Digital Twin–Enabled Multi-Objective Optimisation of Bio-Oil Co-Processing: A Physics-Constrained, Human-in-the-Loop Framework

Ng Sheng Tat

Priority Topic Area: **Digitalisation and Artificial Intelligence**

1 – Advanced Digitalization for Sustainable Bio-Oil Co-Processing

- Renewable co-processing is a major decarbonization trend; however, the variability in bio-feedstocks characteristics destabilizes operation, causing suboptimal operations [1].
- Unlike conventional refineries, bio-refineries often lack Digital Twins to handle the uncertainty [2]. This project delivers a closed-loop, physics-constrained framework to:
  (1) Infer real-time feed properties (TBP) from plant PVs.
  (2) Make predictions with physics-aware surrogates.
  (3) Perform optimization to identify safe and economical set-points for human-validated control.

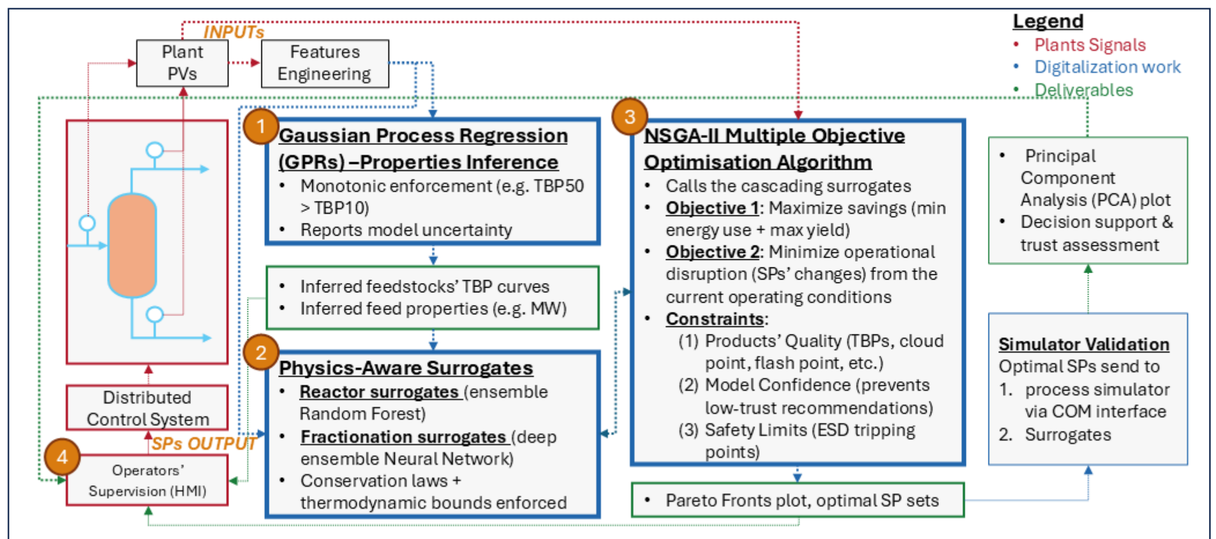
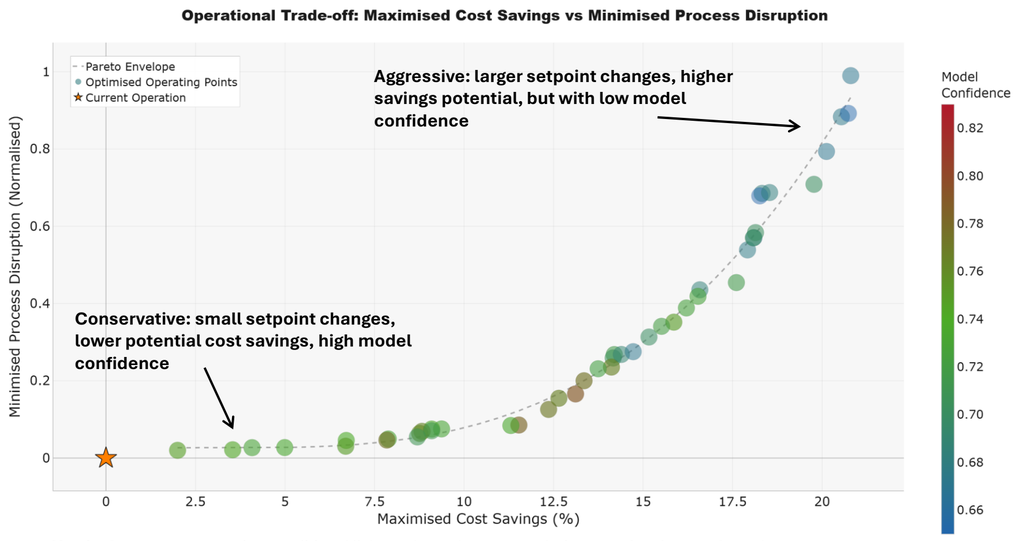


Figure 1. *Closed-loop digital twin architecture*



Y-axis: 0 → current operating condition; Higher value → larger cumulative setpoint changes from the current operating condition
X-axis: Cost savings → derived from minimizing energy consumption and maximizing products' yield

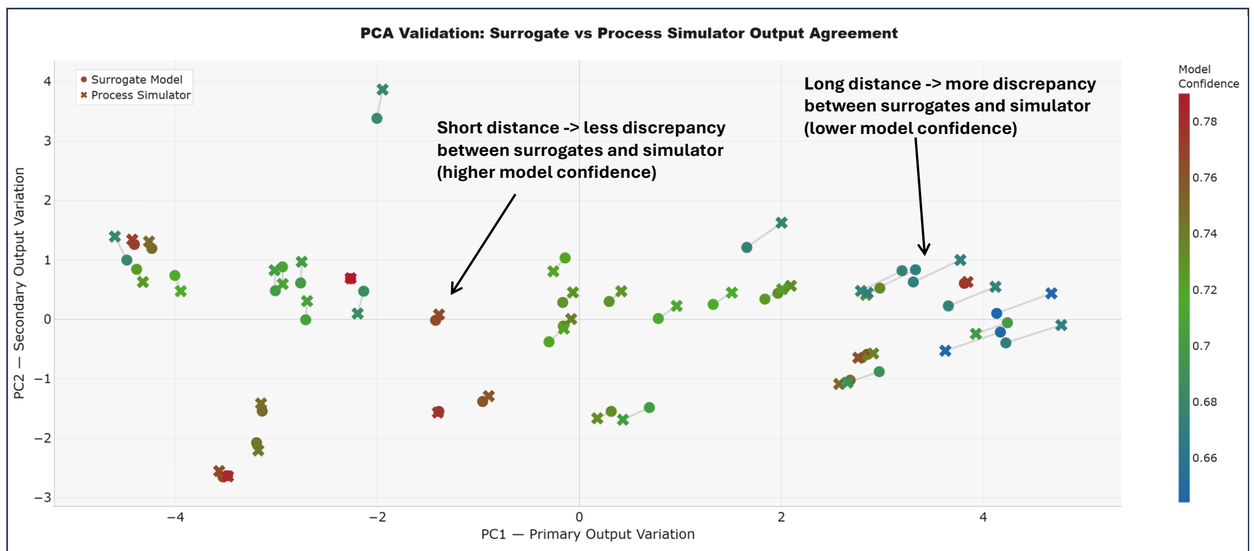
Figure 2: *Pareto-optimal front: cost savings vs. process disruption trade-off.*

2 – Multi-Stages Physics-Aware Surrogate Framework

- <u>Gaussian Process Regression (GPR)</u>, with physics-based monotocity and uncertainty quantification, infers dynamic feedstock properties from plant PVs by learning from combined plant and validated process simulation data
- A cascading reactor surrogate, built using <u>ensemble Random Forest</u> models, predicts intermediate mass and reaction-driven property shifts while respecting conservation laws and feasible trends.
- Outputs propagate to a fractionation surrogate layer, implemented using **deep ensemble Neural Networks**, to predict energy use, yields, and product quality within thermodynamically consistent bounds.
- Ensemble-quantified <u>epistemic uncertainties</u> are condensed into a confidence score that gates NSGA-II optimization, ensuring only high-trust predictions are reported as Pareto-optimal setpoints.

3 – Multi-Objective Optimization & Performance Validation

- Engineered severity indices and catalyst-age features improved surrogate robustness trained with limited data.
- The framework <u>yielded 45 Pareto-optimal</u>, safety-compliant operating points, revealing profit–deviation trade-offs, with <u>5–10% cost savings</u> with small operational adjustment and up to ~20% under larger deviations, guiding the plant's control.
- Surrogate predictions <u>show close agreement</u> with a high-fidelity process simulator.
- Ensemble-derived confidence scores (0.65–0.85) guide operator decisions — a <u>one-click HMI–COM loop</u> triggers simulator validation of selected Pareto points, enabling operator-led optimisation without modelling expertise.



Note: Principal Component Analysis (PCA) maps high-dimensional outputs into two axes
Shorter distances between paired points → closer surrogate-simulator agreement and vice versa

Figure 3. *PCA validation: surrogate and simulator outputs in agreement.*

Figure 4: *Evolution toward an AI-assisted Digital Twin with operator-approved control. (This illustrative figure was generated using generative AI and is intended for conceptual visualization only.)*

4 – Benefit to society

- Enables refineries to maximise yield and minimise energy consumption, thereby <u>increasing profitability, cutting production costs</u> and <u>promoting decarbonisation</u>.
- <u>Reduces GHGs emissions</u> through valorisation of waste-derived bio-feedstocks into fuels, supporting low-carbon fuel production and circular economy principles.
- Promotes the adoption of sustainable technology with optimised efficiency without compromising environmental and public safety.
- Provides an operator-centric digital twin, <u>empowering both experienced and new operators</u> to make informed decisions without modelling expertise.

5 – Next steps

- Model confidence is currently limited by available data; additional data and continued validation with clients are needed to increase confidence.
- Future development may enable continuous learning, direct DCS connection for automated set point propagation after operator approval, and optional integration with a plant virtual assistant to support operator decisions.

References/Acknowledgements

[1] Smith, R. & Jones, M. (2023). "Feedstock Variability in Biofuel Production: Impact of fatty acid composition on refining." Renewable Energy Reviews.
[2] Zhang, L., et al. (2024). "Digital Twins for Industrial Process Optimization: A Review of RTO integration." Journal of Process Control.