*

- ✓ **Classification-based diagnostics and prognostics for ALT datasets.**
  - Leveraged high-fidelity laboratory data from ALT experiments to develop models for classifying degradation states of inverters
  - Asset conditions are defined as a function of time to failure
  - Two models were developed: (i) Single-observation model is used to provide fast prediction, and (ii) multi-observation model uses multiple observations to offer a slower but more accurate prediction of inverter condition
  - Multi-observation prediction model achieved 91-97% prediction accuracy

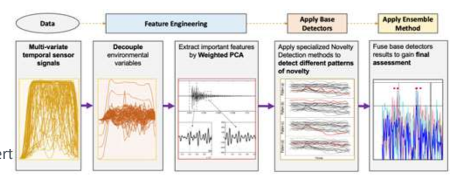


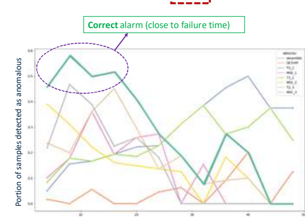
Multi-observation predictions

Single-observation predictions

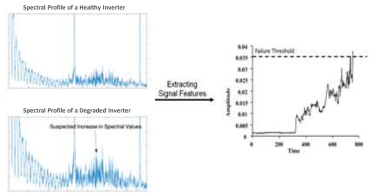


- ✓ **Anomaly detection based diagnostics for industrial datasets**
  - Leveraged sequential industrial data (i.e. weekly patterns) to discover latent signs of degradation in inverters
  - Developed methods to decouple confounding environment impacts on sensor data to enable diagnostics in dynamic operational environments
  - Formulated a ensemble method to fuse the strength of multiple novelty detectors to capture a wide range of degradation behavior
  - Demonstrated statistically significant improvements in novelty detection, to alert operators of evolving degradation processes





- ✓ **Prognostic modeling – mathematical formulations to predict future failure risks**
  - Modeled long-term behavior of degradation using Brownian Motion based degradation models, and use first-passage time to predict future failure risks
  - Used Bayesian statistics to continuously update the degradation parameters, and the associated predictions for the remaining life distribution of inverters
  - Proposed approach captures population-based degradation characteristics and finetunes them using data-driven asset-specific degradation characteristics
  - Showcased significant improvements in prediction accuracy compared to benchmark models based on reliability-based prediction



Step 1 - Degradation model for inverter i

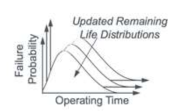
$$D_i(t) = \phi_i(t; k, \theta_i) + \epsilon_i(t; \sigma)$$

Step 2 - Given observations d_i^o, we update the degradation parameters

$$v(\theta_i) = P(\theta_i | d_i^o) = P(d_i^o | \theta_i) \pi_i(\theta_i)$$

Step 3 - Reevaluate the prediction on the remaining life distribution of inverter i

$$P(R_{t_i^o} > t) = \int P\left(\sup_{t_o \le s \le t_o + t} D_i(s) < \Lambda_i | \theta_i, x, y, \delta\right) v(\theta_i) d\theta_i$$



Prognostics-Driven Operations and Maintenance – *from Prognostic Predictions to Fleet Optimal O&M Decisions*

- ✓ **Prognostics-driven decision optimization models for operations and maintenance**
  - Developed a novel decision optimization model that inherently captures inverter degradation and prognostic predictions on failure risks
  - Modeled economical dependencies across assets including maintenance crew routing, spare part logistics, opportunistic maintenance, and operational implications
  - Modeled planned maintenance actions as first stage variables, and corrective maintenance, rerouting and operational decisions as second stage variables
  - Demonstrated significant improvements in maintenance cost (>45.8%), and crew costs (>36.5%) by proactive planning of O&M decisions using prognostic predictions.

$$\max_{y,x,z,v,w,u} \quad \underbrace{a^t \boldsymbol{y}}_{\text{Expected revenue}} - (\underbrace{b^t \boldsymbol{x}}_{\text{maintenance team deployment cost}} + \underbrace{c^t \boldsymbol{z}}_{\text{expected corrective maintenance cost}} + \underbrace{d^t \boldsymbol{v}}_{\text{preventive maintenance cost}} + \underbrace{e^t \boldsymbol{u}}_{\text{expected end of life cost}})$$

s.t.

$A\boldsymbol{x} \le f$ → Maintenance team logistics/deployment constraints

$B\boldsymbol{x} + C\boldsymbol{z} + D\boldsymbol{v} \le g$ → Maintenance (preventive & corrective) coordination constraints

$E\boldsymbol{z} + F\boldsymbol{v} + G\boldsymbol{y} \le h$ → Energy generation and maintenance coupling constraints (operational)

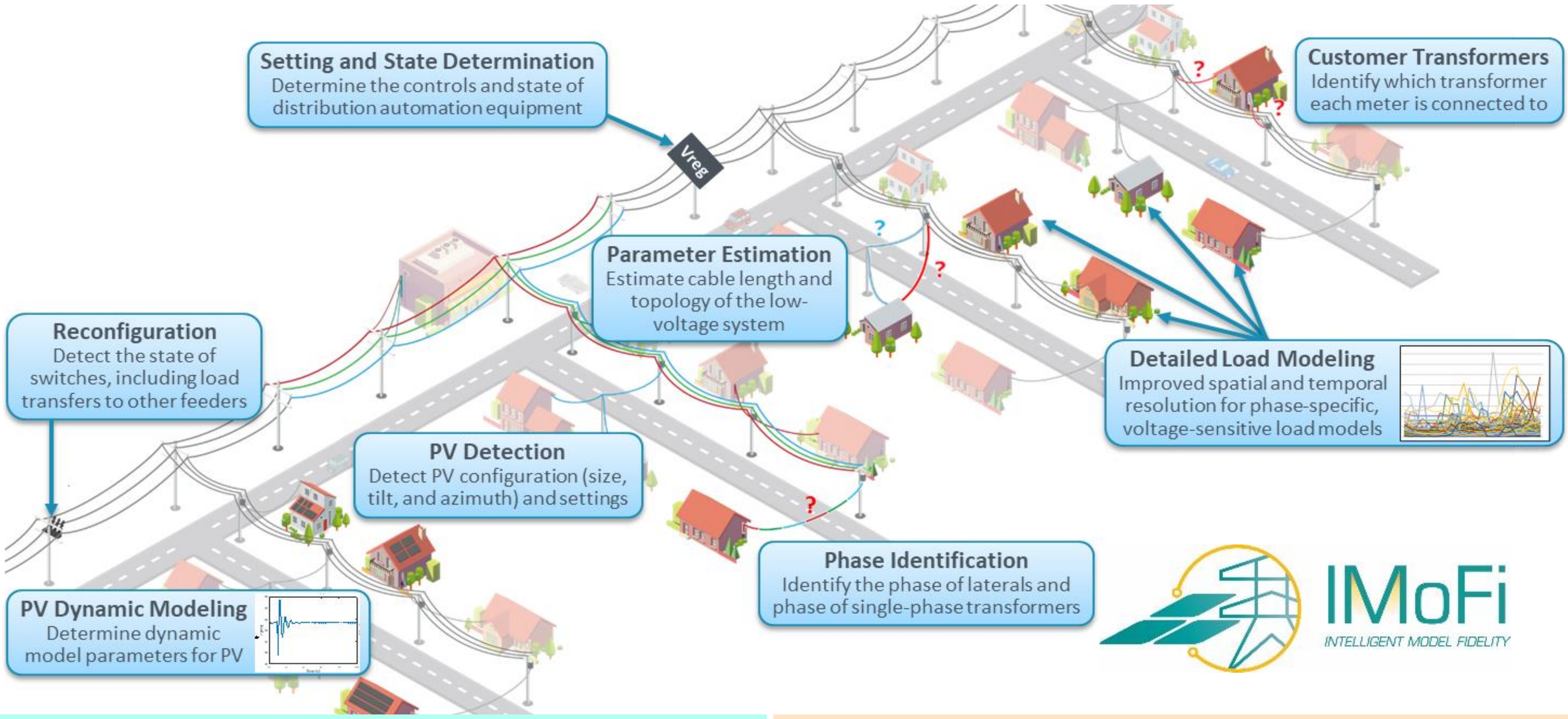
$M\boldsymbol{u} + N\boldsymbol{z} + P\boldsymbol{v} \le n$ → Degradation update constraints

	Prognostics-Based	Time-Based	Δ%
#Preventives	129	153	↓ 15.9%
#Correctives	21	60	↓ 65.0%
#Failures	21	63	↓ 66.7%
#Crew Visits	122	192	↓ 36.5%
Unused Life	18.62	68.55	↓ 72.8%
Availability	97.67	95.03	↑ 2.8%
Maintenance Cost	319,500	589,500	↓ 45.8
Crew Cost	366,000	576,000	↓ 36.5

Intelligent Model Fidelity (IMoFi)

Award # DE-EE0034226: "Physics-Based Data-Driven Modelling to Accelerate Accurate PV Integration"

Principal Investigator: Matthew Reno



Introduction:

- Modern distribution analysis algorithms and tools are continually improving (hosting capacity analysis, QSTS, and DERMS) but use feeder models based on manual data entry that is prone to error and often out of date with little validation or calibration
- The PV hosting capacity is highly sensitive to the feeder model - A few volts difference can result in PV hosting capacity varying by more than 200%
- Improved models provide more accurate interconnection screening (reducing PV interconnection costs)
- Recent additions of Advanced Metering Infrastructure (AMI), or smart meters, provide measurements of each customer's power consumption, and possibly other quantities, such as voltage, that provide new insights and levels of accuracy in distribution system modeling

Key Takeaways:

- Leveraging physics, such as power-flow equations and models, within data-driven methods is key
- Data management is a critical consideration for successful AI/ML implementations
- Data-driven methods provide significant time-saving benefits compared to traditional model calibration methods
- Tools should interface directly in utility GIS systems

[1] M. J. Reno, et al, "IMoFi (Intelligent Model Fidelity): Physics-Based Data-Driven Grid Modeling to Accelerate Accurate PV Integration Updated Accomplishments", Technical Report SAND2022-12669, Sandia National Laboratories, 2022.
[2] L. Blakely, M. J. Reno, "Phase Identification Using Co-Association Matrix Ensemble Clustering," IET Smart Grid, 2020.
[3] B. D. Peña, L. Blakely, M. J. Reno, "Data-Driven Detection of Phase Changes in Evolving Distribution Systems", TPEC, 2022.
[4] L. Blakely and M. J. Reno, "Identification and Correction of Errors in Pairing AMI Meters and Transformers," IEEE Power and Energy Conference at Illinois (PECI), 2021.
[5] L. Blakely, M. J. Reno, "Identification and Correction of Errors in Pairing AMI Meters and Transformers", PECI, 2021.
[6] C. Gomez-Peces, S. Grijalva, M. J. Reno, L. Blakely, "Estimation of PV Location Based on Voltage Sensitivities in Distribution Systems with Discrete Voltage Regulation Equipment", IEEE Powertech, 2021.
[7] K. Ashok, M. J. Reno, D. Divan, L. Blakely, "Secondary Network Parameter Estimation for Distribution Transformers", ISGT, 2020.

Challenges:

- Acquiring and processing the data for these methods is challenging. For example, instantaneous vs average voltage measurements, missing data, time synch between data sources, and insufficient measurement resolution; additionally, utilities often have differing data collection practices

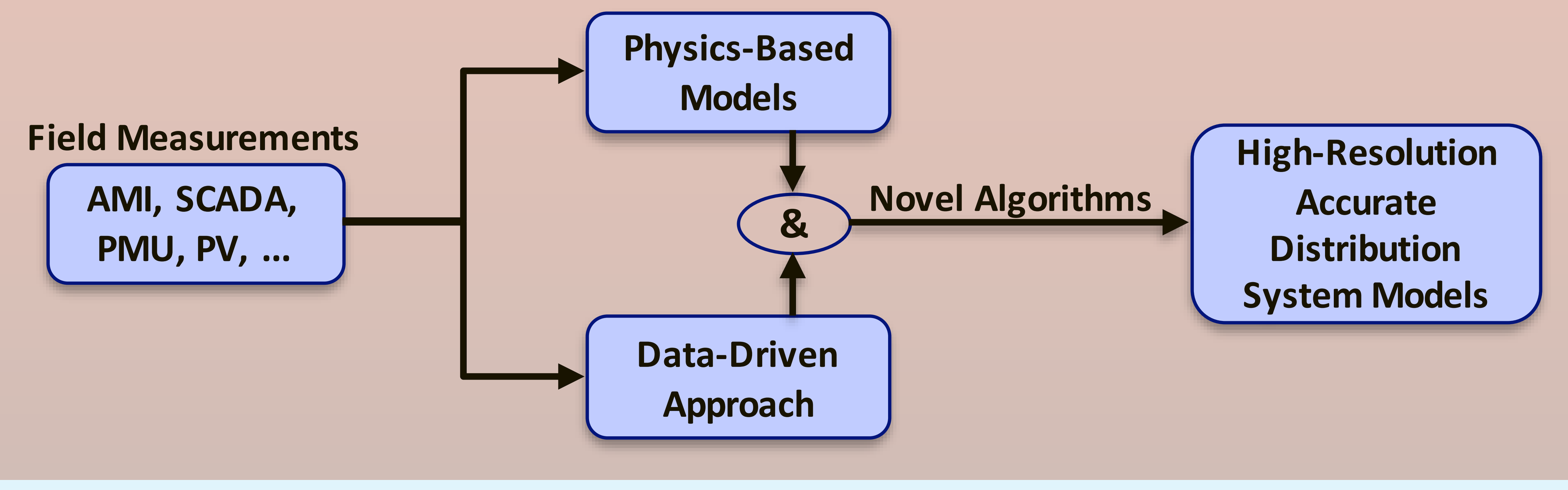
Best Practices:

AMI Data Collection Recommendations

Data Consideration	Recommendation
Measurement Interval	15-minute intervals
Meter Precision	At least one decimal on voltage (240), real power, and reactive power measurements
Meter Bias	Bias does not impact the algorithms developed in this project
Measurement Noise	< 0.35% maximum uniform random noise is recommended
Time Synchronization	For most situations using 15-min intervals will compensate for time synch issues. However correct date/time and daylight savings time settings are required
Missing Data	Developed algorithms are robust to some amounts of missing data
Data Availability	1-3 months is recommended at minimum depending on the algorithm. 6-12 months is recommended as best practice

How is AI/ML used in this project?:

- This project leveraged unsupervised ML (phase identification), regression techniques (parameter estimation & meter/transforming pairing), deep learning (PV detection), and correlation analysis (meter/transform pairing)
- Algorithms published as open-source: https://github.com/sandialabs/distribution-system-model-calibration
- Algorithms published on NRECA Open Modeling Framework: https://www.omf.coop/ (phase ID)



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. ____________

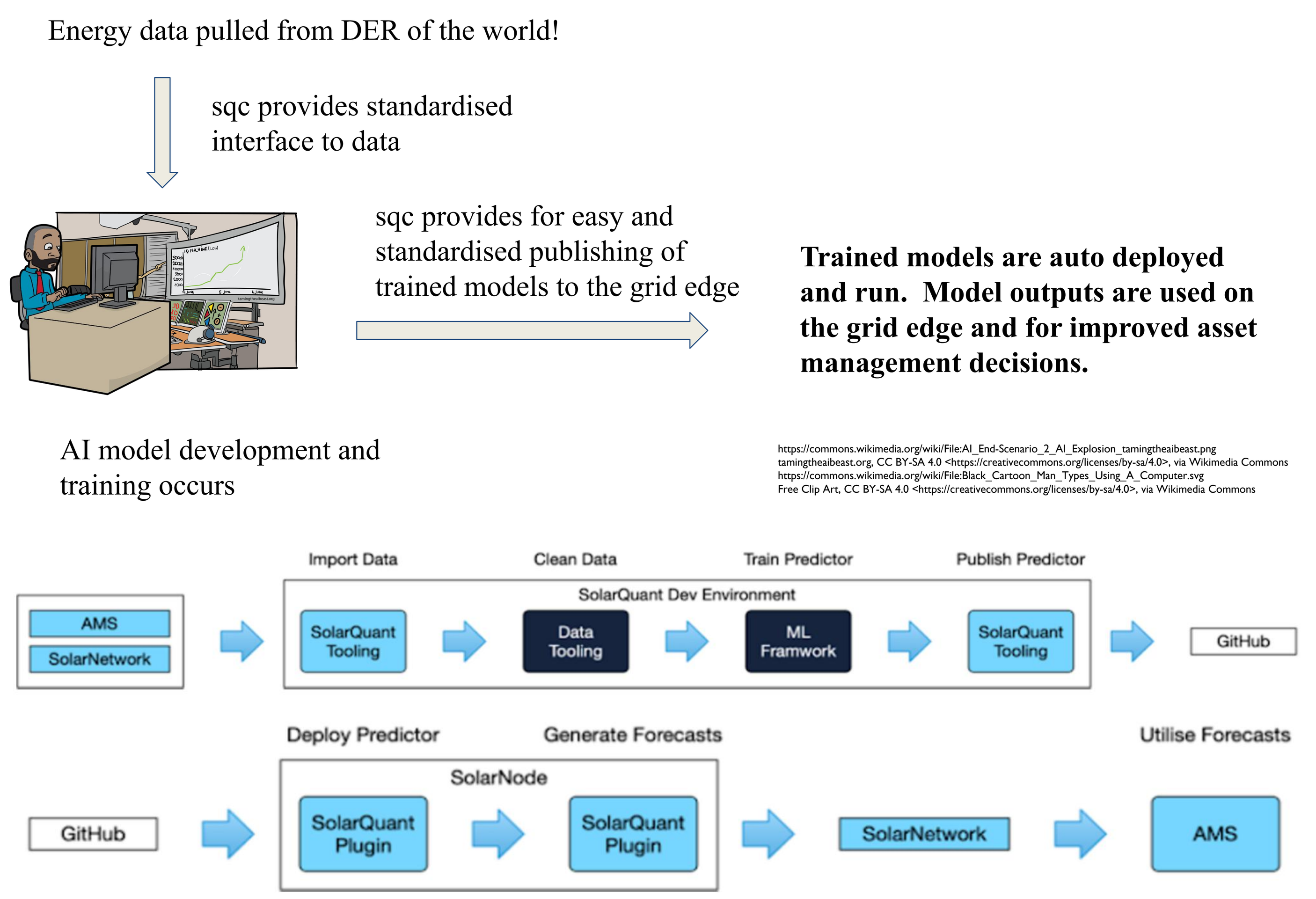
Learning-Assisted Preventive and Corrective Maintenance of PV Systems: Predicting Heterogeneous Failures from Heterogeneous Data

Yue Zhao, Kang Pu, Philip Court, John Gorman, Arun Veeramany, and Meghana Ramesh
Stony Brook University, Ecogy Energy, PNNL

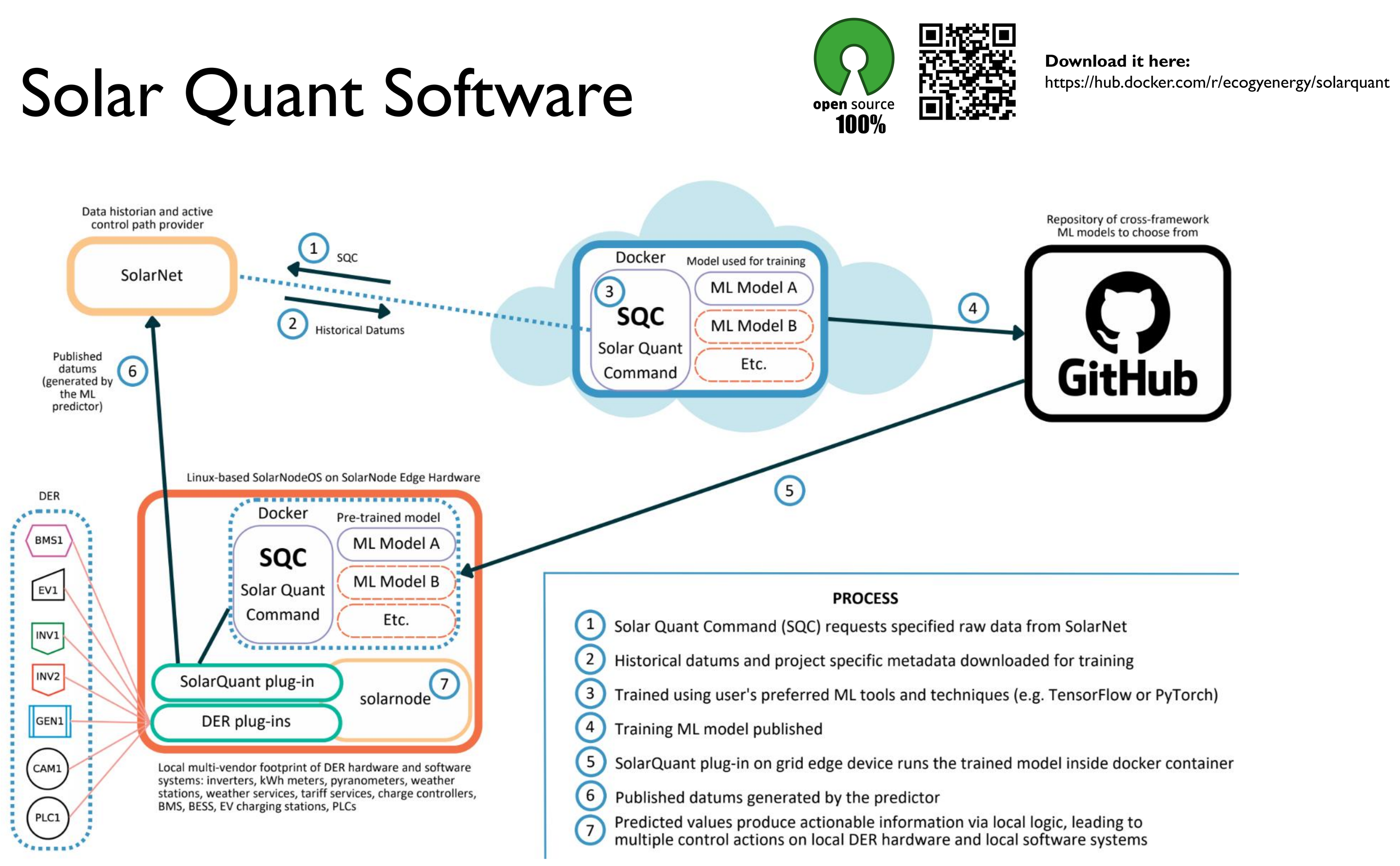
Introduction

- Objective: A family of data-driven machine learning modules that provide real-time detection, classification, and prediction of heterogeneous PV system failures and thereby facilitate efficient corrective and preventive maintenance of PV systems.
- The solutions will enable PV asset managers to a) be aware of hidden problems by detecting incipient failures so that loss of energy are minimized, b) reduce redundant maintenance due to false alarms or over-conservativeness, and c) better optimize preventive maintenance scheduling and improve maintenance efficacy.
- Goal: Reduce the PV asset owner's O&M budget by 20% and annual maintenance costs to below $4/kW per year.

Solar Quant Concept



Solar Quant Software



Challenges

- **Lack of physical models for PV systems.**

---- Primarily data-driven methods needed.

- **PV system failure events are heterogeneous and rare.**
- **Only a limited number of relevant events are logged.**

---- Unsupervised learning methods.

- **Data quality and heterogeneity, missing data.**

---- Difficult to distinguish between data issues and physical anomalies.

- **New project sites with few available data.**

---- Transfer learning.

- **Assets under vastly different weather conditions.**

---- Solar projects exist in a different locations with very different weather patterns.

System Model

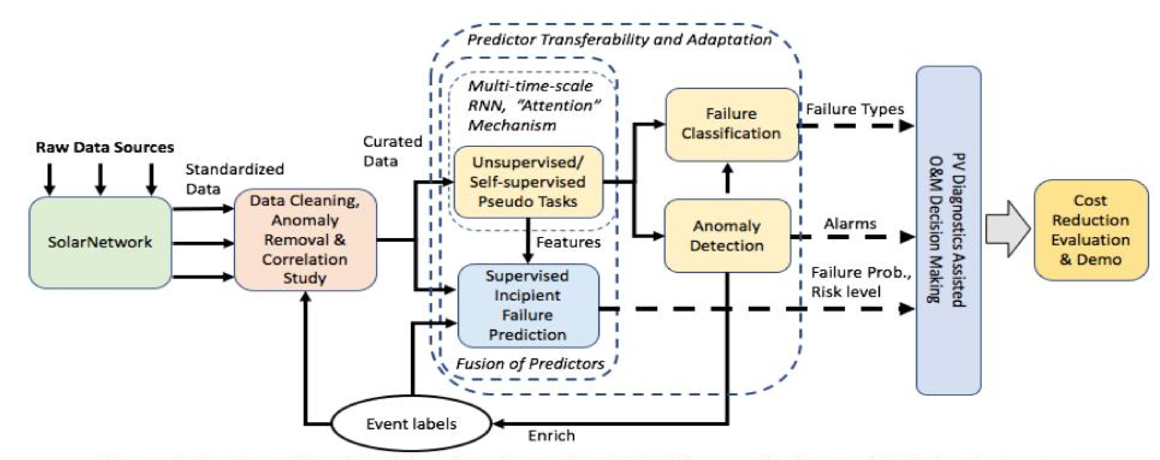


Figure 1. Pipeline of the data-driven learning-assisted PV failure predictions and O&M optimization.

- Unsupervised learning for PV anomaly detection and classification.
- Supervised learning for incipient failure prediction.
- Fusion of different prediction mechanisms.
- Predictor training for transferability and adaptation.
- Data-Driven evaluation and live demonstration.

Methods (Unsupervised Learning)

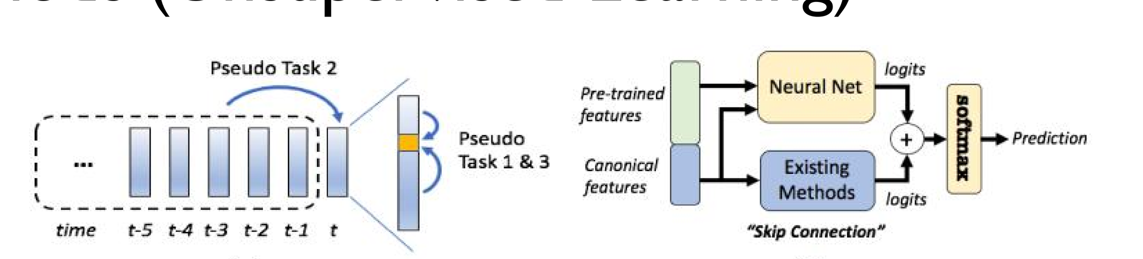


Fig. 2: (a) Unsupervised/self-supervised pseudo tasks for training PV anomaly detectors and learning features.
(b) Residual network architecture for integrating existing methods in predictor training.

- Unsupervised / Self-supervised Pseudo Tasks

  Pseudo Task 1 (cross-measurement correlations)

  Variables: voltage, power, …

  Pseudo Task 2 (temporal correlations)

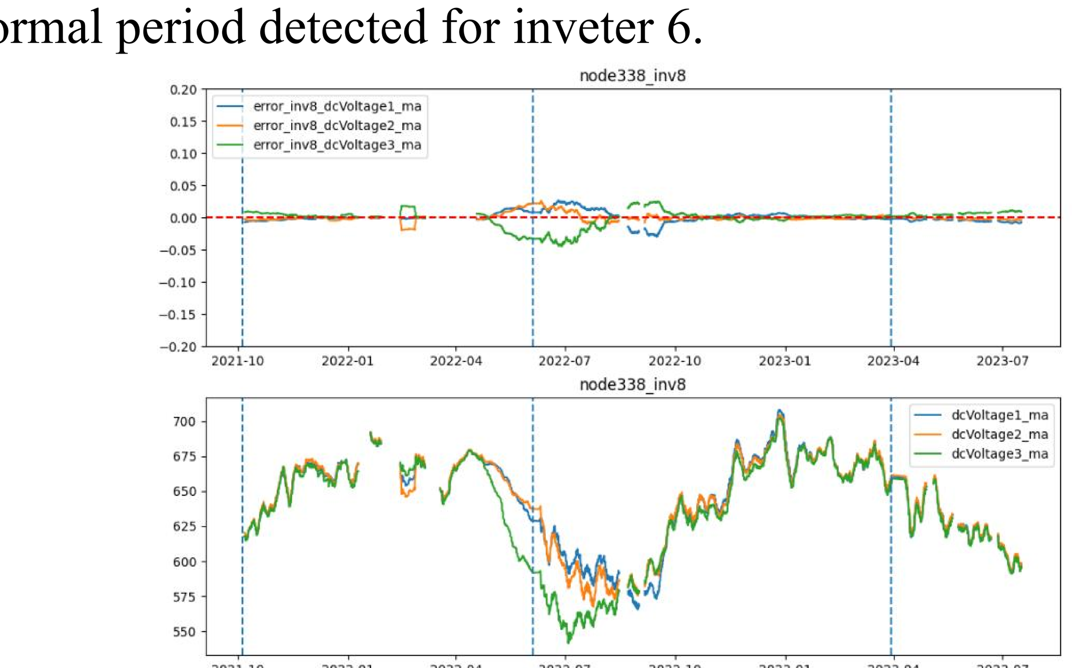
  Pseudo Task 3 (cross-system correlations) Predicting, for a single system (e.g., an inverter or a string of panels), the values of a subset of the data sources based on those of the others.

Preliminary Results

- Pseudo-tasks that train predictors to predict parts of the monitoring data based on other parts.
- If the measured value differs a lot from the predictor outputs, the likelihood that this indicates an underlying anomaly is higher.

$$Prediction\ Error\ Rate\ = \frac{Ground\ truth\ value\ -\ Predicted\ value}{Predicted\ value}$$

- Use other inveters' watts to predict a target inveter's watts.
- Abnormal period detected for inveter 6.



- Use summary variables such as average dcVoltage to predict individual variables such as dcVoltage of each string of solar panels.
- Abnormal trend detected that precedes the first recorded event.

Ongoing and Future Work

- Communicating with asset managers about the detected anomalies and enriching the events set.
- Designing unsupervised predictors that are more effective in indicating asset physical anomalies.
- Predicting incipient failure with the help of pre-trained features from different pseudo-tasks.
- Evaluating the impact of the developed predictors on reducing the O&M cost of solar PV assets, based on historical data and live demonstration with Ecogy.

Machine Learning Opportunities for Flexible Perovskite Solar Cells

Julia W. P. Hsu (jwhsu@utdallas.edu)

Department of Materials Science and Engineering, the University of Texas at Dallas, Richardson, TX, USA

Bayesian Optimization of Photonic Curing Process for Flexible Perovskite Photovoltaic Devices

W. Xu, Z. Liu, R. T. Piper, and J. W. P. Hsu

Solar Energy Mater. Solar Cells **249**, 112055 (2023) 10.1016/j.solmat.2022.112055

Goal:

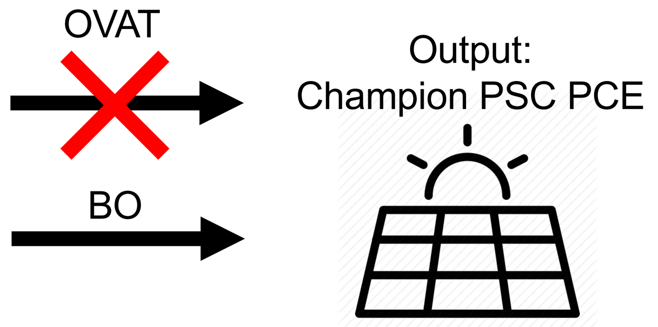
Find the photonic curing (PC) conditions and precursor formulation that produces flexible perovskite solar cells (fPSCs) with the highest power conversion efficiency (PCE).

Method:

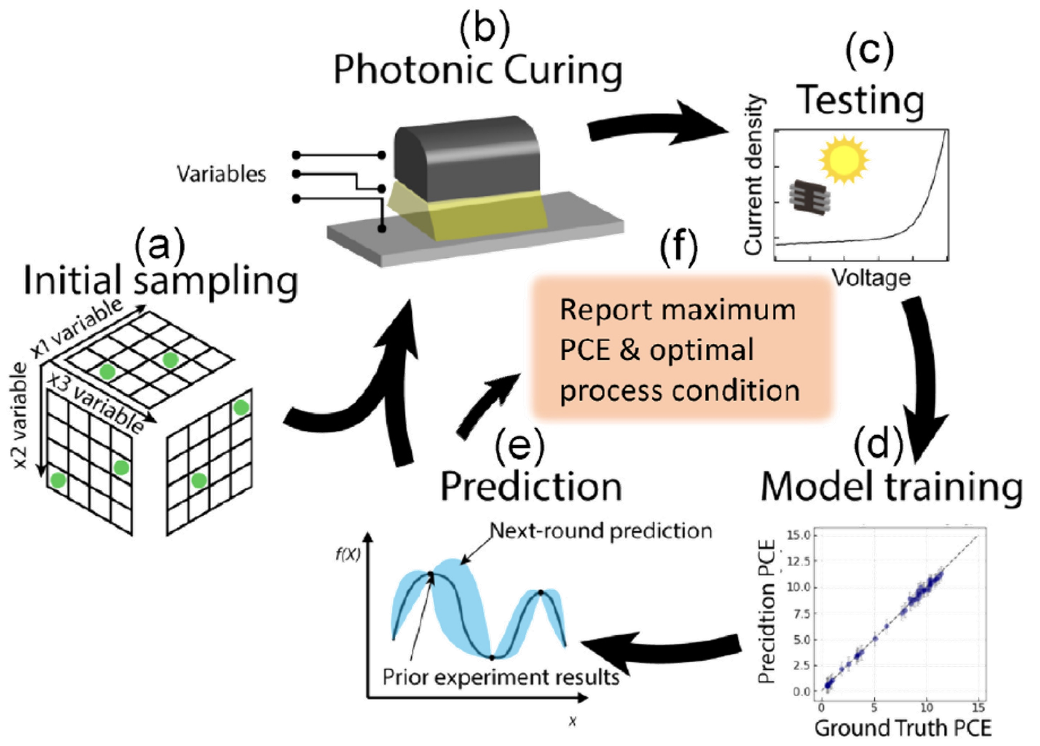
- Choose input parameters and ranges

Input variable	Range (Interval)
MAPI concentration	1.3 – 1.6 M (0.1 M)
CH_2I_2 vol%	0 – 250 µL (50 uL)
Pulse voltage	200 – 440 V (1 V)
Pulse length	1 – 20 ms (0.1 ms)

> 1 million combinations

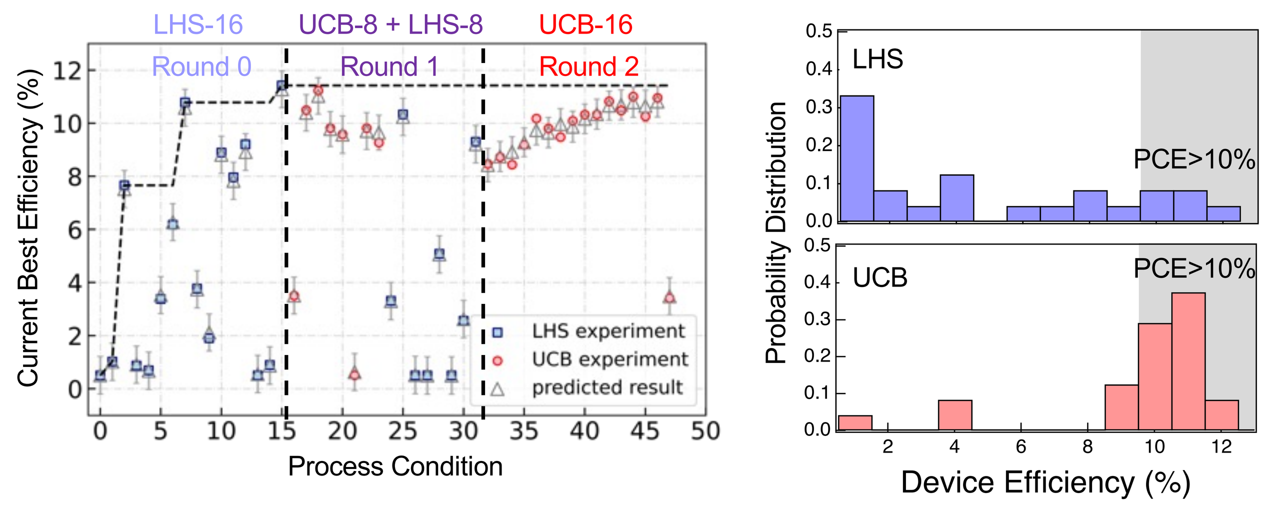


Bayesian Optimization (BO) Workflow



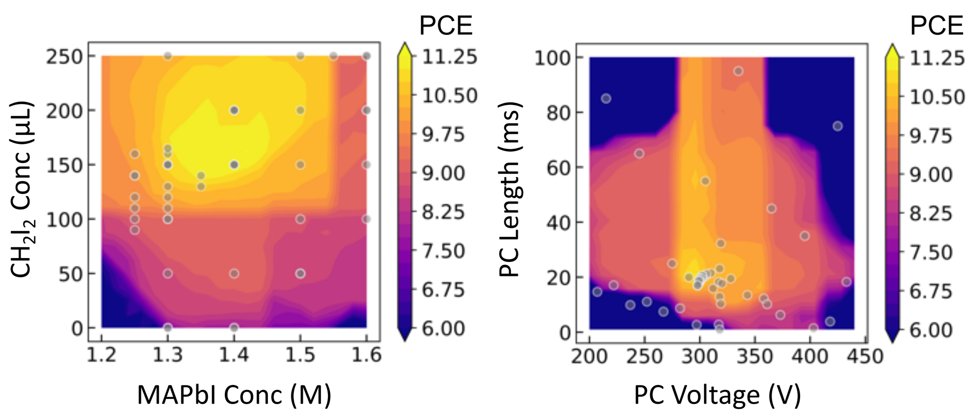
a) Adopt Latin Hypercube Sampling (LHS) to pick 16 initial conditions.
b) Fabricate fPSCs on Willow Glass® with the input parameters selected by LHS or BO
c) Test fPSCs at standard AM 1.5 G condition.
d) Use fPSC PCEs to build Gaussian Process Regression model.
e) Apply Upper Confidence Bound (UCB) as the acquisition policy to select the next round of 16 input parameters.
f) Continue iteration until the input parameters converge or resources are exhausted.

Device Optimization Results



- With the limited experimental budget of 48 conditions, we achieved a champion PCE of 11.42%.
- BO has a higher success rate finding the conditions that produce PCE > 10% than randomly picking.

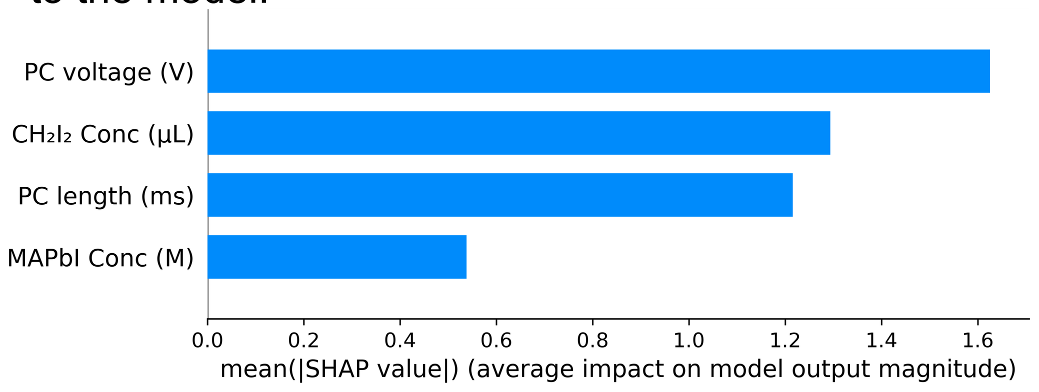
Model (GPR + RF ensemble)



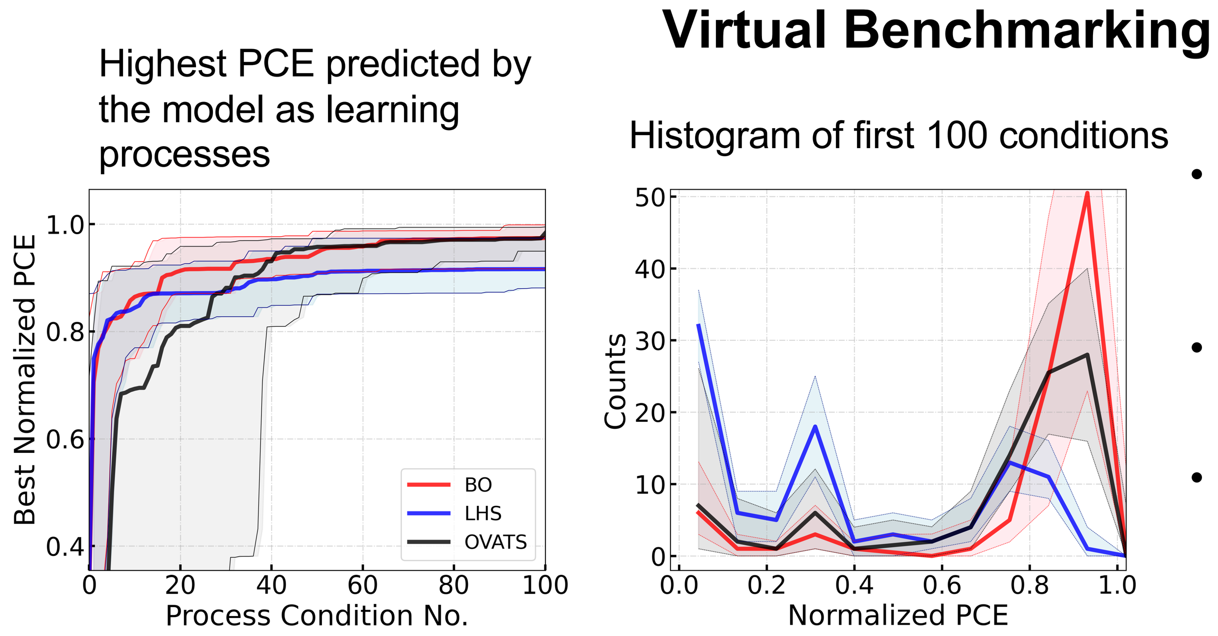
BO is good at building a model to describe a black-box function. But we want to get some insight into the results.

Which Input Parameter Matters the Most?

We apply the SHarply Additive exPlanation (SHAP) framework to analyze the contribution of each variable to the model.



Virtual Benchmarking



- Using the fully trained model shown above, we perform benchmarking with three different optimization methods.
- Improving Device PCE:
  BO > LHS > OVAT
- Success rate finding > 0.9 normalized PCE:
  BO > OVAT > LHS

All three approaches have been randomly iterated for 100 times to get the 5 – 95% confidence intervals.

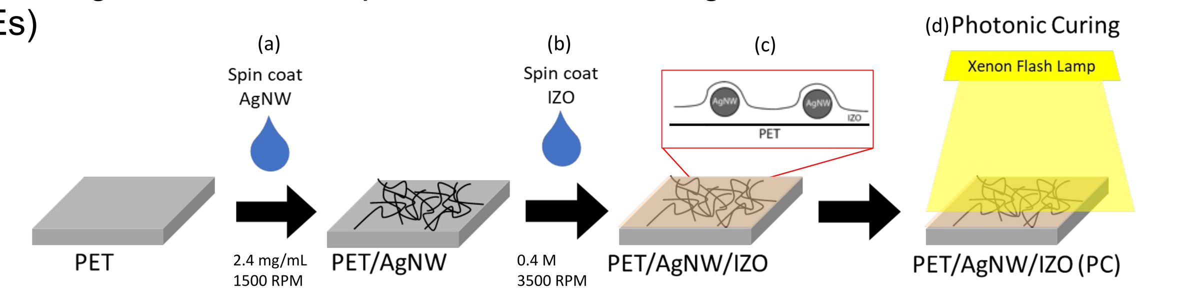
Multiobjective Optimization of Silver-Nanowire Deposition for Flexible Transparent Conducting Electrodes

M. Lee, R. T. Piper, B. Bhandari, and J. W. P. Hsu

ACS Appl. Nano Mater. **6**, 17364 (2023) 10.1021/acsanm.3c03599

Making TCEs on PET substrates is challenging:

- Can't exceed 120 ℃ during processing → ITO properties on PET are inferior to those on glass
- Must ***simultaneously*** have high transmittance and high conductivity, two competing physical quantities
- Combine Ag metal bus bars printed on PET with AgNWs + IZO to fabricate flexible transparent conducting electrodes (fTCEs)



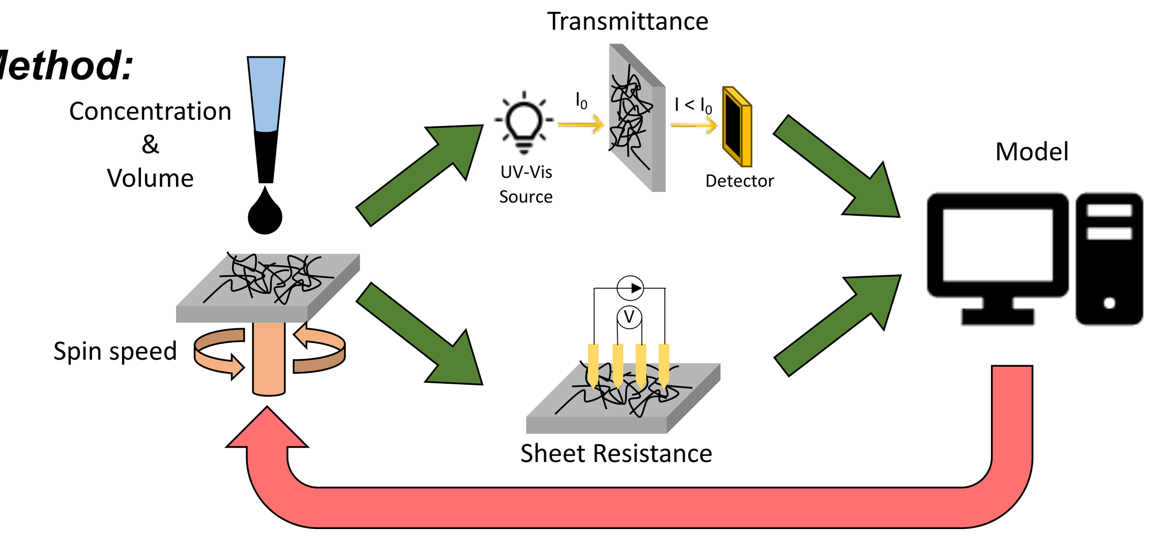
R. T. Piper, W. Xu, and J. W. P. Hsu, "Silver Nanowire-Indium Zinc Oxide Composite Flexible Transparent Conducting Electrodes Made by Spin-coating and Photonic Curing," MRS Adv. **8**, 177-182 (2023) https://doi.org/10.1557/s43580-022-00478-x

- Transmittance (T) and sheet resistance (R_{sq}) are determined by the AgNW layer

Goal:

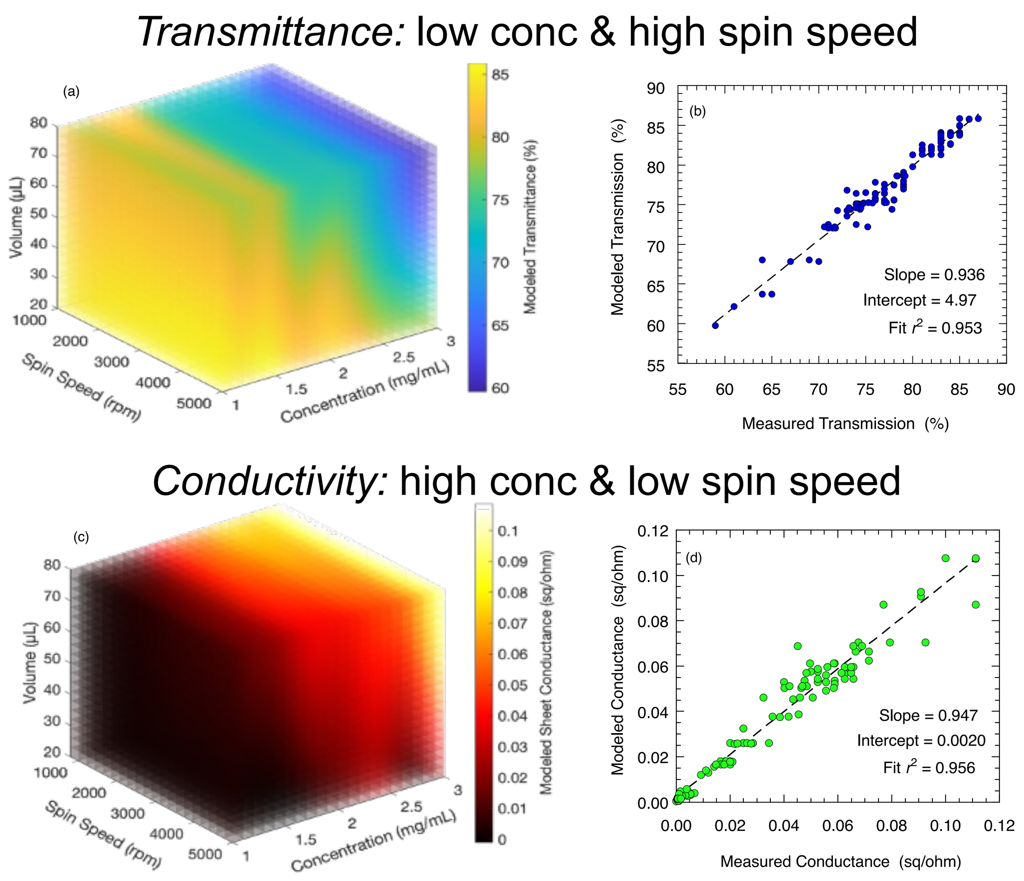
Find the AgNW concentration, spin speed, and dispense volume that produces fTCEs with ***both*** T_{avg} ***≥ 75% and*** R_{sq} ***≤ 15 Ω/sq***

Method:



a) Adopt Latin Hypercube Sampling (LHS) to pick 36 initial conditions.
b) Spin-coat AgNWs on PET and measure their transmittance and sheet resistance.
c) Input the results to build 2 GPR models: one for T and one for G_{sq}.
d) Use the models to analyze scalar figure of merit (FOMs) or build Pareto front.
e) Perform additional experiments to achieve desired fTCE specs.

GPR Model + Scalar FOMs



TCE field often uses scalar FOMs as the metric. Two commonly used ones are:

$$FOM_{unitless} = \frac{188.5\Omega}{R_{sq}(T^{-1/2} - 1)}$$

$$FOM_{T10} = \frac{T^{10}}{R_{sq}}$$

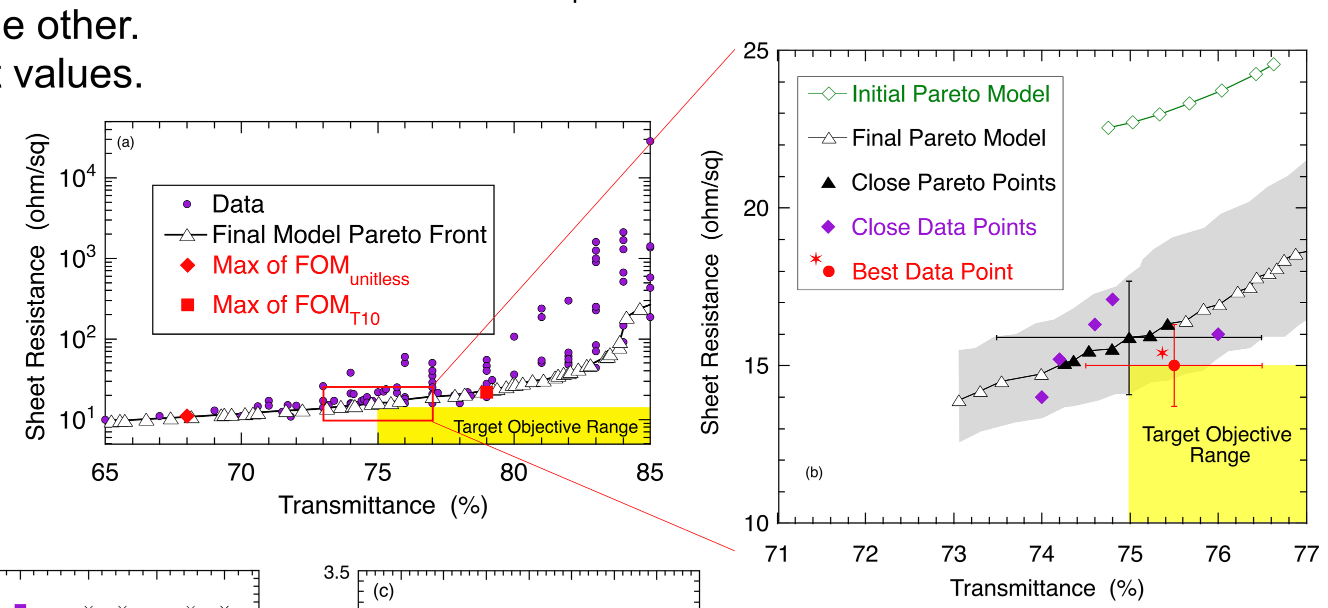
	AgNW Concentration (mg/mL)	Spin Speed (rpm)	Dispense Volume (µL)	T (%)	R_{sq} (W/sq)	Max FOM value
$FOM_{unitless}$	3.0	800	40	68 ± 2	11 ± 1	82 ± 8
FOM_{T10} (1/Ω)	2.3	1200	28	79 ± 2	22 ± 3	(4.2 ± 1.0) x 10^{-3}

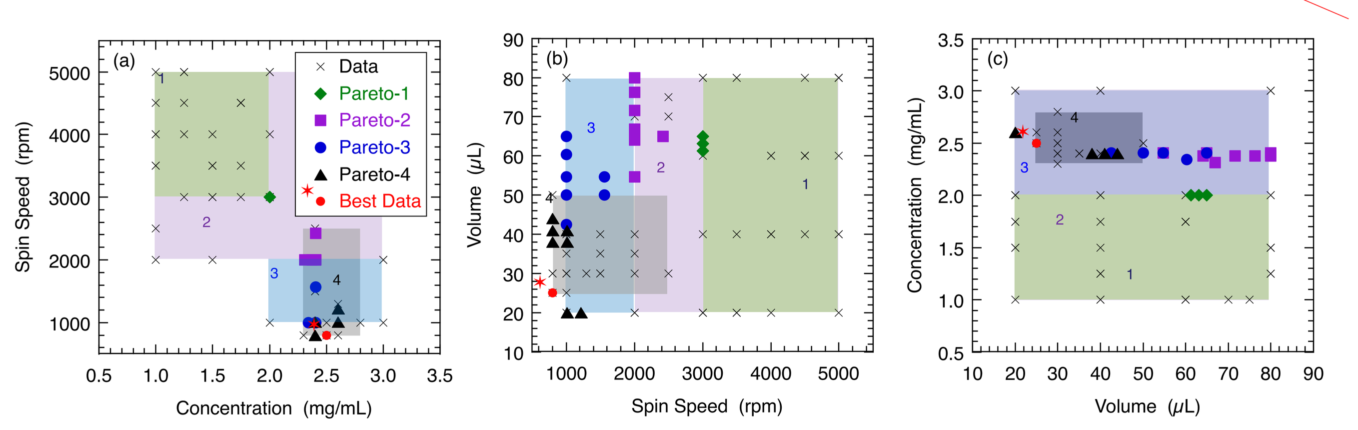
Optimizing processing conditions using either FOM would produce fTCEs that satisfy the desired spec!

Pareto Analysis

Alternative method:

- Construct a ***Pareto front***, the boundary in a modeled T vs R_{sq} plot on which neither objective can be improved without degrading the other.
- Use it to select improved input values.
- T ≥ 75% & R_{sq} ≤ 15 Ω/sq is at the edge of possible, as revealed by Pareto analysis.
- An ML/Pareto analysis is applicable to optimization problems when two or more independent criteria must be simultaneously satisfied.





Further improvement will need to come from other factors, e.g., different AgNWs, coating methods, or photonic curing conditions.

Funding Acknowledgement

NSF CMMI-2135203 and DE-EE0009518. J.W.P.H. acknowledges the Texas Instrument Chair in Nanoelectronics and Simons Foundation Pivot Fellowship.

Machine Learning-based Analysis of Solar Cell Degradation from Femtoseconds to Gigaseconds

Gergely Zimányi, A. Diggs, Z. Zhao, D. Unruh, C. Hansen, A. Goga, Z. Crawford

Department of Physics, University of California, Davis CA

Mariana Bertoni, Stephen Goodnick, R. Meidanshahi

School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe AZ

1. Project Description, Main Products & Results

Our project is focused on determining the structural dynamics of Si Heterojunction and TOPCon solar cells starting at femtoseconds, in order to predict their degradation up to gigaseconds.

We used Machine Learning to create SolDeg, the most precise and efficient Si structural dynamics simulator

With SolDeg, we discovered that the degradation of HJ cells is driven by H drifting away from the cSi/aSi interface

We proposed forming a reversed Si gradient at the cSi/aSi interface to reduce degradation by up to 80%. This idea has been adopted by Trina, Meyer-Burger and Huasun since

We discovered that TOPCon cells degrade by forming pinholes. We found that pinhole formation is reversible

We developed a Machine Learning algorithm that detects electronic defect formation in solar cells from structural information alone

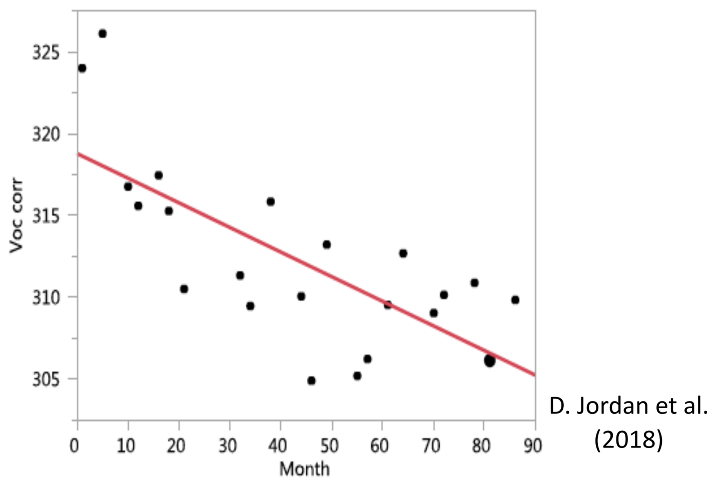
Supported by SETO DE-EE008979 and DE-EE009835, and by the NERSC supercomputer center

2. Motivation, SolDeg platform

1. Si Heterojunction cells hold Si world record efficiency of 26.7%

2. Efficiency degradation of 1%/yr reported, twice the usual. 0.5%/yr, was attributed to Voc.

3. Eliminating 0.5%/yr degradation equivalent to increasing the efficiency by close to 2% in terms of LCOE, based on the System Advisor Model



4. Bertoni: Defect generation at interface drives degradation in c-Si/a-Si stacks.

We developed the SolDeg platform to analyze the formation of defects at the c-Si/a-Si interface. This requires

1. Connecting extreme time scales from femtoseconds to gigaseconds (30 years)

2. Simulation of large number of large samples with extreme precision

1. Create c-Si/a-Si stacks

2. Generate initial-final state pairs of H (1) breaking out of a Si-H bond, or (2) diffusing between interstitial locations

3. Check that the induced defects are neutral electronic defects

4. Determine the energy barriers that control the generation of these electronic defects

5. Determine distribution of thousands of these barriers

6. Determine the defect generation dynamics from the energy barrier distribution, analytically and/or numerically

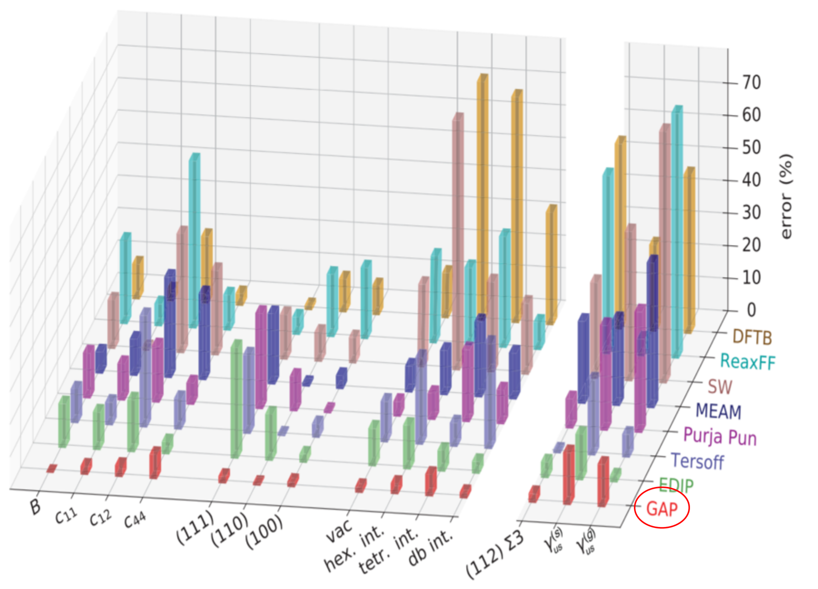
D. Unruh, et al., *ACS Applied Materials & Interfaces* **13**, 32424 (2021)

3. Precise Structures by Machine Learning-developed Potentials

1. Create c-Si/a-Si stacks using Molecular Dynamics with Machine-Learning (ML)-developed Si-Si potential: Gaussian Approximation Potential Si GAP. -- Timestep: femtoseconds

We constructed Si-H GAP potential via 27 rounds of training on wide variety of structures. – with G. Csanyi

GAP reproduces DFT much better than other interatomic potentials.



Big problem: No Si-H GAP potential

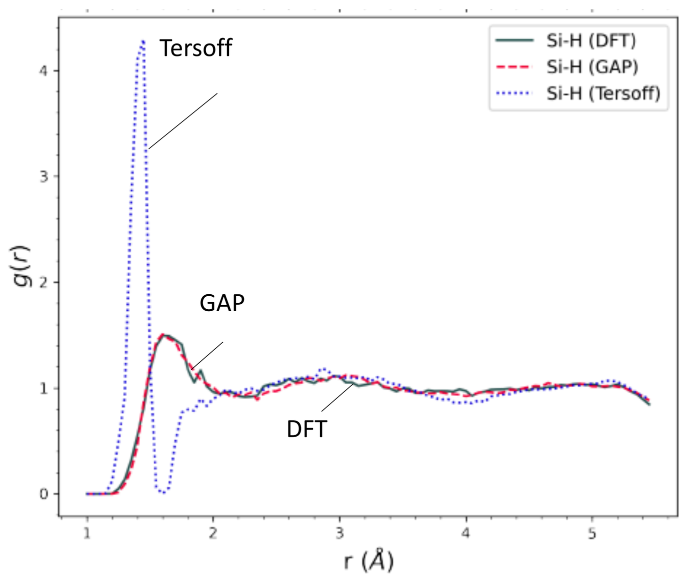
Iteration	Structure Type
1	Optimized structures (all phases)
2	Optimized structures (all phases)
3	Low T anneal of a-Si:H
4	High T anneal of liq-Si:H
5	High T anneal of liq-Si:H
6	Med T anneal (1100K) of a-Si:H
7	Heating a-Si:H from 500K to 800K at 10^{13} K/s
8	Heating a-Si:H from 800K to 1100K at 10^{13} K/s
9	Heating a-Si:H from 1100K to 1400K at 10^{13} K/s
10	Heating a-Si:H from 1100K to 1400K at 10^{13} K/s
11	Heating a-Si:H from 800K to 1400K at 10^{12} K/s
12	Added new a-Si:H structures
13	Add new a-Si:H structures
14	Added c-Si/a-Si:H interface structures
15	Added c-Si/a-Si:H interface structures
16	Added new c-Si divacancy structures
17	Added new liq-Si:H structures
18	Added new c-Si vacancy structures
19	Added new c-Si interstitial structures
20	Low T anneal of c-Si/a-Si:H interface structures
21	Optimization of c-Si/a-Si:H interface structures
22	NPT high T anneal of liq-Si:H structures
23	NVT high T anneal of liq-Si:H structures
24	Quenching liq-Si:H from 2000K to 1500K at 10^{13} K/s
25	Annealing quenched liq-Si:H structures at 1500K
26	Quenching liq-Si:H from 1500K to 1400K at 10^{12} K/s
27	Added hydrogen passivated c-Si surface (100) and c-Si surface (111) structures

D. Unruh, et al., *Phys. Rev. Mat.* **6**, 065603 (2022)

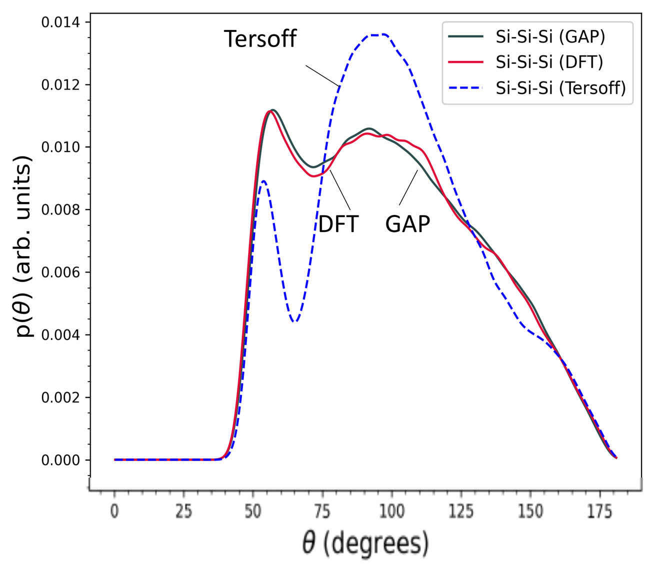
4. 27 Rounds of ML-training Created Precise, Efficient Si-H GAP

1. Energies: reproduces DFT within 4 meV/atom

2. Radial correlation function



3. Bond angle distribution function



4. Accuracy of reproducing DFT forces improved by 35%.

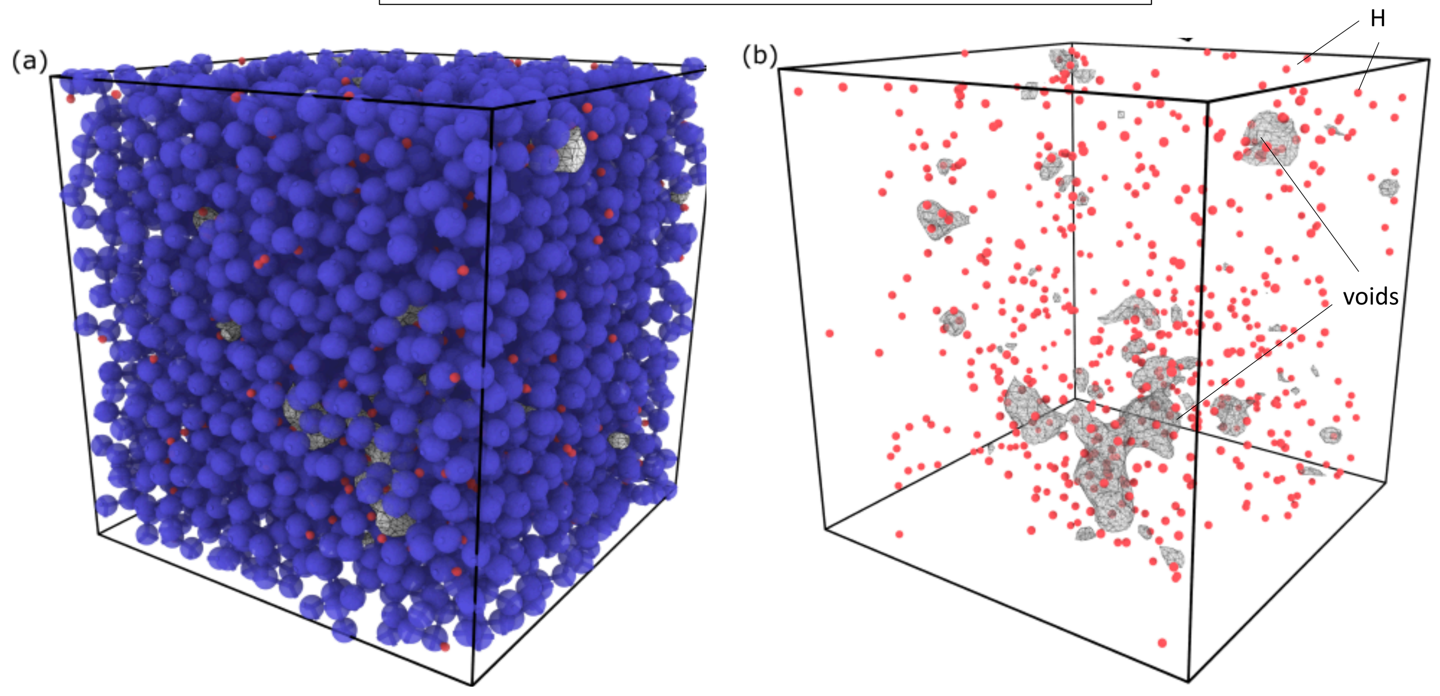
Overall, our Machine-Learning-trained Si-H GAP is the most accurate interatomic potential for Molecular Dynamic simulations of hydrogenated Si solar cells.

5. Si-H GAP: Reaching the Unreachable in Size and Precision

Run time scaling: GAP O(N); DFT O(N³)

DFT can simulate 400-500 atoms

Si:H GAP simulation: 4,096 Si and 558 H atoms

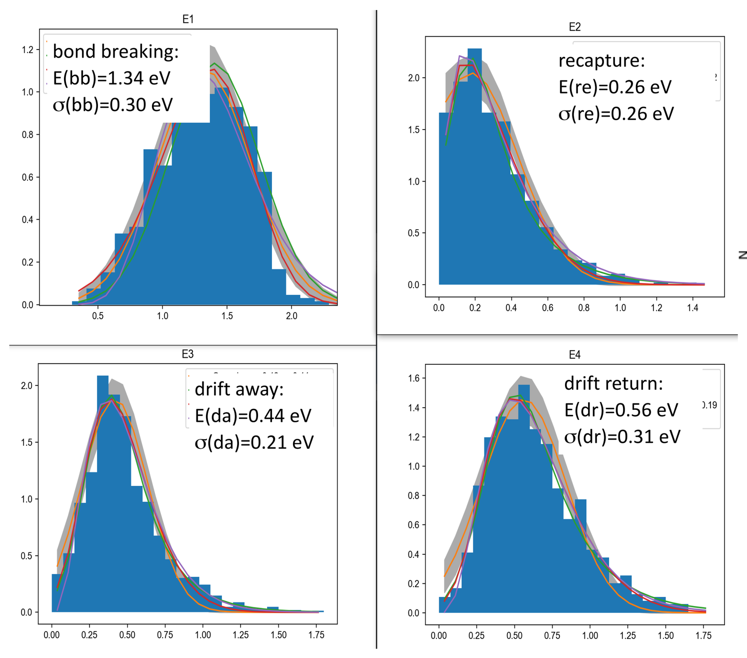


Si:H GAP Molecular Dynamics simulations can reach unparalleled sizes and number of realizations

6. Defect Generation with SolDeg energies reproduces data well

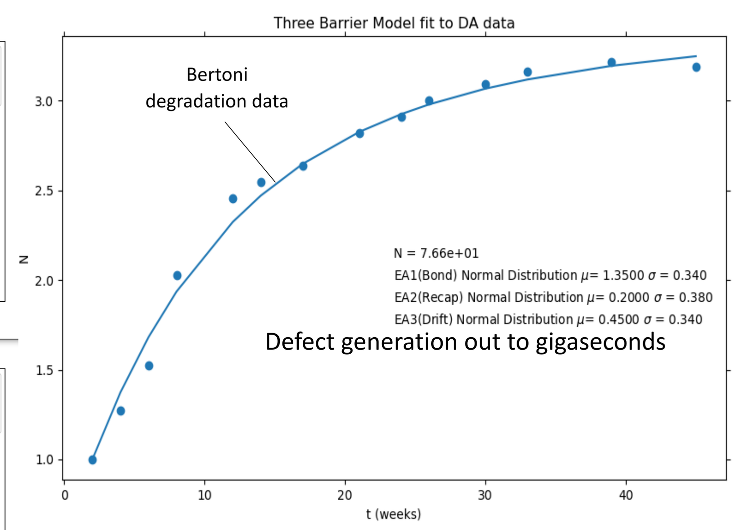
Determined the barriers with the Nudged Elastic Band method

Determined the barriers in more than 1,500 instances, and created distribution of barriers for all four relevant processes



Modelled N(t) with SolDeg-simulated energies, averaged over SolDeg-determined distributions (line). Not a fit.

Correspondence with measured values (dots) is remarkable.



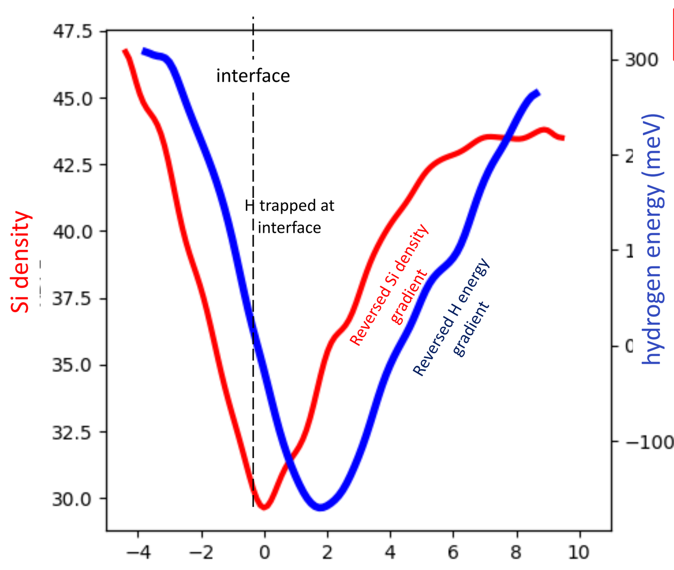
A. Diggs, et al., *Nature Comm. Mat.* **4**, 24 (2023)

7. How to stop H-driven degradation? Reverse Si density gradient!

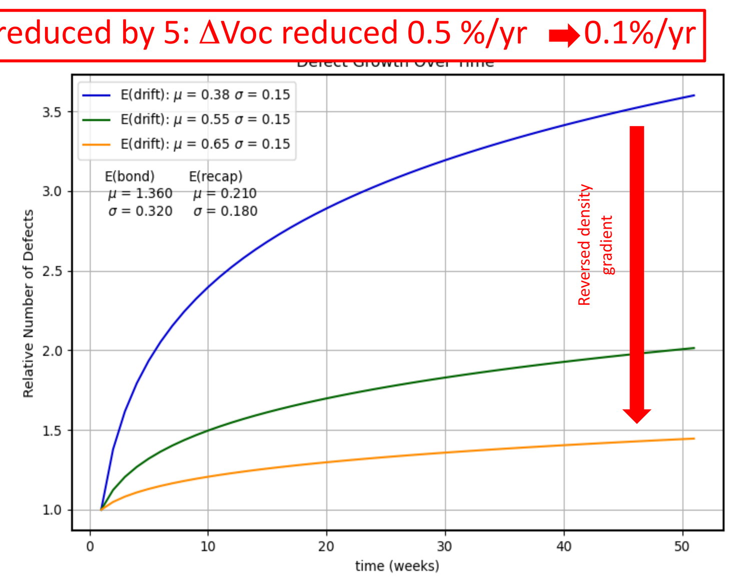
The H energy gradient was created by the Si density gradient.

Idea: Reverse the H energy gradient by reversing the Si density gradient! This will create a density minimum at interface that traps the H and stabilizes the passivation.

We simulated reverse gradient HJ cells. We found that the H energy gradient indeed has changed into a minimum!



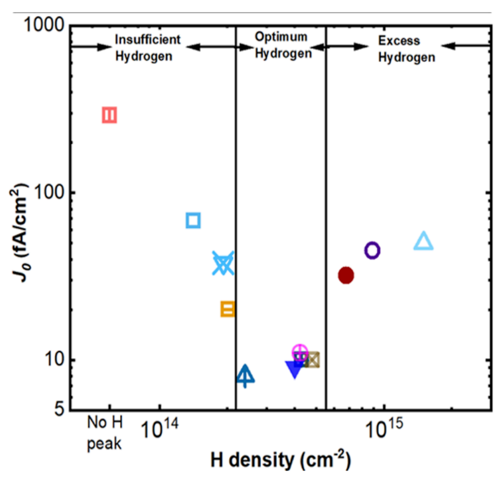
ΔN reduced by 5: ΔVoc reduced 0.5 %/yr ➡0.1%/yr

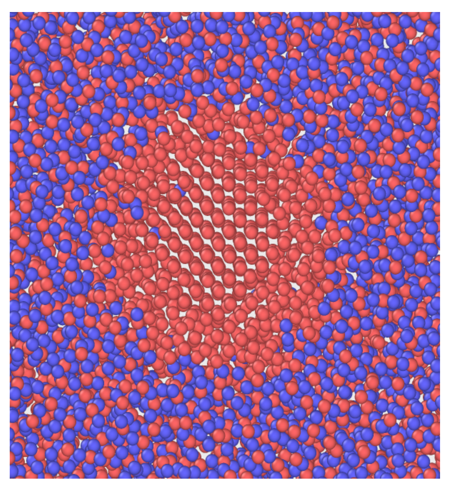


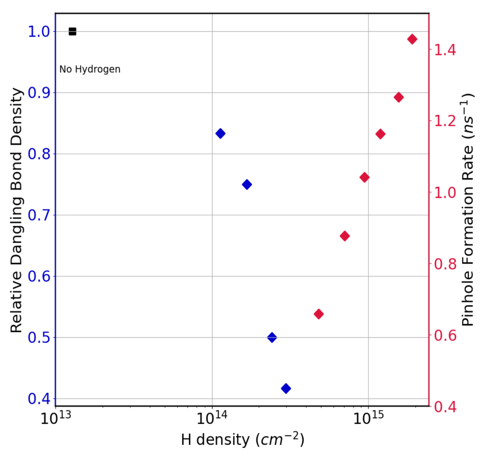
Adopted and used by Trina, Meyer Burger and Huasun in 2023

8. Hydrogen Massively Impacts TOPCon Cell Performance & Degradation

Adding Hydrogen to TOPCon cells showed dramatic improvement of J0 at an optimal concentration. Reason unknown







We found that pinholes form in the SiOx layer above a critical H concentration [H].

At low [H], hydrogen increasingly passivates defects at the interface: J0 decreases.

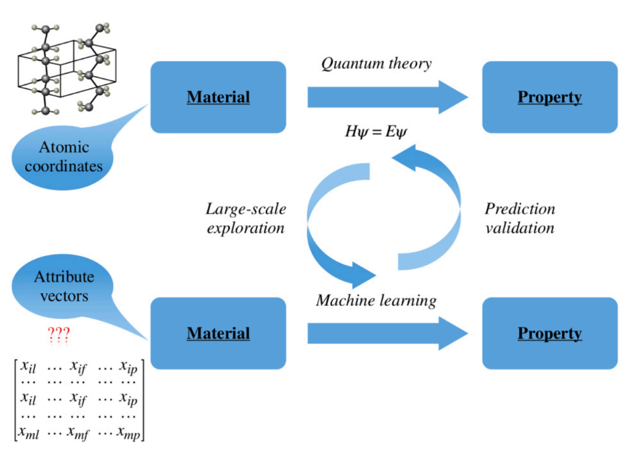
At high [H], H induces nucleation of pinholes that act as recombination centers at the c-Si/SiOx interface: J0 grows.

We found that pinhole formation was reversible.

A. Diggs , et al., *ACS Applied Materials & Interfaces* **2023 submitted**

9. Predicting Electronic Defects from Structural Info by Machine Learning

DFT is much slower and computer intensive than Molecular Dynamics (MD) that is used for Structural Simulations



	Training/validation accuracy		
	cSi	cSi/aSi	cSi/aSi:H
Bond lengths only	68%/68%	61%/61%	60%/60%
3-body angles only	97%/94%	85%/78%	77%/77%
Bond lengths + subset of 3-body angles	98%/95%	75%/75%	76%/76%
Bond lengths + 3-body angles	98%/97%	90%/81%	78%/78%
Bond lengths + 3-body angles + 4-body angles	99%/97%	93%/83%	80%/78%

We ML-trained algorithms on cSi; cSi/aSi interfaces; and cSi/aSi:H structures to find the structures that corresponded to electronic defects

We achieved 78%-97% accuracy to predict quantum electronic defect properties from structural info alone

Z. Zhao , et al., *Phys. Rev. B* **2023 submitted**

Sandia National Laboratories

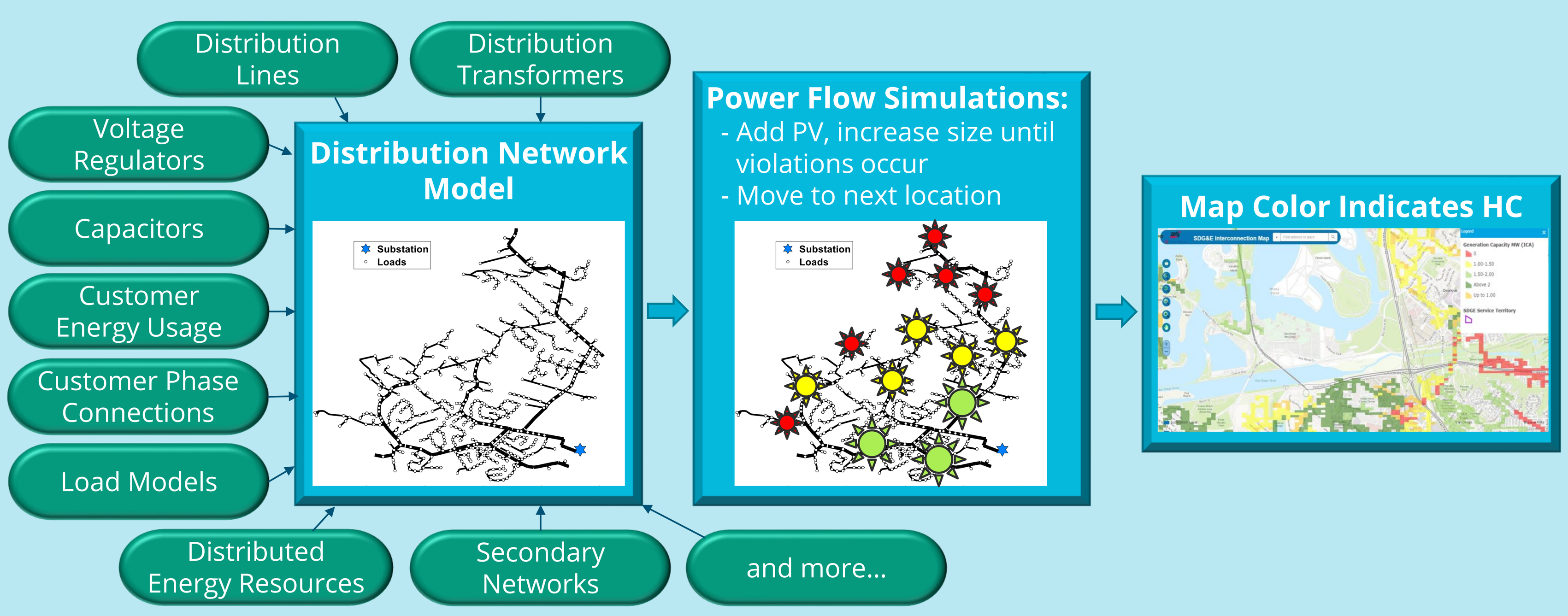
Model-free Hosting Capacity Analysis (MoHCA)

Award # DE-EE0038426: "Smart Meter Data: A Gateway for Reducing Solar Soft Costs with Model-Free Hosting Capacity Maps"

Principal Investigator: Matthew Reno

Introduction:

- Solar photovoltaic (PV) system costs are now dominated by non-hardware or "soft" costs
  - E.g., customer acquisition, permitting, and interconnection costs
- Public-facing **hosting capacity (HC)** maps have been key factors to reducing solar soft costs
  - HC maps are visual representations of the maximum amount of solar that can be installed at various locations without adverse effects on the distribution network (i.e., locational HC)
- However, **hosting capacity analysis (HCA)** must be performed to generate the data for these maps



- Conventional **model-based** HCA methods are time-consuming and computationally intensive due to iterative power flow analyses, and the accuracy of the results depend on the accuracy of the distribution system model, which are complex and prone to error [1]

[1] J. Azzolini, S. Talkington, M. J. Reno, S. Grijalva, L. Blakely, and D. Pinney, "Improving Behind-the-Meter PV Impact Studies with Data-Driven Modeling and Analysis", *IEEE Photovoltaic Specialists Conference* (PVSC), 2022.

Challenges:

- Validating the accuracy of the algorithms still requires grid models and model-based analyses
- Smart meter reactive power measurements are not widespread

Best Practices:

- Developing consistent and generalized data cleaning functions can save a lot of time and effort
- Testing on a wide variety of datasets and feeders can highlight edge cases where the algorithms may struggle

Key Takeaways:

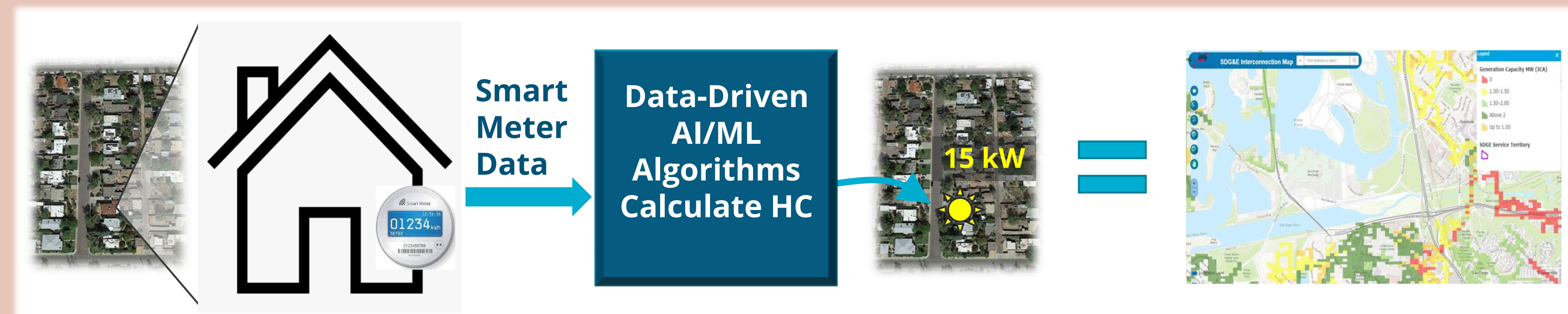
- Our data-driven algorithms that leverage AI/ML have been shown to be highly accurate compared to model-based analyses
- NN-based methods can struggle with extrapolation outside the range of training data, but incorporating physics-based methods into the learning process can mitigate that risk

Future Work:

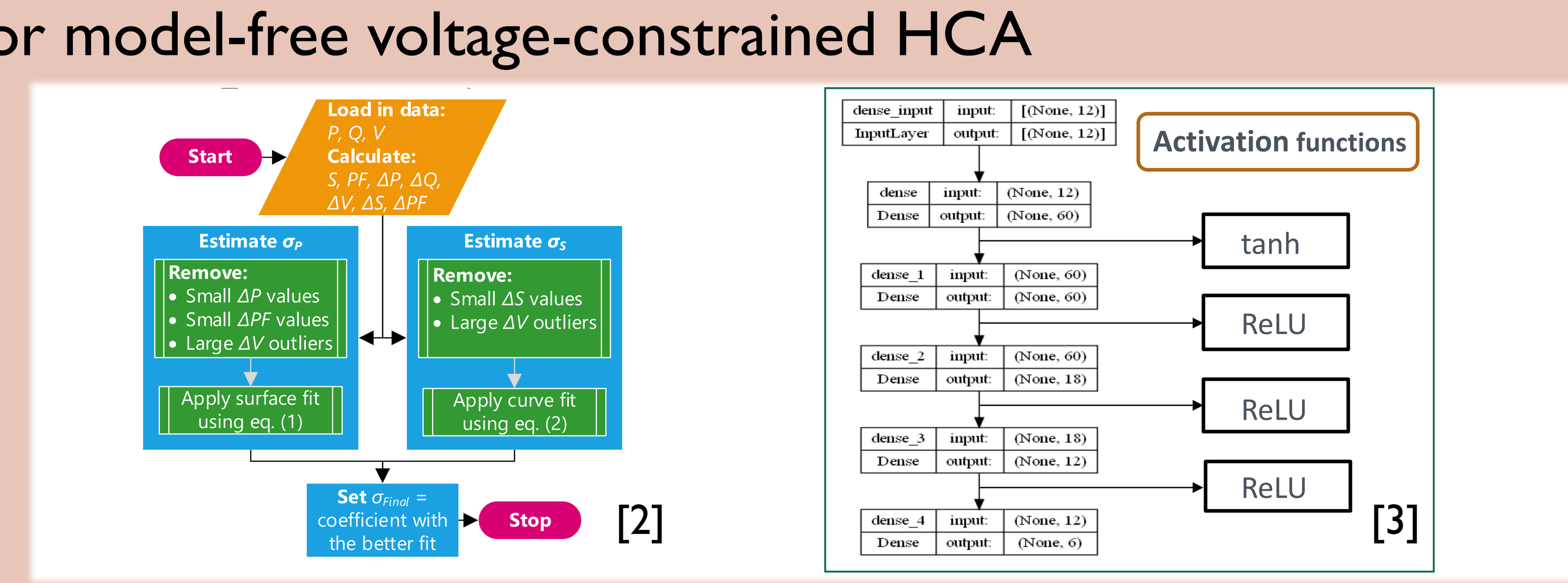
- Fine-tuning the developed algorithms and combining them into a comprehensive open-source tool
- Develop additional functionality that can evaluate the impacts of advanced inverter functions

How is AI/ML used in this project?

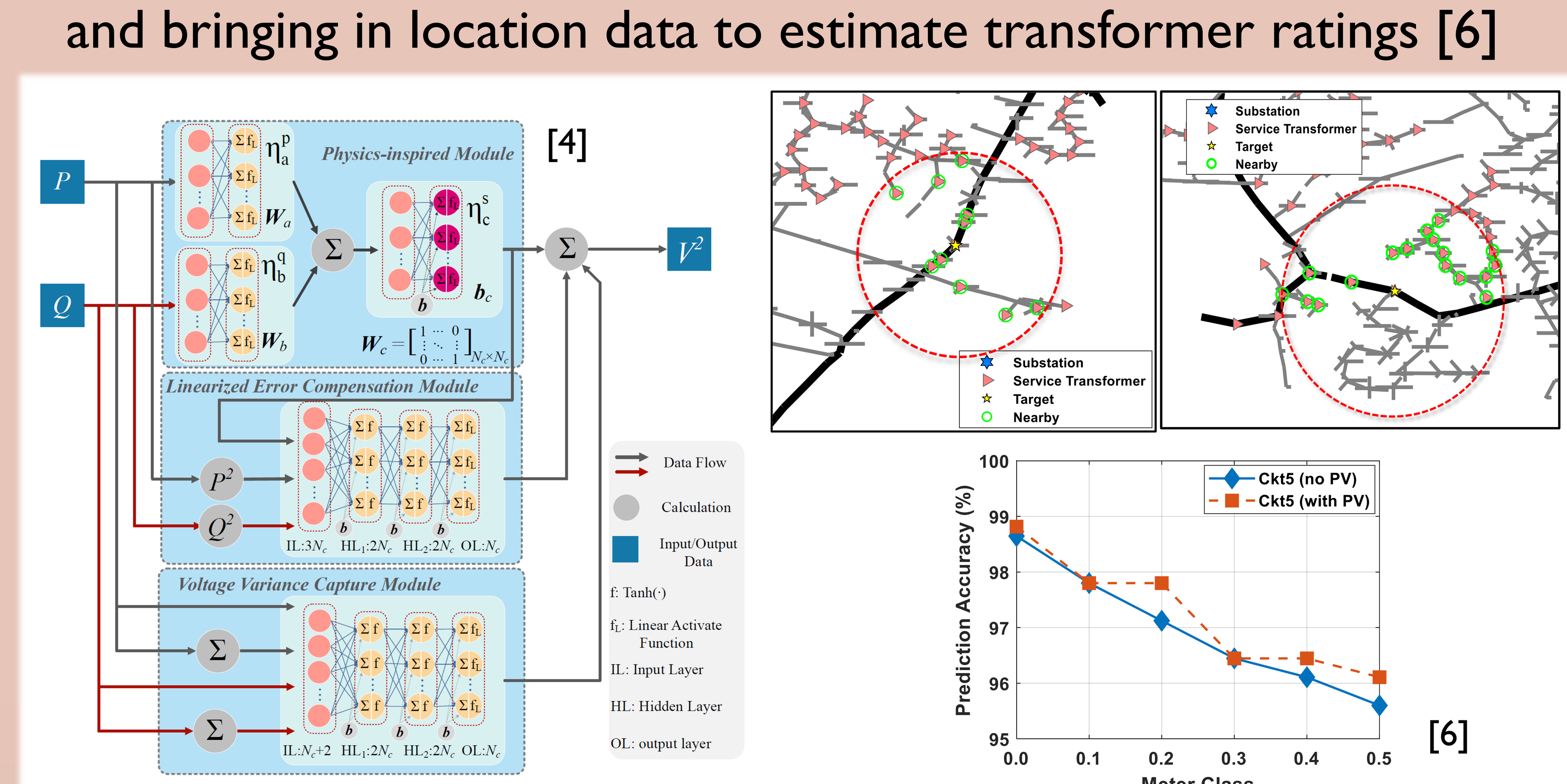
- This project is developing entirely data-driven algorithms to calculate locational PV HC by leveraging widespread smart meter deployments and applying data analytics and AI/ML techniques



- Various AI/ML algorithms are utilized in this project to extract actionable information out of smart meter measurements, e.g.:
  - What is the relationship between a customer's power usage and their voltage?
  - Can we use AI/ML to learn those relationships and extrapolate what the grid impacts would be if that customer installed rooftop PV?
  - What can we infer about the available capacity of the transformer serving a group of customers?
- Regression-based algorithm [2] and a deep neural network (DNN) algorithm [3] to learn the voltage-power relationships for model-free voltage-constrained HCA



- Improving the extrapolation capabilities of the DNN approach by implementing a physics-inspired neural network (PINN) that incorporates the power flow equations in the learning [4]
- Extracting customer-transformer information and groupings [5] and bringing in location data to estimate transformer ratings [6]



- Applying the Adaptive Boosting algorithm using Random Forests when the quality or quantity of smart meter data is limited [7]

[2] J. A. Azzolini, M. J. Reno, J. Yusuf, S. Talkington, and S. Grijalva, "Calculating PV Hosting Capacity in Low-Voltage Secondary Networks Using Only Smart Meter Data," in IEEE Innovative Smart Grid Technologies (ISGT) North America, 2023. doi: 10.1109/ISGT51731.2023.10066372.
[3] J. Yusuf, J. A. Azzolini, and M. J. Reno, "Predicting Voltage Changes of Low-Voltage Secondary Networks Using Deep Neural Networks," in *IEEE Power and Energy Conference at Illinois* (PECI), 2023.
[4] L. Liu, N. Shi, D. Wang, Z. Ma, Z. Wang, M. J. Reno, J. A. Azzolini, "Voltage Calculations in Secondary Distribution Networks via Physics-Inspired Neural Network Using Smart Meter Data" *IEEE Transactions on Power Systems*, 2023. *Under Review*.
[5] L. Blakely and M. J. Reno, "Identification and Correction of Errors in Pairing AMI Meters and Transformers," IEEE Power and Energy Conference at Illinois (PECI), 2021.
[6] J. A. Azzolini, M. J. Reno, J. Yusuf, "A Model-free Approach for Estimating Service Transformer Capacity Using Residential Smart Meter Data," *IEEE Journal of Photovoltaics,* 2023. Under Review.
[7] J. Yusuf, J. A. Azzolini, M. J. Reno, "PV Hosting Capacity Estimation in Low-Voltage Secondary Networks Using Statistical Properties of AMI Data," *IEEE Innovative Smart Grid Technologies Latin America* (ISGT-LA), 2023.

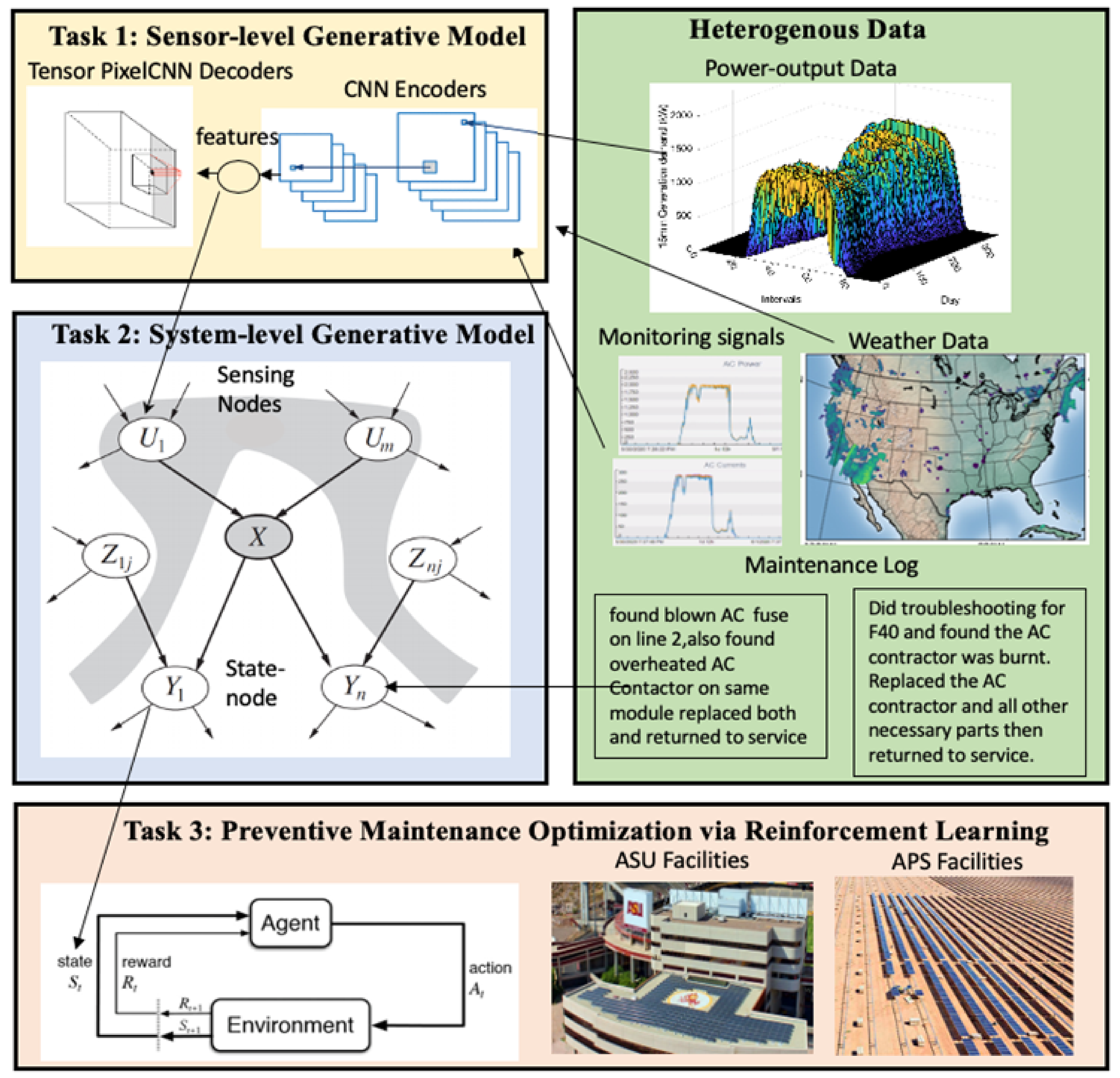
Photovoltaic Plant Predictive Maintenance Optimization under Uncertainties Using Probabilistic Information

Jiuyun Hu[1], Jiayu Huang[1], Nan Xu[1], Qiongfang Zhang[1], Yongming Liu[1], Hao Yan[1,*]

[1] Arizona State University, [*] haoyan@asu.edu

Introduction & Challenge

Data	Objective	Challenge
• Solar panel output	• Failure detection	• High dimensional data
• Solar energy monitoring	• Failure prediction	• Complex data structure
• Weather	• Maintainance schedule	• Collaboration of tasks

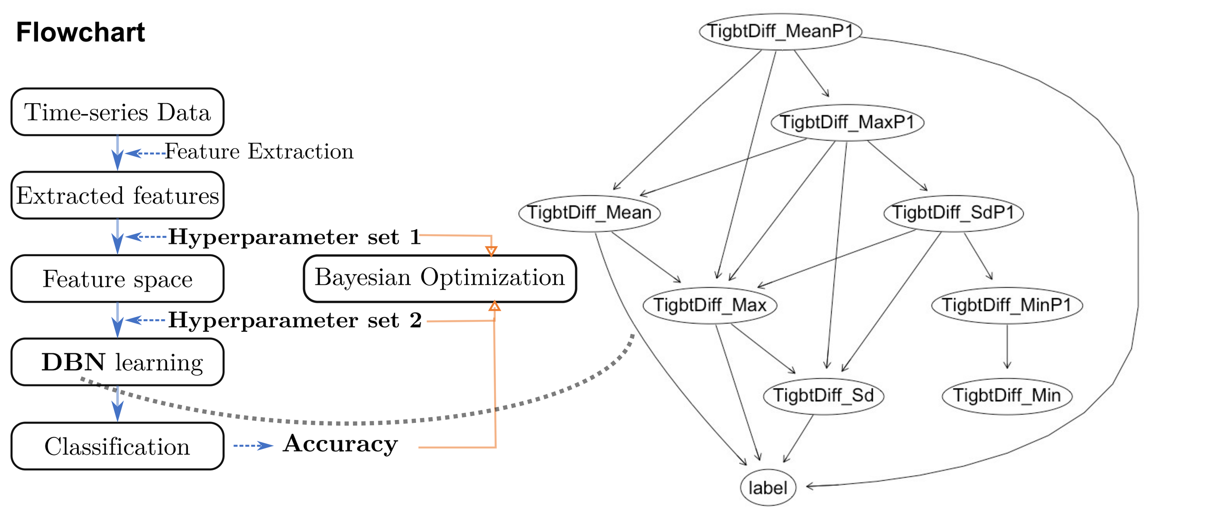


AI/ML Methods & Best Practice

- AI/ML can be used to automatically detect and predict the failure/events of the solar PV inverters/modules from the real-time sensing data
- CNN based method for failure detection and prediction
- Tensor reconstruction and manifold learning for unsupervised feature extraction

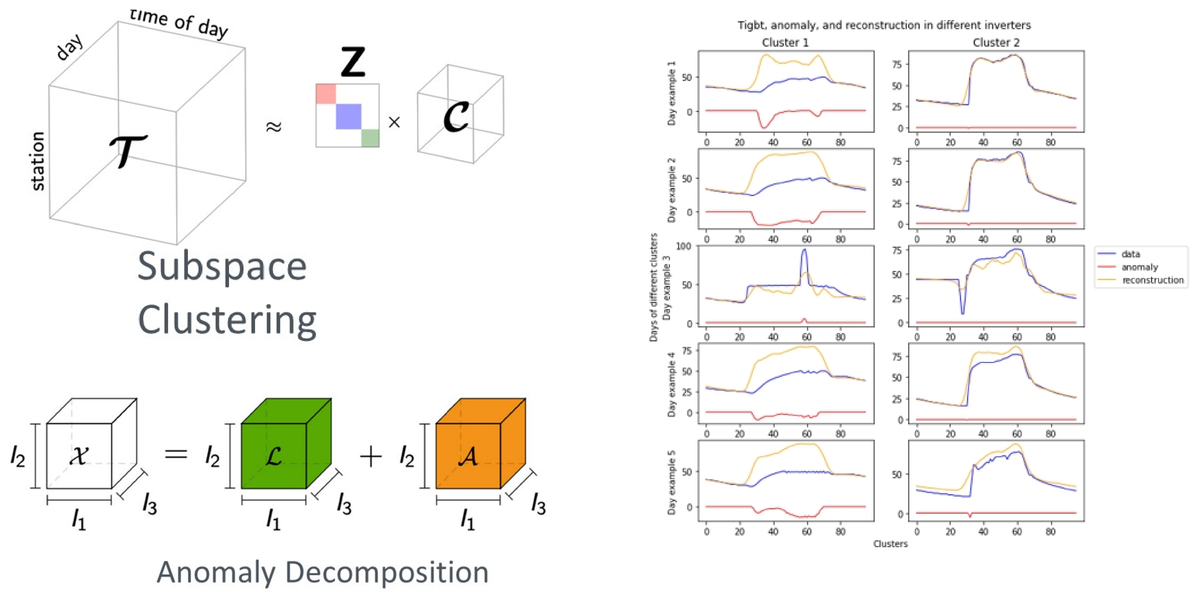
Bayesian Dynamic Network Fault Detection

- Data: Processed sensor time series data.
Sensitivity	Specificity	Precision	F1 Score
0.8122	0.8672	0.5555	0.6597



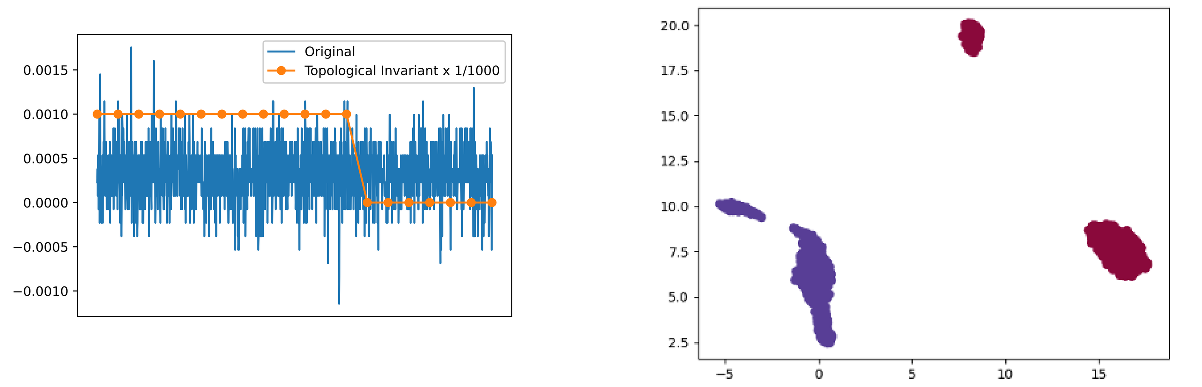
Tensor Clustering for Anomaly

- Decompose the data into anomaly and low rank part
- Self regression on low rank part to cluster the days
- Fault cluster accuracy: 0.64; F1 score: 0.484



Anomaly Clustering Using Manifold Learning

- Our goal: to find the latent embedding of high dimensional spatial-temporal solar data
- Intra-series temporal patterns: Fractal Dimension (FD), Euler Characteristic (EC) (**Left**)
- Inter-series spatial patterns: Manifold Learning



SDE Based Fault Prediction

Use Stochastic Differential Equation (SDE) to model system dynamics of inverter signals

$$dx = a(x,t,\theta)dt + b(x,t,\theta)dw$$

- a is drift function, b is diffusion function, θ unknown parameter vectors
- Dimension reduction and feature fusion by the EC curve of multiple sensing signals for solar inverters
- The dynamics and uncertainty of feature representation x can be modeled by SDE
- When the feature x exceeds a certain threshold, the failure happens.
- Predict the failure event of the inverters, Accuracy: 80.1%

CNN-Based Multi-label Fault Detection and Prediction

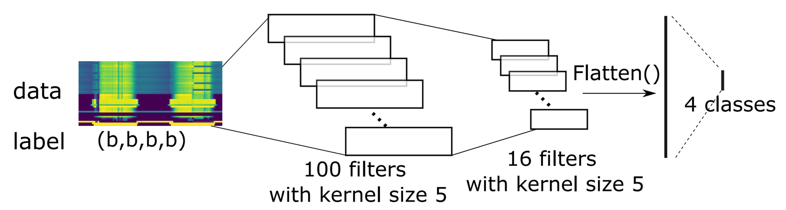
- Selected Important Faults:

Fault	Display	Counts
55	AC CONTACTOR	31765
67	BUS DIFF VOL	48783
95	DCFuse	30788
100	other	-

- Multi-label clarification:

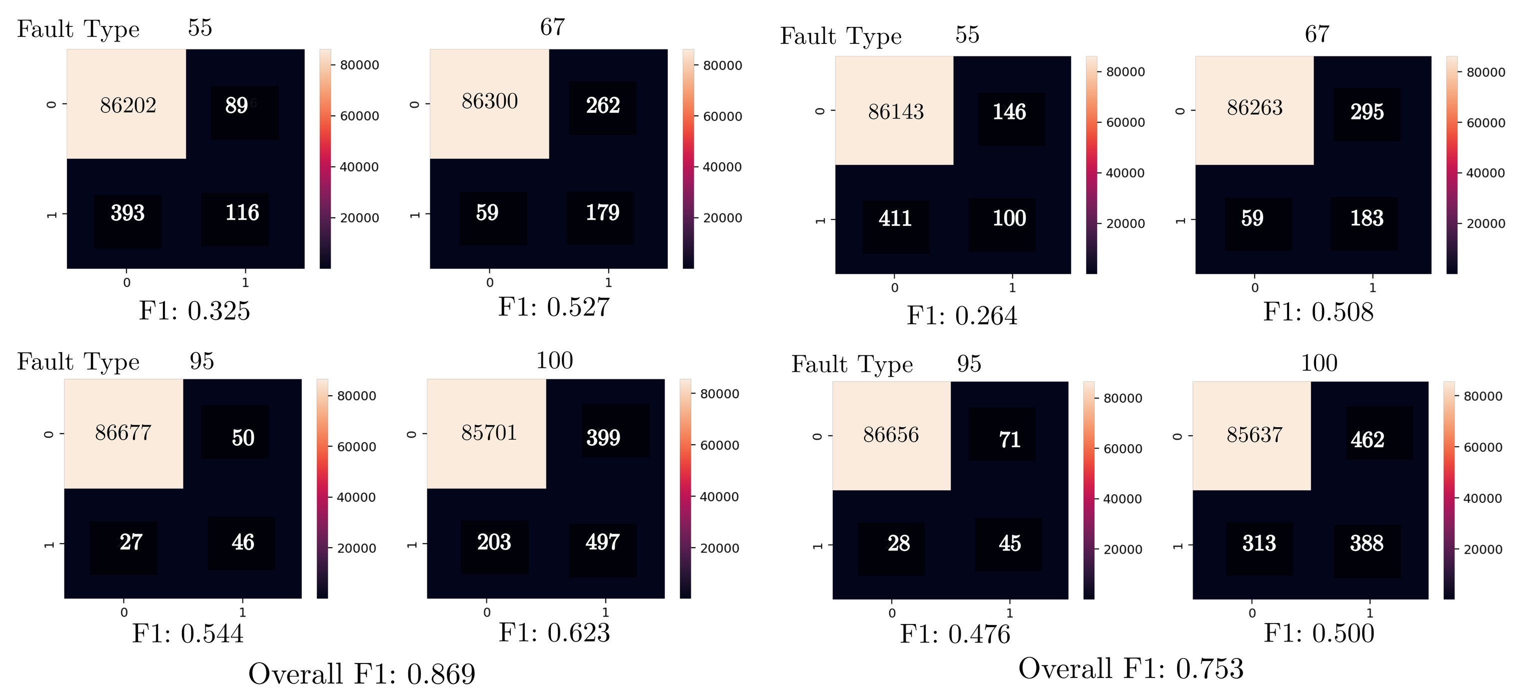
c=3	Multi-class			Multi-label		
F01 F02 F03	Data			Data Precessing: Sliding window		
Label	F01	F03	None	F01,F02	F01,F02,F03	None
Vector	(1,0,0)	(0,0,1)	(0,0,0)	(1,1,0)	(1,1,1)	(0,0,0)

- CNN Model:



Results:

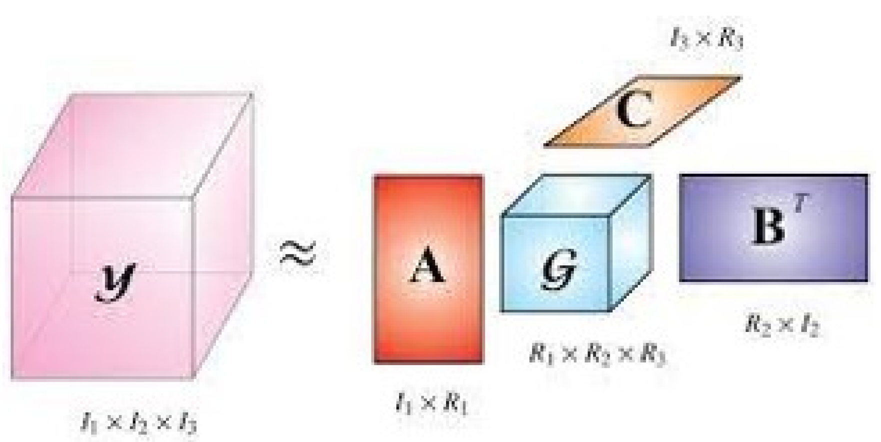
- Detection:
- Prediction:

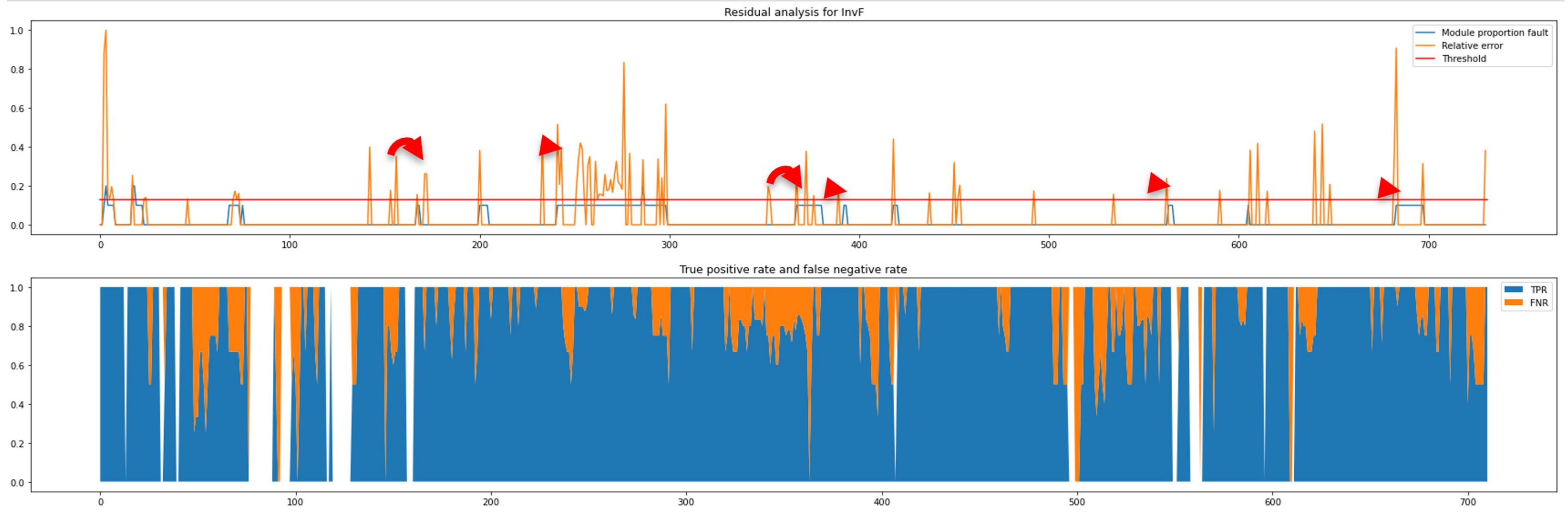


Tensor Based Module-level Faults Prediction

- Find main features and reconstruct data
- Large error indicates potential fault
- Fault Prediction F1-score > 0.7 for a 2-week prediction horizon.
- Mean Remaining Useful Life (RUL) difference less than 3 days.

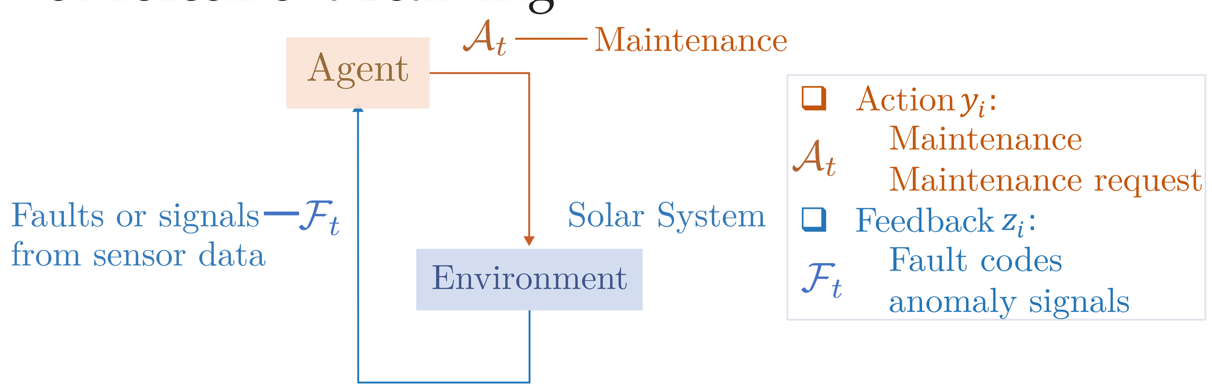
$$\mathcal{E} = \mathcal{X} - \mathcal{G} \times_1 U_1 \times_2 U_2 \times_3 U_3 \times_4 U_4$$



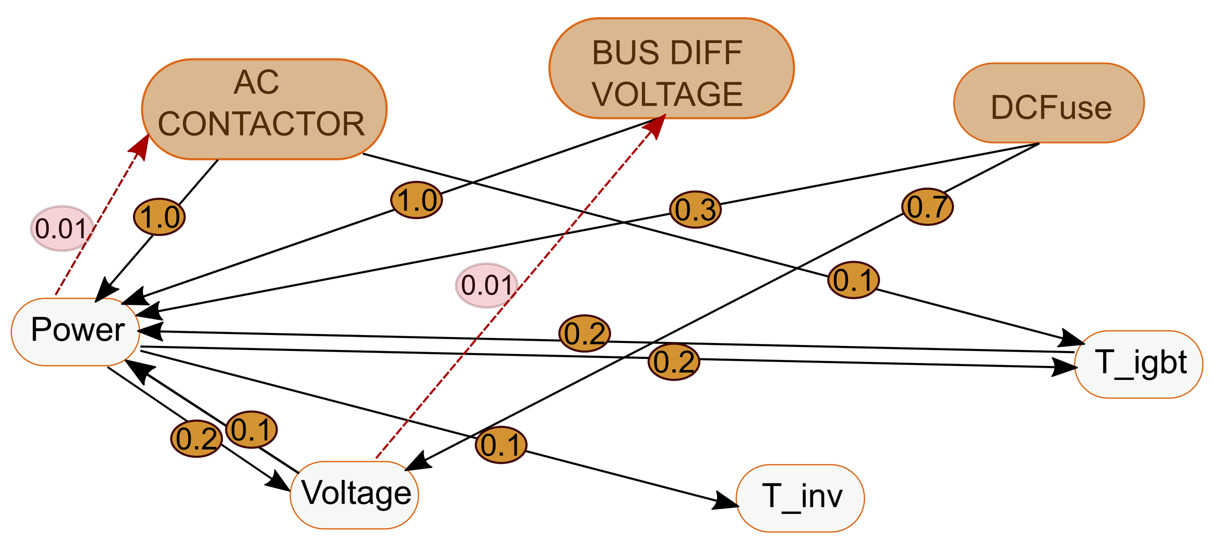


Future Work

Reinforcement Learning



- $\mathcal{F}_t$ and $\mathcal{A}_t$: Hawkes Process - Capture self-and mutually-triggering patterns.



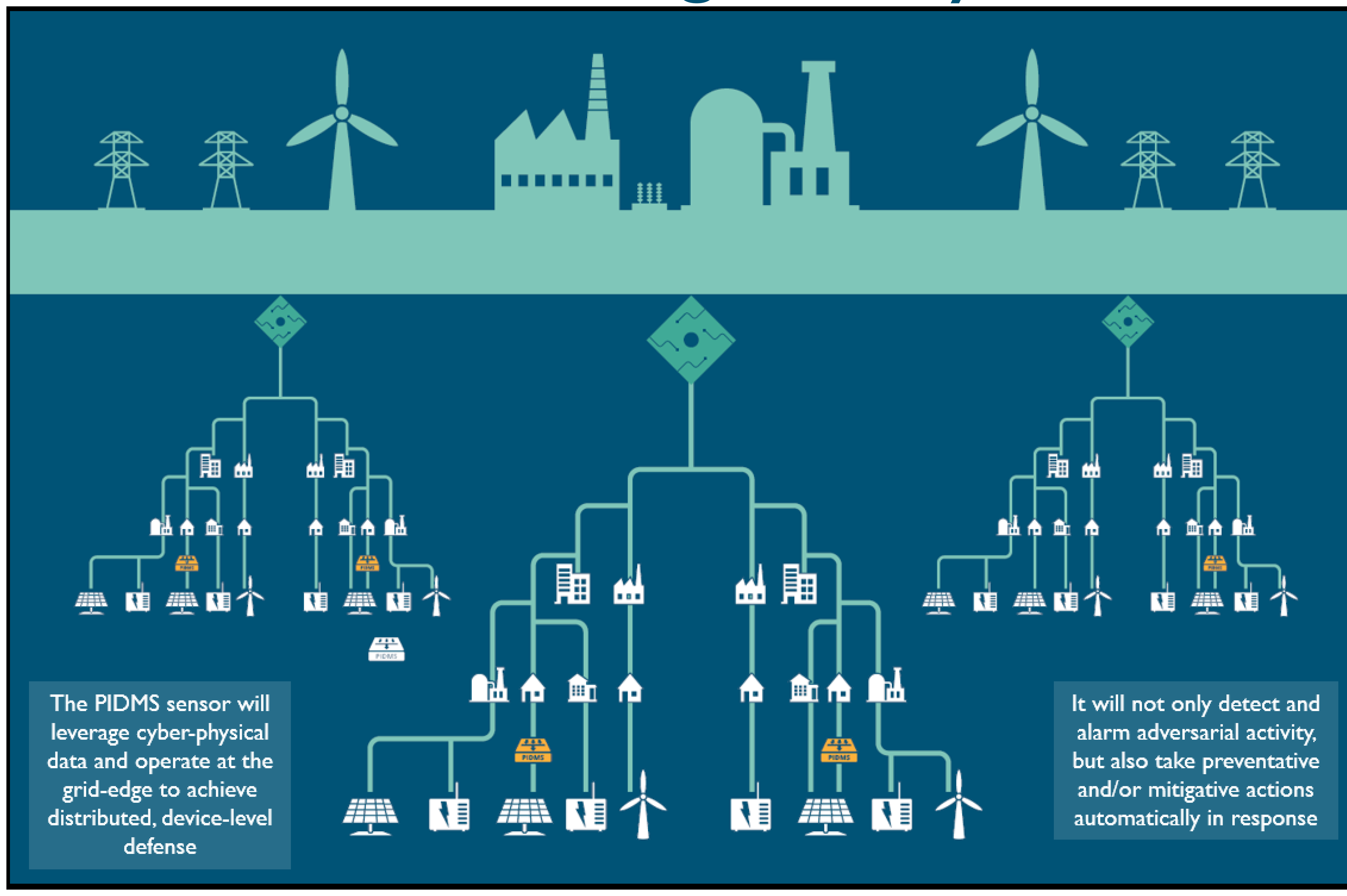
Acknowledgement

The research reported in this paper was supported by funds from the Department of Energy (DOE) (Contract No. DE-EE0009354, PI: Hao Yan). Specifically, we would like to thank Tassos Gonas and Marie Mapes from DOE for their guidance. We would also like to thank Isaac, Chirs, and Jeremey from APS for the helpful discussion.

Proactive Intrusion Detection and Mitigation System

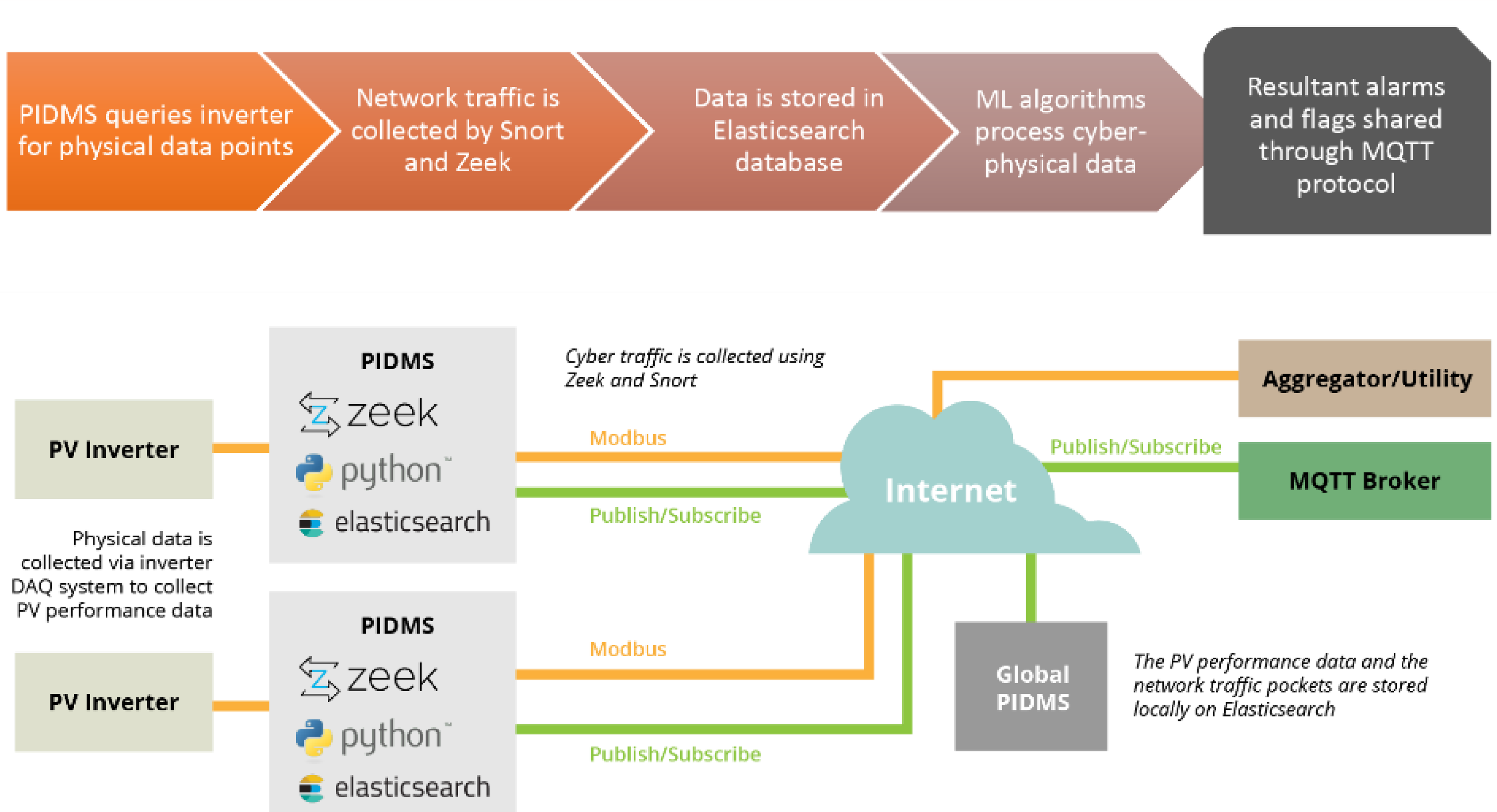
Introduction

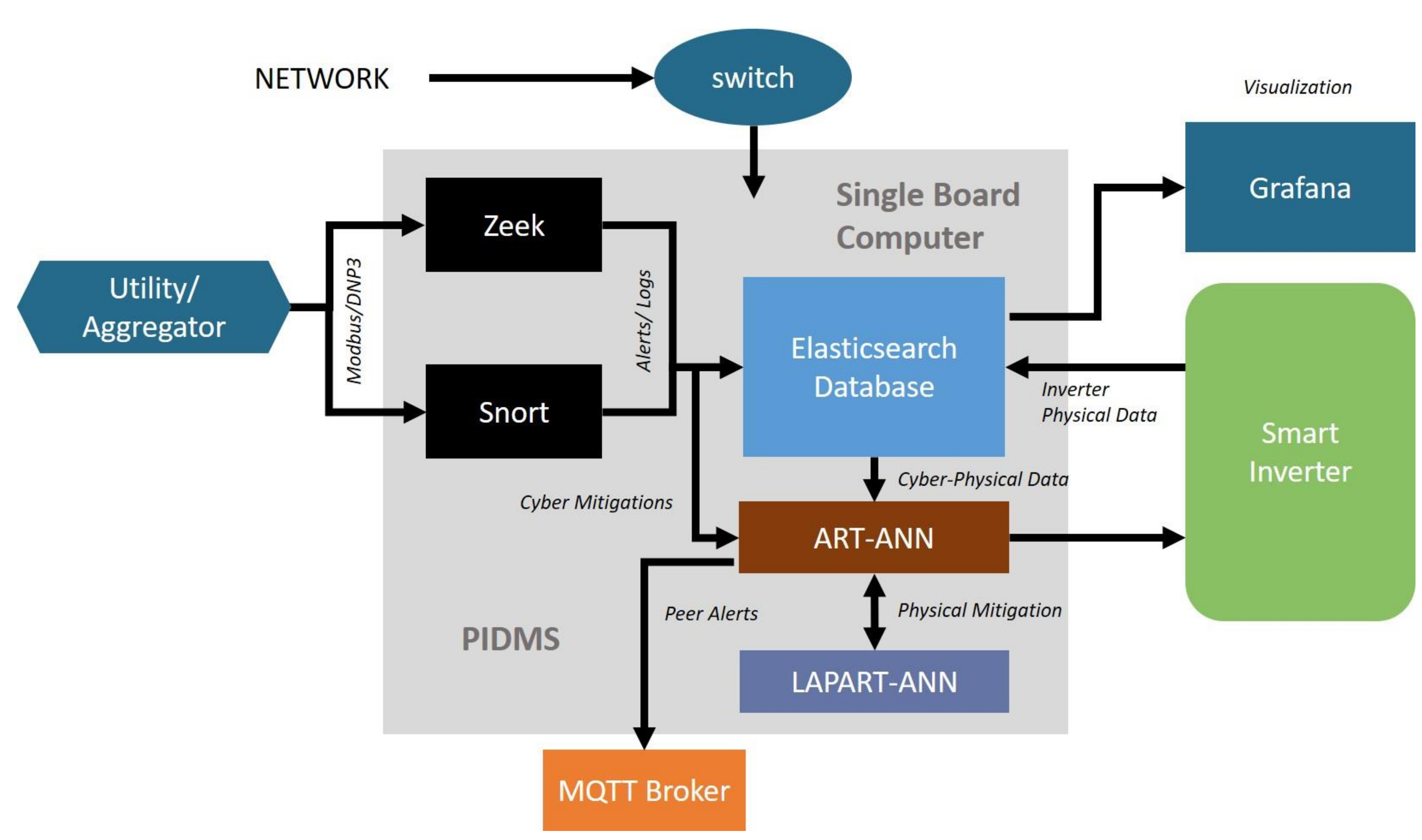
The electric grid is undergoing rapid, revolutionary changes. Recent developments include the addition of advanced, smart, two-way communication technologies, growing penetrations of distributed energy resources (DER), and increasing reliance on interconnectivity and communications to external networks. Unfortunately, these trends are expanding the attack surfaces of the grid with new internet-connected DER interfaces and third-party cloud applications used for DER aggregation and control. In response to this significant, emerging gap in power systems cybersecurity, we have developed a proactive intrusion detection and mitigation system (PIDMS) to secure grid-edge PV smart inverter and other DER equipment in DER systems.



AI/ML Usage

The PIDMS leverages the adaptive resonance theory artificial neural network (ART-ANN) algorithm that analyzes the cyber-physical data for intrusion detection analysis; the ART-ANN implementation performs simultaneous online learning and training such that performance can continue to be improved even after field installation.





Challenges and Best Practices

- Distributed, single-board computer implementation requires careful consideration of computational burden of different ML algorithms; it is important to assess RAM, CPU, temperature, etc. impacts when selecting ML approach and implementation architecture.

Key Takeaways

The PIDMS is a distributed, flexible, bump-in-the-wire (BITW) solution for protecting PV smart inverter communications:

- Automatically processes both cyber (network traffic) and physical (power system) measurements using network deep packet inspection tools and custom machine learning algorithms to detect abnormal events and correlate cyber-physical events.
- Not only detects abnormal events, but also automatically deploys mitigations to limit or eliminate system impact.
- Increased situational awareness and alerting capabilities are also achieved with peer-to-peer communication between PIDMSs at different locations; includes front-end dashboard for alerting and visualization.

Current and Future Work

- 2022 R&D100 winner and market disruptor special recognition (silver)
- Technology maturation efforts underway with DHS Commercialization Accelerator Program funding (FY24)
- Planned field testing at partner utility sites (FY24)

2022 R&D 100 WINNER

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.
SAND No. SAND2023-11481PE

DOE SETO Project #34234, FY18-FY20
Project Partners: EPRI, OPAL-RT, DNK

Design and Optimization of Energy Systems (DOES) Laboratory

Reinforced Hierarchical Probabilistic Solar and Netload Forecasting Based on Dynamic Multi-model Selection

Jingyi Yan[1], Feng Cong[2], Qi Wang[3], Kevin Chen[3], and Jie Zhang[1]

[1]The University of Texas at Dallas; [1]National Renewable Energy Laboratory; [3]Altitude Grid

Introduction

Research Background

- ❖ Exponential growth of solar energy
- ❖ Uncertainty involvement accordingly
- ❖ Increased need for accurate forecasting

Methodology Introduction

- ❖ **Description:** A reinforcement learning based dynamic model selection methodology with **local awareness** to avoid over-reliance on global accuracy.
- ❖ **Reinforcement Learning Method: Q-learning**
- ❖ **Reward Criterion:** The rank of performance
- ❖ **Workflow:** In forecasting period, **Q-Learning agent1** predicts the best DSF from DMP; and **Q-Learning agent2** predicts the best PSF from PMP.

DSF, PSF: deterministic/probabilistic solar forecasting

Research Motivation

- ❖ **Limitation I:** Most of the research focuses on the overall accuracy of the objective function without considering the local performance.
- ❖ **Limitation II:** There are numerous methodologies, while universally best model does not exist.

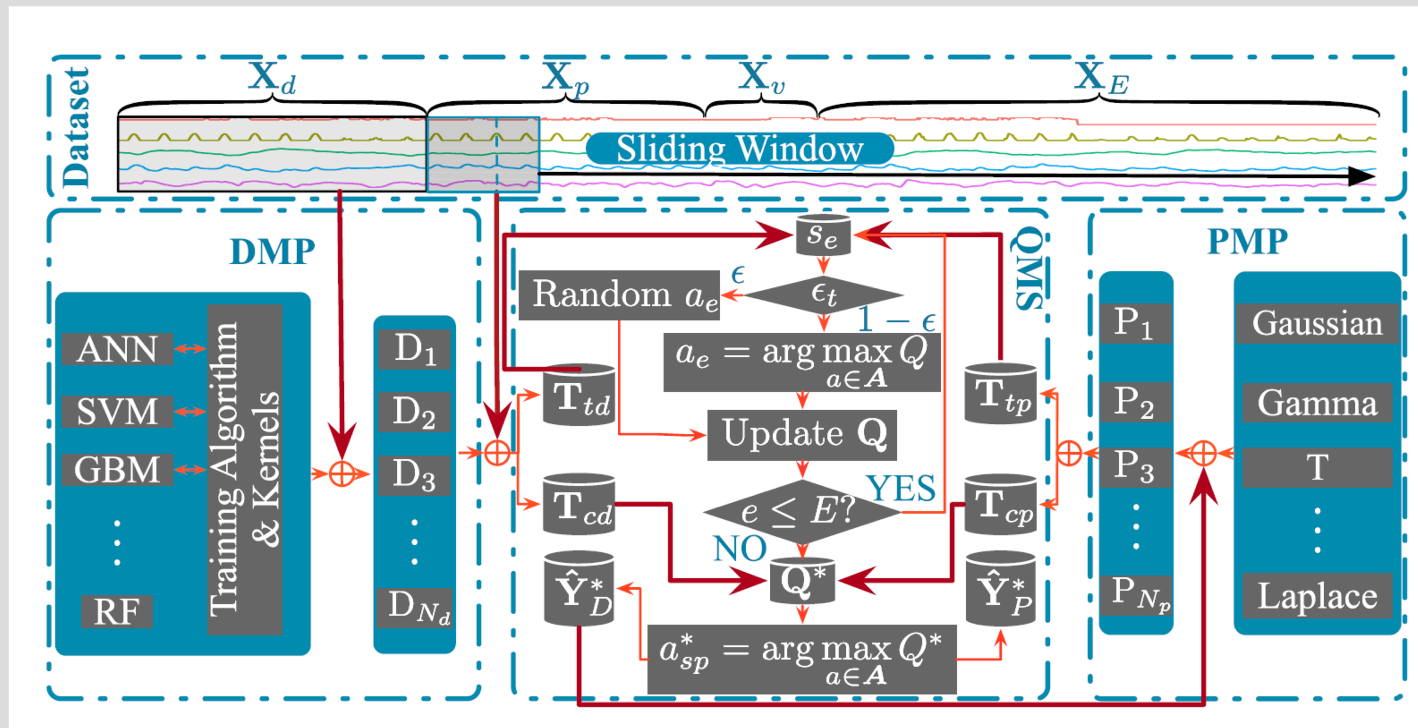


Fig. 1. Framework of Reinforced Hierarchical Probabilistic Forecasting Model Based on Dynamic Multi-model Machine Learning

Challenges and Best Practices

Solar Forecasting

- ❖ Improve overall performance through successive local optimization

Table 2. Deterministic Forecasting Comparison

Criteria	D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	D_{10}	D_Q
MAE (kWh)	0.39	0.40	0.41	0.38	0.42	0.40	0.40	0.40	0.40	0.40	0.37
RMSE (kWh)	0.46	0.47	0.47	0.45	0.47	0.49	0.45	0.46	0.47	0.47	0.44

- ❖ Leverage ensemble numerical weather prediction from ECMWF

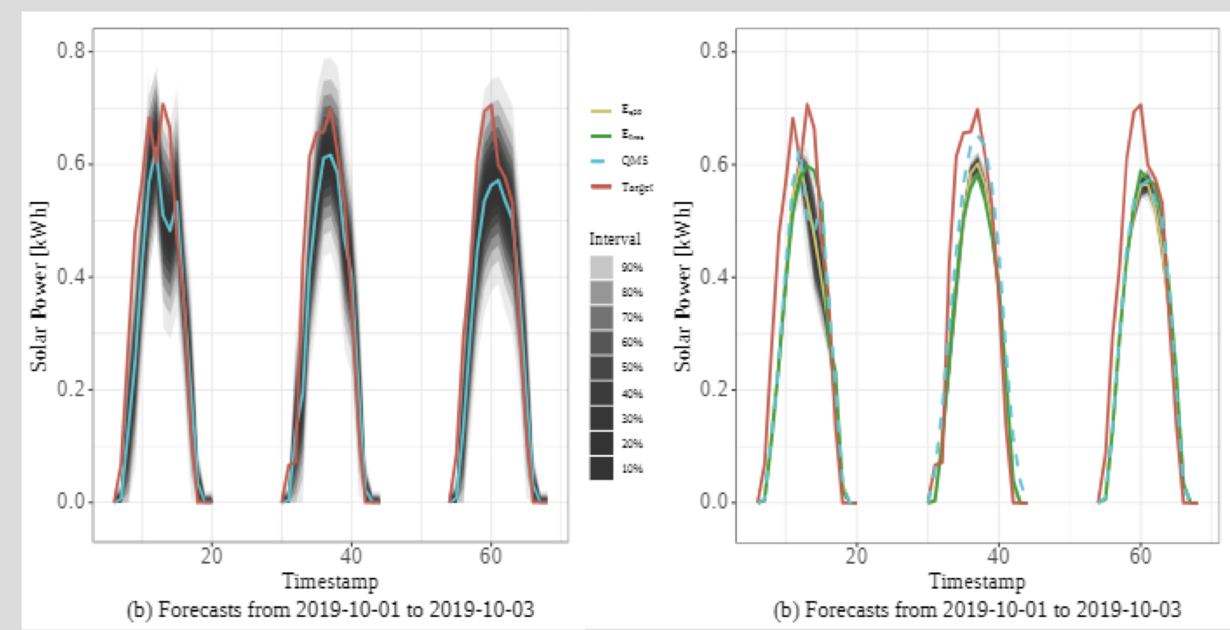


Fig. 2. Target power, DSF and PSF time series of 1) proposed method, and 2) the benchmark

Netload Forecasting

- ❖ Developed a QMS method that has reduced DLF errors by over 50% and improved PLF accuracy by nearly 60%.

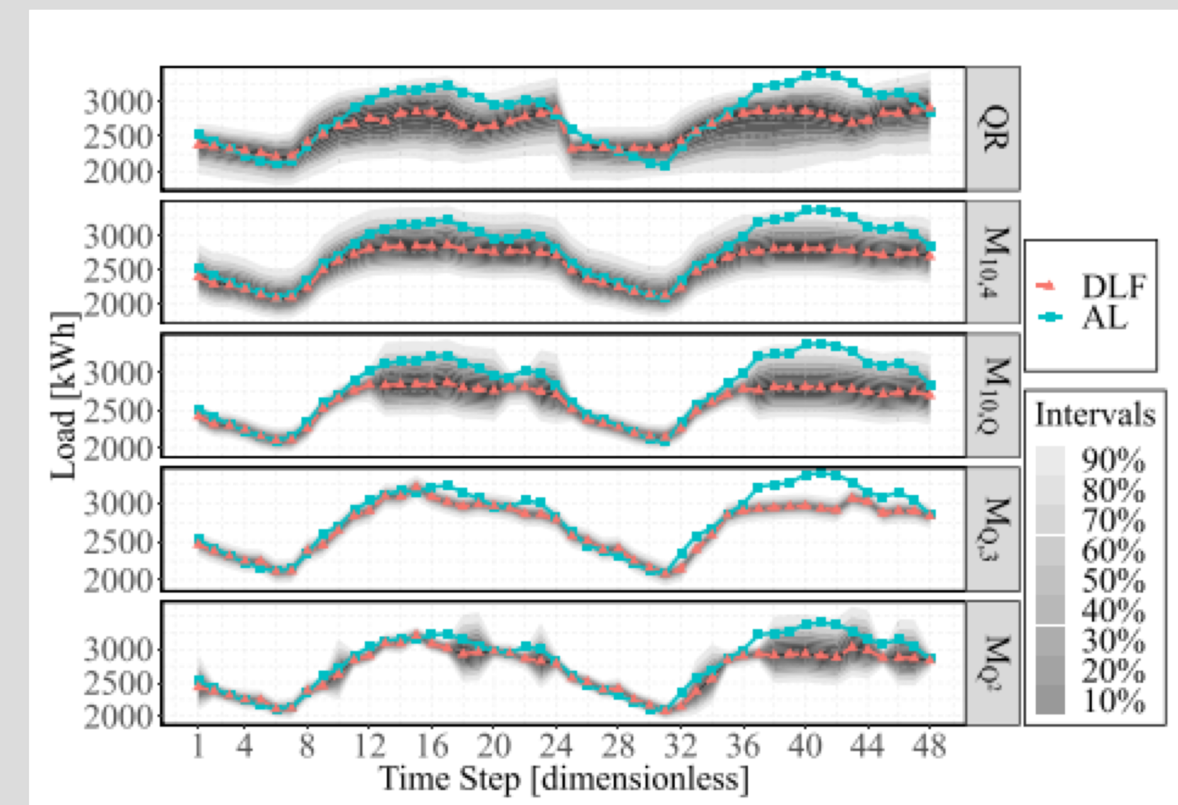


Fig. 3. Actual Load, DLF and PLF time series of 1) baseline models, 2) the best global behavior model of both DMP and PMP, 3) & 4) one best global behavior model with the QMS model, and 5) totally QMS model

DLF, PLF: deterministic/probabilistic load forecasting

RL Framework

Major Components

- ❖ Deterministic Forecasting Model Pool (DMP)
- ❖ Probabilistic Forecasting Model Pool (PMP)
- ❖ Q-Learning Based Dynamic Model Selection (QMS)

Table 1. ML Based Model in DMP and PMP

Model Type	Algorithm / Function	Model	Training algorithm or kernel function
DSF	ANN	D_1	Standard back-propagation
		D_2	Momentum-enhanced back-propagation
		D_3	Resilient back-propagation
	SVR	D_4	Linear kernel
		D_5	Polynomial kernel
		D_6	Radial basis function kernel
	GBM	D_7	Squared loss
		D_8	Laplace loss
		D_9	T-distribution loss
	RF	D_{10}	CART aggregation
PSF	CDF	P_1	Normal
		P_2	Gamma
		P_3	Laplace
		P_4	noncentral-t

Algorithm 1 Q-Learning Based Dynamic Model Selection

Requirements:
Number of steps, N_{sp}, in a QMS procedure
Model pool dimension N_m, which is either N_d or N_p
Q-learning training dataset T_{td} or $T_{tp} \in R^{N_q \times N_m}$
QMS process dataset T_{cd} or $T_{cp} \in R^{N_{sp} \times N_m}$
Learning rate α, discount factor γ , number of episodes N_e
Ensure: Select the best model from N_m models at each step in T_{cd} or T_{cp}

1. Initialize Q
2. **for** $e = 1$ to N_e do
3. With the probability of ϵ select a random action a_e, otherwise select $a_e = \arg\max_{a\in A} Q_e(s_e, a)$
4. Calculate R by Reward Function
5. Update Q by
$$Q_{e+1}(s_e, a_e) = (1-\alpha)\, Q_e(s_e, a_e) +$$
$$\alpha[R_e(s_e, a_e) + \gamma \max_{a\in A} Q_{e+1}(s_{e+1}, a)]$$
6. $\epsilon \leftarrow \epsilon - \frac{1}{N_e}$
7. **end for**
8. **for** sp = 1 to N_{sp} do
9. Take action $a^*_{sp} = \arg\max_{a\in A} Q^*(s_{sp}, a)$
10. **end for**

Discussion

Takeaways

- ❖ Case studies from both netload and solar forecasting present the accuracy and robustness of the proposed RL methodology, which chooses next forecasting model based on the performance of current forecasting model.
- ❖ This revisable reinforced hierarchical model can be utilized to promote forecasting in different scenarios, with different targets, time scales, and spatial scales.
- ❖ With ensemble weather forecasts disseminated by major operational weather prediction facilities, training dataset for both deterministic and probabilistic forecasting can be reduced to less than 20 days.

Future Work

- ❖ FusionalAI: A Deep Transformer-based Energy Forecasting Methodology
- ❖ Real-Time Outage Management in Active Distribution Networks Using Reinforcement Learning over Graphs

Acknowledgement

- ❖ **Funding:** DOE STTR Phase I (DE-SC0022928)

Reference

[1] Feng, C., Sun, M. and Zhang, J., Reinforced Deterministic and Probabilistic Load Forecasting via Q-Learning Dynamic Model Selection, IEEE Transactions on Smart Grid, Vol. 11, Issue 2, 2020, pp. 1377-1386.

[2] Yan, J., Feng, C., Sun, M. and Zhang, J., Enhancing Short-term Solar Ensemble Forecasts via Reinforcement Learning with ECMWF Integration, 2023 (under preparation)

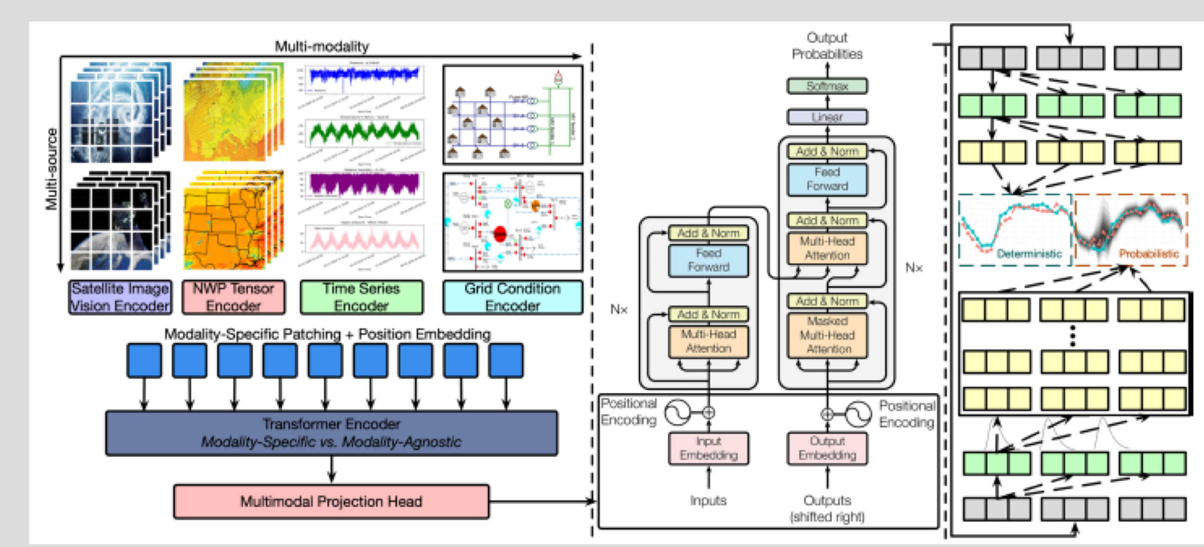


Fig. 4 The FusionalAI framework for multimodal learning-based solar/net load forecasting

SolarAPP+: Solar Automated Permit Processing Plus

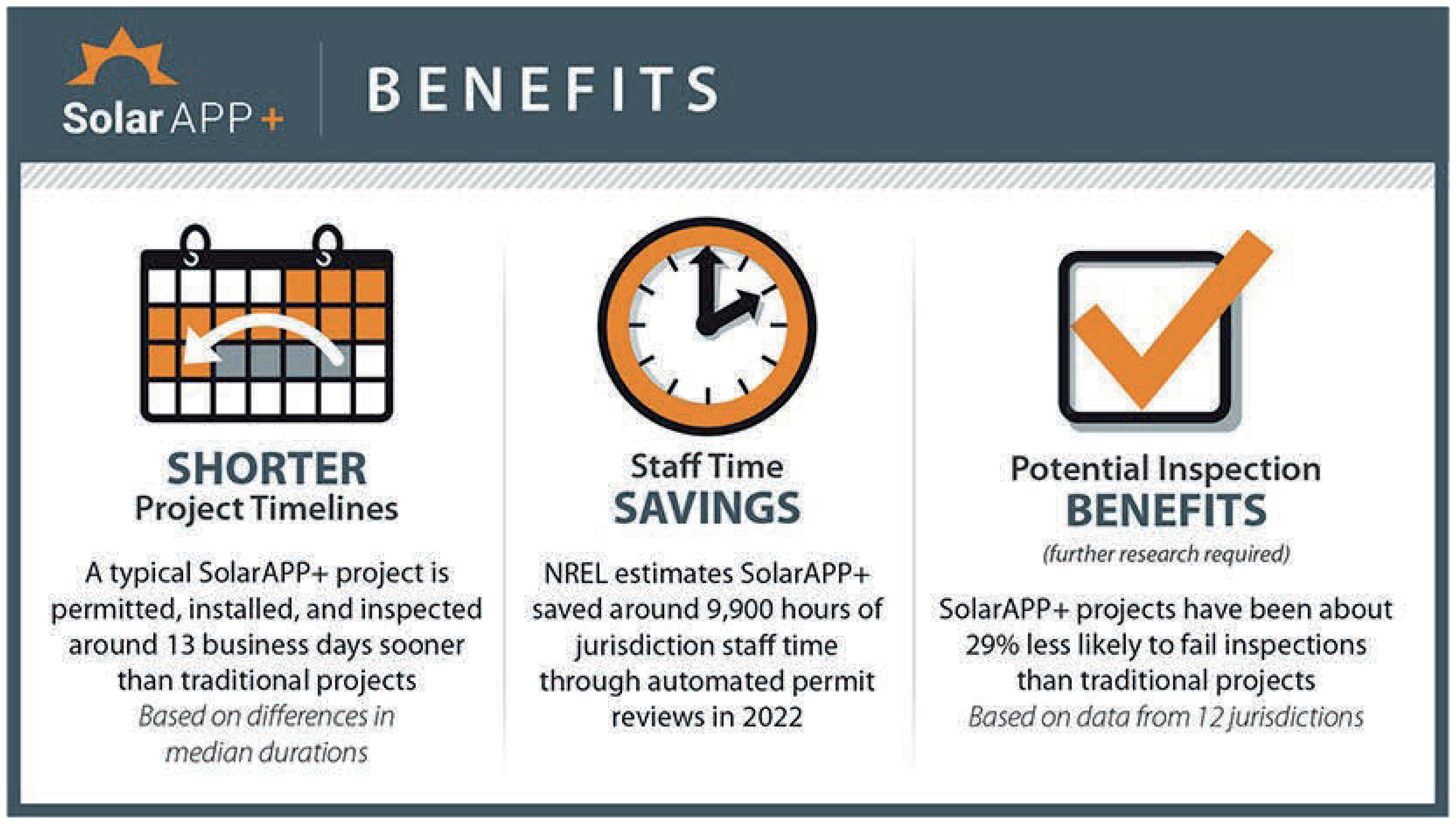
Jeff Cook and Emily Dalecki, NREL

27,000+
Residential rooftop PV permits approved to date, including 4,500+ revisions

2,500+
Permits approved to date for solar-plus-storage projects

140 AHJs across 9 States
Are piloting or have fully adopted the tool

29,500+ hours of staff time saved
Through automated processing of permits



Introduction

To address residential solar photovoltaic (PV) permitting resource constraints and streamline solar permitting processes among authorities having jurisdiction (AHJs), the National Renewable Energy Laboratory (NREL) led a collaborative effort to develop the Solar Automated Permit Processing Plus (SolarAPP+), a solar permitting software solution provided at no cost to AHJs. The SolarAPP+ tool is an online portal that automates permit plan review, thereby enabling an instant permit approval process for code-compliant residential PV systems. Based on national model building, electrical, and fire codes, SolarAPP+ automatically performs a compliance check of permit inputs against code requirements and produces an inspection checklist that can be used to verify installation practices, workmanship, and adherence to the approved design.

How is AI/ML used in this project?

NREL has implemented AI in the following ways:

- Automate review of model building, electrical, and fire codes to incorporate into the SolarAPP+,
- Automated Programming Interfaces (APIs) between solar design, SolarAPP+, and government permitting software with less human touch/change for errors,
- To conduct robust testing of the code compliance checks completed within SolarAPP+ to detect errors before release, and
- NREL continues to consider pathways to increase the use of AI to help build internal databases of equipment and perform more robust compliance checks.

Challenges and Best Practices

1. **Establishing Trust in Automation:** One key challenge with SolarAPP+ adoption has been establishing a level of trust with AHJs in an automated process. The SolarAPP+ team have utilized a variety of outreach techniques to overcome this challenge.
2. **Building and Updating Internal Databases:** Other challenges for SolarAPP+ include creating and maintaining databases which the automated code compliance checks rely on, such as eligible combinations of racking systems and modules per UL 2703.
3. **Building and Maintaining APIs:** Developing APIs with existing databases and existing software providers has also proved challenging due to time and budget constraints, while the system and underlying technology is rapidly changing.
4. **AI is critical to ensuring long-term viability of SolarAPP+.** AI can allow for more testing of the safety of SolarAPP+ than is humanly possible. This is critical to the long-term trust of users in the platform.

Despite these challenges, a typical SolarAPP+ project already completes the full project timeline (permit submission to passed inspection) about 13 business days before a project submitted and reviewed through the traditional route.

Key Takeaways and Future Work

In the future, SolarAPP+ plans to incorporate more features and technologies into the current portfolio. These features include solar-plus-storage permits for new home construction, roofing-integrated PV products, all future versions of the National Electrical Code (NEC) and I-Codes, reroof permits, solar thermal permits, residential electric vehicle charging permits and more. While utilizing AI has helped address some challenges with SolarAPP+ to date, opportunities to utilize AI may expand as new features and technologies are incorporated into the SolarAPP+ tool.

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 Solar Energy Technologies Office. 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.

Solar Applications of Artificial Intelligence and Machine Learning Workshop
Washington, D.C.
October 31 – November 1, 2023
NREL/PO-6A20-87837