INL/MIS-24-80685-Revision-0

A Causal Approach to Model Validation and Calibration

September 2024

Changing the World's Energy Future

Ronald Louis Gonzales

Idaho National Laboratory

INL/MIS-24-80685-Revision-0

A Causal Approach to Model Validation and Calibration

Ronald Louis Gonzales

September 2024

Idaho National Laboratory
Idaho Falls, Idaho 83415

http://www.inl.gov

Prepared for the
U.S. Department of Energy
Under DOE Idaho Operations Office
Contract DE-AC07-05ID14517

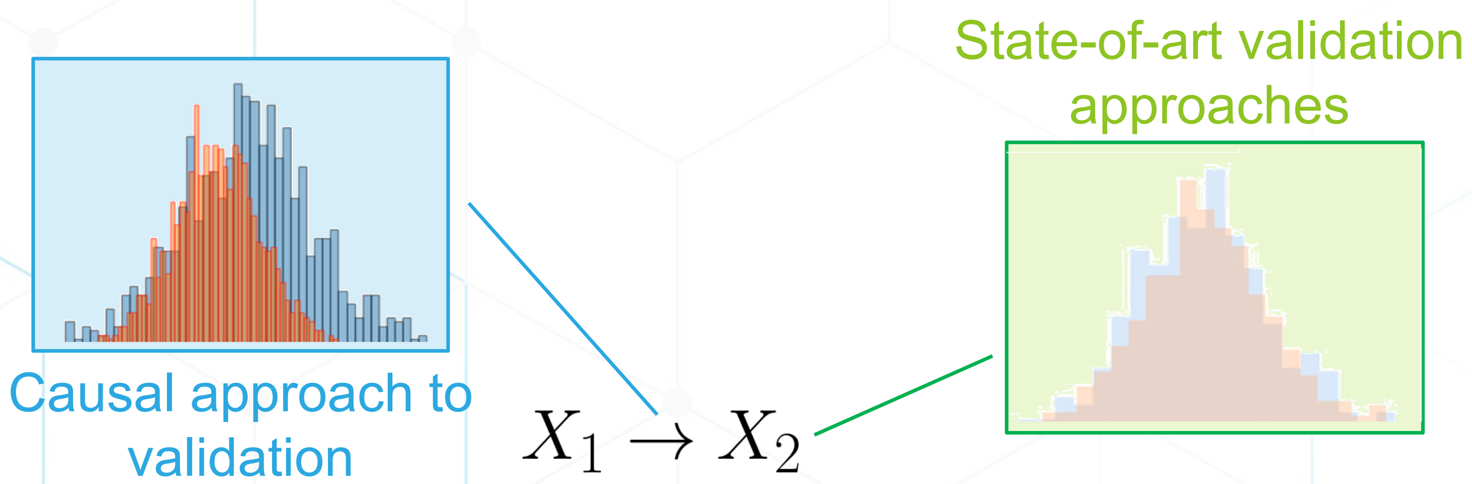
A Causal Approach to Model Validation and Calibration

Context

- Machine learning (ML) and artificial intelligence (AI) have spread in almost all science and technology fields, but the employed methods have focused on “fitting” rather than “understanding” the data
- This project has brought back a science-based mindset to AI/ML methods with the goal of discovering and quantifying causal relationships between observed data elements

Model validation

- Current validation methods employ standard statistical analysis or machine learning methods to quantitatively measure the difference between simulated and experimental datasets
- We propose an alternative approach based on causal inference
- **Goal:** Identify and quantify the causal relationships between data elements rather than looking at their associations

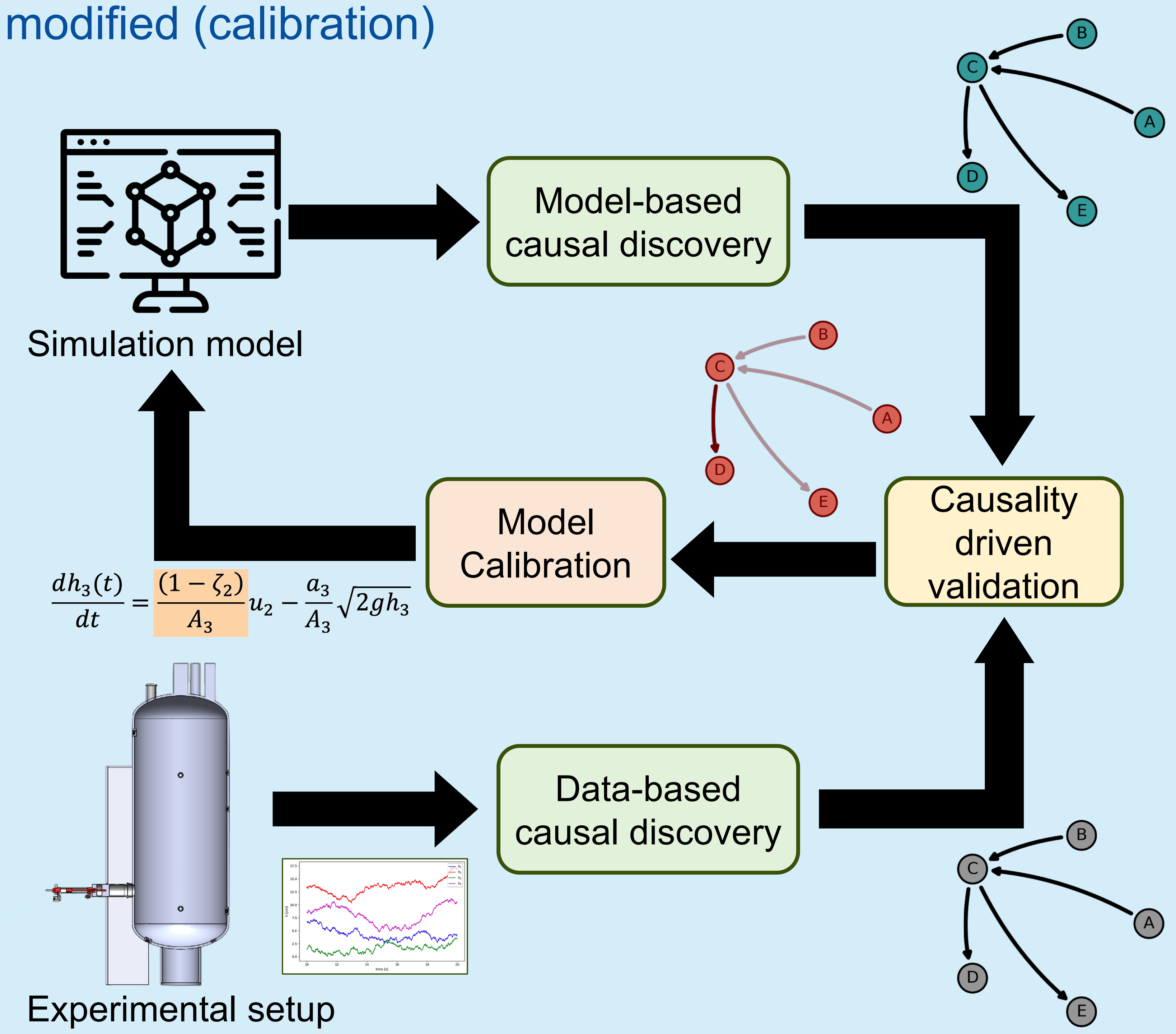


- Comparison of simulated and experimental setups is framed in terms of generation and comparison of causal models

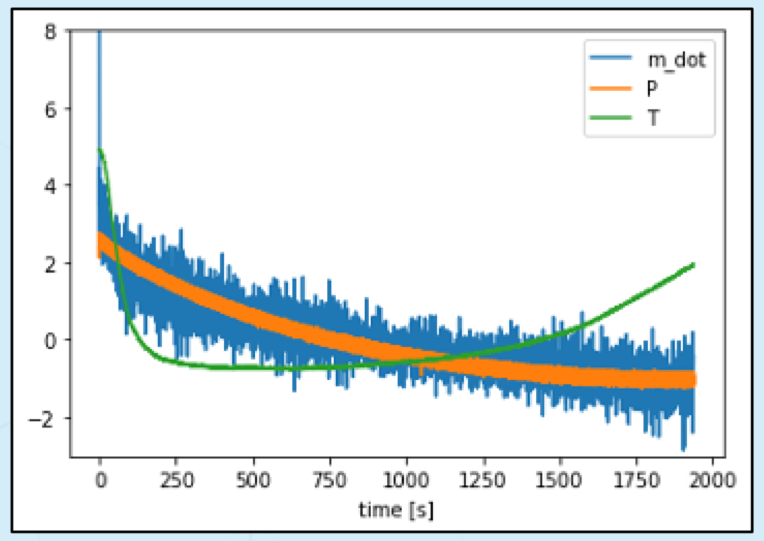
Structural causal models (SCMs): Causal models able to model cycles, interventions, and counterfactuals

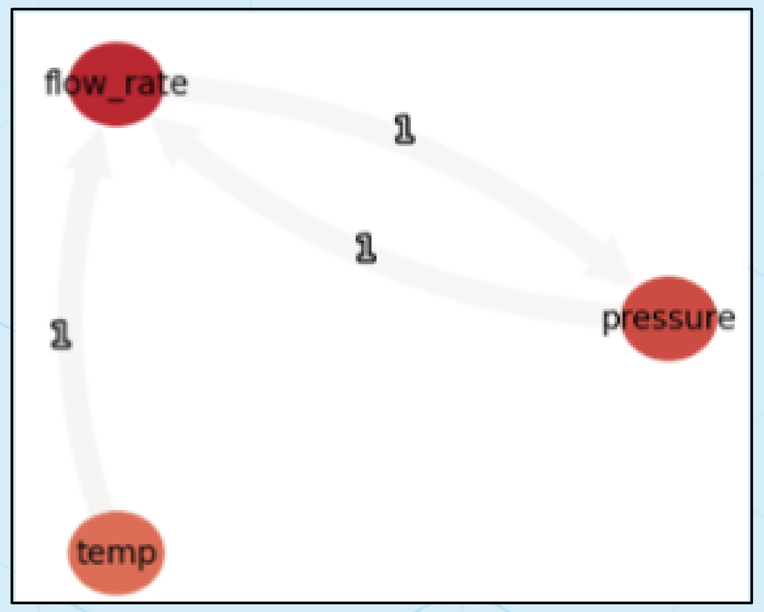
- Given N variables x_n , an SCM consists of a set of structural equations of the form $x_i := f_i(x_j, N_x)$
- SCMs can be visualized as directed graphs

Validation through a causal lens: Causal discovery methods applied to simulation models and observed experimental data are designed to generate SCMs which are then compared to help analysts identify the equation parameters of the simulation model that needs to be modified (calibration)



HAIRE Experiment





Developed AI/ML methods embrace causality at their core by identifying the physical relation among data elements, and then by quantifying it in the form of causal models

Our methods blend statistical independence testing methods (to discover physical relations) with regression models (to quantify the amount of these causal relations) to construct causal models

Model-based causal discovery

- Intervention driven method
- Requires simulation model
- Relies on Maximum Mean Discrepancy (MMD)

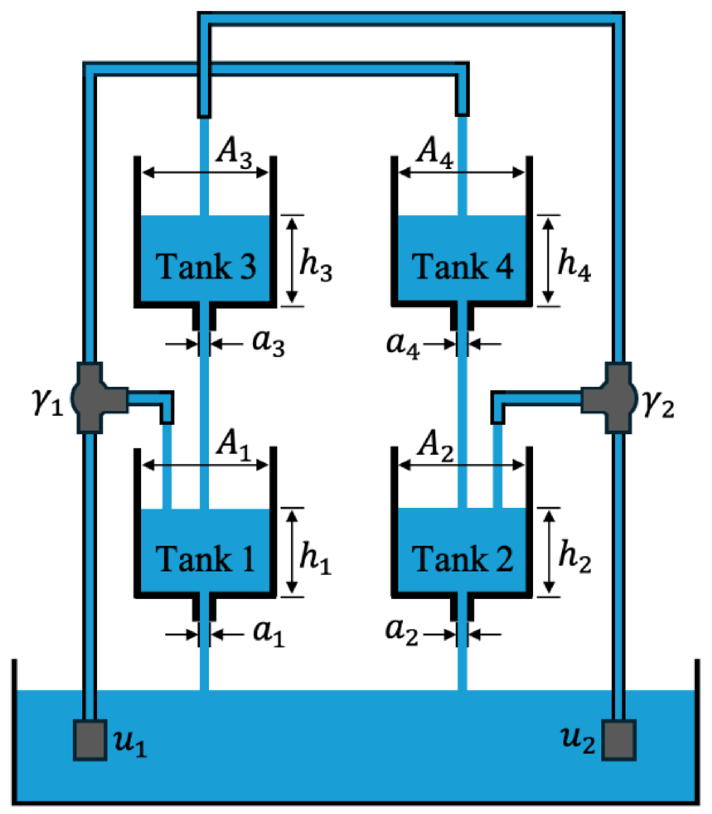
Data-based causal discovery

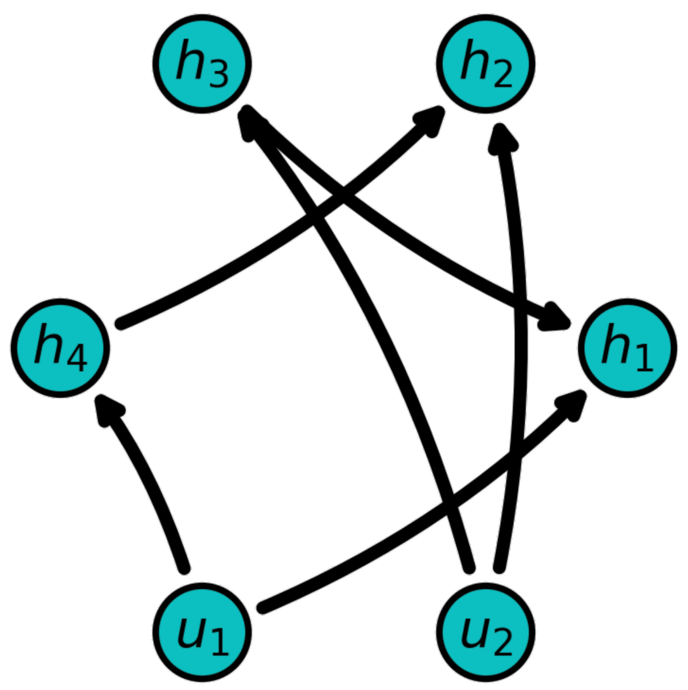
- Input: Temporal profile of N variables
- Conditional independence testing: PC, PCMCI
- Regression methods: SINDY, MMD

Assumptions

- Causal sufficiency
- Causal stationarity
- Contemporaneous effects
- Markov + faithfulness properties

Quadruple tank model





R. Gonzales, D. Mandelli, C. Wang, M. Abdo, P. Balestra, S. Qin

Project Number: 23A1070-144FP

LRS Number: INL/MIS-24-80685

Work supported through the INL Laboratory Directed Research & Development (LDRD) Program under DOE Idaho Operations Office Contract DE-AC07-05ID14517."