National Aeronautics and
Space Administration

Intern Presentation

Incorporating Weather Data for PUE Prediction in LaRC Data Centers

Sam Briggs

01/16/24

CENTER OPERATIONS DIRECTORATE

About Me

- **Machine Learning Engineer Intern here at NASA Langley Research Center (LaRC) for Fall 2023**
- **Master's student in Computational Linguistics @ University of Washington (UW)**
  - Machine learning (ML) applications for natural language
  - Intersection between Linguistics and Computer Science
  - Branch of linguistics that includes natural language processing (NLP)
- **Previous:**
  - Bachelor of Arts in Mathematics & Linguistics @ UW
  - NLP/ML engineer Intern at Customer Relations Management start-up
- **Next:**
  - NLP/ML engineer Intern at the AMES Research Center for NASA
  - Graduating Spring 2024

Langley Research Center

1.

Fig 1. University of Washington. "The purple Block W logo of the University of Washington." 4 Feb, 2018, https://en.wikipedia.org/wiki/File:University_of_Washington_Purple_Block_W_logo.svg. Retrieved 3 Jan. 2024

Agenda

- **Project Introduction**
- **Weather Data:**
  - Sources
  - Automation
- **Power Usage Effectiveness (PUE) Modeling:**
  - Data engineering
  - Models
  - Model Optimization
  - Results

Project Introduction

- **My project had two main parts:**
  1. Collecting LaRC local environmental weather data
  2. Model/estimate PUE using Machine Learning (ML) models
     - Incorporate weather data into ML models

1.

Fig. 1: Image from Microsoft Version 2208

PUE

- **What is PUE?**
  - **Power Usage Effectiveness (PUE):** a metric used to measure the efficiency of a data center's energy consumption.
  - Ideally, all the energy of a data center would be used to power the computing equipment, rather than for the cooling, heating, lighting, etc.
  - Ideal PUE is 1.0

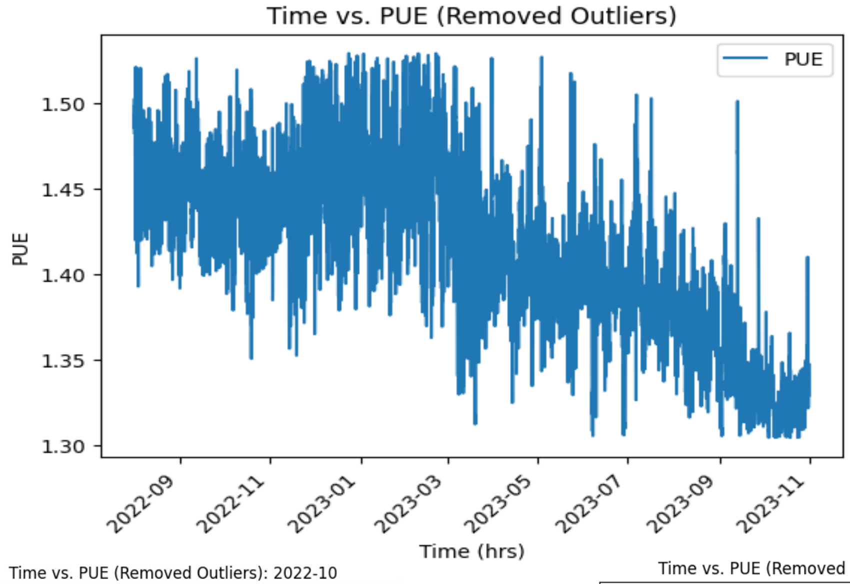
$$PUE = \frac{Total\ Facility\ Energy}{IT\ Equipment\ Energy} = 1 + \frac{Non\ IT\ Facility\ Energy}{IT\ Equipment\ Energy}$$

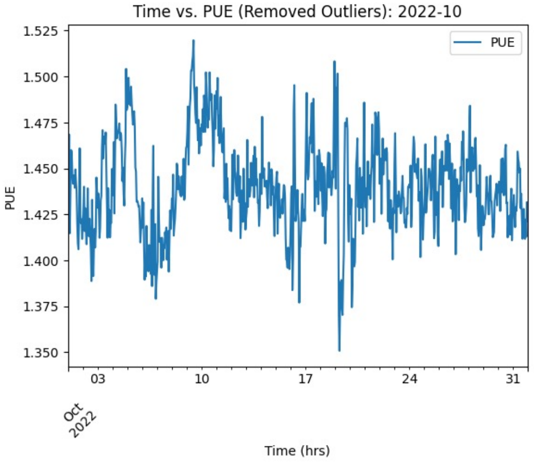
1.

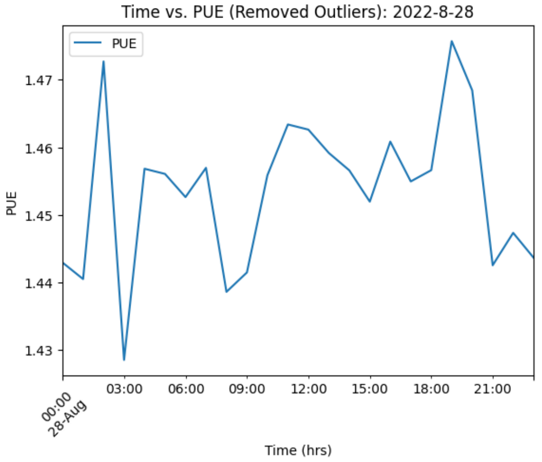
Fig. 1: Image from Microsoft Version 2208

Problem Motivation

- **Decrease PUE to create efficient data centers**
  - How do we best use pumps, chillers, place thermal tiles, etc.?
- **PUE fluctuates/varies. A lot!**
  - PUE is highly dependent on environmental factors
  - How do we flatten the fluctuations?
- **Data center consolidation**
  - Decisions about which data centers to shut down, which to expand, etc.

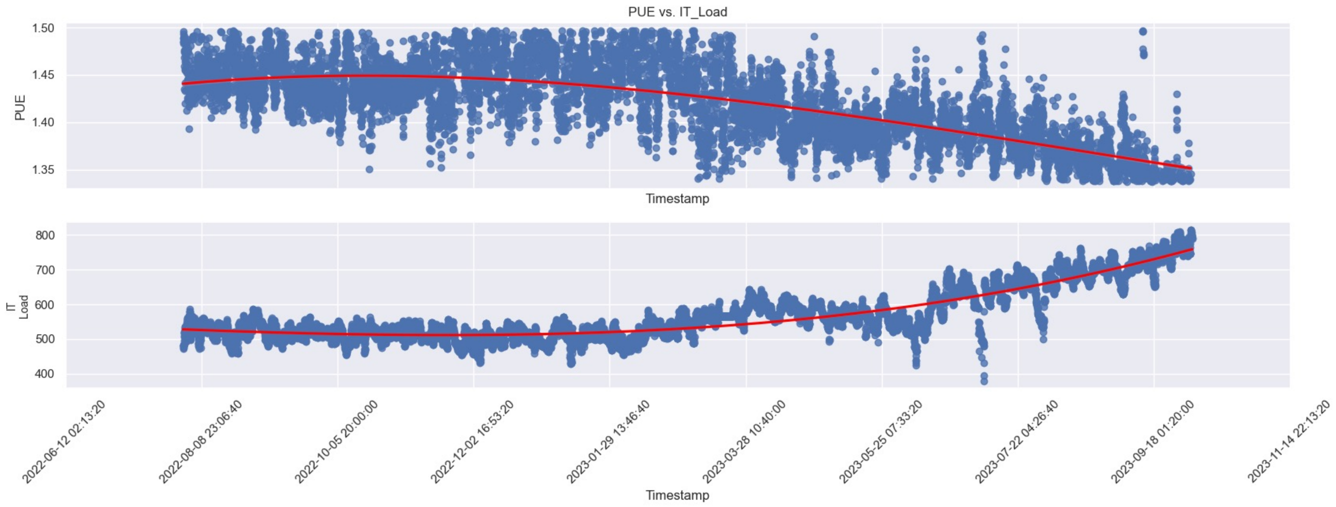


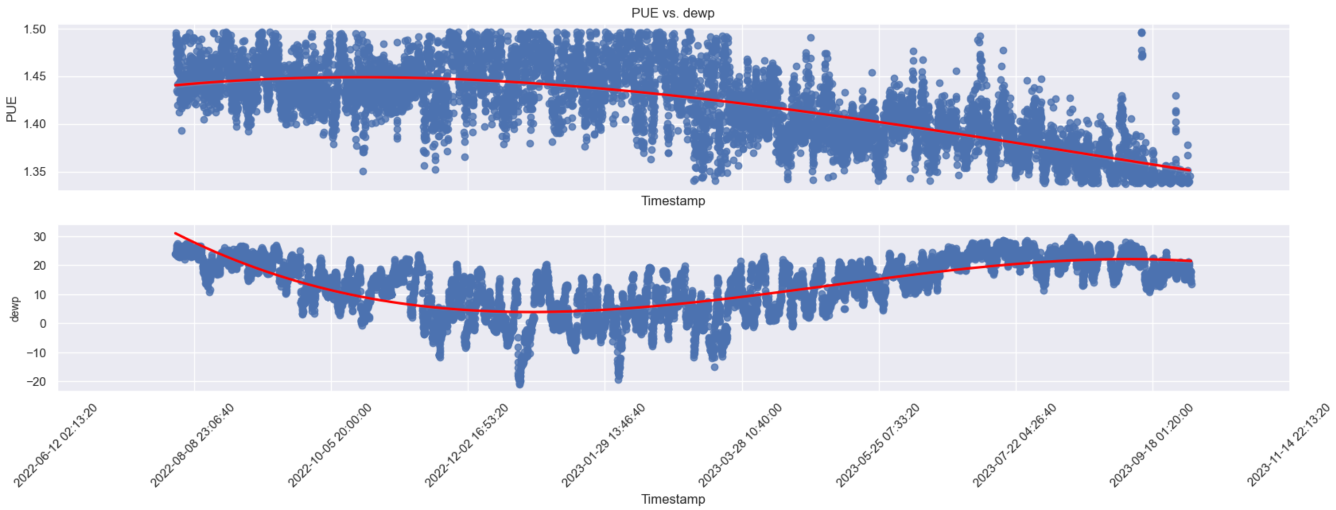




Problem Context

- ***Why is this a hard problem?***
  - PUE data fluctuates and varies greatly over time
  - Limited PUE and data center data
    - Only have data from 08/01/2022
  - Data center configuration constantly changing
    - IT Load correlated to PUE
    - Even thermal tile placement greatly impacts the PUE
  - PUE highly subject to the current environment, e.g., weather





Agenda

- Project Introduction
- **Weather Data Collection:**
  - Sources
  - Automation
- **PUE Modeling:**
  - Data
  - Model engineering
  - Model Optimization
  - Results

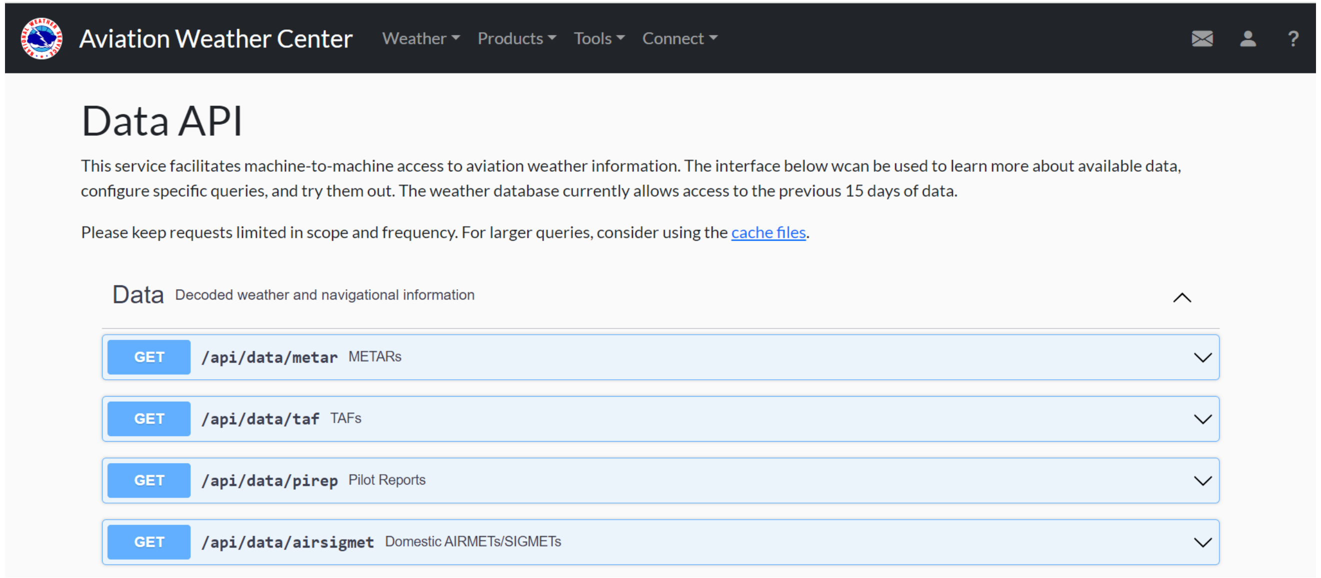
Weather Data Sources

- **Langley Air Force Base (LAFB)** (ICAO: KLFI)
  - Just next door to LaRC!
- **METeorological Aerodrome Report (METAR)**
  - Encoded **hourly weather data** pertinent to flying, e.g., temperature, dewpoint, wind speed, pressure, precipitation, etc.
  - **Automatically** produced weather encodings
- **Terminal Aerodrome Forecast (TAF)**
  - Issued about every **6 hours**
  - Forecast good for the **next 24 to 30** hours for about a **5 mile radius** of LAFB
  - Created **manually** by a weather forecaster

Fig. 1: Image from Microsoft Version 2208

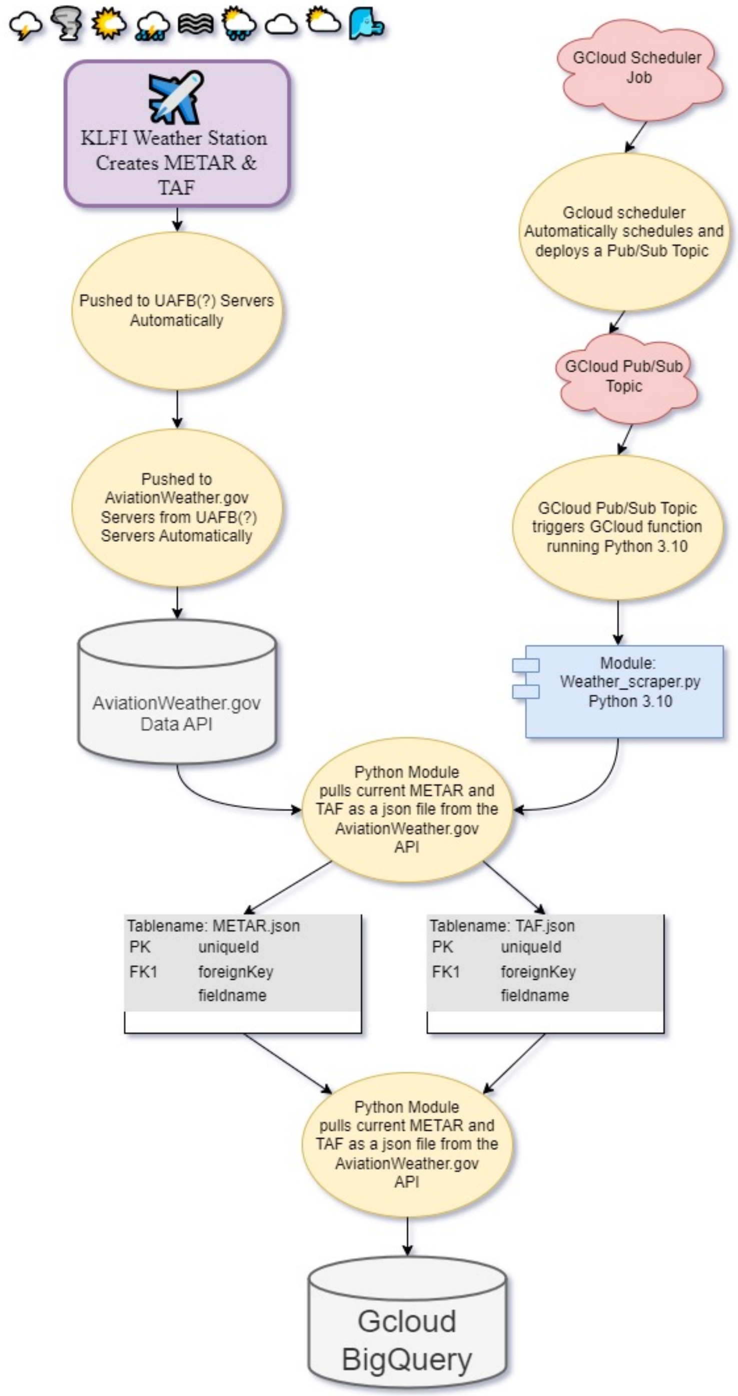
Aviation Weather Center

- **Aviation Weather Center (AWC) provides weather updates for the purpose of aviation.**
- **Data API**
  - The Aviation Weather Center hosts their own data API @ aviationweather.gov/data/api
  - Pull aviation related weather for **the last 15 days** using **Python code**!
  - Returns **Json** files!



Automated Weather Data Collection

- **Gcloud Scheduler**
  - Creates the Gcloud pub/sub topic once every hour
- **Gcloud pub/sub topics**
  - Triggers the Gcloud function
- **Gcloud functions**
  - To run python script which pulls weather data Json file from aviationweather.gov data API
- **Gcloud Source Repositories**
  - For versioning and version control
- **Gcloud BigQuery**
  - To store the weather data
- **TODO: Deployment!**

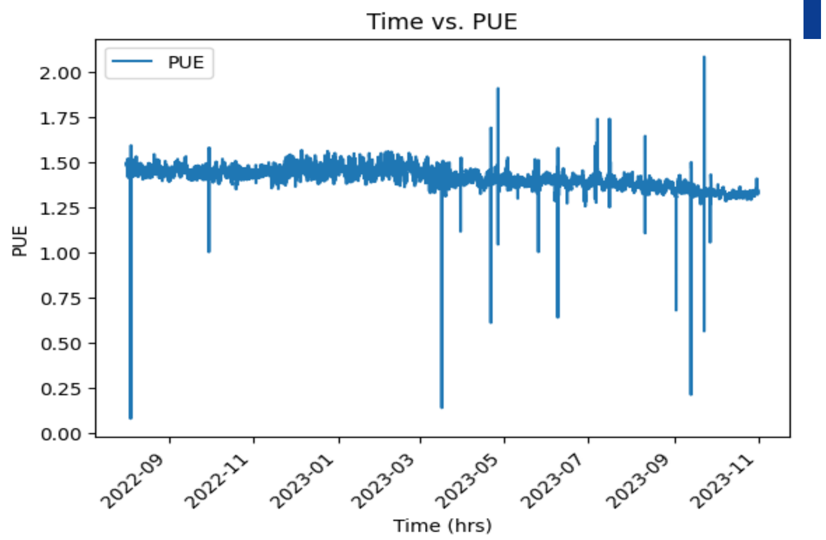


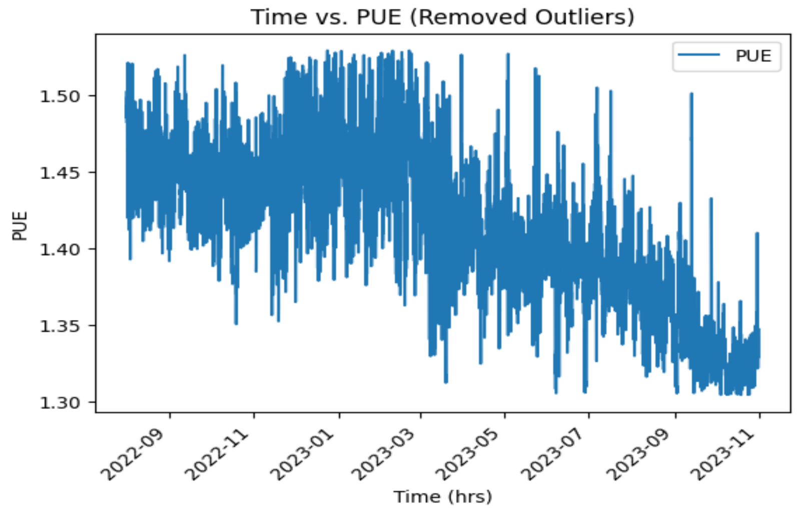
Agenda

- Project Introduction
- Weather Data Collection:
  - Sources
  - Automation
- **PUE Modeling:**
  - Data Engineering
  - Model engineering
  - Model Optimization
  - Results

Data Center Energy Consumption Data

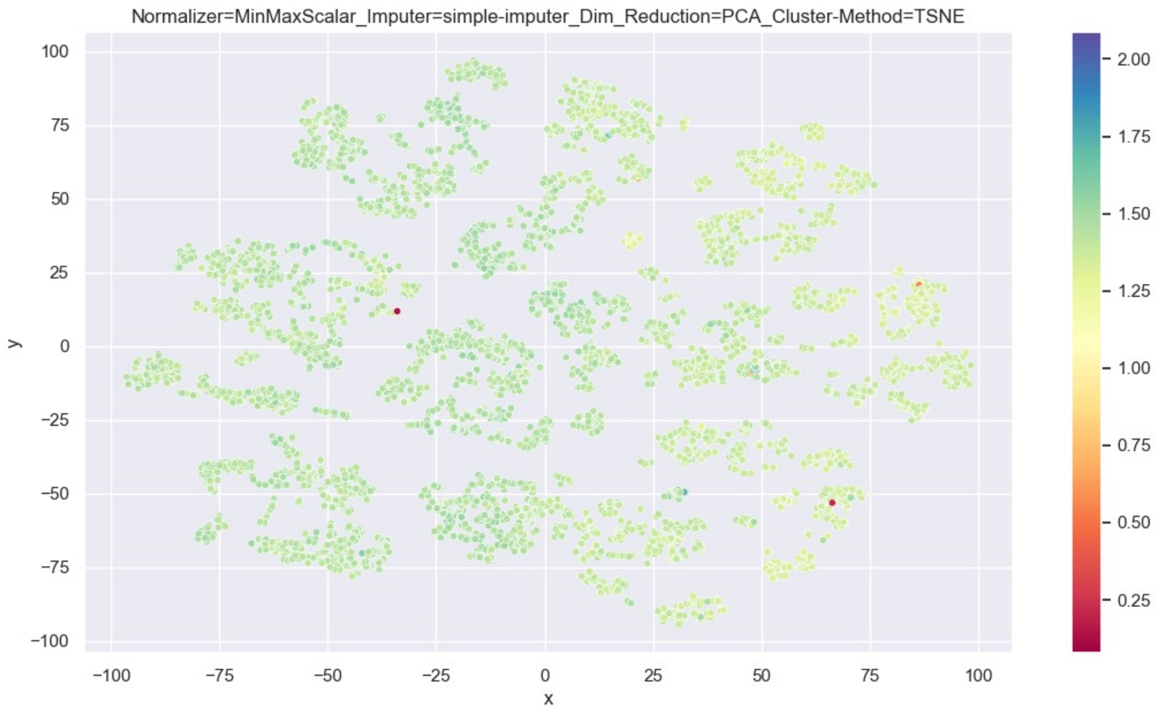
- **Just over 1 year of data center energy consumption data**
  - From 08/01/2022 – 10/31/2023
- **Data includes energy consumption and data center configuration**
  - water pumps, chillers, cooling fans, heaters, etc.
  - Most importantly, includes the PUE!
- **Sample hourly readings of PUE and data center conditions.**
- **Joined with the hourly weather METAR data!**
- **Data corresponding to PUE outliers in the bottom 2% and top 2% were removed**





Data Visualization

- **Can we make sense of the joint dataset?**
  - Dimensionality Reduction & Data visualization techniques…

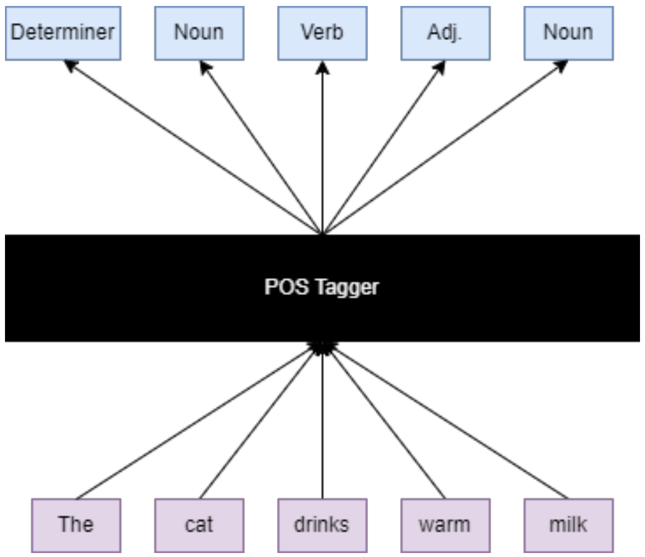


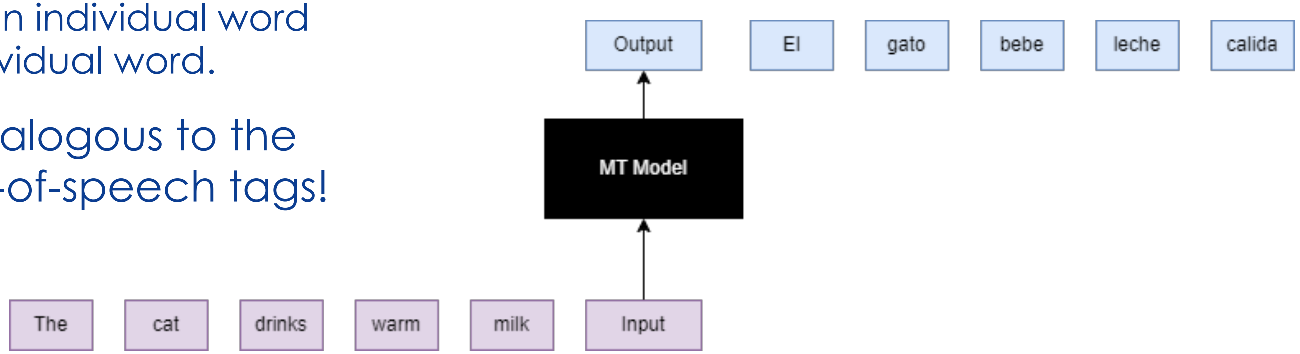
Feature Engineering

- **This data does NOT contain information about changes to the data center, e.g.**
  - IT Load
  - Replacement of old inefficient hardware
  - Reconfigured thermal tiles
- **Problem: How do we get a model to expect such changes in infrastructure, or usage patterns and continue to be able to accurately predict the PUE?**

Approach

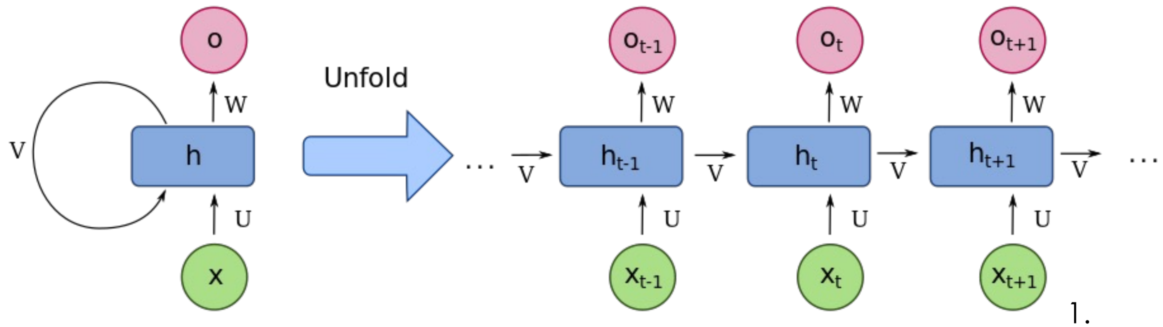
- **Problem: Given *n* timesteps worth of data, can we predict the next *n* timesteps of PUE?**
- **As a computational linguist…**
  - I thought of this as a **supervised** problem, we already have the PUE!
  - I thought of this problem as a **machine translation (MT)** problem, or **part-of-speech (POS)** tagging problem
  - One time stamp of data (both data center consumption & weather data) is analogous to a **word embedding**
    - Word embedding: A vector representation of an individual word capturing semantic information about that individual word.
  - The sequence of next n timesteps of PUE is analogous to the translated sentence, or the sequence of part-of-speech tags!



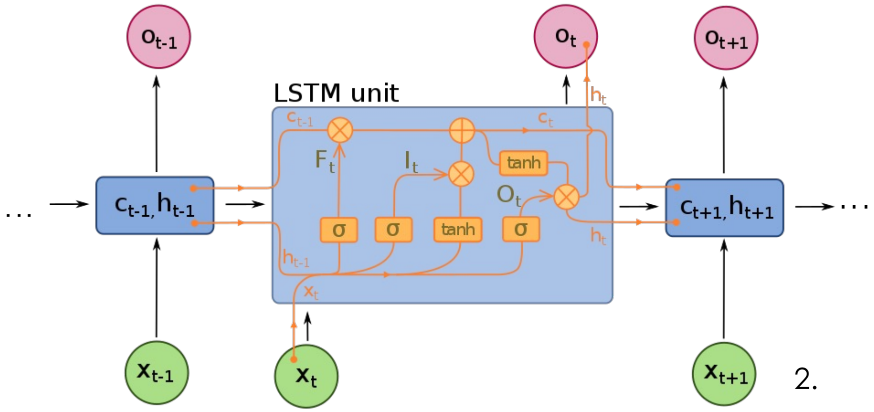


RNNs

- **Recurrent Neural Networks (RNN) work with sequential data**
- **Two types of RNNs: Vanilla RNN vs. Long Short-Term Memory (LSTM)**



1.



2.

Fig 1. fdeloche, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons, https://upload.wikimedia.org/wikipedia/commons/b/b5/Recurrent_neural_network_unfold.svg
Fig 2. fdeloche, CC BY-SA 4.0 <https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons, https://upload.wikimedia.org/wikipedia/commons/6/63/Long_Short-Term_Memory.svg

Models

- **Vanilla Recurrent Neural Network (RNN)**
  - One weight matrix
  - Con: Vanishing/Exploding Gradient Problem
  - Con: Only able to rely on the previous timestamp
- **Sequence to Sequence Long Short-Term Memory (Seq2Seq LSTM)**
  - Seq2Seq architecture contains an encoder and a decoder (two weight matrices instead of one)
  - LSTM: Solves Vanishing/Exploding Gradient Problem
  - LSTM: Memory cell retains information from much earlier timestamps
- **Both**
  - Recurrent Neural Network architectures
  - Input: timeseries sequence of N timestamps of data
  - Output: corresponding predicted sequence of PUE
  - Loss: Used Mean Squared Error

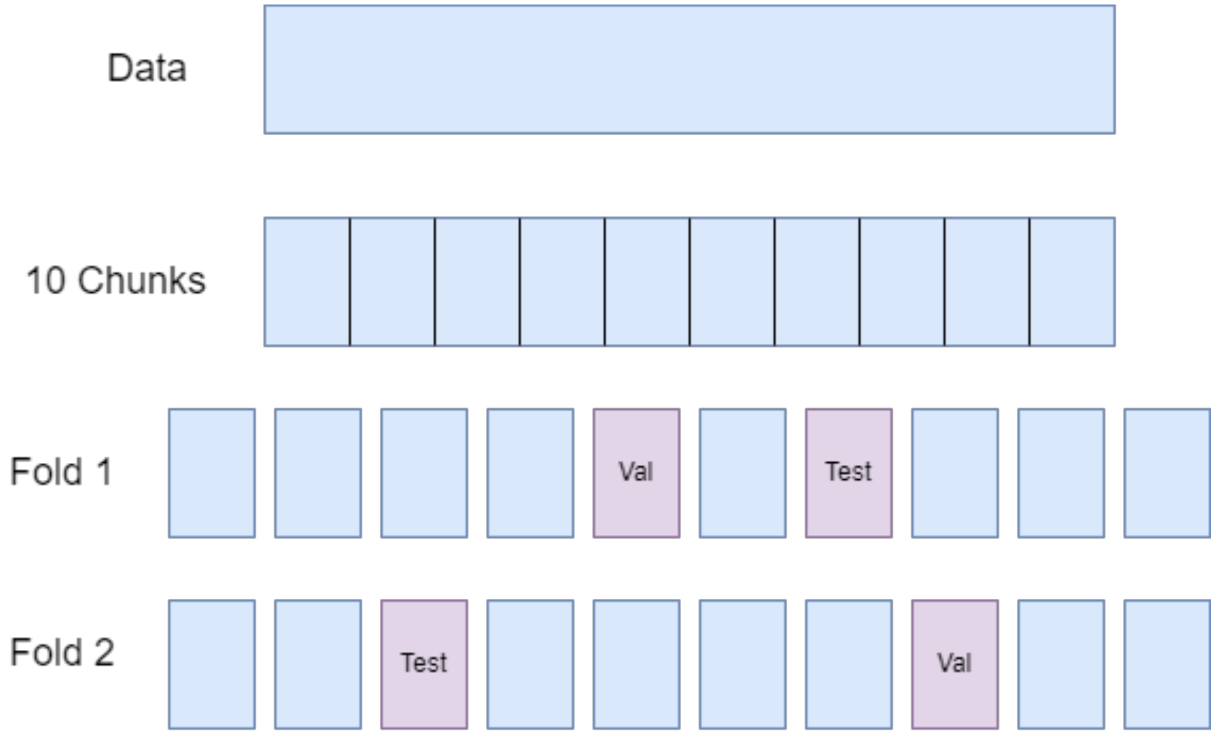
Vanishing Gradient:

$$\mathrm{w}' = w - \alpha \nabla L(w)$$

$$\nabla L(w) = 0$$

Hyper-parameter Tuning: Validation

- **Testing different hyper-parameter combinations**
- **K-Fold Cross Validation (K = 5)**
- **Split the data into 10 even chunks**
  - Choose one chunk as the **test** set
  - Choose a different chunk as the **validation** set
  - Remaining 8 chunks concatenated as **train** set
- **For each fold:**
  - Trained on train set
  - Inferenced on the validation and test set
  - Loss scores between folds were aggregated (mean) to account for run variability

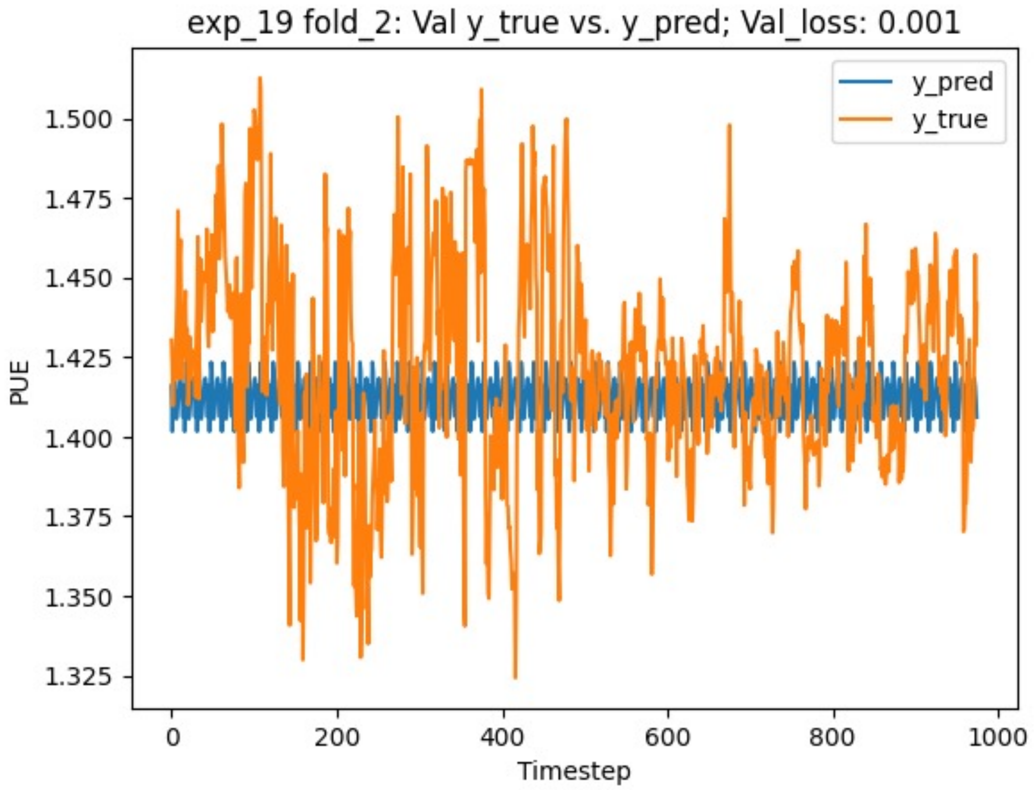


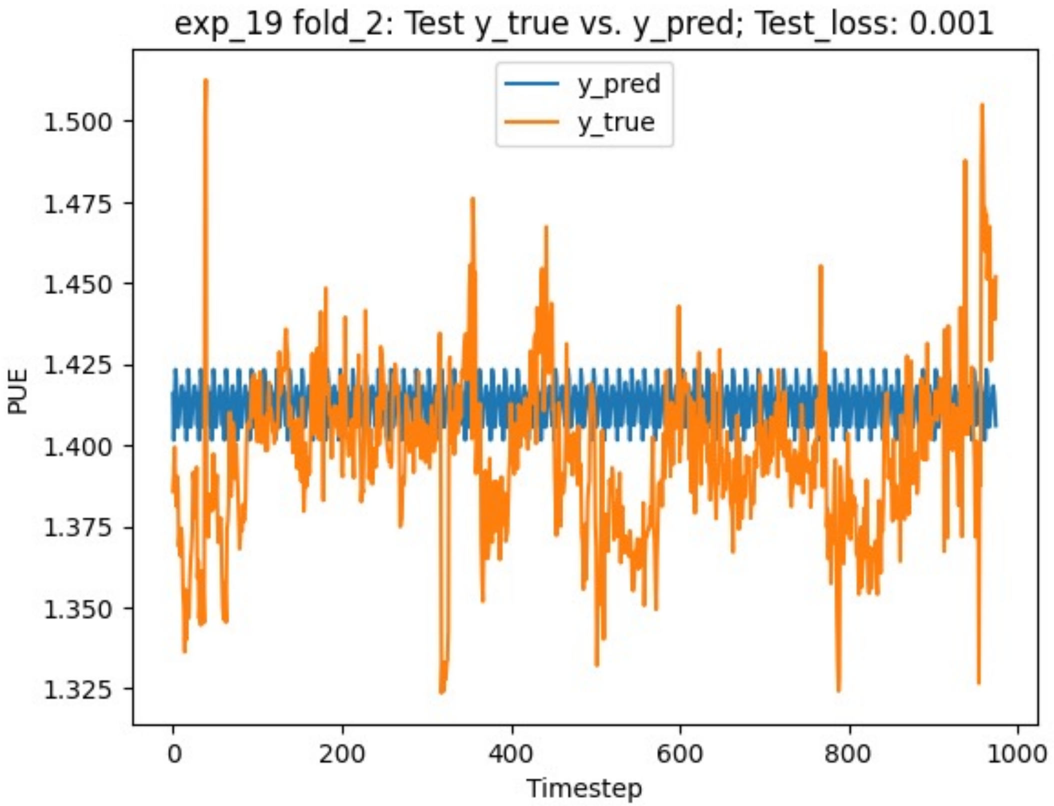
Hyper-parameter Tuning: Random Search

- **Why Random Search**
  - Hyper-parameter search space is very large
  - Using K-Fold Cross Validation (multiplied # of combinations by K)
  - Only used CPU
- **Random search randomly picks a subset of hyperparameter combinations**
  - Vanilla RNN: tried 20 different models
  - Seq2Seq LSTM: tried 88 different models

Results

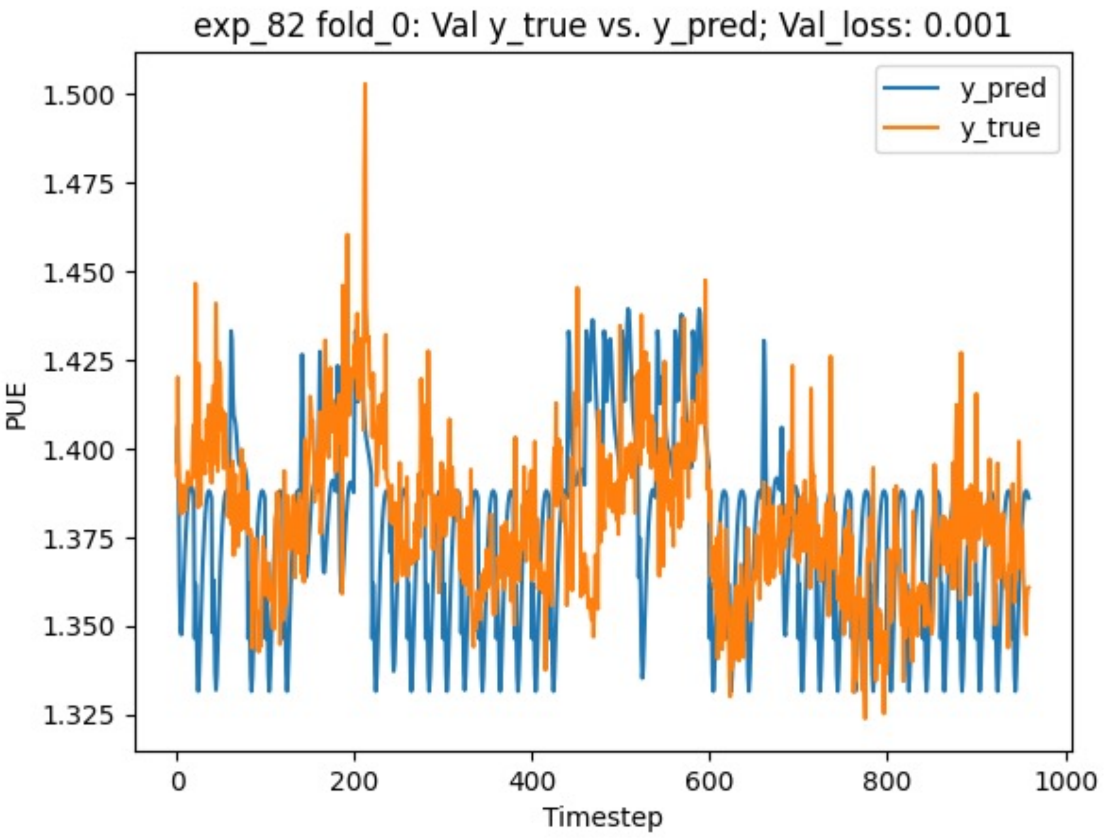
- **Sometimes good loss scores do not mean good results**
  - Model would perform well when only predicting the mean PUE score over the whole val or test set

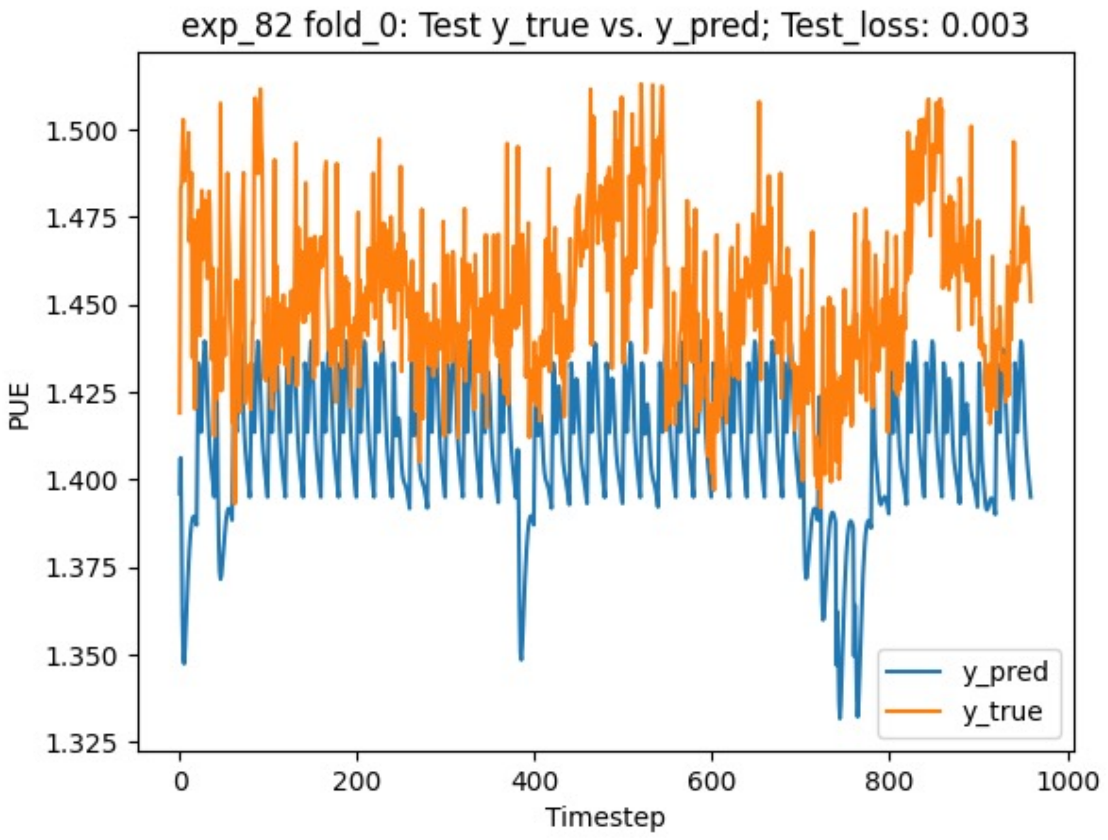




Results II

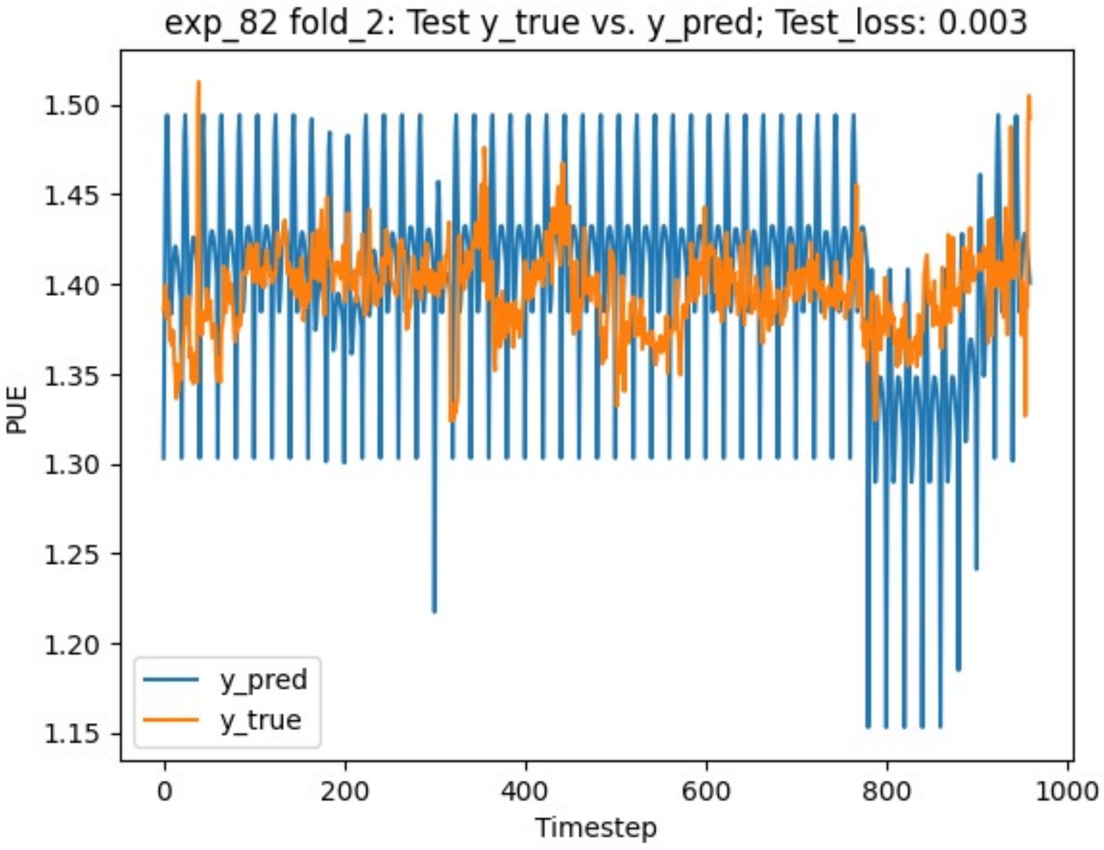
- **Sometimes results varied very widely between folds**

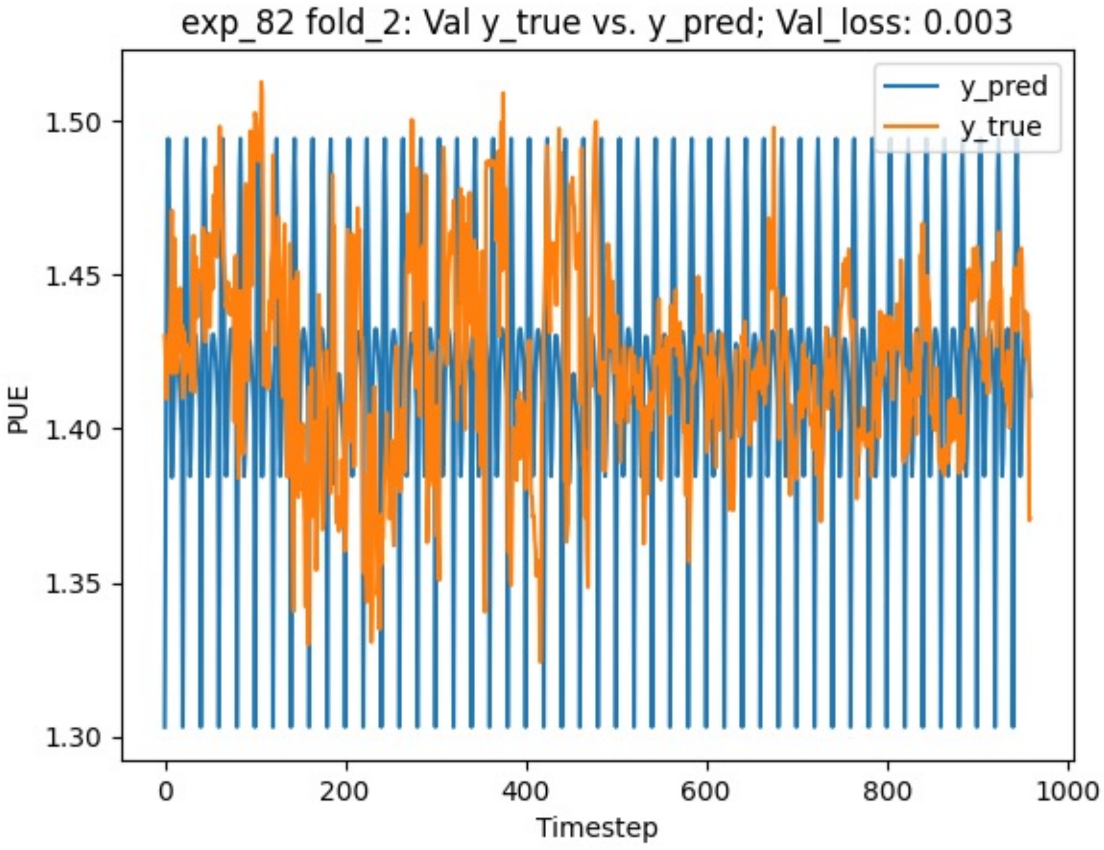




Results II

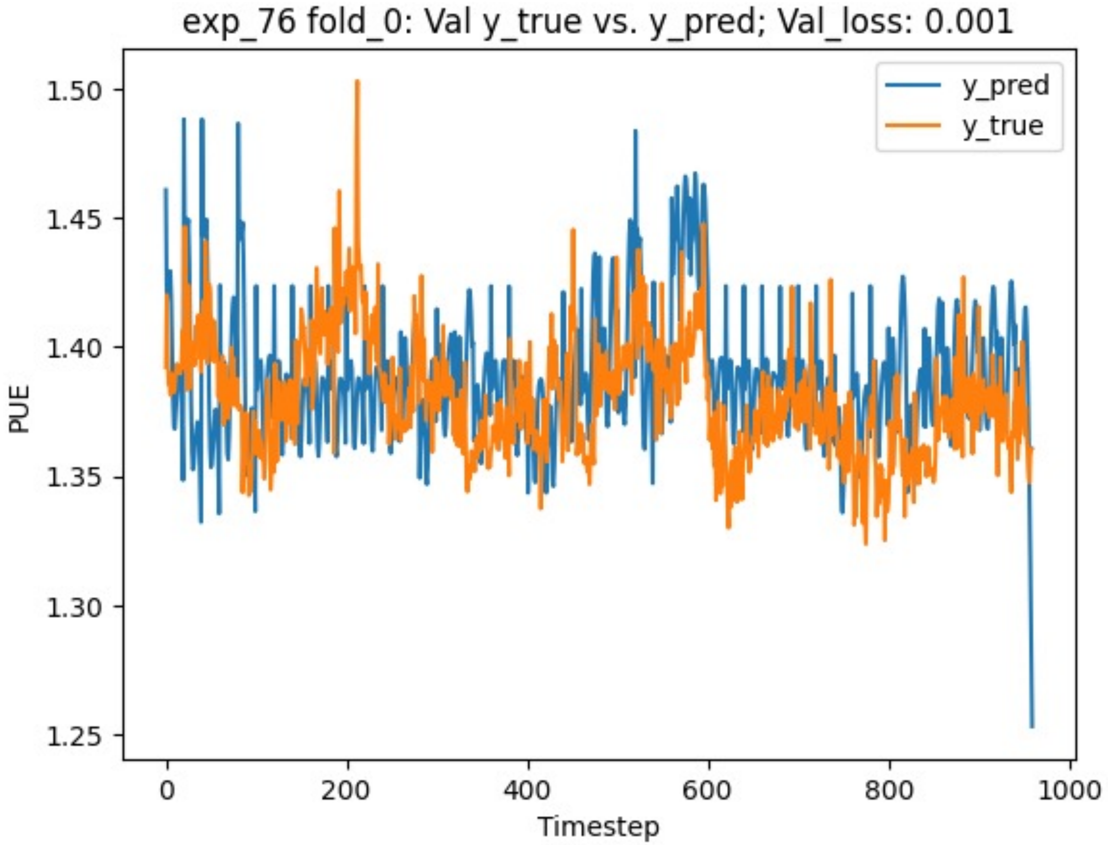
- **Sometimes results varied very widely between folds**

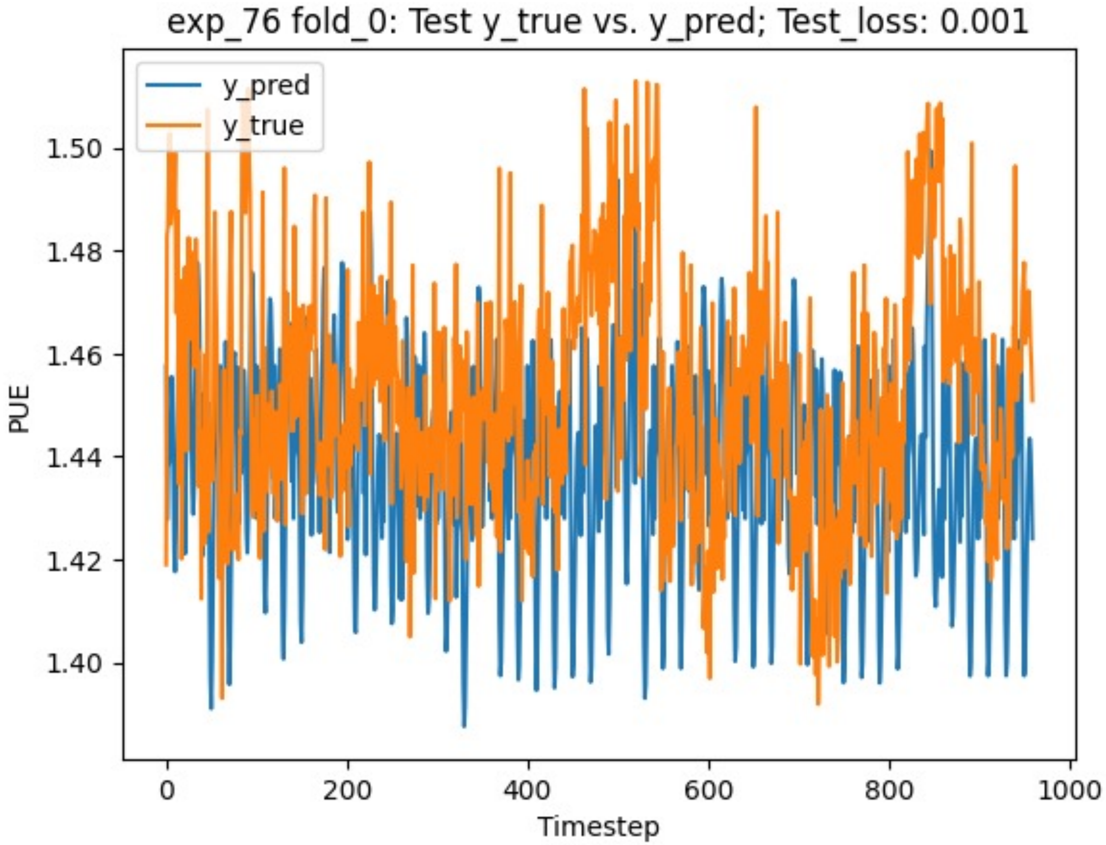




Results III

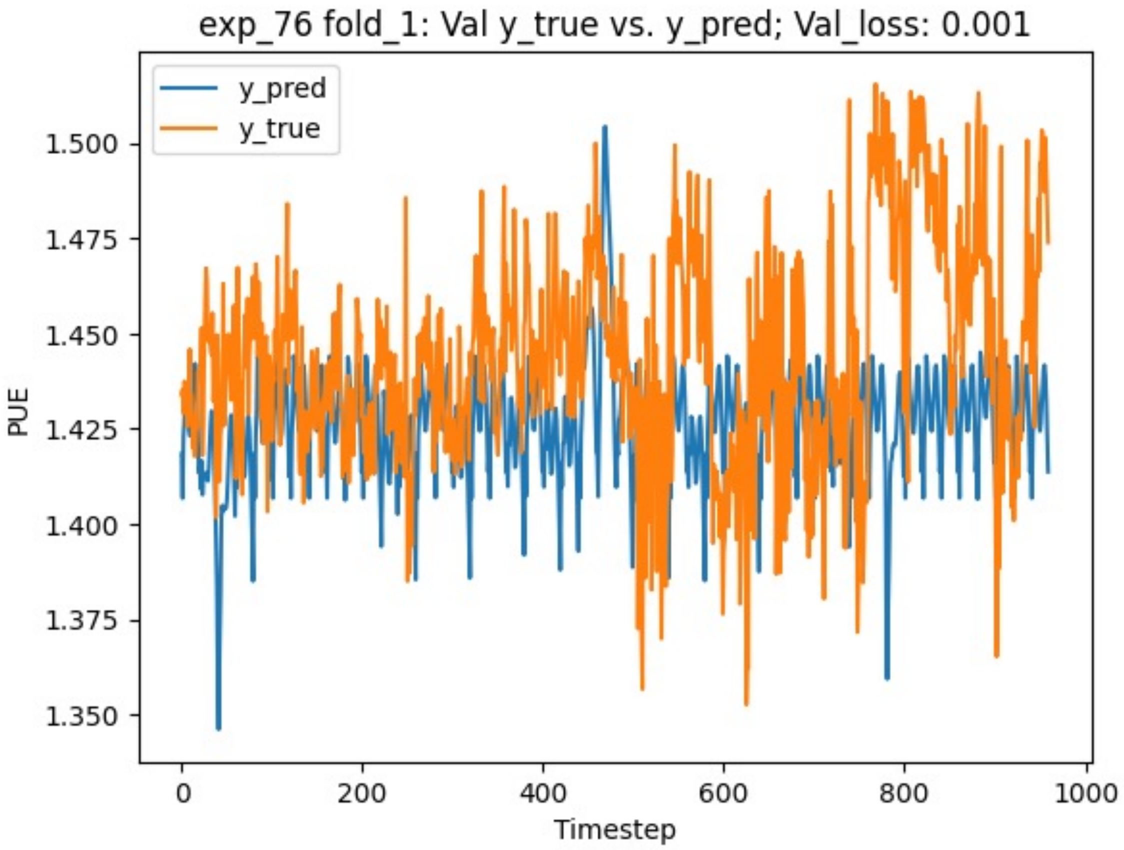
- **Sometimes results look pretty good**

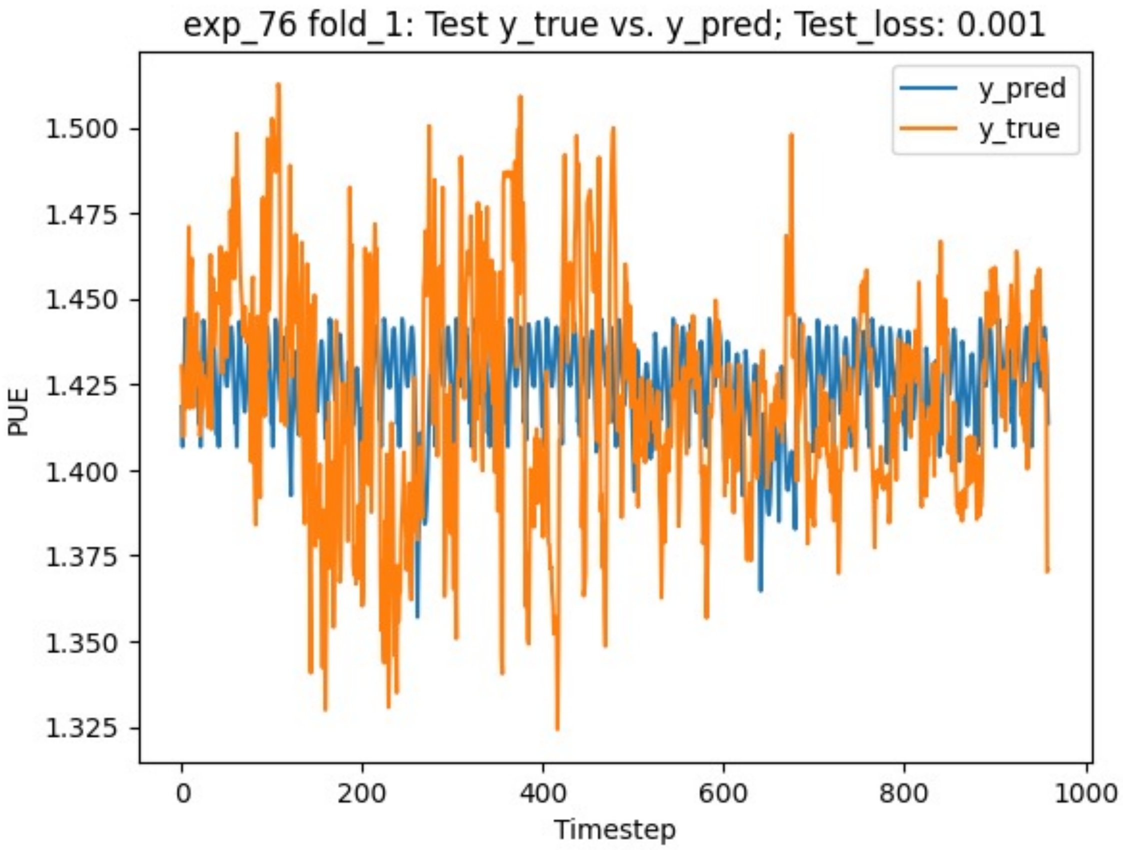




Results III

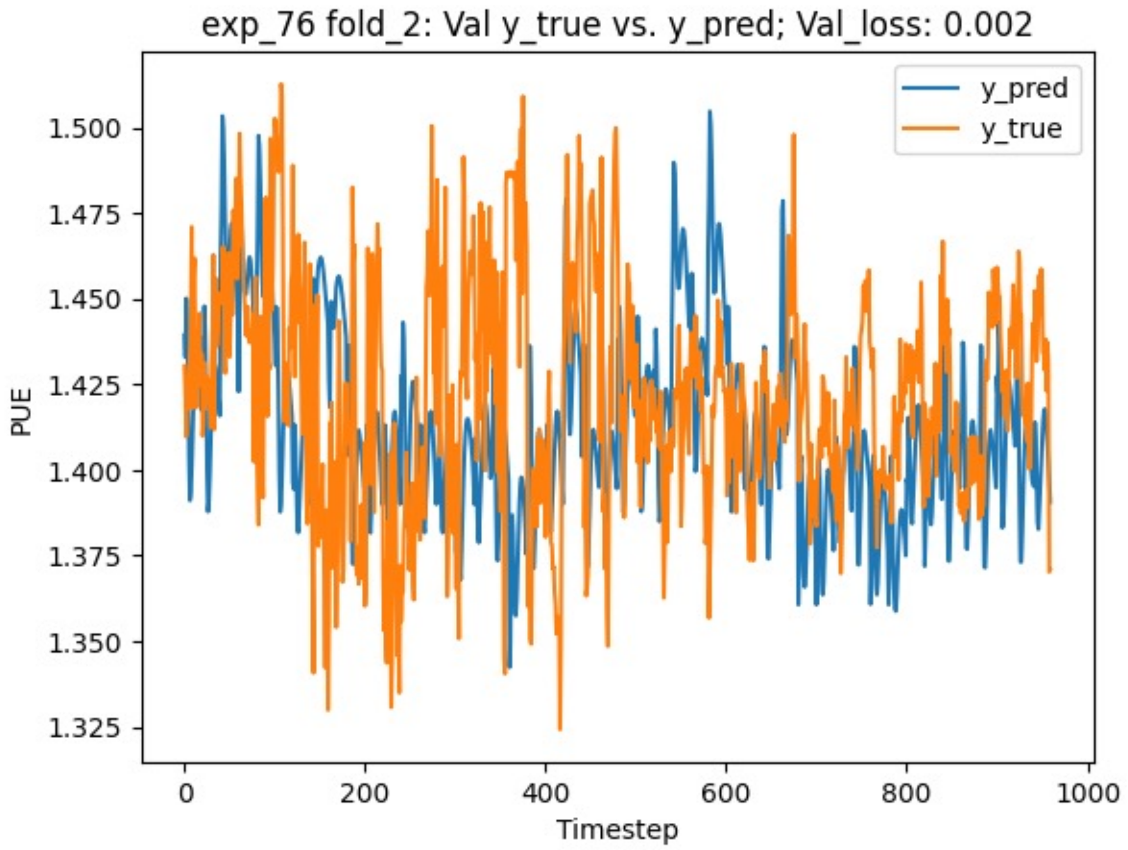
- **Sometimes results look pretty good**

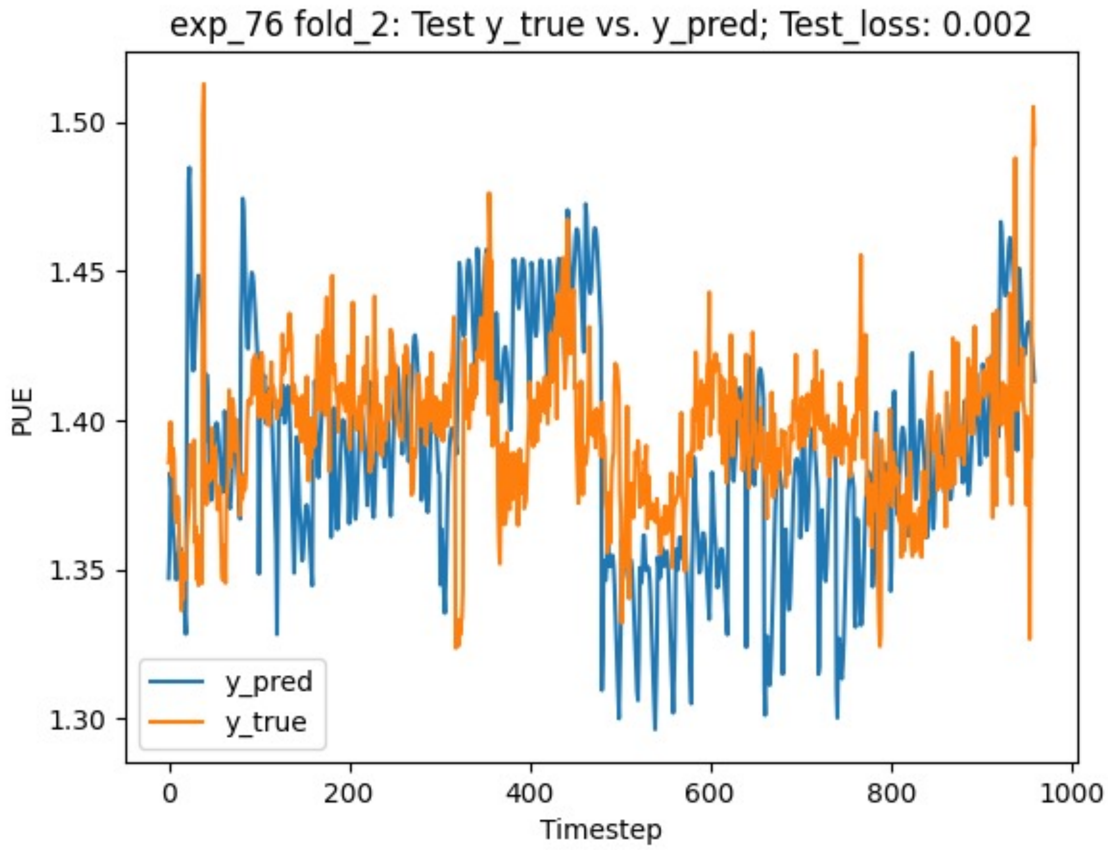




Results III

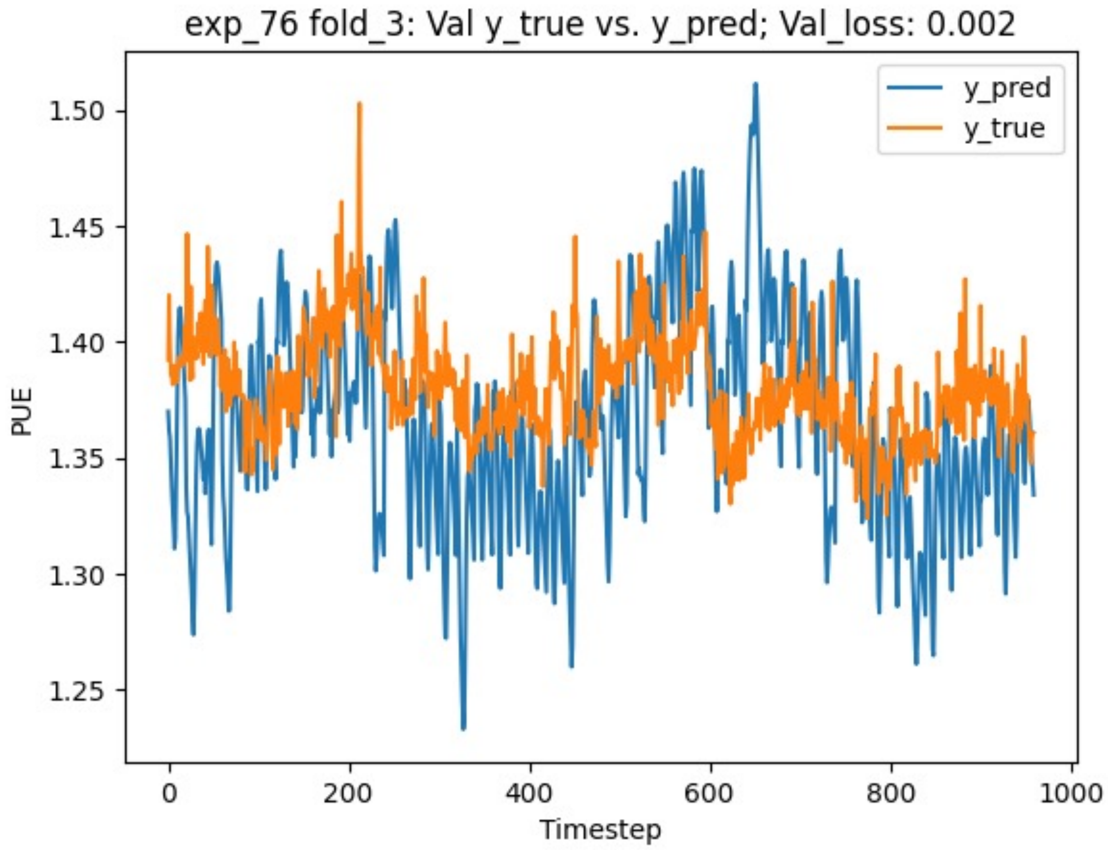
- **Sometimes results look pretty good**

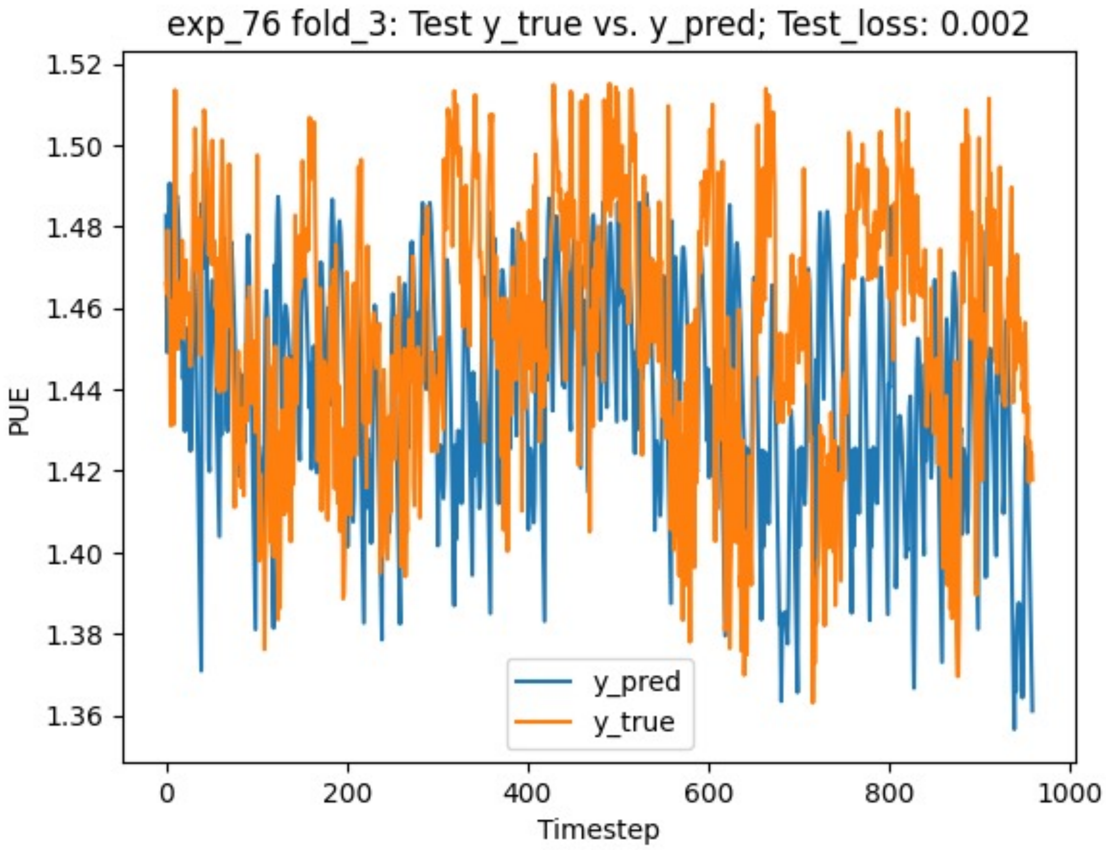




Results III

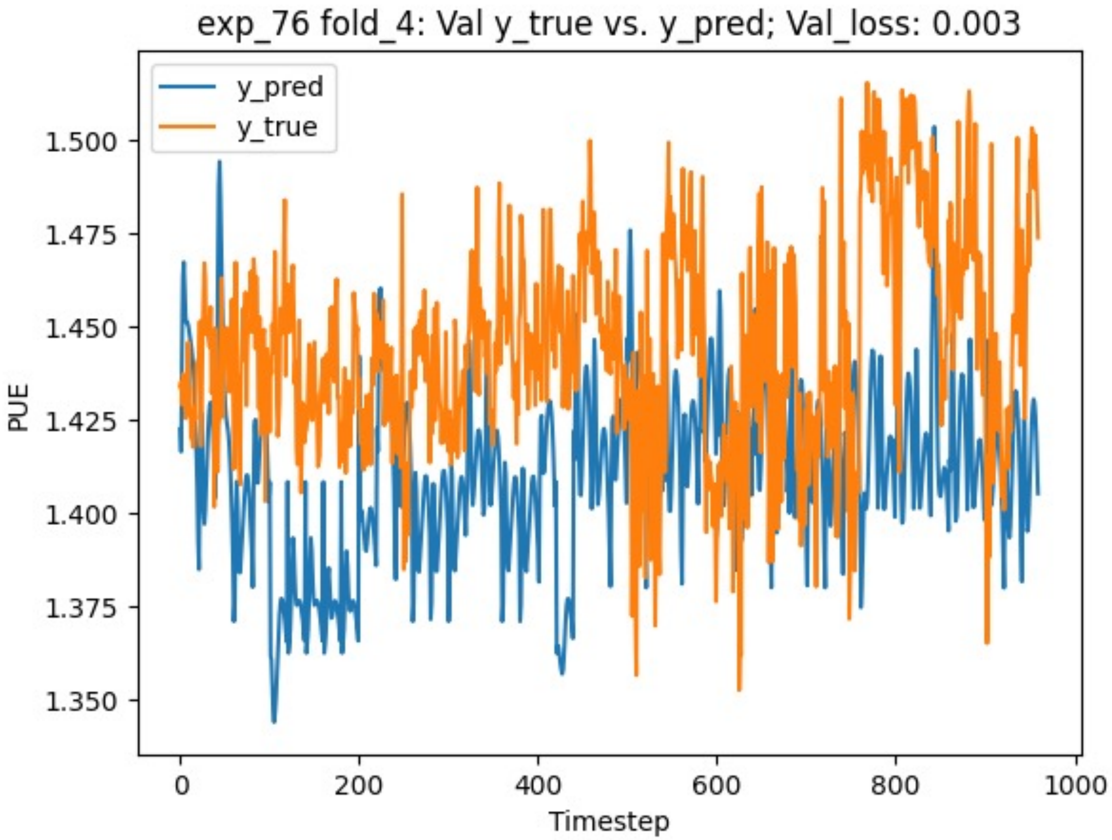
- **Sometimes results look pretty good**

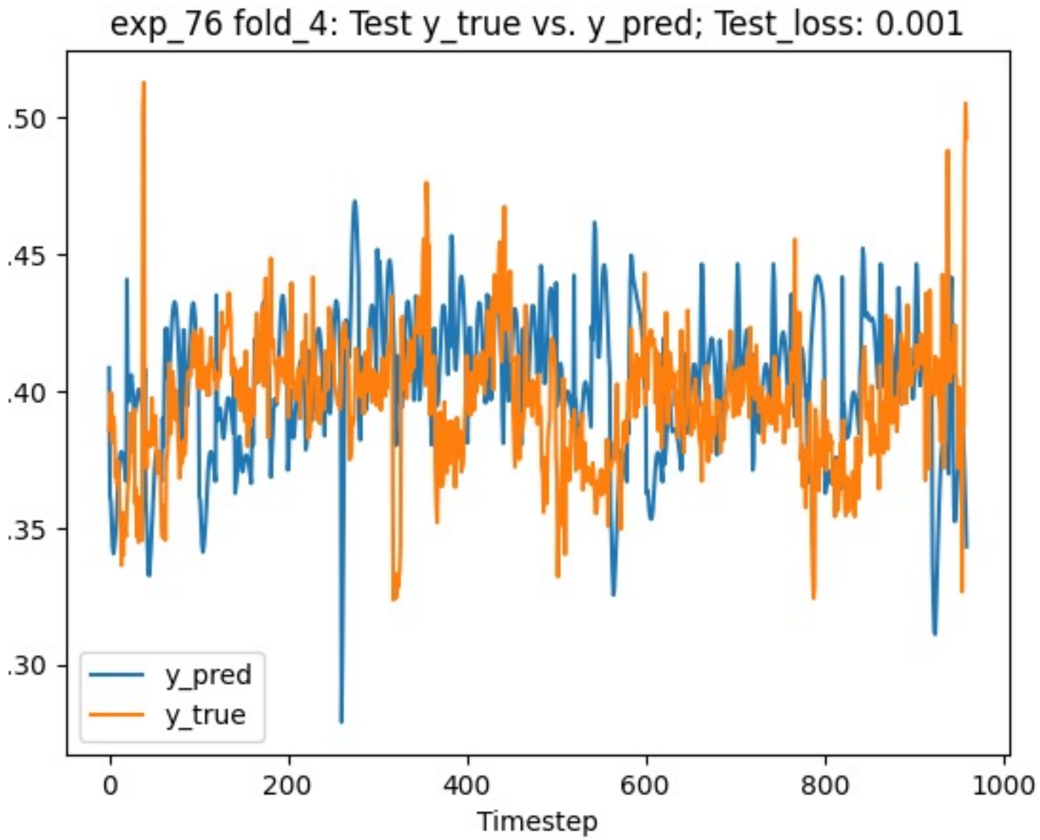




Results III

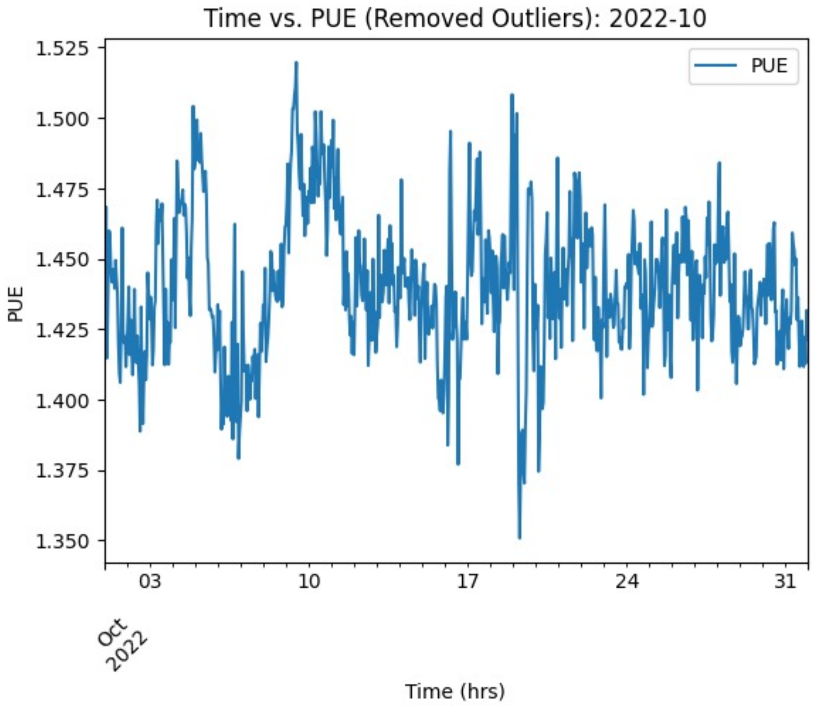
- **Sometimes results look pretty good**

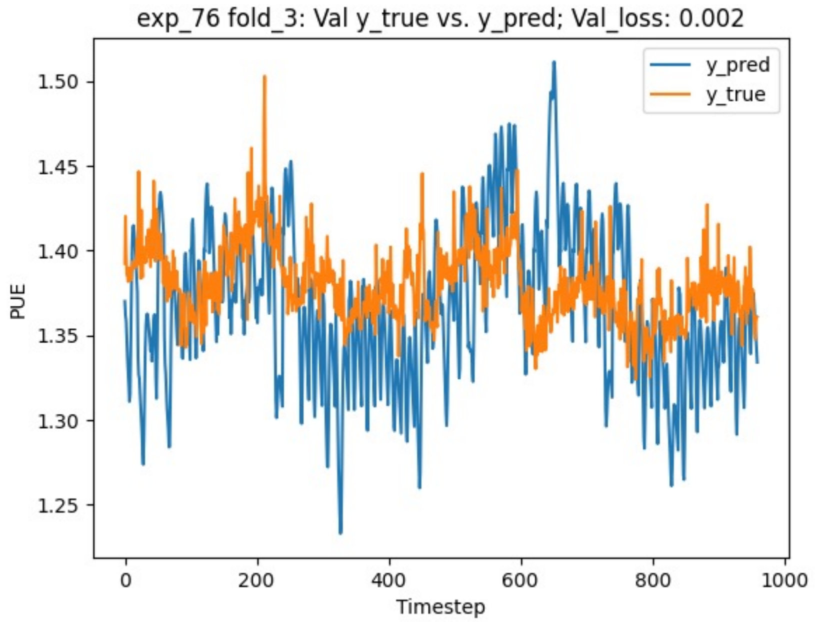




Results IV

- You may have noticed the predicted PUE fluctuates quite a bit and is very *"spiky"*
- Hypothesis: The PUE data fluctuates quite a bit → model has learned this behavior
- This could be a problem





Future Work

- **Do more thorough examination of salient and pertinent weather features**
- **Other inputs**
  - Is it a work day? Is it evening?
- **Figure out which hyper-parameters affect the model performance the most**
- **Try other models such as a transformer model**
- **Consider new model features to account for data center reconfiguration**
  - e.g., What happens to the PUE if we add a new data server?

Acknowledgements

- **Many thanks to…**
  - Charles Liles, Project Mentor
  - Langley Transformation Initiative
  - LaRC Center Operations Directorate
  - LaRC Integrated Operations Center
  - Guardians of Honor
  - Mark Pugh, Stephen Fehr, Danny Morris for meeting Charles and I to discuss the data center energy consumption and how to better model PUE
  - Mike Mueller and his son Thomas Mueller for getting me a working NASA laptop after everything that happened

Feel free to reach out!

- ***99samb@gmail.com***
- ***briggs3@uw.edu***
- ***Linkedin.com/in/samuelbriggs31***