Generation of Synthetic Multi-Resolution Time Series Load Data via Generative Adversarial Networks

Andrea Pinceti, Lalitha Sankar, and Oliver Kosut

Arizona State University

MOTIVATION

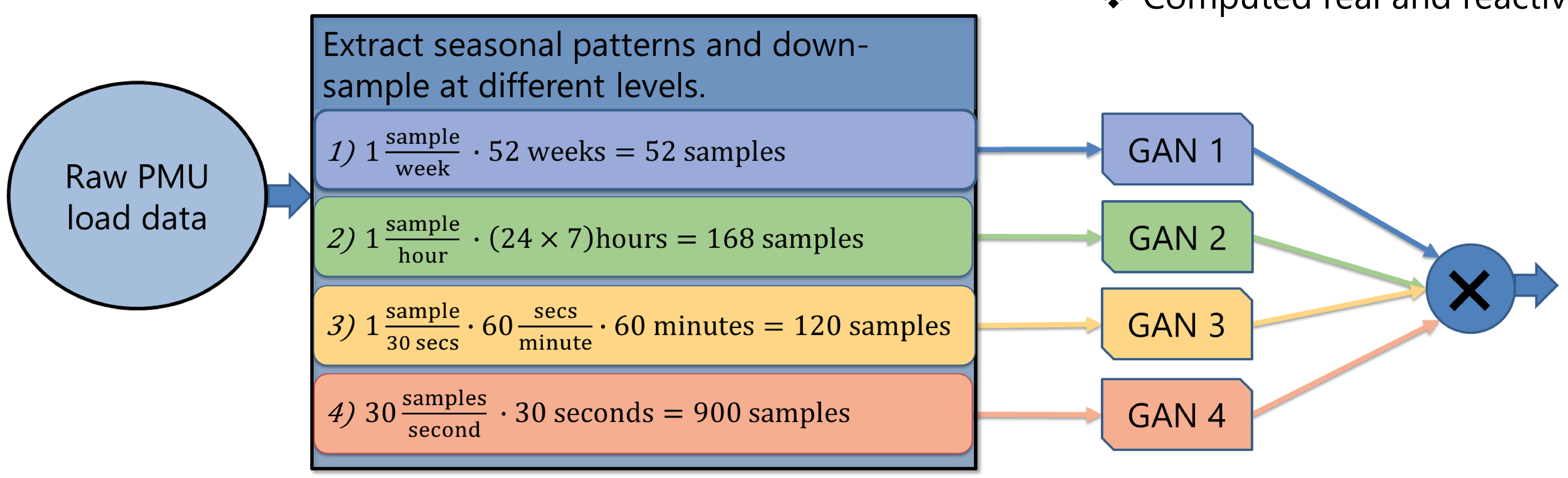
- A large portion of power system research focuses on the development of machine learning-based applications
- This represents a new paradigm: moving away from physics-based models to a data-centric analysis of system behaviors
- Power system datasets are scarce and not readily available

OBJECTIVE

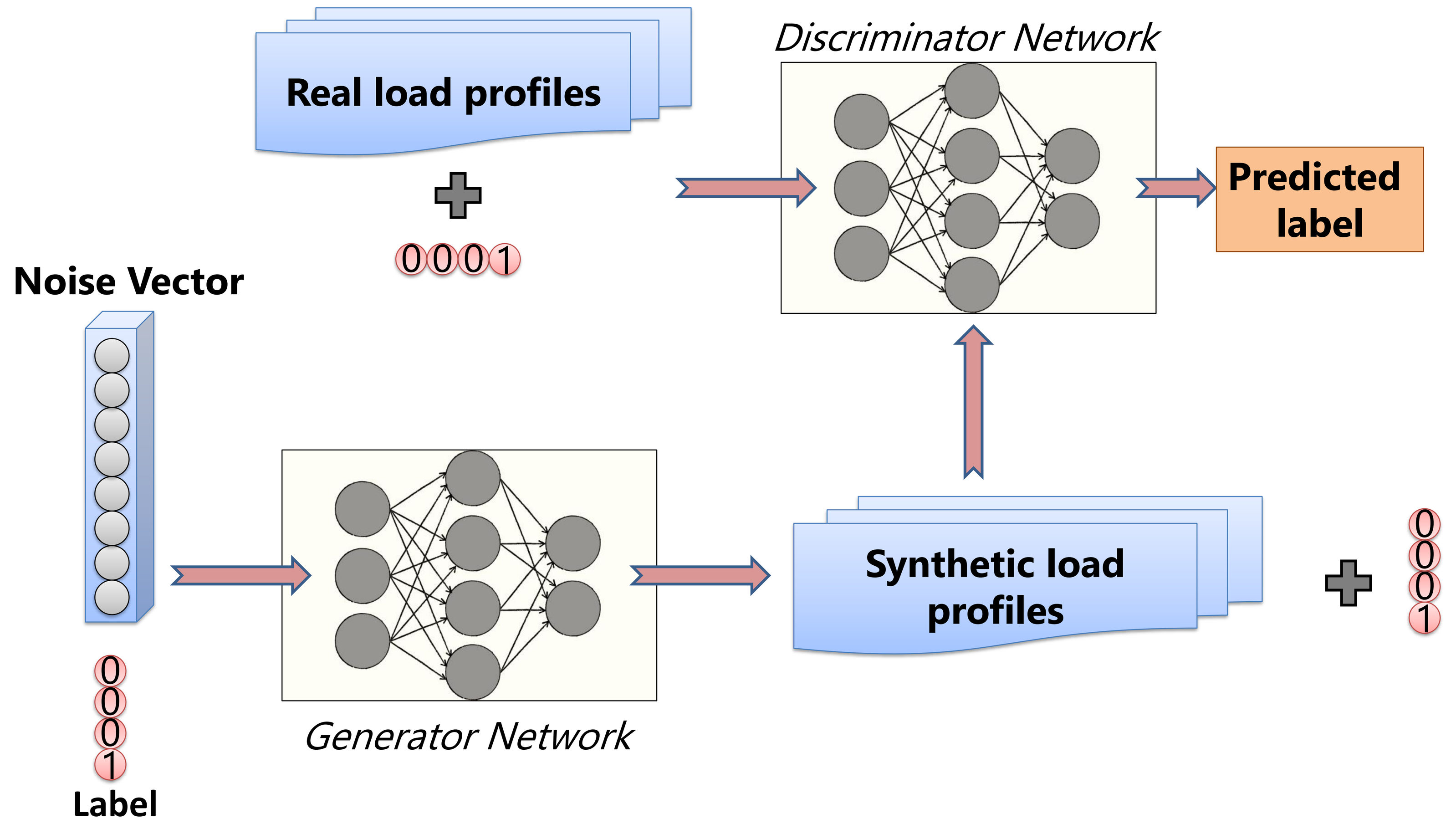
The goal of this project is to develop an open source tool for the generation of synthetic time-series load data at varying sampling rates and for different time lengths.

OVERVIEW

- Focus on system loads as the latent variable driving power system behaviors: loads depend on phenomena outside of the power system itself (consumer behaviors, weather, etc.)
- Leverage a proprietary PMU dataset to learn and model the behavior of system loads
- Use the learnt models to generate realistic load data at different time-resolutions and time-horizons
- The process consists of four steps:
  1. Raw PMU data (voltages and currents) is used to compute loads at different buses
  2. The load data at PMU speeds is down-sampled at four different levels: 1) year profile, 2) week profile, 3) hour profile, and 4) minute profile
  3. A generative adversarial network (GAN) model is trained individually for each level
  4. The fully trained GANs can be shared and used by researchers to generate data according to their needs



GENERATIVE ADVERSARIAL NETWORK

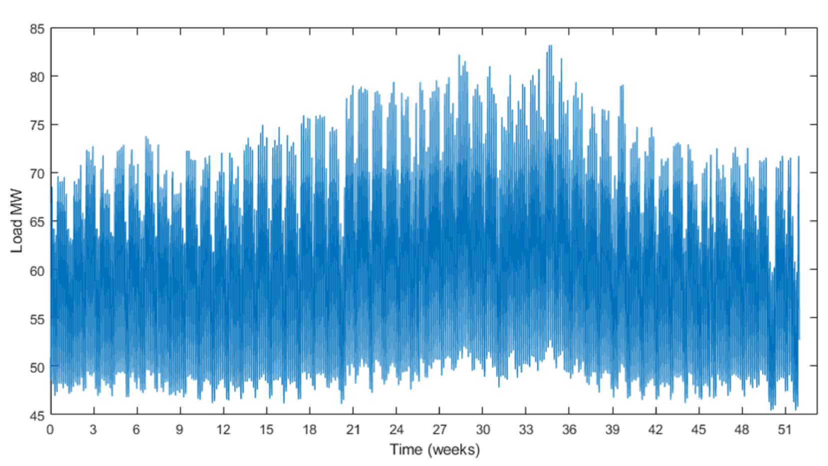


- GANs are a powerful algorithm for training generative models
- The generator G is trained to output realistic data by making it compete against a discriminator D whose job is to distinguish between real and generated samples
- Generation is conditioned on a label (load type, season, etc.)
- G and D are 1D convolutional neural networks able to capture the temporal correlation of load profiles
- The architecture of G and D is similar across the four time-resolution levels:
  - Two 1D convolutional layers with 4 to 16 filters
  - Two fully connected layers with 32 to 256 nodes
- Quality of the generated profiles is measured using multiple metrics:
  - Convergence of GAN training
  - Wasserstein distance
  - Power spectral density

DATASET DESCRIPTION

- 100 TB of raw PMU data
- 500 PMUs
- 3 years of data (2017, 2018, and 2019)
- Computed real and reactive power for 20 loads

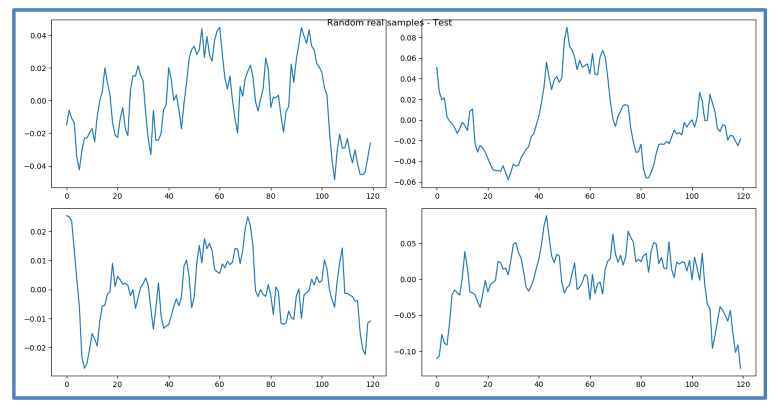
1 year of data at 30 samples/sec



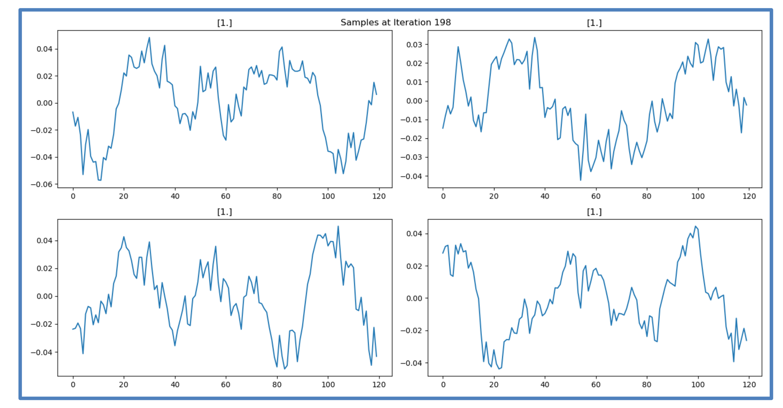
RESULTS

Hour profiles

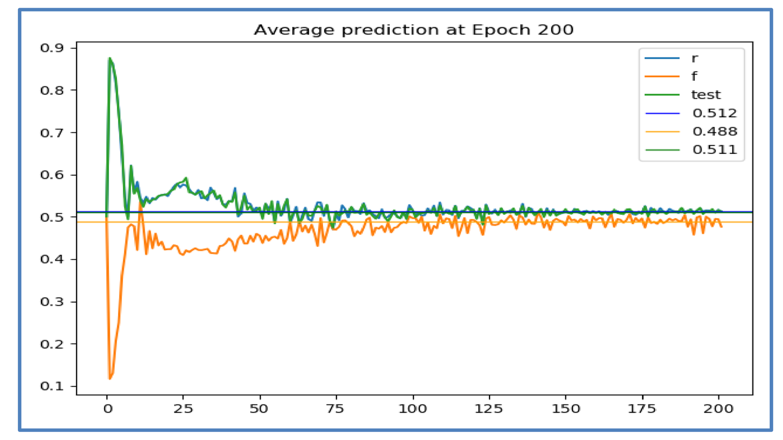
Real



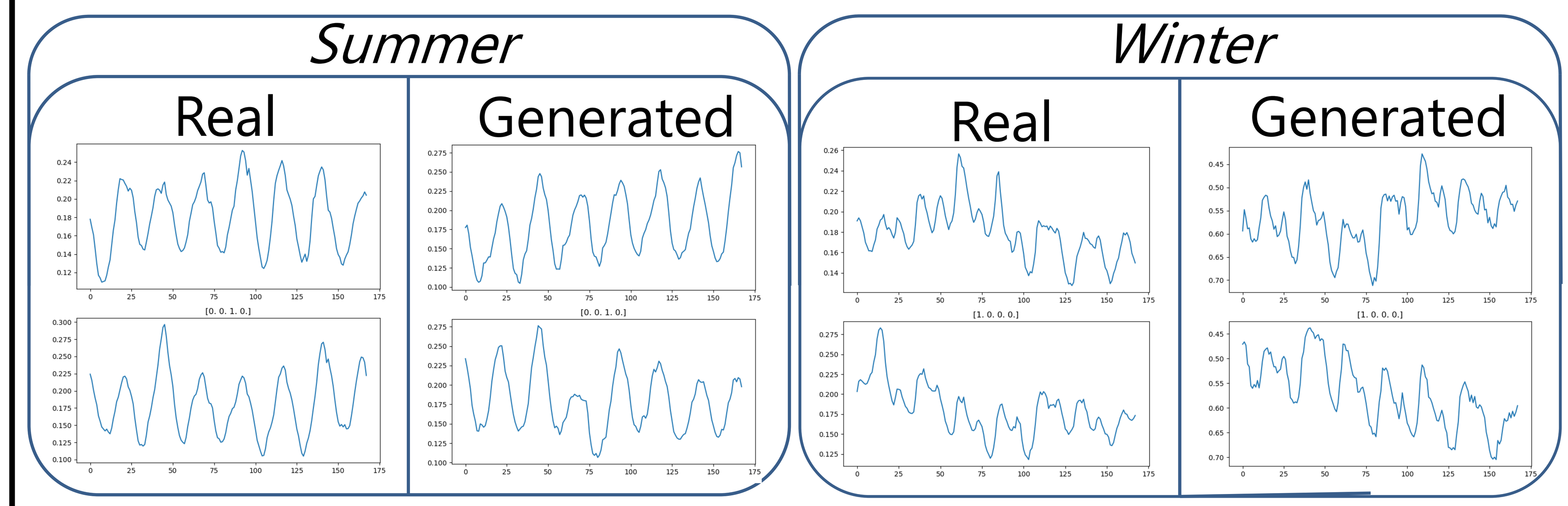
Generated

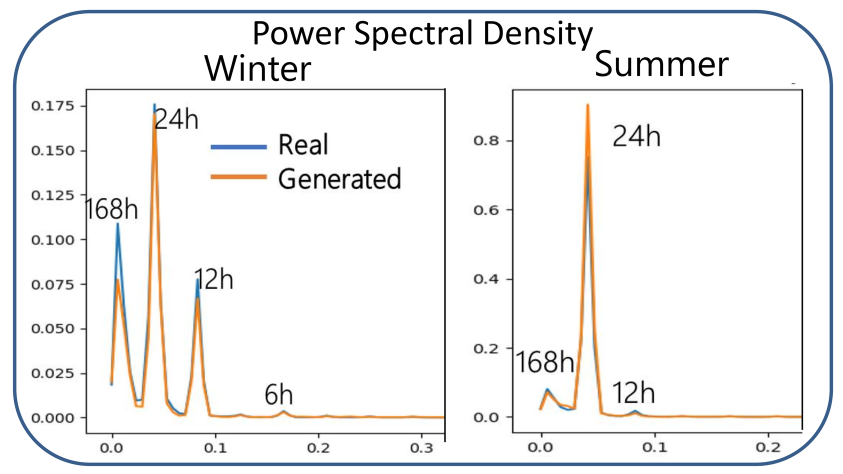


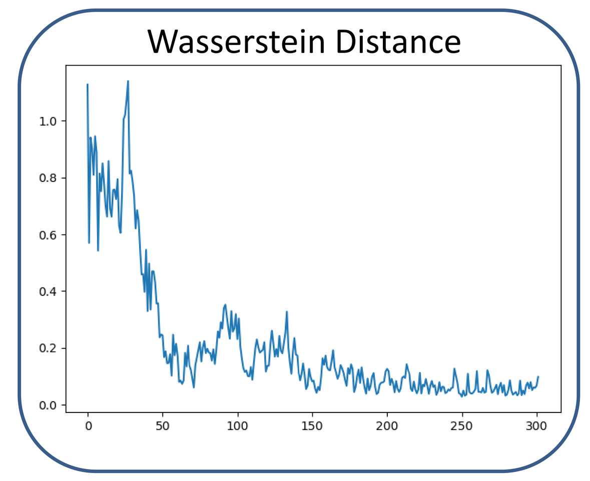
- Real and generated profiles are visually indistinguishable
- Generated profiles are unique
- Discriminator predictions show training convergence



Week profiles







- Generation of profiles conditioned on season
- Power spectral density of real and generated profiles matches closely
- Wasserstein distance between distributions of real and generated data converges to zero

BENEFITS AND POTENTIAL USES

- Automatic, easy-to-use generative model creates realistic load data for any grid model
- This approach allows for the generation of detailed load data which can be openly shared and used by the community

This material is based upon work supported by the National Science Foundation under Grants No. CNS-1449080 and OAC-1934766 and the Power System Engineering Research Center (PSERC) under projects S-72 and S-87.