OPF-Learn: A software package for efficiently creating AC optimal power flow machine learning datasets

22ISGT1137

Trager Joswig-Jones, *NREL & University of Washington;* **Ahmed Zamzam**, *NREL;* **Kyri Baker**, *University of Colorado: Boulder*

Introduction

A significant power systems research area is addressing challenges operators face when determining how to **fulfill electric power demand at the lowest cost**, referred to as **AC optimal power flow (AC-OPF)**

+ Renewable energy resources are a key factor in the transition to a more sustainable power grid.
− However, these resources make operating the grid economically more challenging, due to their variable and distributed nature.

AC-OPF is central in determining how to best operate the grid but increasing levels of **distributed renewable generation make this problem computationally challenging.**

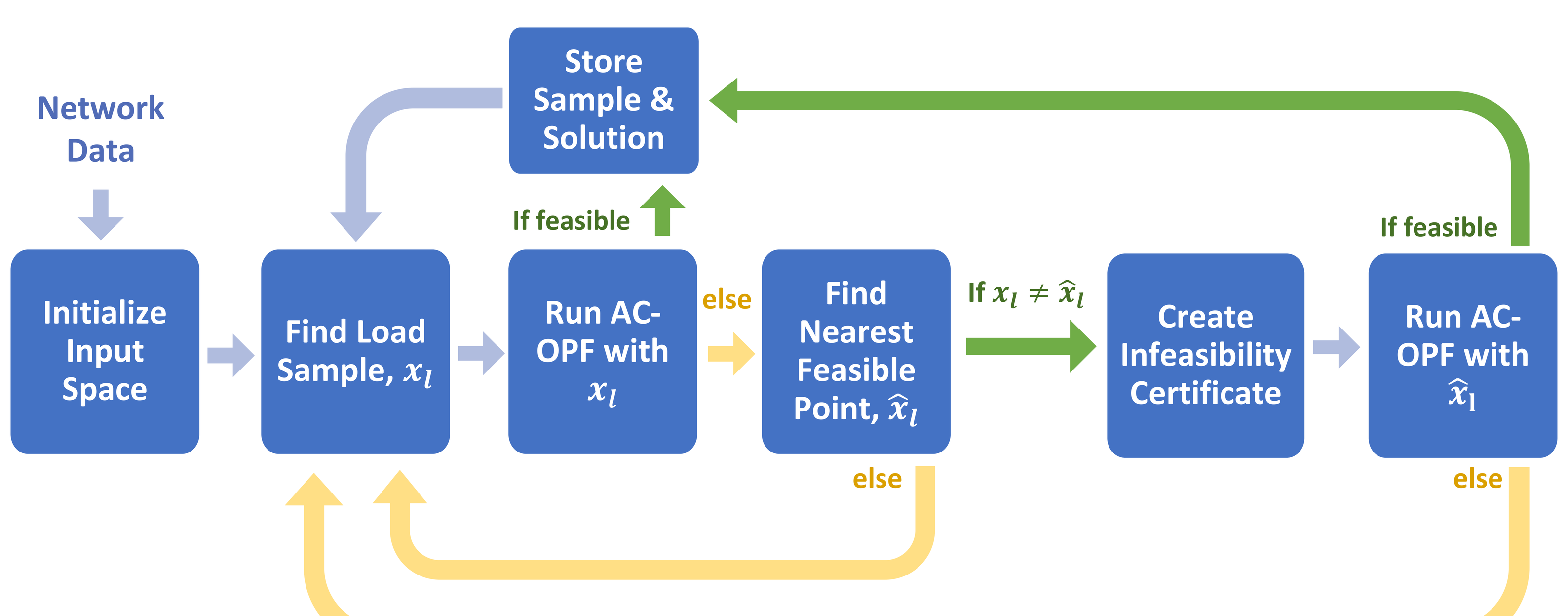
+ One possible solution is applying **machine learning (ML) methods** to solve AC-OPF problems more efficiently.
− A **lack of disciplined dataset creation** and benchmarking prohibiting useful comparisons between ML methods.

Research Question

How can datasets that are representative of the entire AC-OPF feasible input space be efficiently created for machine learning applications to solving AC-OPF?

OPF-Learn Software Package

OPF-Learn efficiently creates datasets mapping input load profiles to optimal generator setpoints for machine learning applications.



OPF-Learn software flow chart for creating AC-OPF samples.

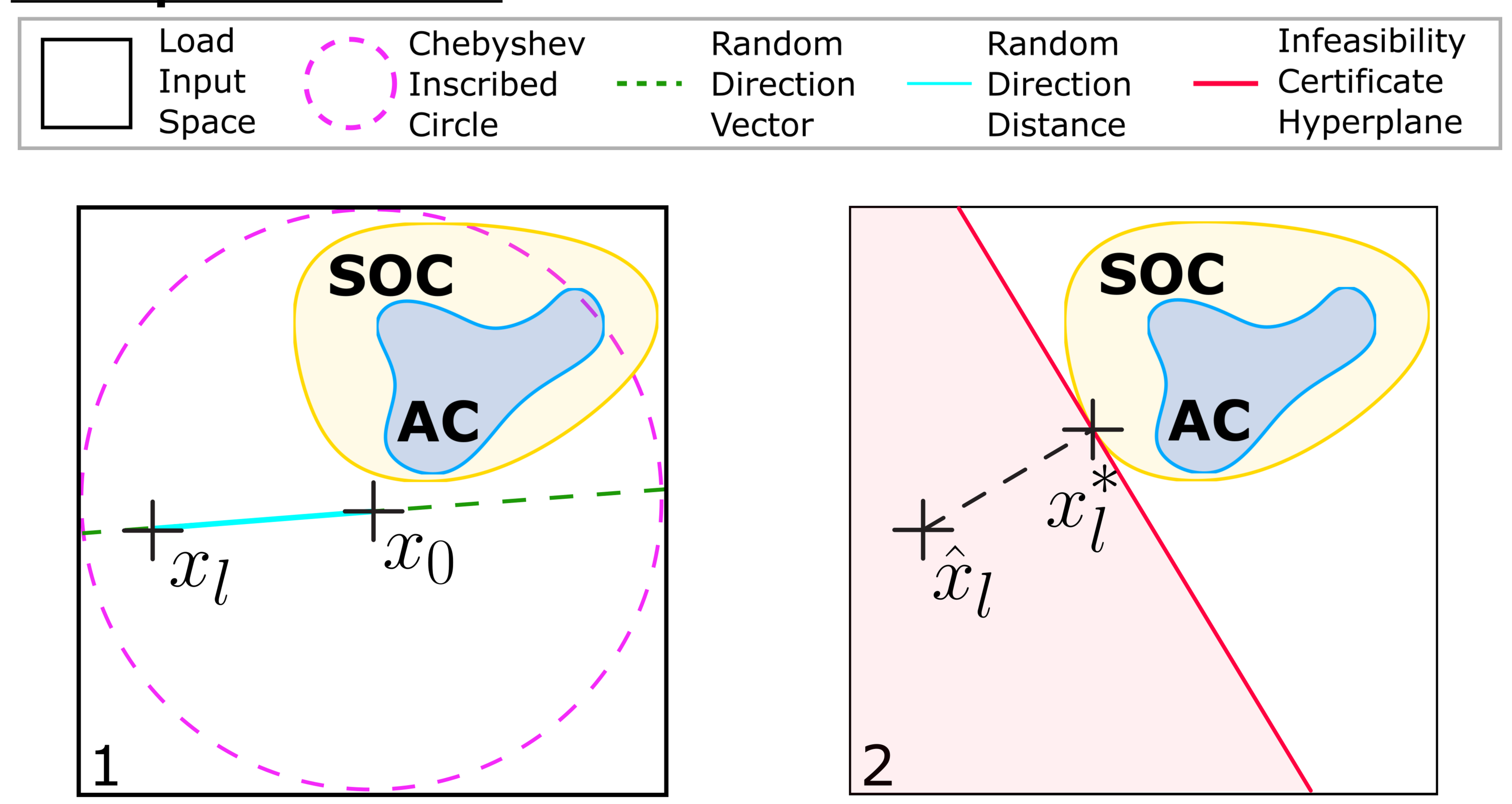
Typical AC-OPF datasets contain only a small subset of the AC-OPF feasible loads space. OPF-Learn improves upon this using the following features,

- **Uniformly Samples Input Space:** The input space, containing the entire AC-OPF feasible region, is uniformly sampled.
- **Reduces Input Space with Infeasibility Certificates:** Infeasibility certificates are created to reduce the input space as specified in. Hyperplanes are created using a convex AC-OPF formulation, the second-order cone (SOC) relaxation, to find the nearest feasible point.

AC-OPF Dataset Creation

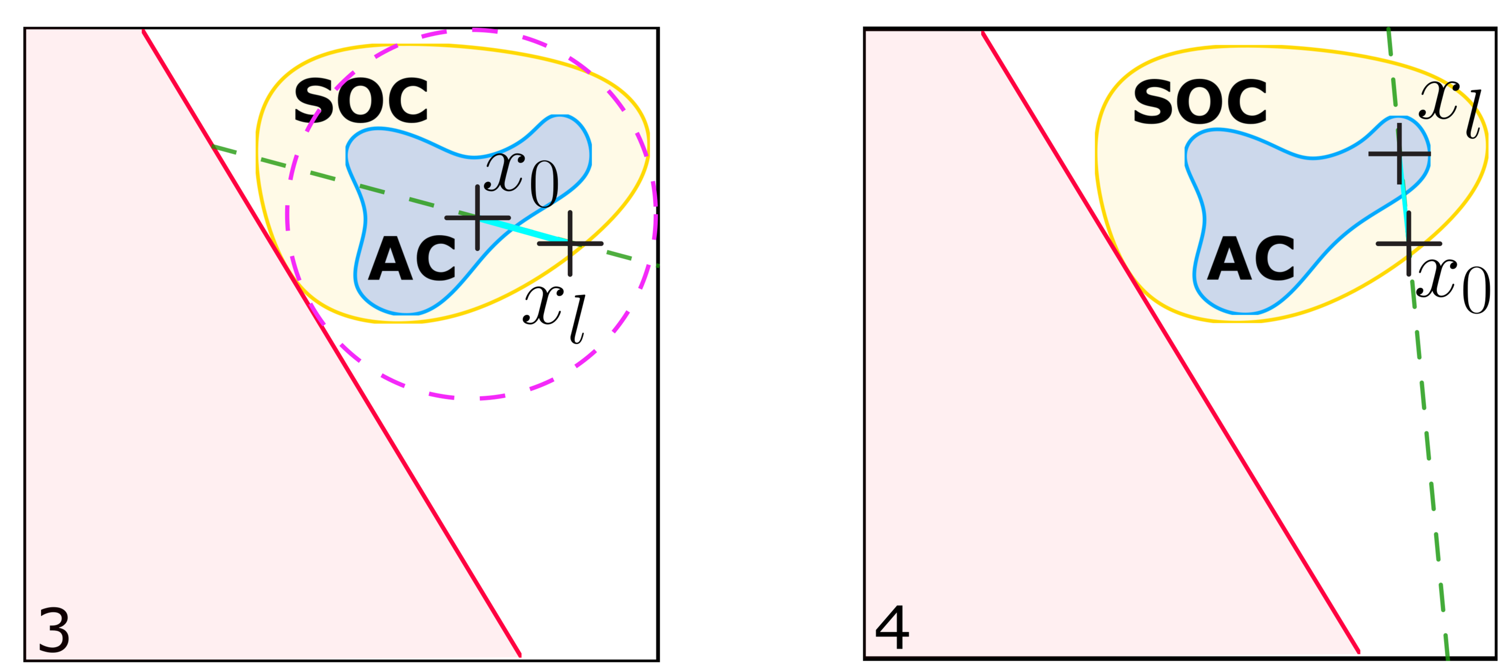
An example of the methodology used to find load samples can be seen below,

Example Iterations



1. Find the Chebyshev center, x_0. Generate a random direction vector and travel a random distance along this vector to find a new load sample, x_l.

2. Check if x_l is AC-OPF feasible. If not feasible, find the nearest SOC feasible point, x_l^*. Since $\hat{x}_l \neq x_l^*$, define a new infeasibility certificate at x_l^* with normal, $\vec{n} = \hat{x}_l - x_l^*$.

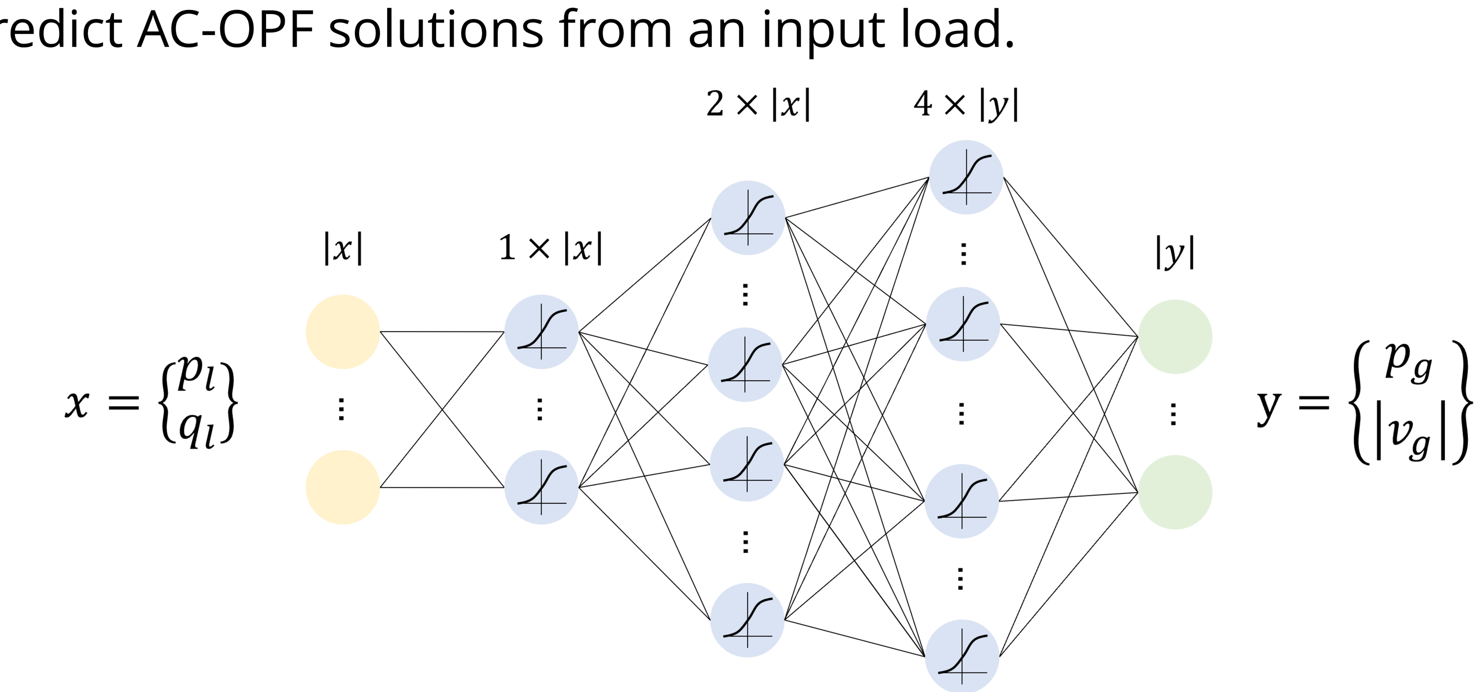


3. Gather a new sample, x_l, as in step 1. Check if the new x_l sample value is AC-OPF feasible. Here it is not, so the nearest SOC feasible point is found. $\hat{x}_l = x_l$, so discard this sample.

4. Sample a new load profile, x_l, as in step 1, but starting from the last point, now x_0. Check if x_l is AC-OPF feasible. x_l is AC-OPF feasible, so store x_l and its AC-OPF optimal solution.

Training Neural Networks

To test datasets generated using OPF-Learn, neural networks (NN) were trained on OPF-Learn datasets as well as "typical" datasets to predict AC-OPF solutions from an input load.



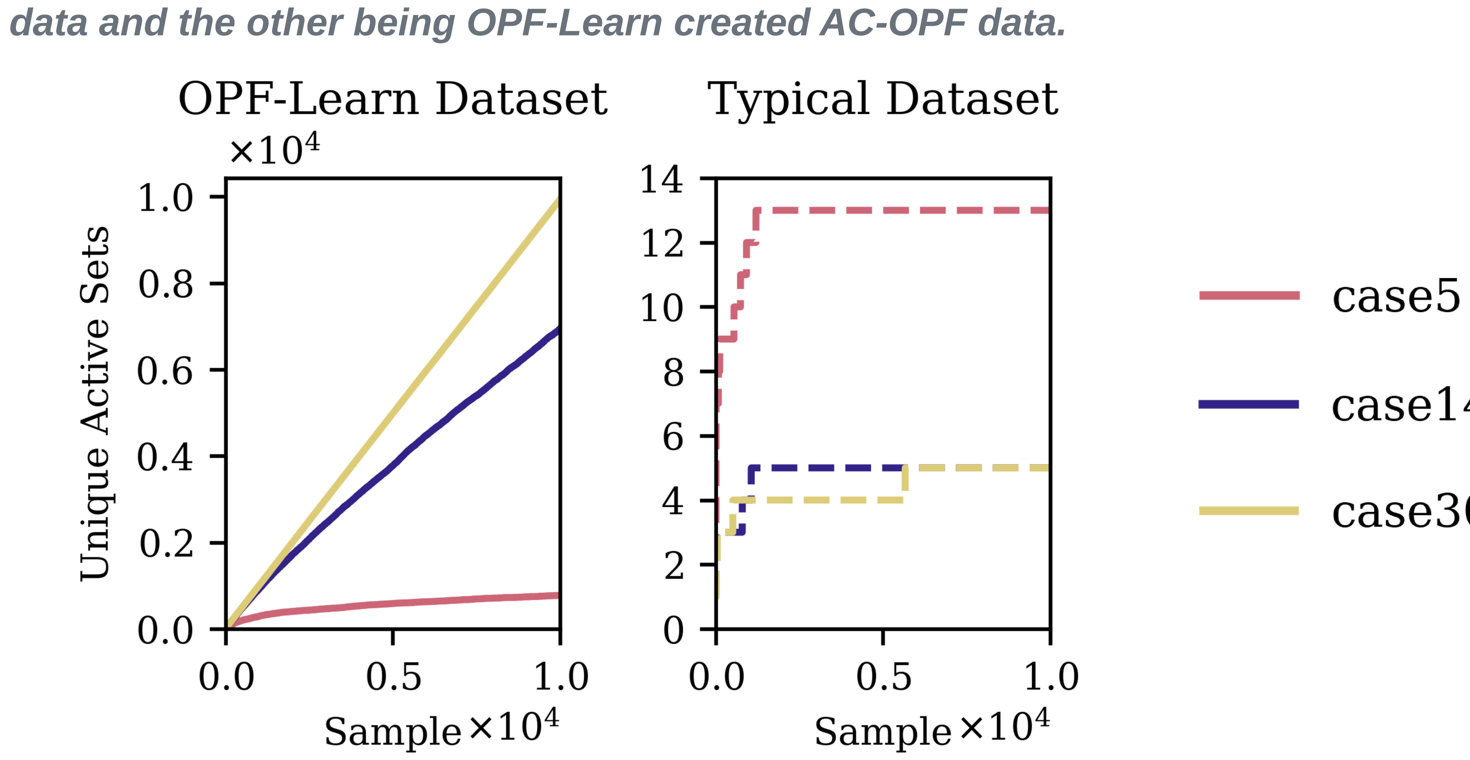
Neural network model with three hidden layers and sigmoid activation functions.

Results

- OPF-Learn is a publicly available Julia & Python package to create representative AC OPF datasets
  - Able to find more unique active sets of constraints in load samples than typical dataset creation methods.
  - NN models trained on OPF-Learn data have less mean squared error (MSE) when tested on representative test sets compared to the same model trained on a typical dataset.

	OPF-Learn Dataset Trained Model		Typical / OPF-Learn	Typical Dataset Trained Model		OPF-Learn / Typical
Test Dataset:	OPF-Learn	Typical		OPF-Learn	Typical	
case5	2.17E-2	1.86E-3	**8.57E-2**	1.33E+0	9.08E-6	**1.46E+5**
case14	2.75E-4	1.01E-4	**3.67E-1**	3.94E-2	9.41E-7	**4.19E+4**
case30	1.55E-4	5.46E-4	**3.52E+0**	8.17E-3	1.60E-8	**5.11E+5**
case118	6.97E-2	2.35E-1	**3.37E+0**	4.47E-1	4.47E-3	**1.00E+2**

NN mean squared errors for the OPF-Learn dataset trained and Typical dataset trained P_g NN models. Each NN was tested with two test datasets, that the models had not seen during training, with one dataset containing typically created AC-OPF data and the other being OPF-Learn created AC-OPF data.



Plot of unique active sets found over time with a typical dataset creation methods and the OPF-Learn dataset creation method. Note the difference in the y-axis scale

Conclusion & Future Work

Creating more representative AC-OPF datasets allows for the benchmarking and comparison of machine learning approaches over the entire AC-OPF feasible load space.

Future work:

- Understand the mapping of load profiles to optimal generator setpoints within unique active sets.
- Comparison of proposed machine learning approaches using OPF-Learn generated datasets

GitHub

Acknowledgements: I would like to thank my mentors Ahmed Zamzam and Kyri Baker for their support and guidance throughout this project. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI).