Constrained Bayesian Methods for Sensor Validation

Stellenbosch University Iyunivesithi Universiteit

W. Horak & Prof T.M. Louw, Prof S.M. Bradshaw
19272685@sun.ac.za

Sensor Validation: Making process data more reliable

Detection
- Process monitoring strategies aim at detecting faults and improving data quality.
- Reliable process data are required for safe, efficient and profitable operation of industrial processes.
- Fault detection makes use of statistical methods.

Diagnosis
- Inaccurate measurements are detrimental to performance and result in economic losses or safety issues.
- Fault diagnosis aims at determining the root cause of a fault.

Reconstruction
- Replacing erroneous measurements with reconciled estimates ensured good data reliability.

1. INTRODUCTION

Aim

Integrate chemical engineering domain knowledge with innovations in machine learning to develop a model that can be used for process monitoring.

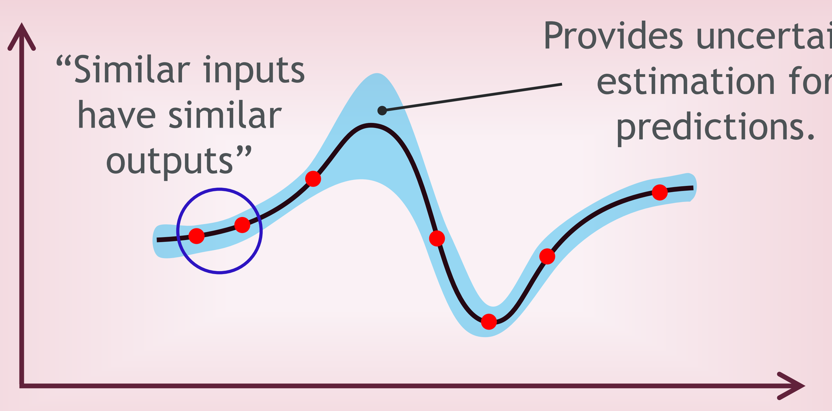
Objectives

1. Extend GPR to satisfy linear and nonlinear constraints based on mass and energy balances.
2. Evaluate and compare method performance with traditional methods.
3. Evaluate method performance to perform sensor validation.
4. Apply to industrial case study using real process data.

2. LITERATURE

Gaussian Process Regression & Data Reconciliation

- Probabilistic machine learning method.
- Can model autocorrelation in data.
- Has seen a rise in popularity for process monitoring applications.



Similarity determined by user-defined covariance function.

$$\left\{ f \right\} = \sigma_s \exp\left\{-\frac{1}{2l}(x - x_*)^2\right\}$$

- Reconstruct measurement data to adhere to physical laws.
- Applied in process monitoring and process control.

Steady state data reconciliation
- Uses high confidence models (i.e., mass and energy conservation.
- Disregards data autocorrelation.
- Can only be applied at single time instance.

Dynamic data reconciliation
- Uses dynamic process models
- Dynamics models are expensive and difficult to develop.
- Parameter uncertainty can affect reconciled estimates

A Novel Combined Approach

- Combine the strengths of Gaussian process regression (GPR) and data reconciliation (DR).
  - Constraints are built into the covariance function.
    - Model multi-output GPs on a lower-dimensional constraint plane.
      - Project latent GPs into the measurement space via linear transformation.
        1. Performs data reconciliation: Predictions necessarily satisfy conservation law constraints.
        2. Leverages GPR strengths by giving both a prediction and uncertainty estimation for ground truth.

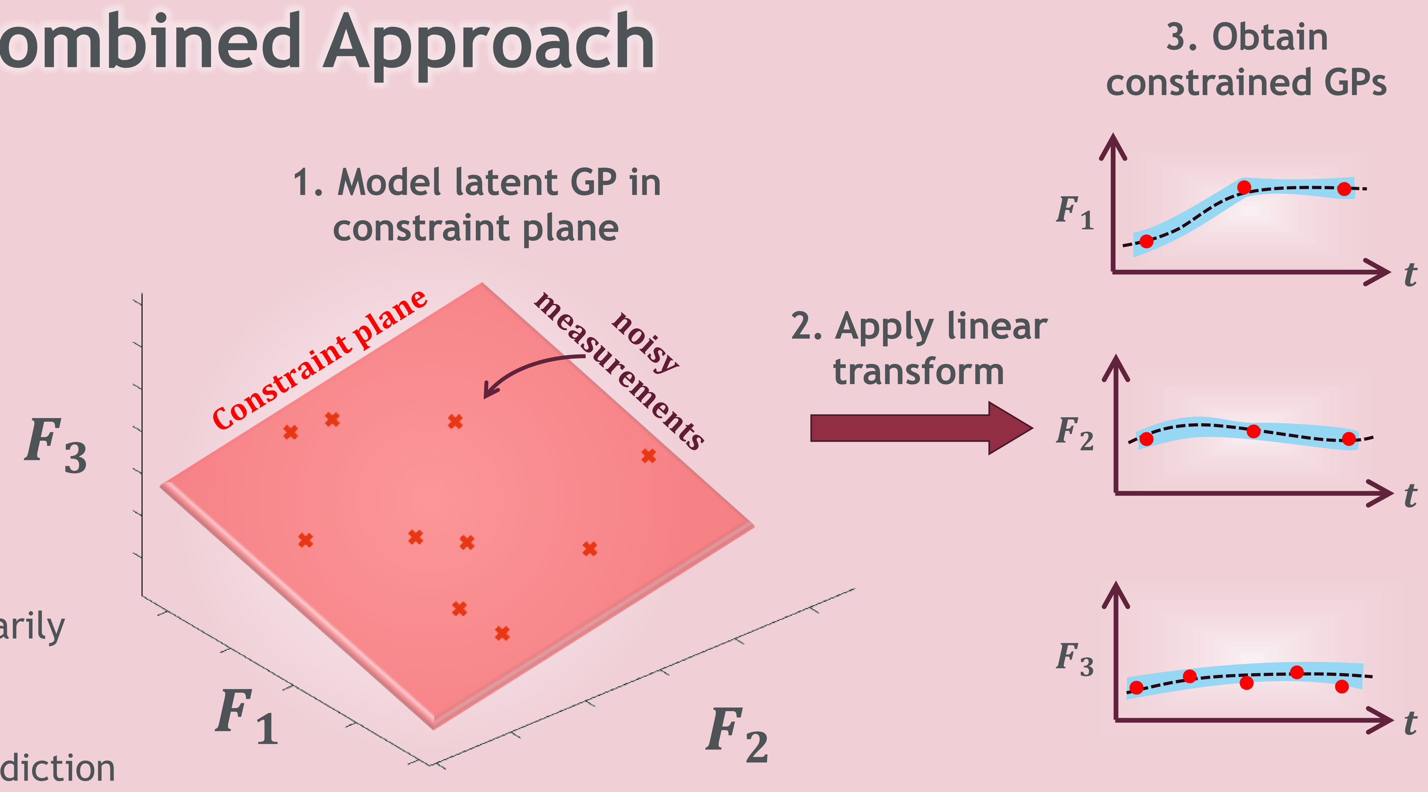


Diagram 1: By modelling a multi-output GP in the plane defined by the conservation constraints, such that it best fits the noisy measurements, a linear combination of the latent GPs yield correlated Gaussian processes whose predictions satisfy the constraints at every point in time.

3. RESULTS

Establishing a baseline

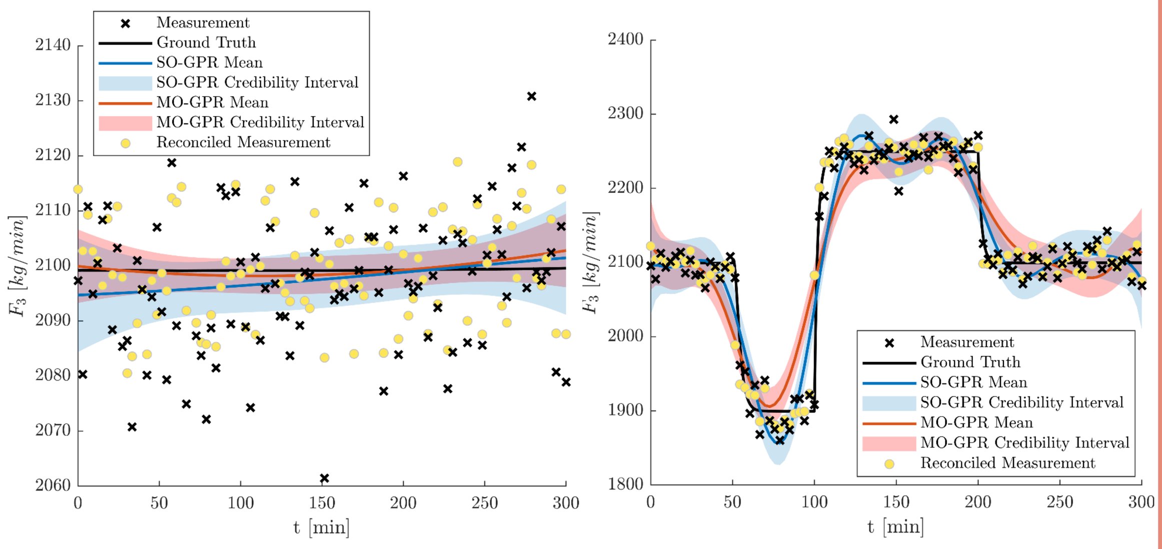


Figure 1: SO-GPR and C-MO-GPR outperforms DR in a steady-state simulation of a blending tank, making accurate and confident predictions close to the ground truth. DR performs better when the process is highly dynamic.

Different sampling rates

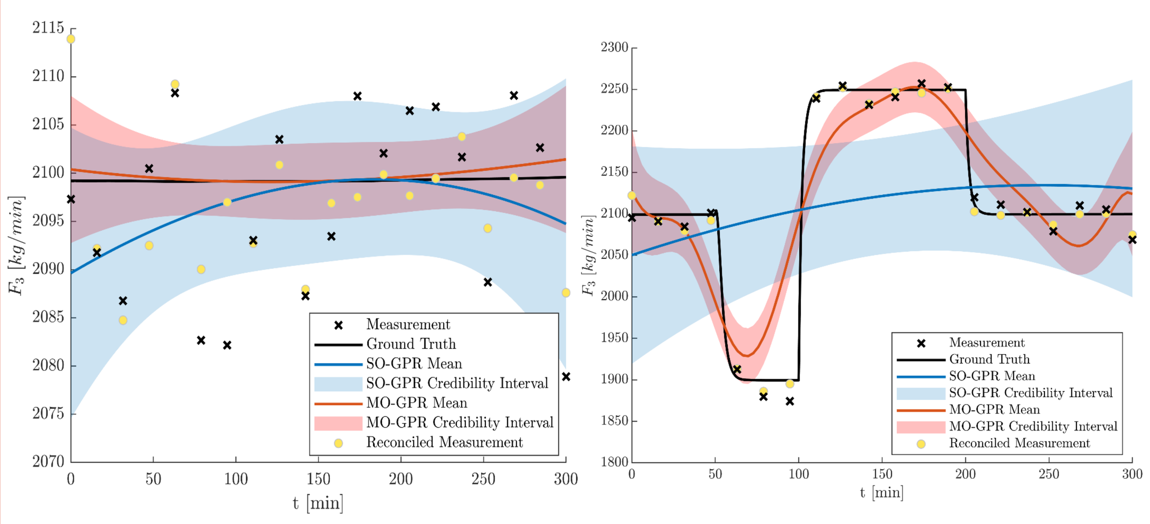


Figure 2: C-MO-GPR and DR can leverage information gained from the process model and other measurements to still make accurate predictions, while SO-GPR suffers in performance.

Table 1: Performance comparison of data reconciliation (DR), single-output Gaussian process regression (SO-GPR) and constrained multi-output Gaussian process regression (C-MO-GPR) in various scenarios.

	RMSE		
	DR	SO-GPR	C-MO-GPR
Baseline			
Steady-state	5,8	1,6	1,0
Open loop	15,9	174,6	167,6
Different sampling rates			
Steady-state	5,7	2,7	1,1
Open loop	12,4	168,0	64,9
Precision Degradation			
Steady-state	21,9	2,6	1,3

Fault! Precision degradation

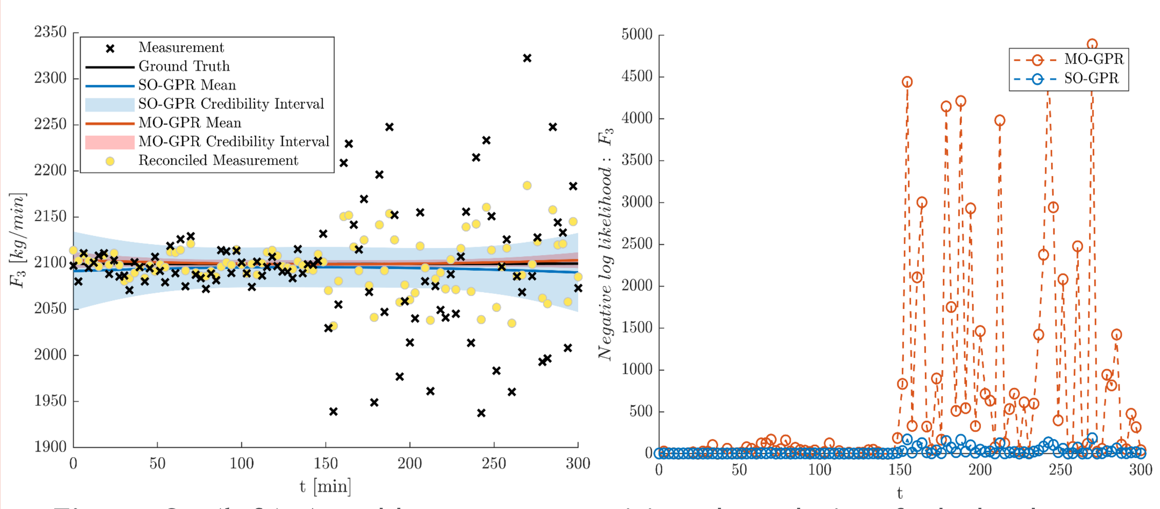


Figure 3: **(left)** A sudden sensor precision degradation fault leads to a decrease in DR performance; gross error detection and removal is required as a prior step. In contrast, the GPR models' predictions remain close to the ground truth. **(right)** The negative log loss for the predictions of the constrained GPR model may prove useful to identify faulty scenarios.

And the winner is...
- Novel GPR model outperforms DR for NOC data.
- DR better when process dynamics are included.

4. FUTURE WORK

1. Incorporate nonlinear constraints:
   - Current model limited to linear constraints.
   - Extension to include nonlinear constraints is important for chemical engineering applications.
2. Incorporate sensor validation:
   - Model should be able to be used to detect sensor faults and reconstruct faulty measurements
3. Address computational concerns:
   - Computation time scales with $\mathcal{O}(N^3)$.
   - Sparse-approximation techniques required to scale up

Postgraduate Symposium 2023

Chemical Engineering