



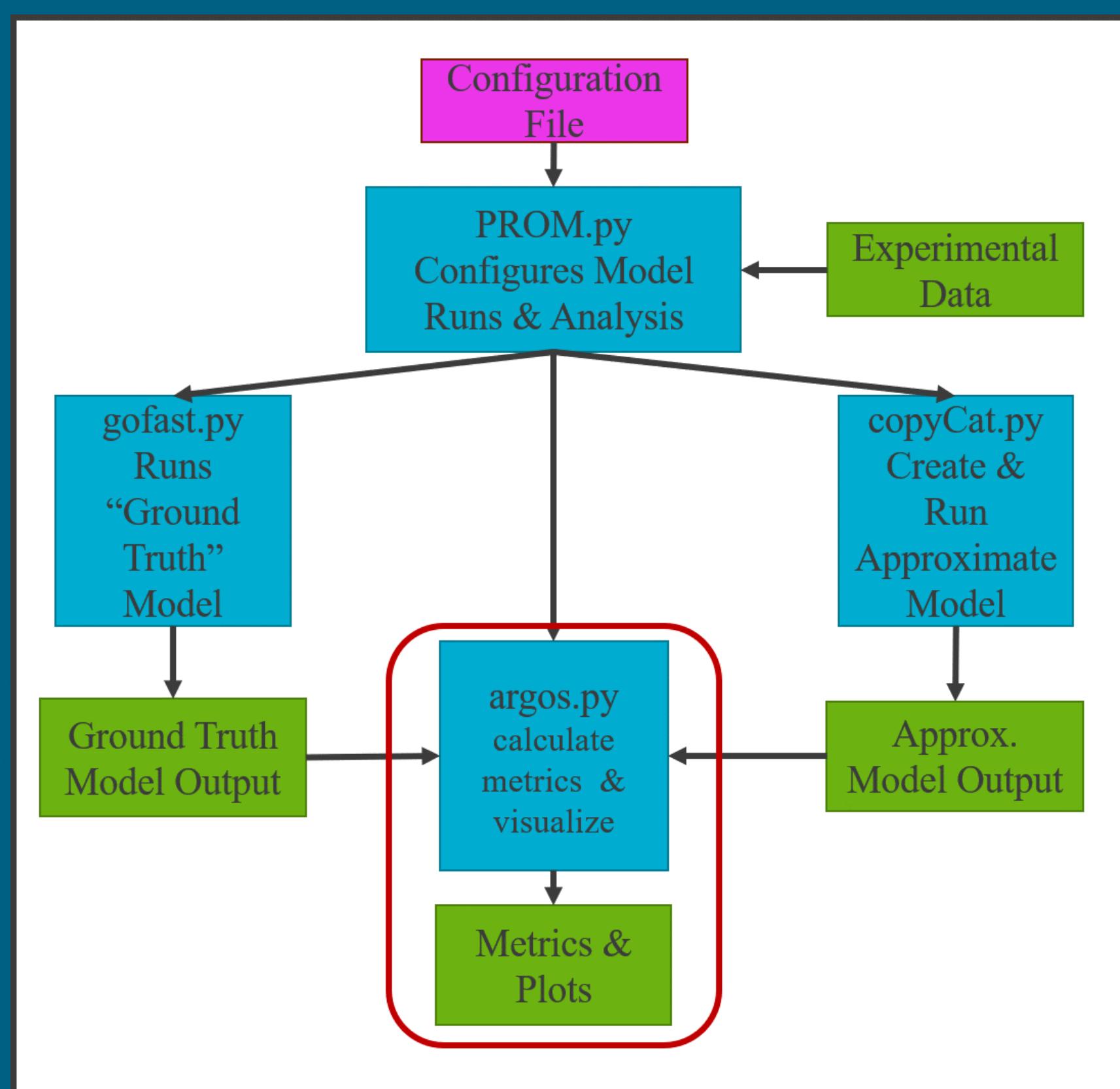
Project Mentors: Ray Fasano, Org. 5621, Eric Vugrin, Org. 5600

Problem Statement:

- PROM project builds math models of physical processes in industrial control systems
- System Identification (SID) algorithms build the models by processing data from SCADA servers.
- **Research Question:** Is the SID model “good enough” to be useful?

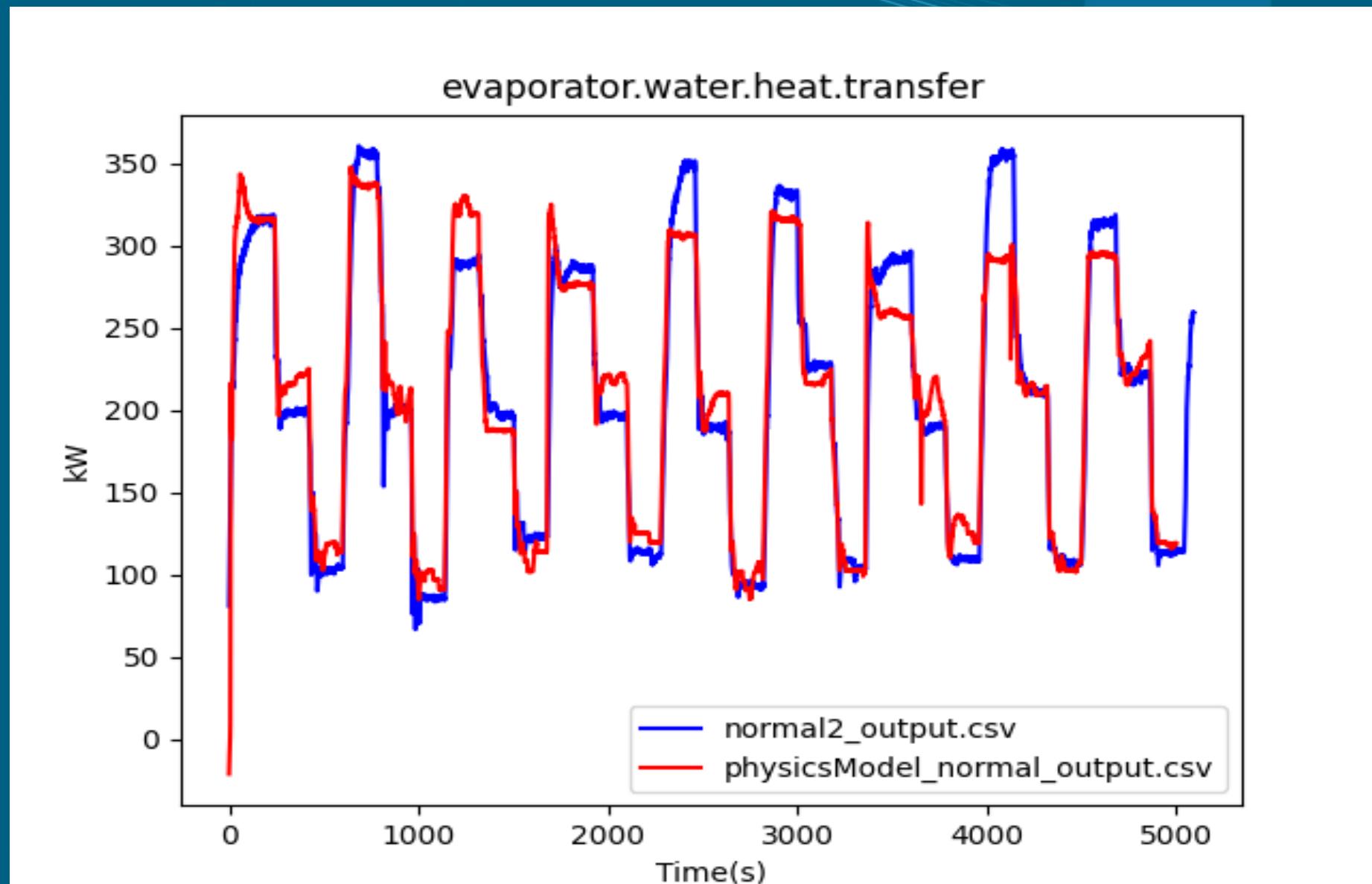
Objectives and Approach:

- Example system: industrial chiller
- Goal: evaluate the suitability of the SID model for cybersecurity investigations



- **Requirement:** select, develop, and code metrics to compare collections of time series data

Results:



Visual comparison for one variable
 $R^2 = 0.92, RMS = 23.54$

- Implemented 9 metrics in Python
- Most promising metrics included
 - R^2 metric: fraction of “difference variance” attributed to linear model

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Potential challenges applying R^2 to highly nonlinear systems

- Root-Mean Squared Error (RMS): spread of differences between datasets

$$RMS = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}$$

- Future work: investigate acceptance criteria

Impact and Benefits:

- Automated visual and statistical comparison of models
- Enhanced confidence in resulting models and analyses