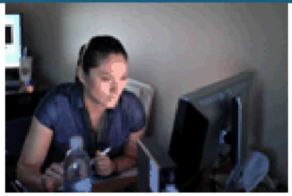
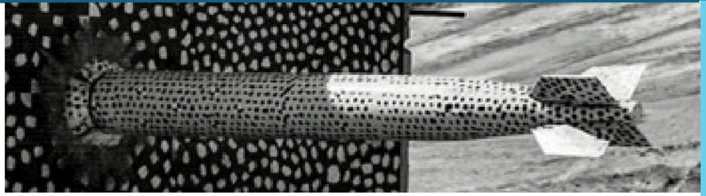
SAND2020-5942C

Advances in Machine Learning for Safeguarding a PUREX Reprocessing Facility





Presented By

Nathan Shoman and Benjamin Cipiti

SAND-

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.

Outline

- Motivation
- PUREX process model
- Machine learning background
- Machine learning framework
- Current results
- Conclusions and future work

Elimination of on-site laboratories at reprocessing facilities is a long-standing goal of the IAEA

Traditional safeguards for large throughput facilities are challenging

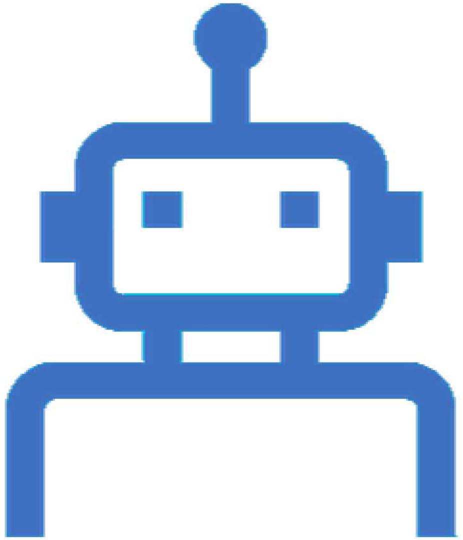
- Require small measurement uncertainties
- Expensive and time consuming
- Can require onsite laboratory
- Mature and proven

Data science driven approaches could improve safeguards, but have challenges of their own

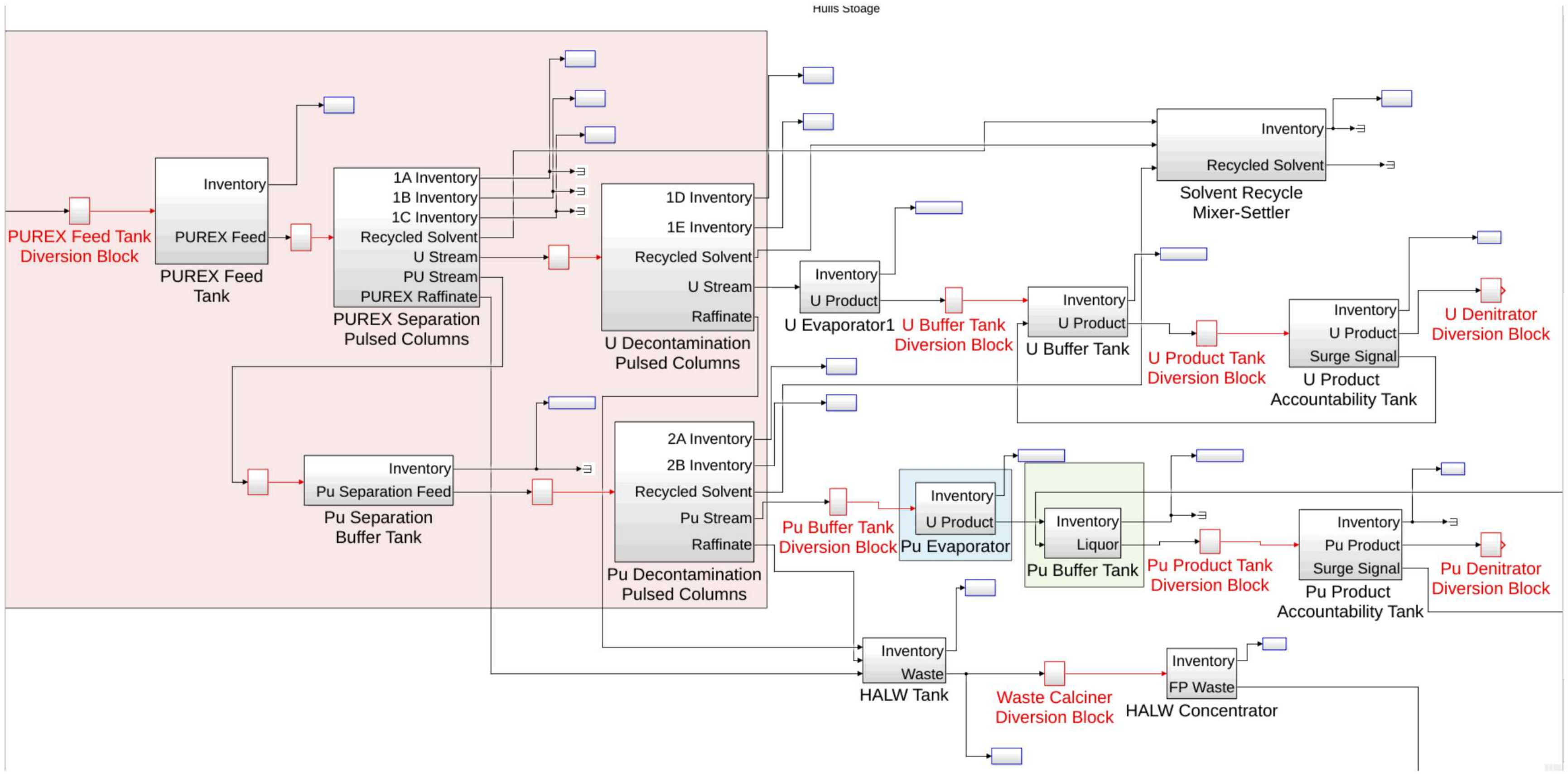
- Can leverage large streams of data
- Could utilize unattended monitoring systems
- Compliments existing safeguards
- Can be difficult to interpret and require more R&D

Hypothesis: An arbitrary function learned under normal conditions will poorly represent anomalous facility behavior

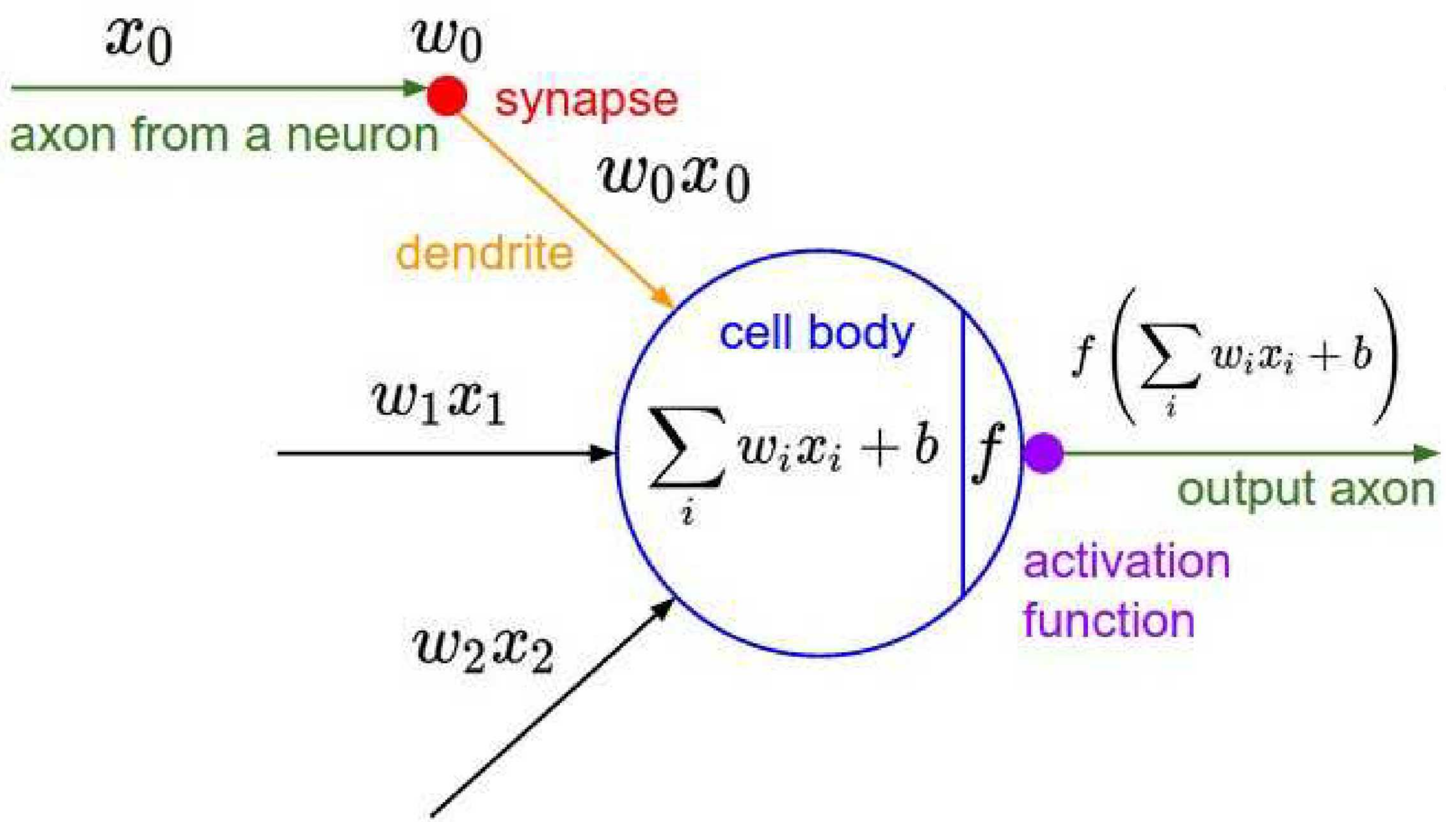
- Bulk throughput facilities often have a normal operating regime
- How well could an entire facility be modeled by a function?
- How well could individual unit operations be modeled by a function?
- Would anomalous behavior be represented by a different function than normal behavior?



Detailed facility model is used to generate training data

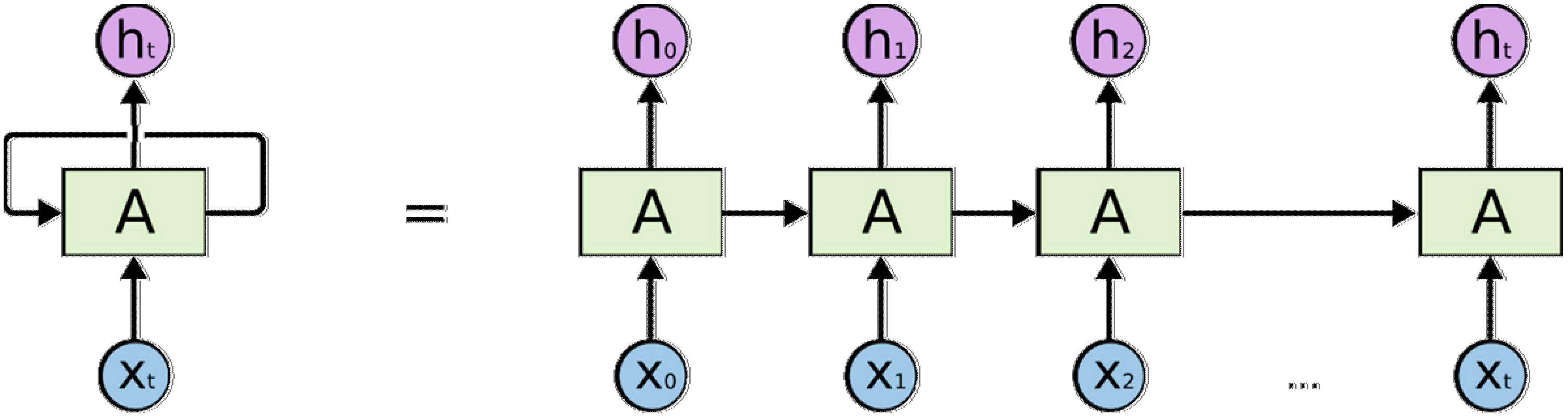


Feedforward neural networks are powerful tools that can learn any continuous function



http://cs231n.github.io/neural-networks-1/#feedforward

- Neurons receive signals, send output to connected neurons
- Activation functions allow for learning of non-linear functions
- Weights are adjusted during training to more closely match desired output
- Neurons are arranged in a network



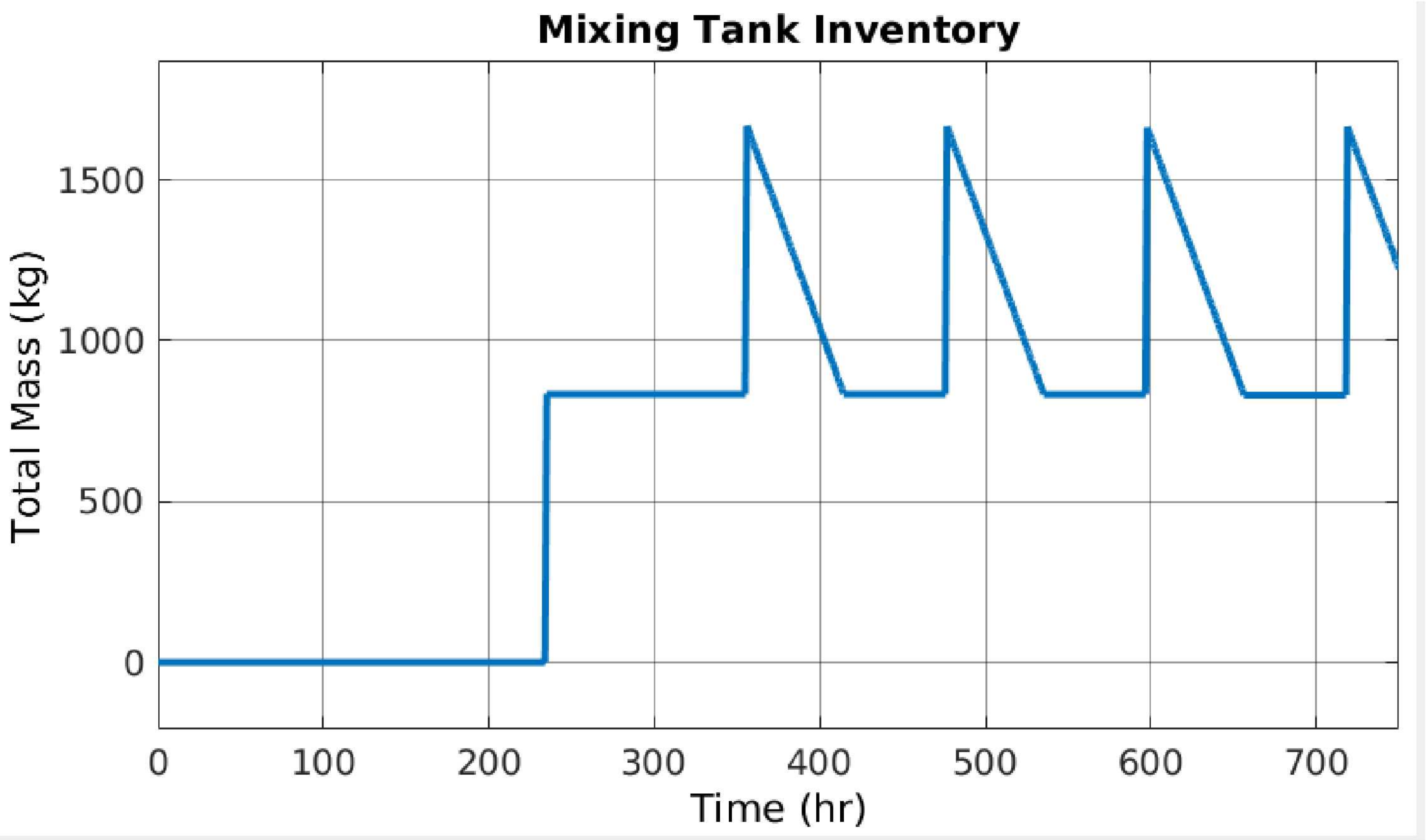
https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Recurrent neural networks can address temporal dependencies in data

- Recurrent neural networks learn time-dependent behavior by passing information to other parts of the network
- Very effective in other areas of machine learning such as speech recognition or language translation
- Long-Short Term Memory (LSTM) networks, a type of recurrent neural network, is used in this work

LSTM networks are used to predict the future output of a unit operation

- Train LSTM network to predict the future
- $\widehat{y_{t+1}^{n}} = f(\overrightarrow{x_t^n})$
  - $\widehat{y_{t+1}^{n}}$ is the prediction of the LSTM network for feature n at time (t+1)
  - f() is the learned function
  - $\overrightarrow{x_t^n}$ = $[x_{t-199}^{n}, \dots, x_t^n]\ for\ 199\ \leq t,$ is some historical window of data
  - y_{t+1}^{n} is the observed value for feature n at time t+1
- The LSTM is trained on some historical input to a unit process, $\overrightarrow{x_t^n}$, and attempts to predict the output of the process, $\widehat{y_{t+1}^{n}}$
- The reconstruction error, $\widehat{y_{t+1}^{n}}$ - y_{t+1}^{n} is used to determine if anomalous behavior is occurring



Learning the facility function – unit operations

- Certain areas of the facilities require special consideration
- Mixing/buffer tank outputs are a function of their entire history
- Tank may also have non-uniform output sizes
- Feature representing running average of inventory concentrations required
- Feature representing bulk level measurement required

Reconstruction error is arbitrary and must be converted to a useful metric

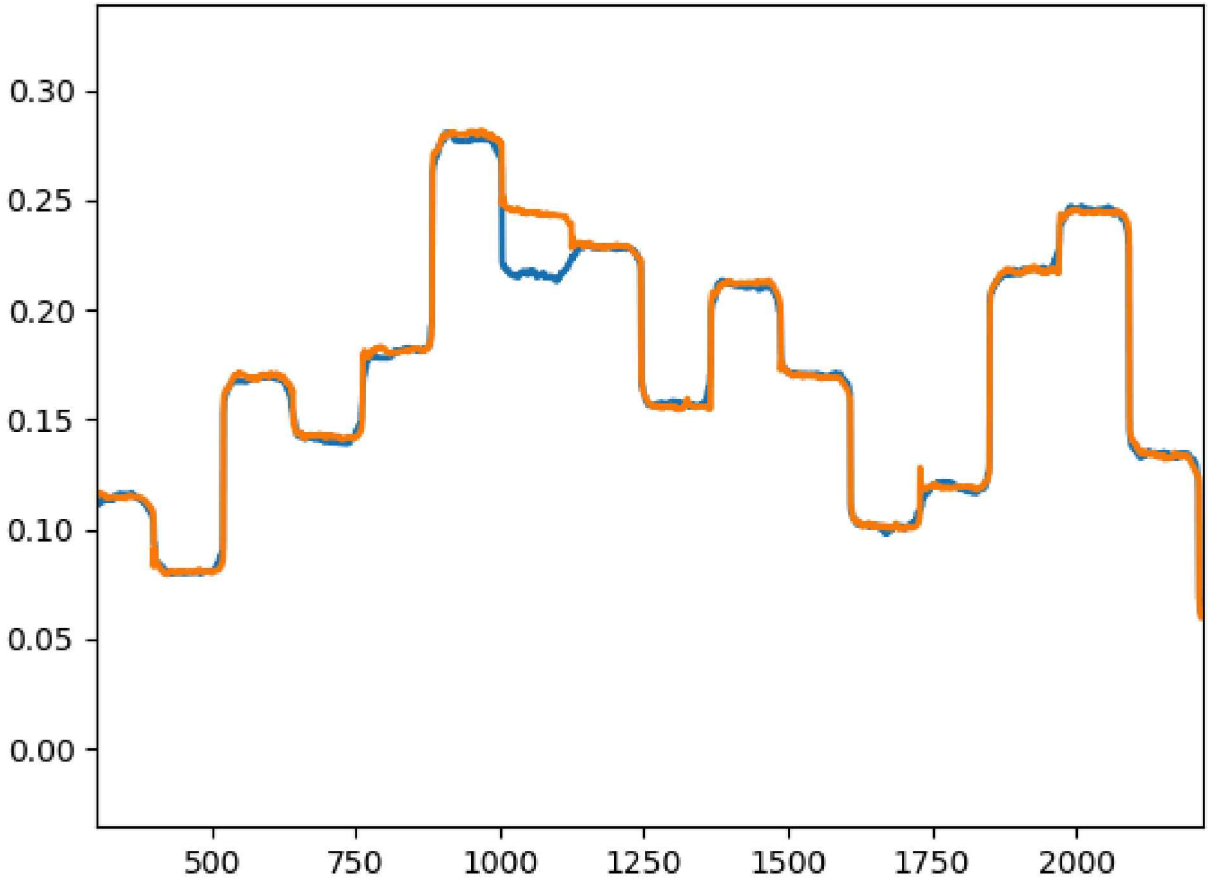
Under normal operation the difference between the LSTM prediction and observed value should be small

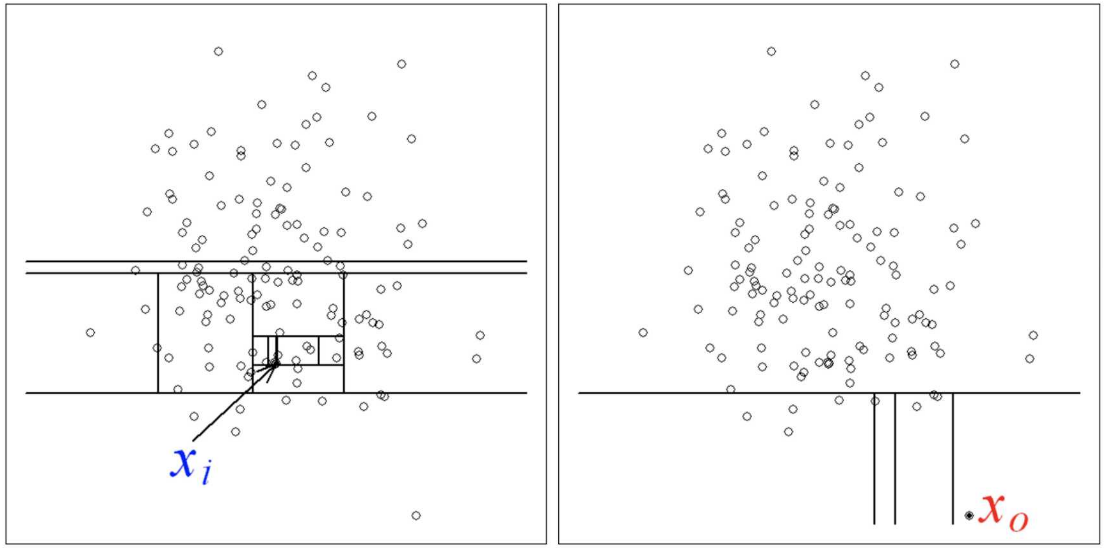
Anomalies should have larger differences than normal conditions

Changes in reconstruction error can be subtle and occur across multiple features making anomalies difficult to detect

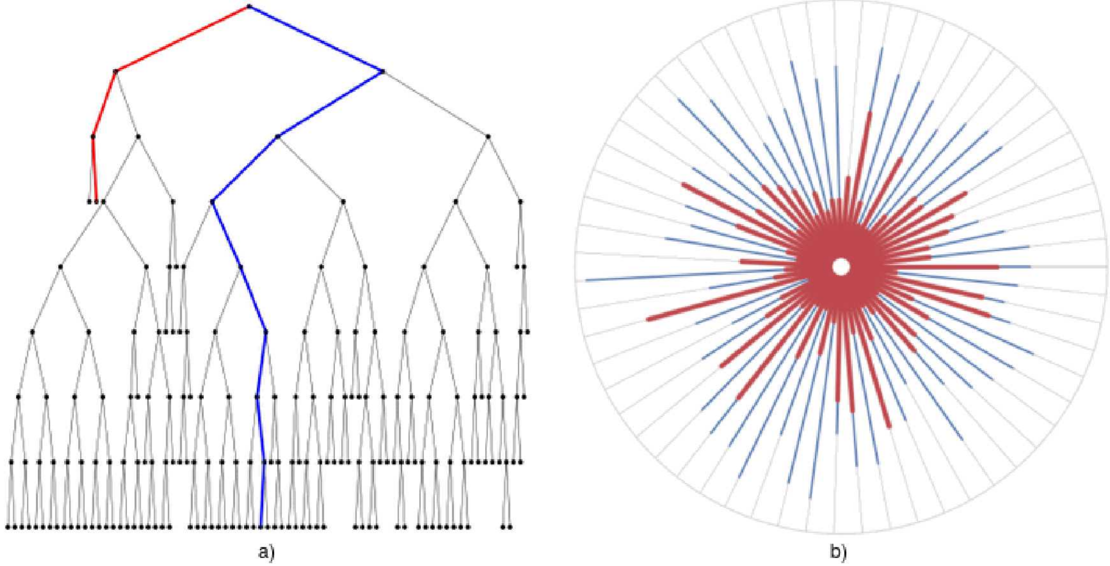
Isolation forest is used to translate arbitrary reconstruction errors into a measurement of normality

Calculating the residual





Fei Tony Liu, et al. *Isolation Forest*, IEEE, 2008

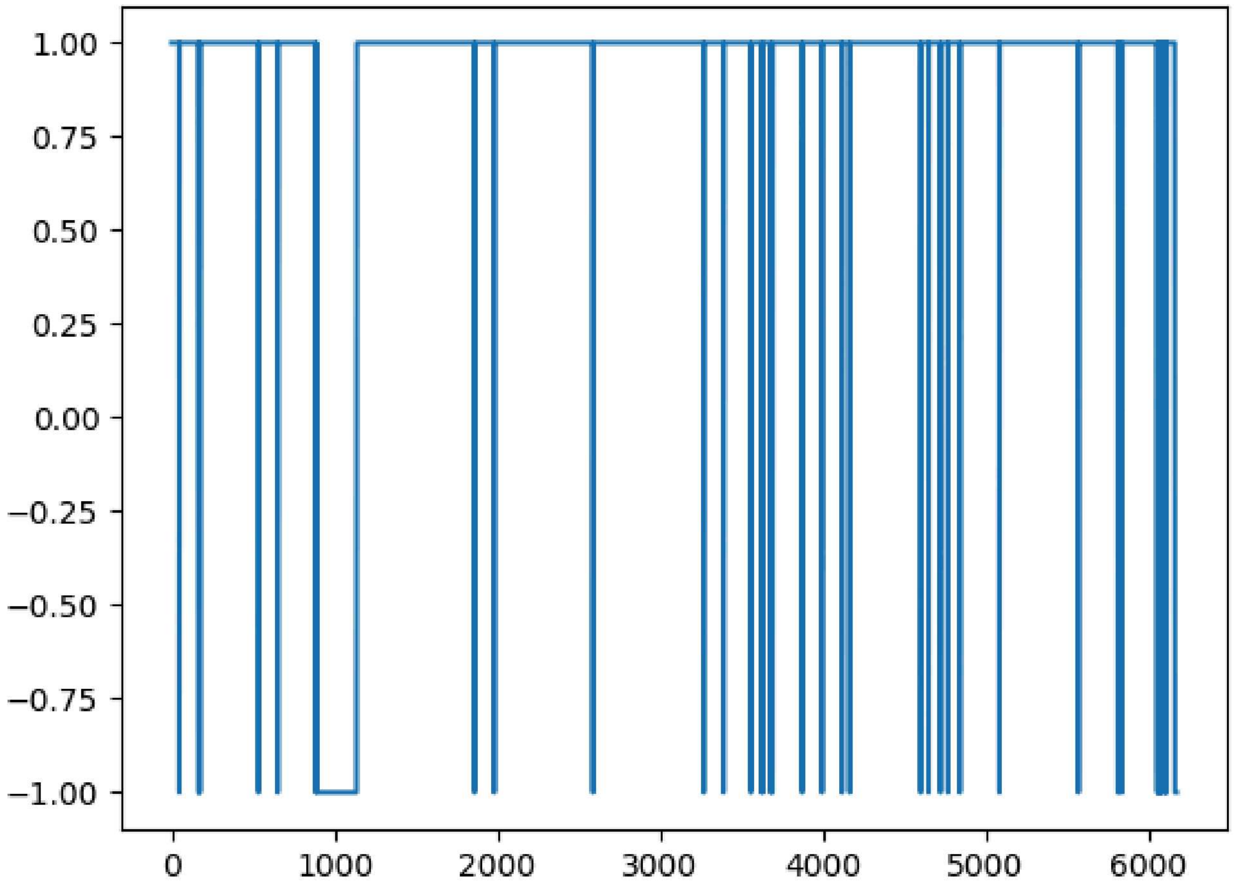


https://github.com/sahandha/eif

Anomalies should be few and different when compared to normal data

Isolation forest output is used to set thresholds for alarms

- Isolation forest is trained to partition a part of the data as anomalous
  - Required since in practice there are no or very few examples of anomalies
  - Intuition is that the most extreme normal cases are labeled as anomalous
- Occasionally and sparsely normal data will be labeled as an anomaly
- A true anomaly will have a dense clustering of anomalous labels



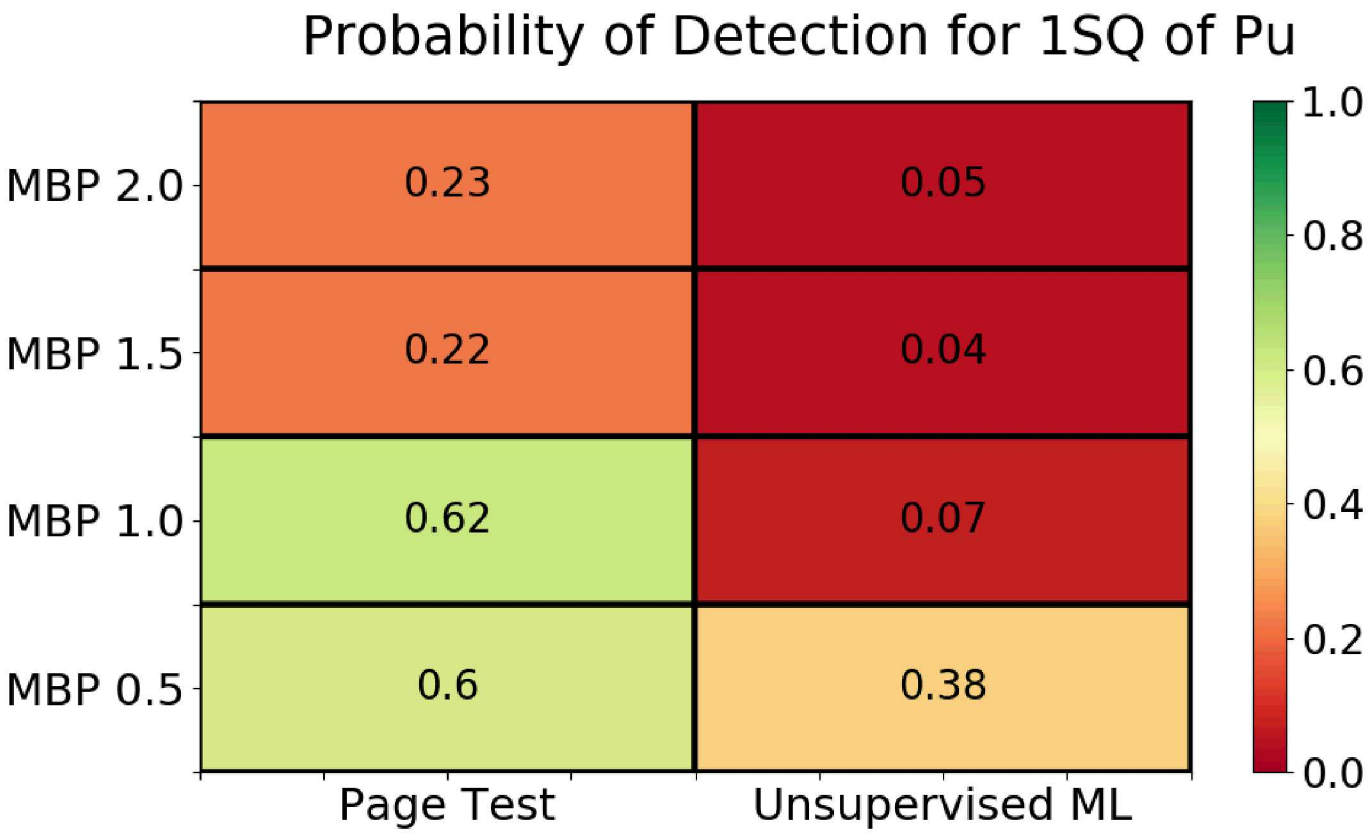
Putting it all together…

1. Train algorithm to predict output of unit operation given historical data
2. Calculate the difference between the ML prediction and observed values, aka the “residual”
3. Classify the residual as normal or off-normal using isolation forest
4. Determine if alarm should be raised based on frequency of off-normal classification

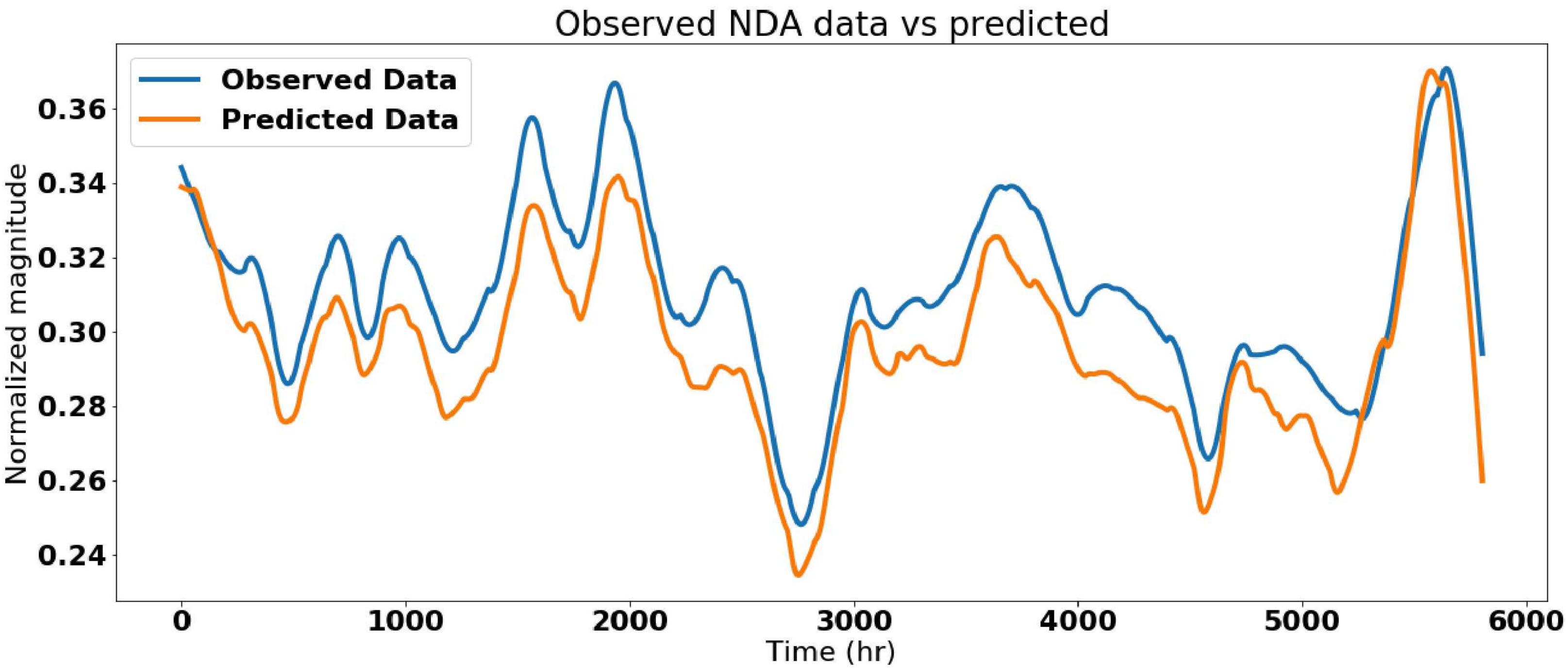
Benchmarking setup

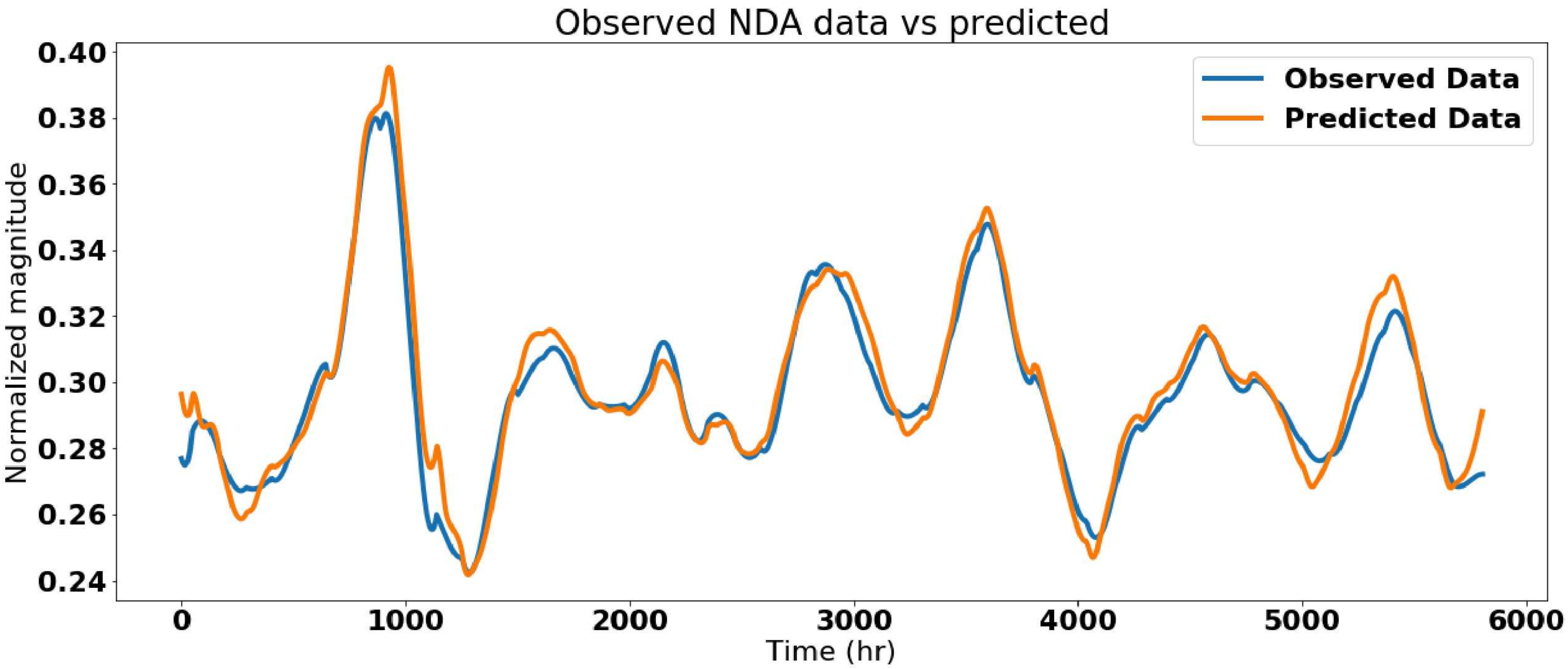
- Substitution material loss at a generic PUREX facility
- Removal of 1SQ of Pu over various lengths of time, expressed as multiple of MBP
- Masses from process model used as inputs to machine learning models and traditional safeguards tests
  - In practice the machine learning approach would utilize NDA features such as gamma counts, but mass was used to reduce computational overhead
- A 1% random and systematic error was added to all measurements

Initial machine learning results exhibit poor results

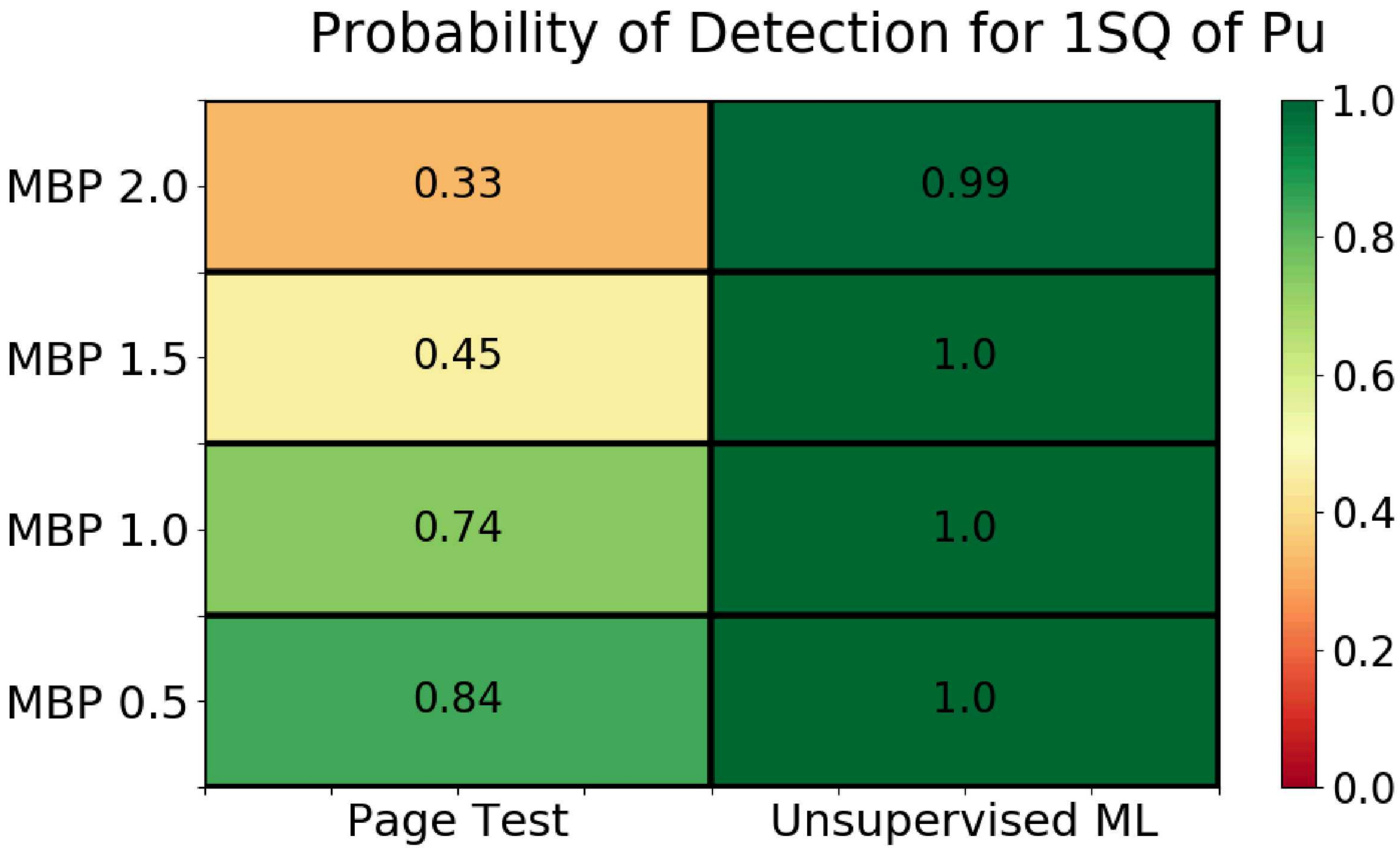


Mismatches in systematic error causes biases in residuals





Results improve drastically with cross-calibration of sensors which reduces differences in systematic biases



ML can enhance safeguards when applied appropriately

- Results of current approach work very well when systematic error is controlled
- Good ML performance requires understanding of underlying process
- Working to resolve systematic error issues
- Work with IAEA to discuss ML based approaches

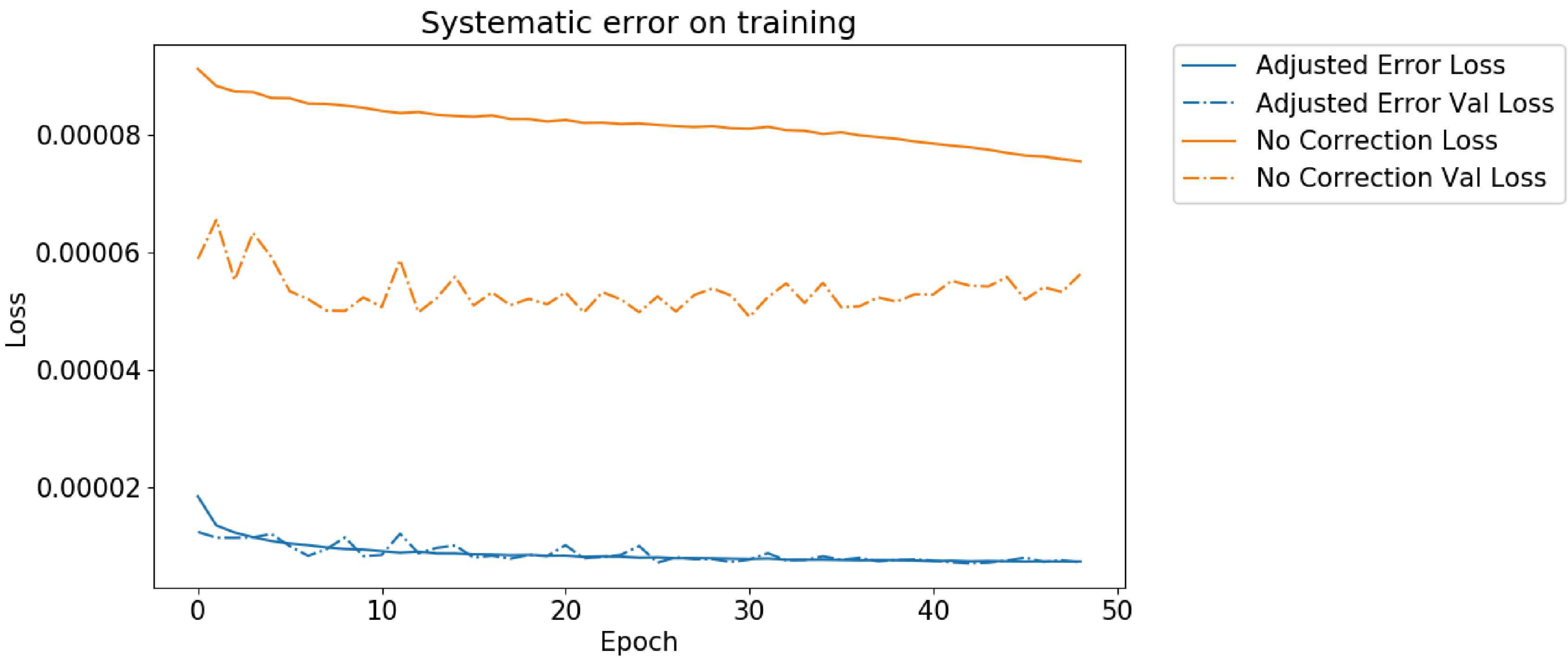
Acknowledgements

- The authors would like to acknowledge fellow collaborators Benjamin Wilson and Thomas Grimes at Pacific Northwest National Laboratory.
- The authors would also like to acknowledge the National Nuclear Security Administration's Office of Nuclear Safeguards (NA-241) for funding this work
- There are various open source packages that made this work possible including scikit-learn and Tensorflow.
- Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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.

Supplementary - Key improvements for FY20

- LSTM regression coverage increased to entire MBA
- Improved anomaly detection through isolation forest
- Improved LSTM architecture
- Determined systematic error must be resolved systematically
- Examined performance of supervised methods in light of systematic biases
- Conducted parametric studies to determine required training data

Supplementary – validation error on cross-calibrated data vs raw data



Supplementary slide – non-linear feedback in bias correction / estimation