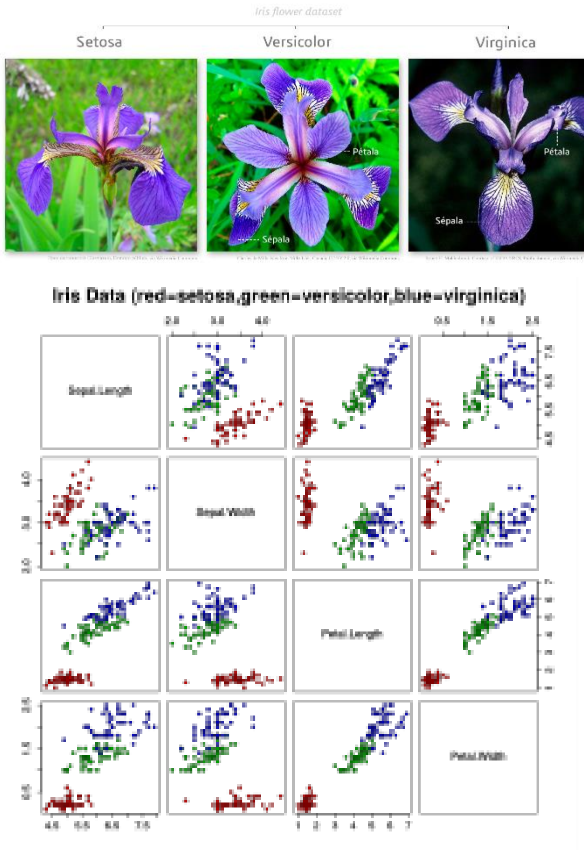
Batch Manufacturing Datasets – Open-source Data for Academia and Industry

Daniel Palaci,

Philippe Neyraval, Benjamin Katz, Mattia Vallerio, Carlos Perez, Francisco Navarro (f.navarro@imperial.ac.uk)

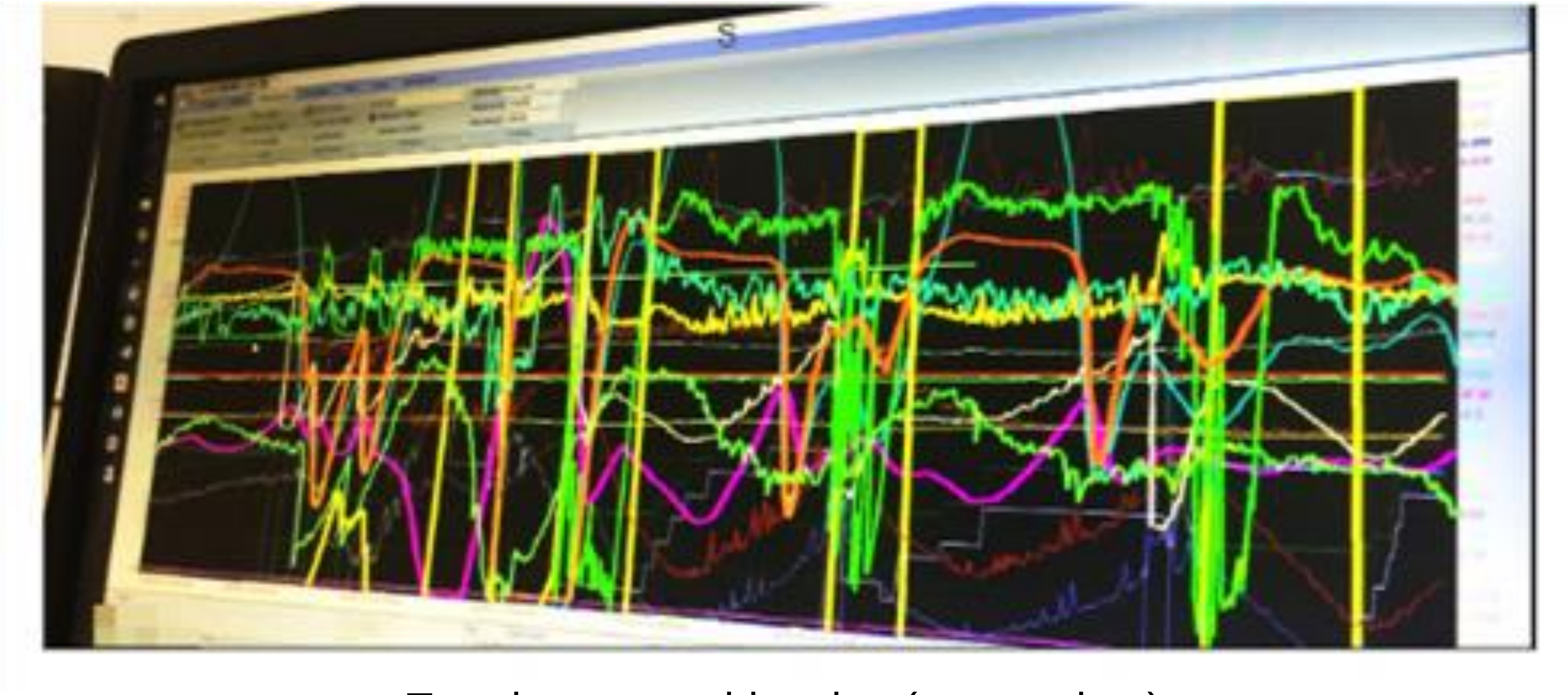
ENBIS 2023 conference, September 10th-14th, 2023

Machine learning is (often) taught with irrelevant datasets



101 ML example:
Types of flowers

?

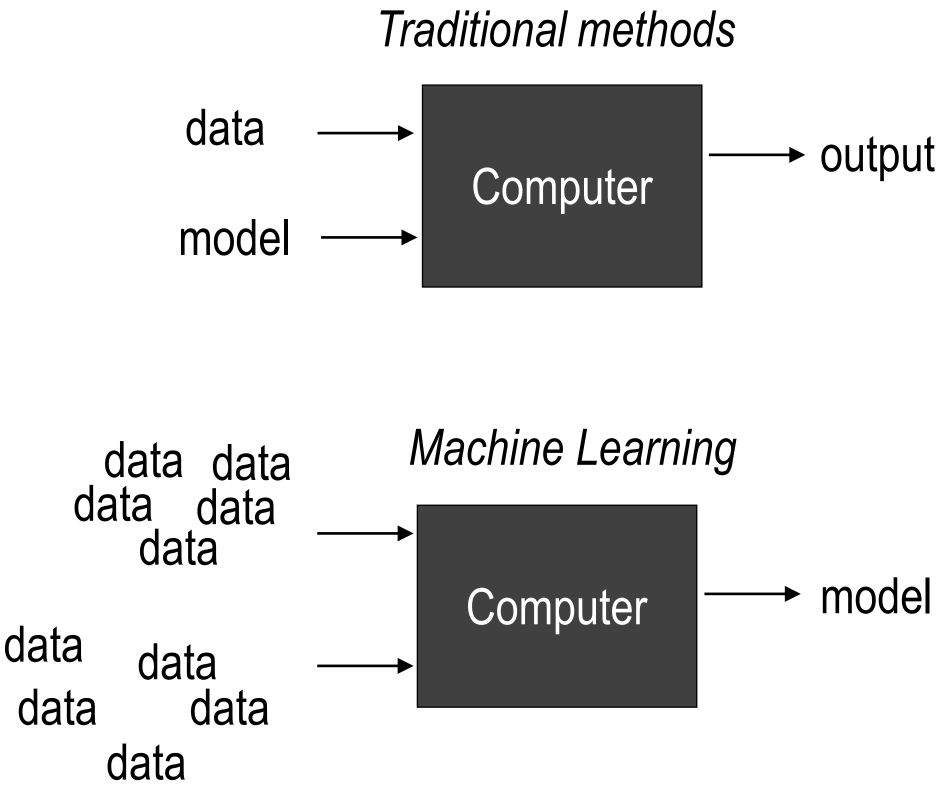


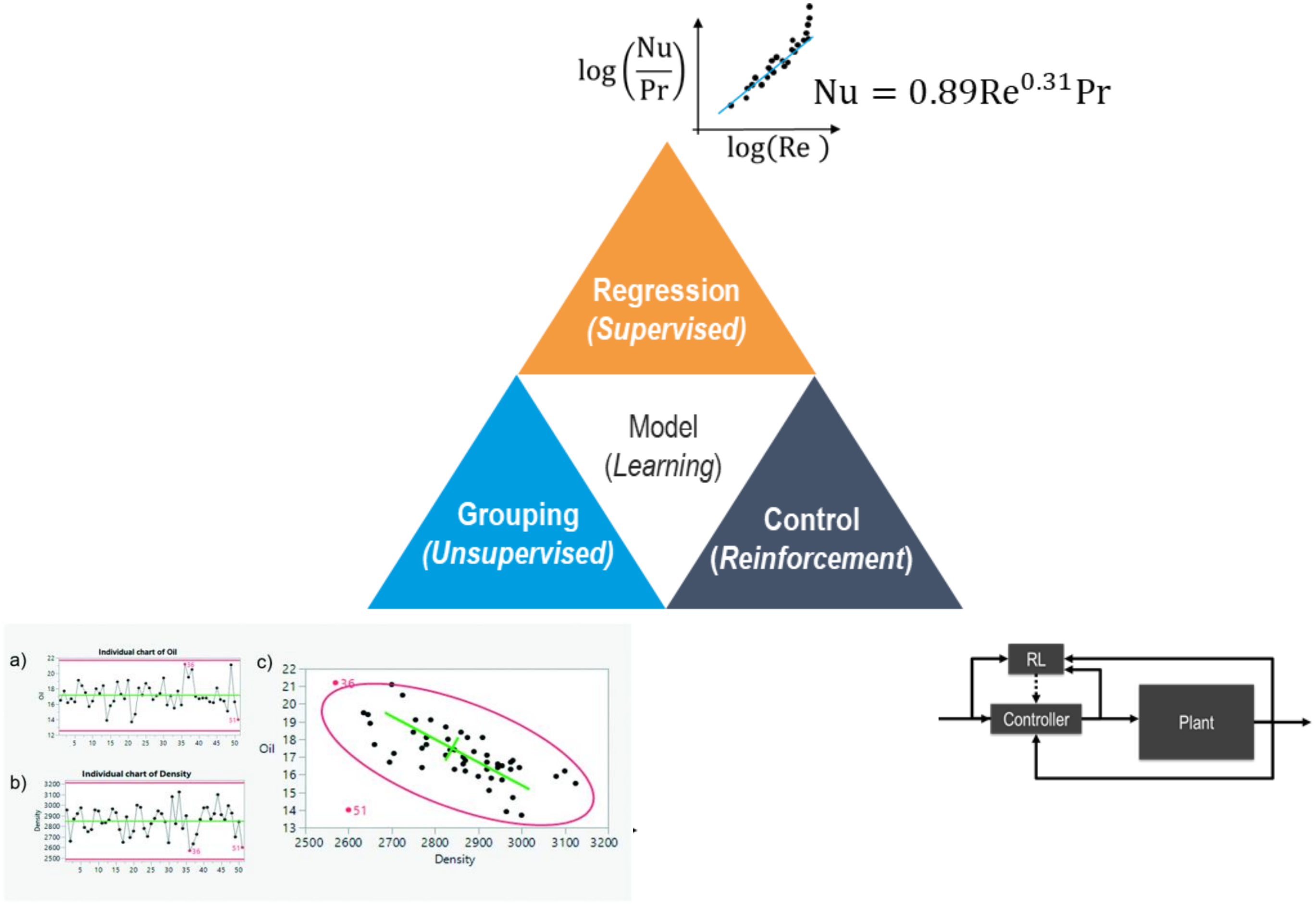
Tags in process historian (sensor data)

Image credit:
Scaling up the Use of Machine Learning in Chemical Process Industries (2023-EU-30MP-1346)

Introduction

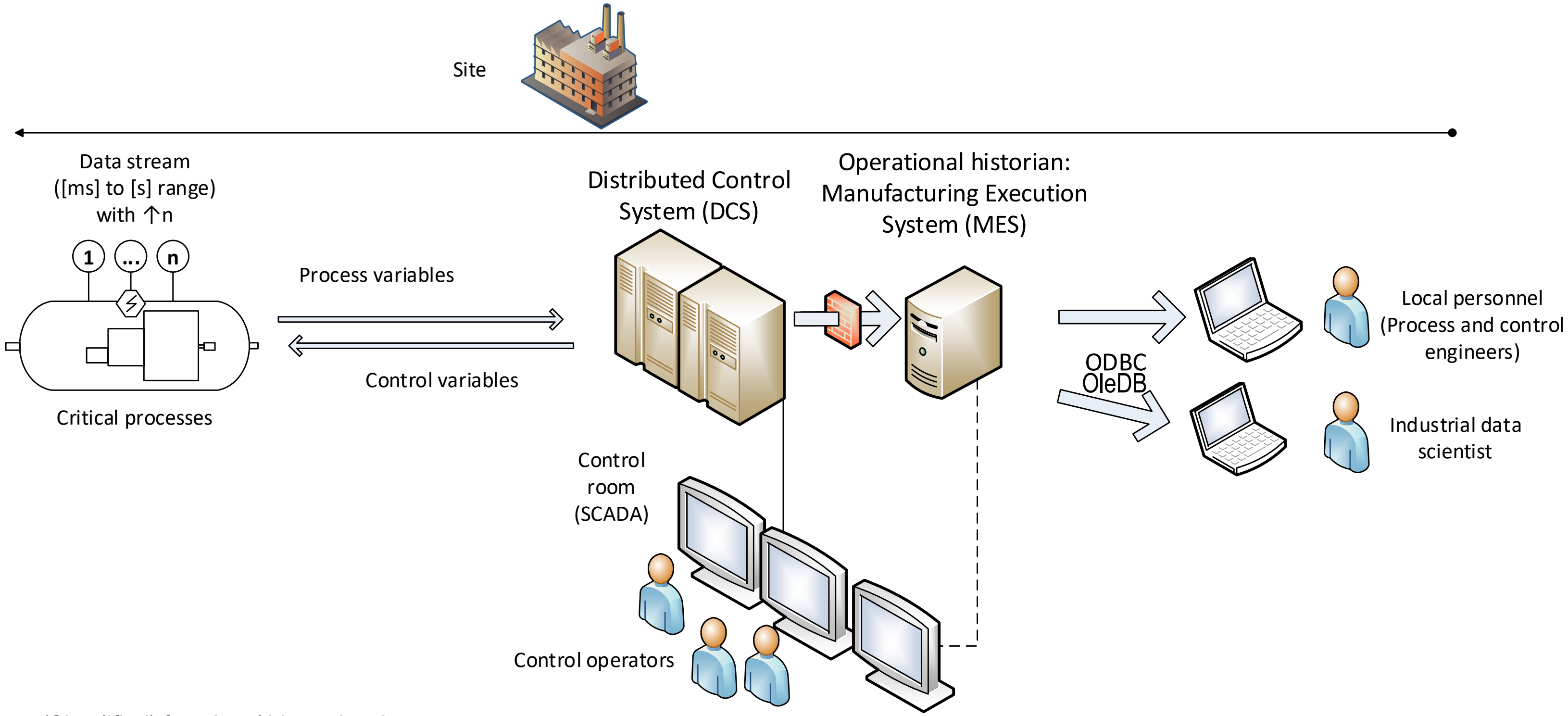
Machine Learning is not new for Chemical Engineers



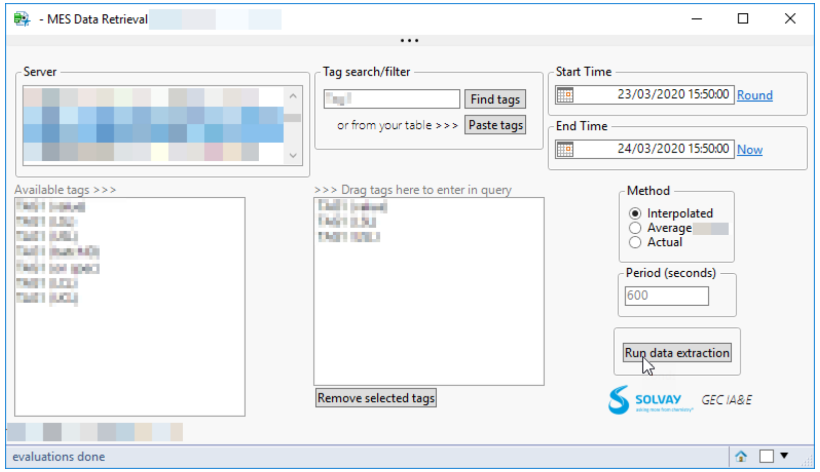


Source:
[Industrial data science – a review of machine learning applications for chemical and process industries](#) React. Chem. Eng., 2022, 7, 1471-1509

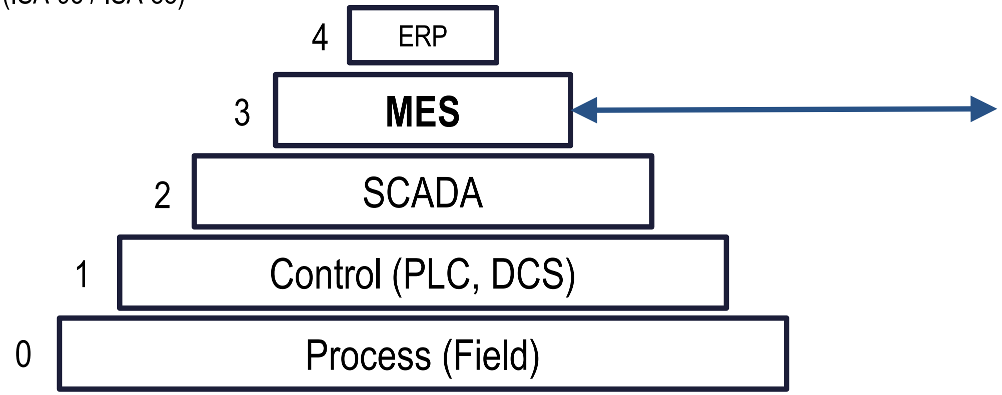
Self-service access to process data (Aspentech IP.21 and Osisoft PI)



- Open source MES data JMP add-in
- Tagreader (Python)



(Simplified) functional hierarchy view
(ISA-95 / ISA-88)



Enterprise Resource Planning

4- Establishing the basic plant schedule - production, material use, delivery, and shipping. Determining inventory levels.

Manufacturing Execution System

3- Work flow/recipe control to produce the desired end products. Maintaining records and optimizing the production process.

Manufacturing Control

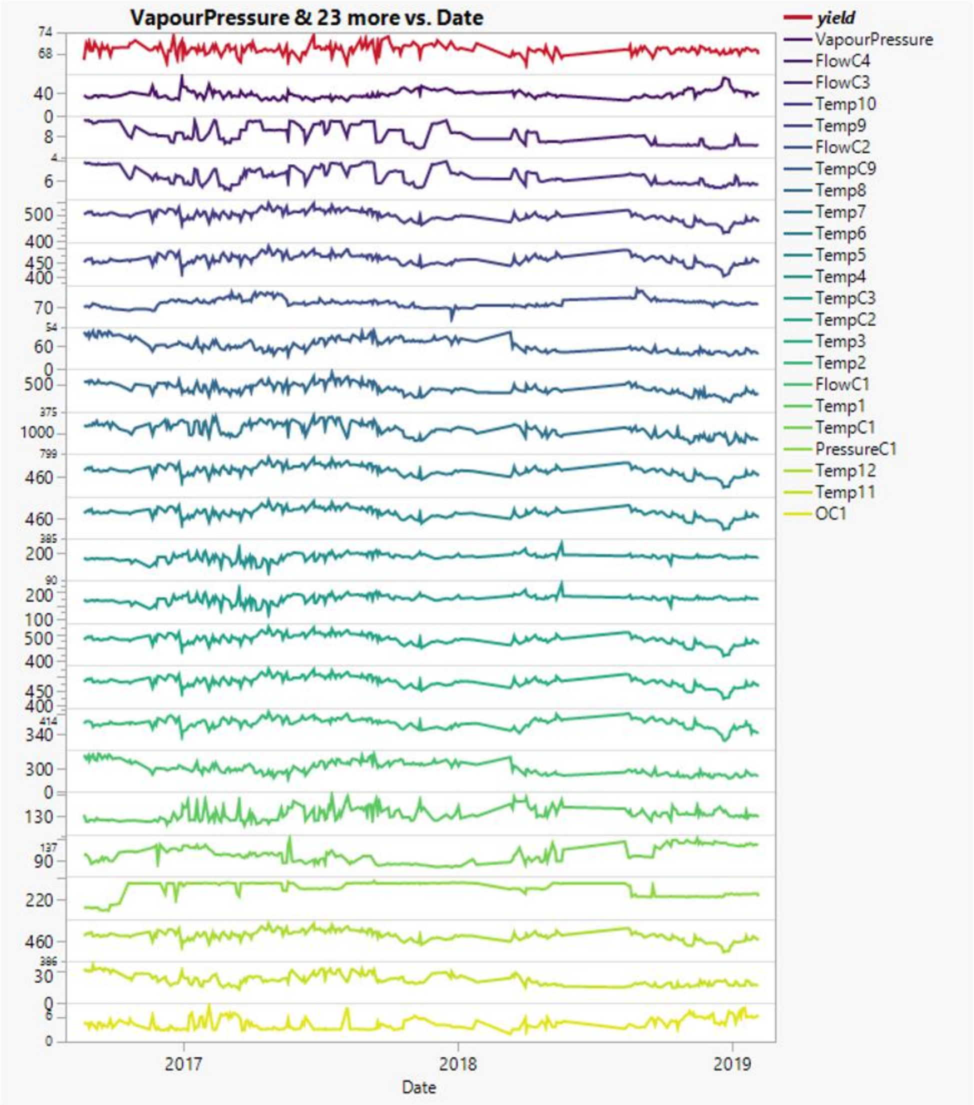
2- Monitoring, supervisory/automated control of the process

1- Sensing and manipulating the process

0- The physical production process

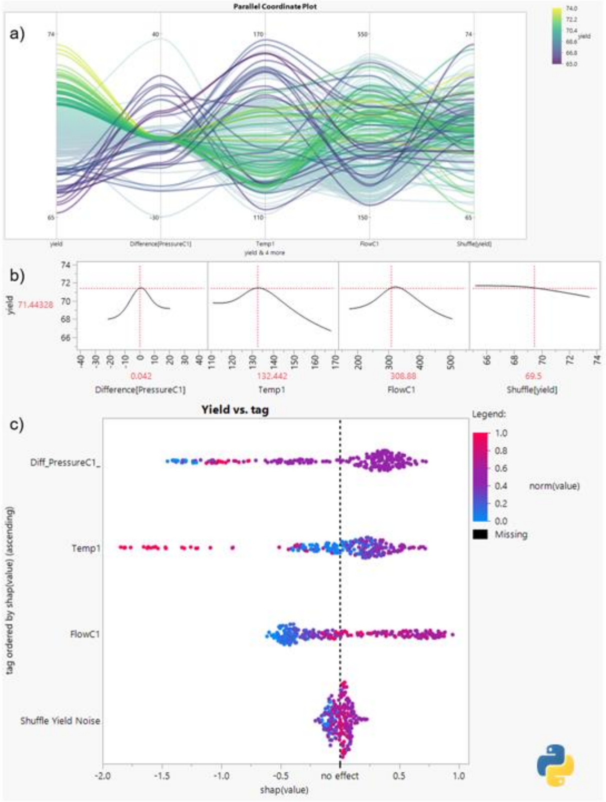
Automated ML for process engineering

Predictor Explainer JMP+Python add-in



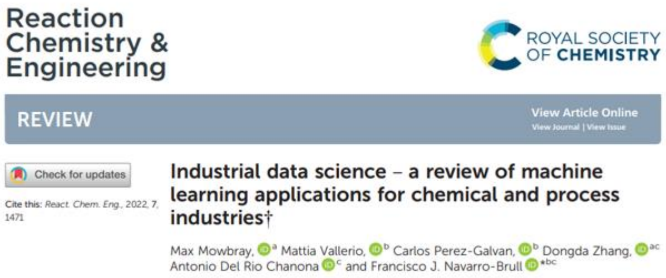
Predictor Explainer

Automated screening and selection of tags (sensors) using interpretable ML.



Open-source add-in with relevant datasets (continuous and batch processes)

Ref:


Reaction Chemistry & Engineering

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online

Check for updates

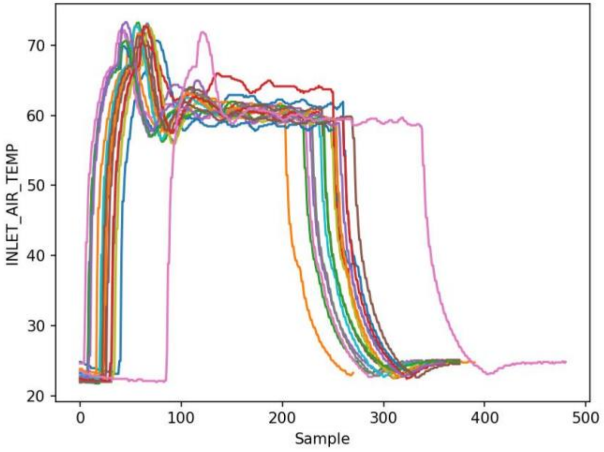
Cite this: React. Chem. Eng., 2022, 7, 1471

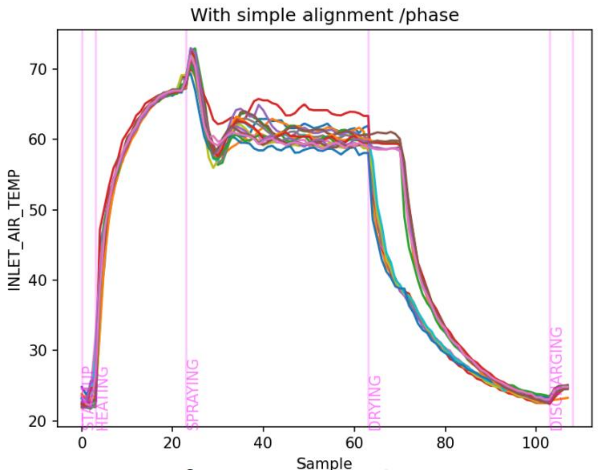
Industrial data science – a review of machine learning applications for chemical and process industries†

Max Mowbray, Mattia Vallerio, Carlos Perez-Galvan, Dongda Zhang, Antonio Del Rio Chanona and Francisco J. Navarro-Brull

Python package with applications to industrial batch data

PyPhiBatch – Multivariate Analysis of Batch Processes





python™

Years of industrial research made open-source
Phi toolbox for multivariate analysis by Sal Garcia
https://github.com/salvadorgarciamunoz/pyphi

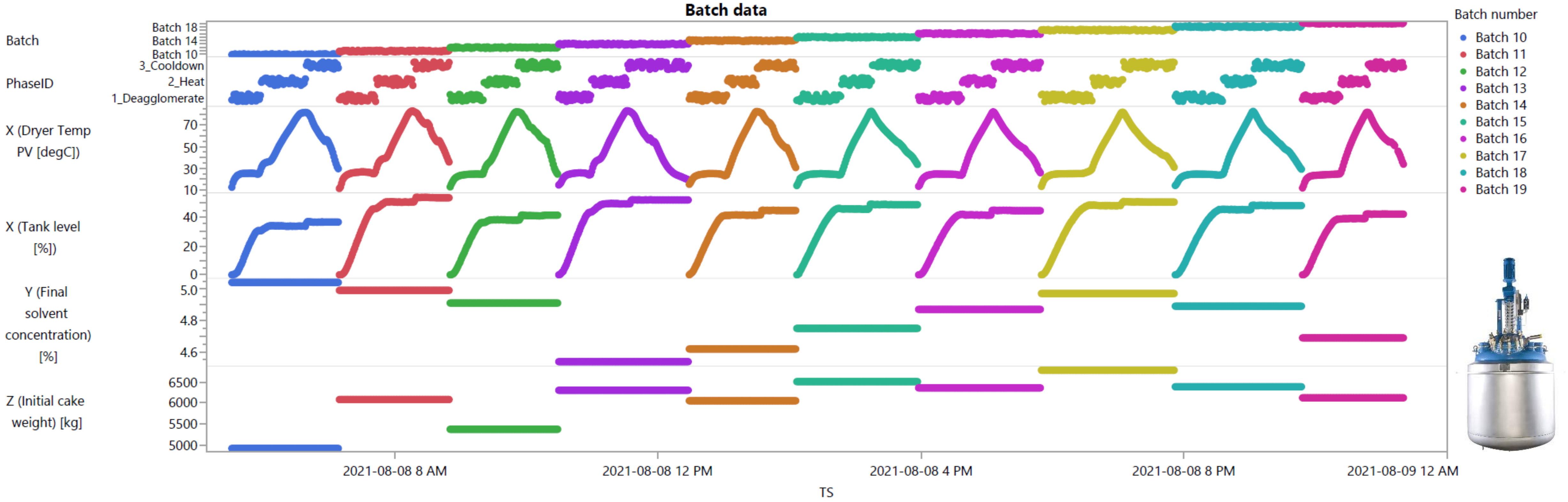
Process Analytics Course

Sargent Centre for Process Systems Engineering

https://www.imperial.ac.uk/process-systems-engineering/courses-and-seminars/workshops-and-courses/process-analytics-course-/

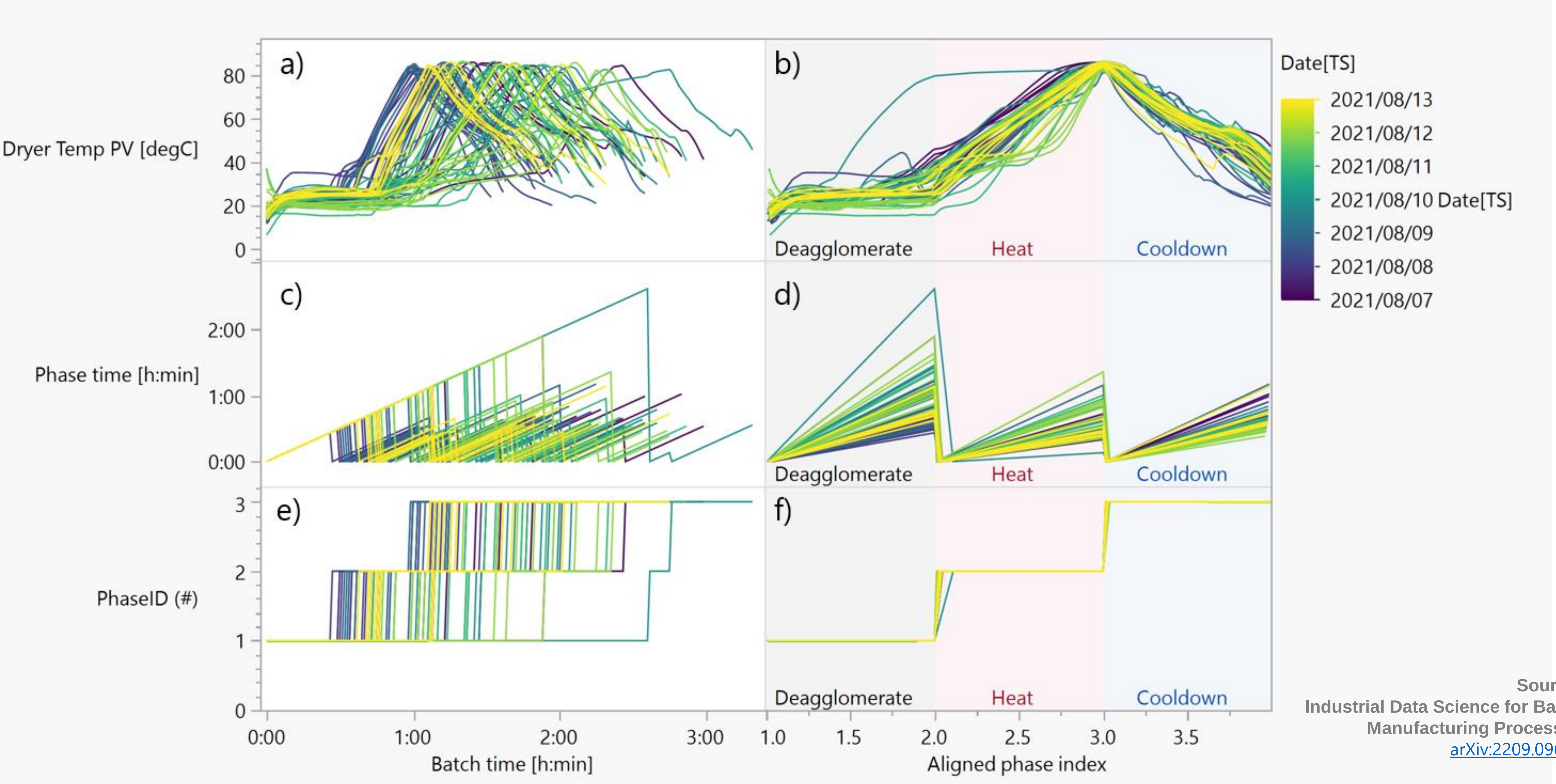
Imperial College London

Batch data analysis (dryer)



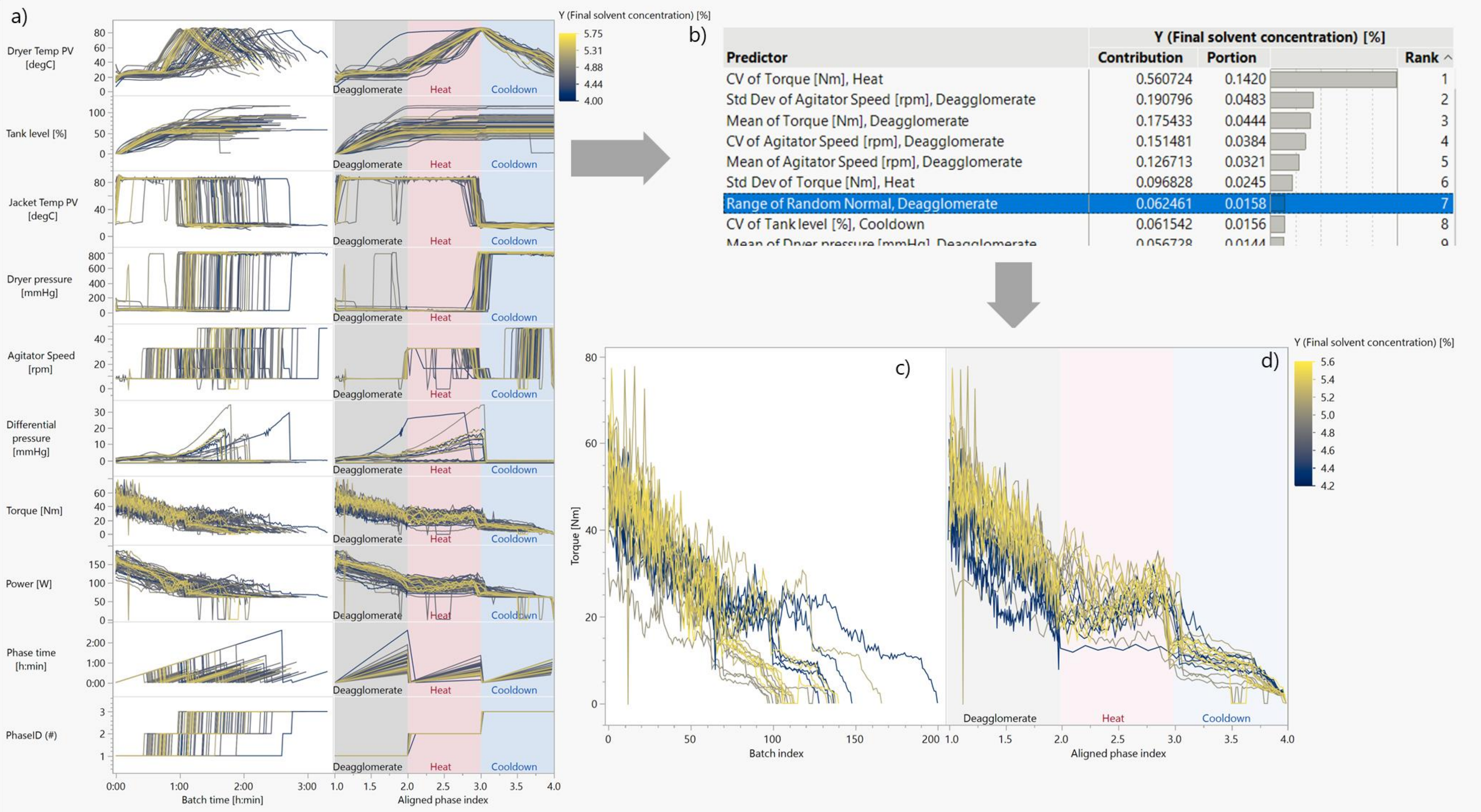
Source:
Industrial Data Science for Batch Manufacturing Processes
arXiv:2209.09660

Batch data analysis



Source:
Industrial Data Science for Batch Manufacturing Processes
arXiv:2209.09660

Batch data analysis

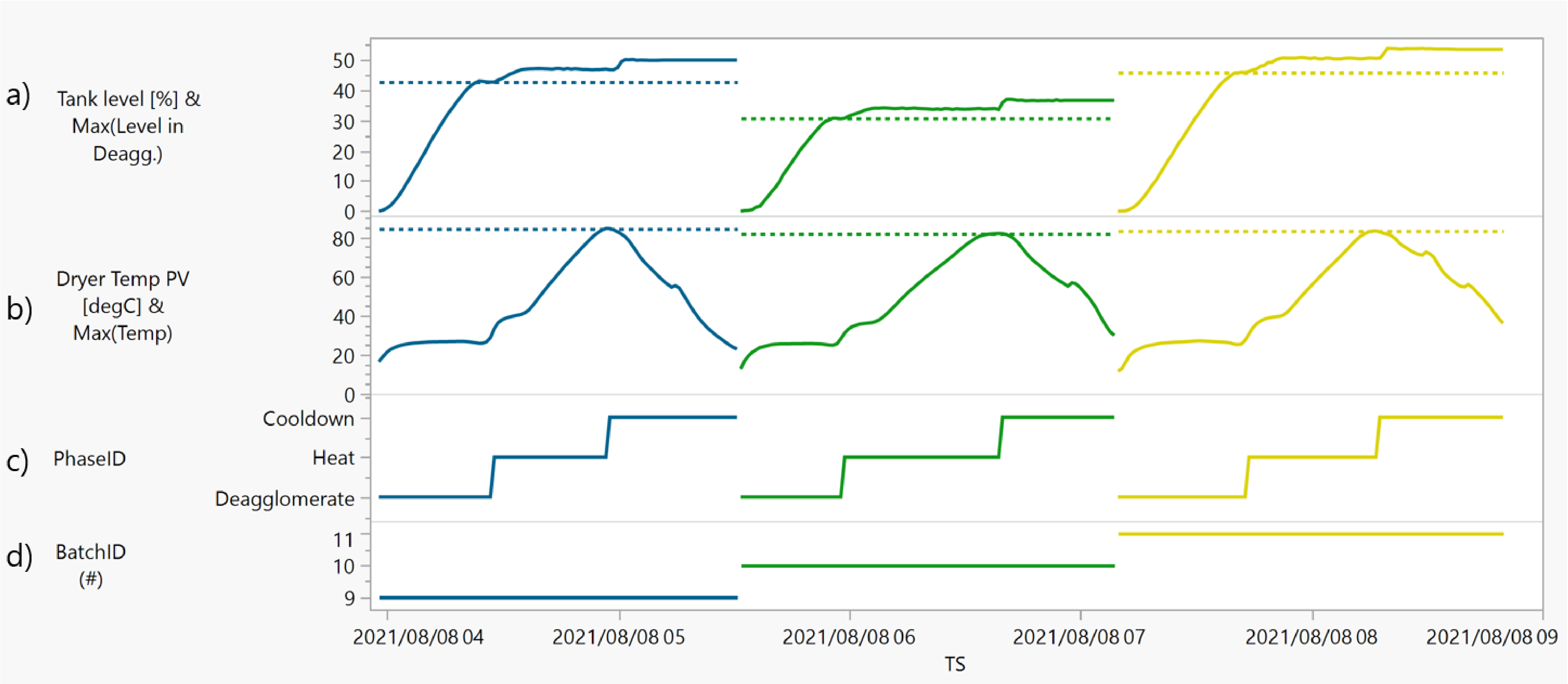


An automated analysis can be used to identifying leading correlations that can explain or predict variability seen in final solvent concentration.

Source:
Industrial Data Science for Batch Manufacturing Processes
arXiv:2209.09660

Batch monitoring

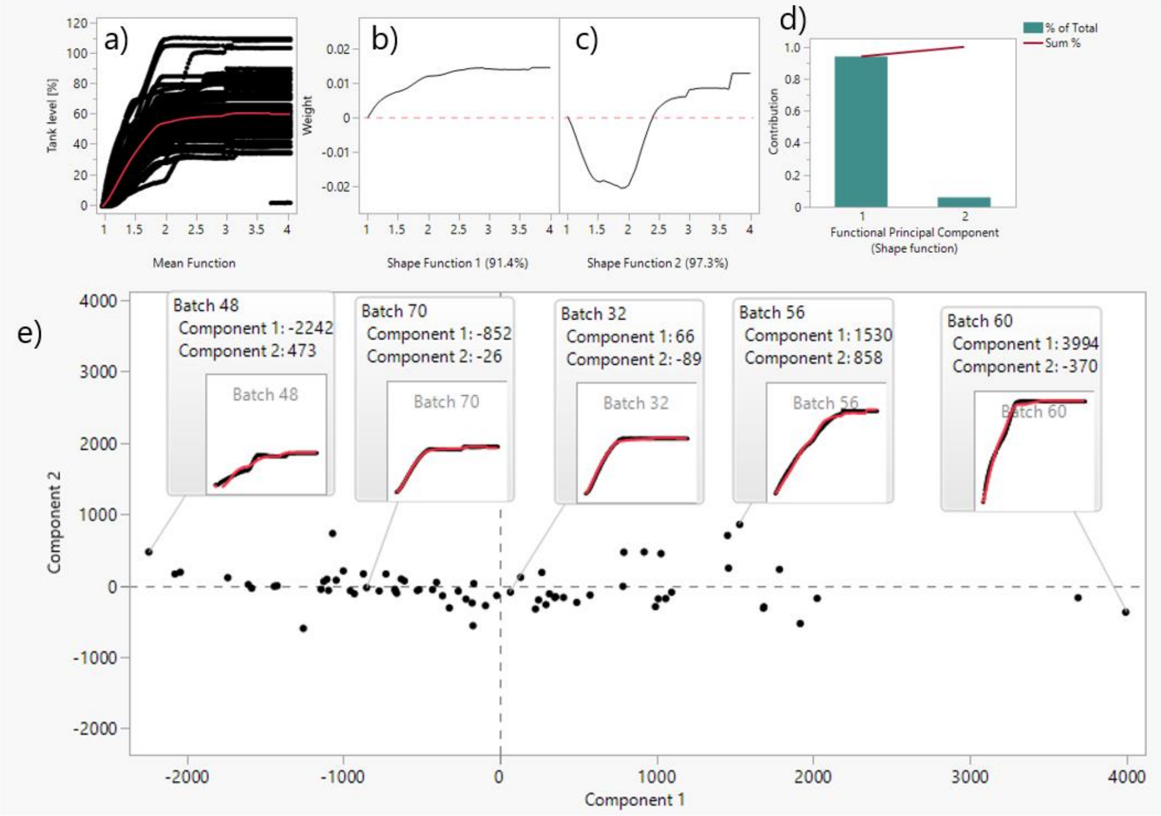
Monitoring batch processes

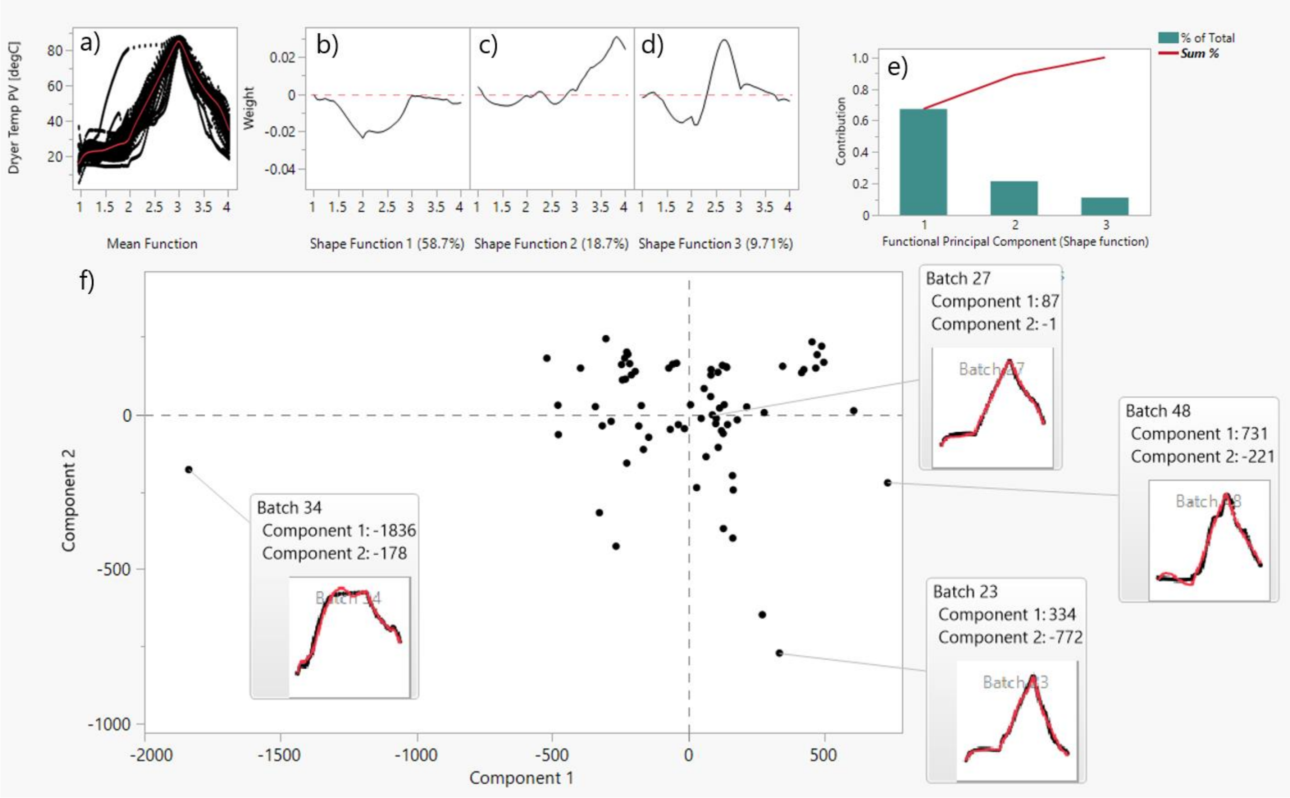


Model inputs for batch processes can be generated by summarizing the information into statistics. Here, the tank level at the end of the deagglomerate phase (a) and the maximum temperature reached during the drying process (b) can be used to capture information in specific phases (c) and summarize it per batch (d). These single values are then used for either correlation analysis or statistical process control monitoring.

Source:
Industrial Data Science for Batch Manufacturing Processes
[arXiv:2209.09660]

Functional PCA

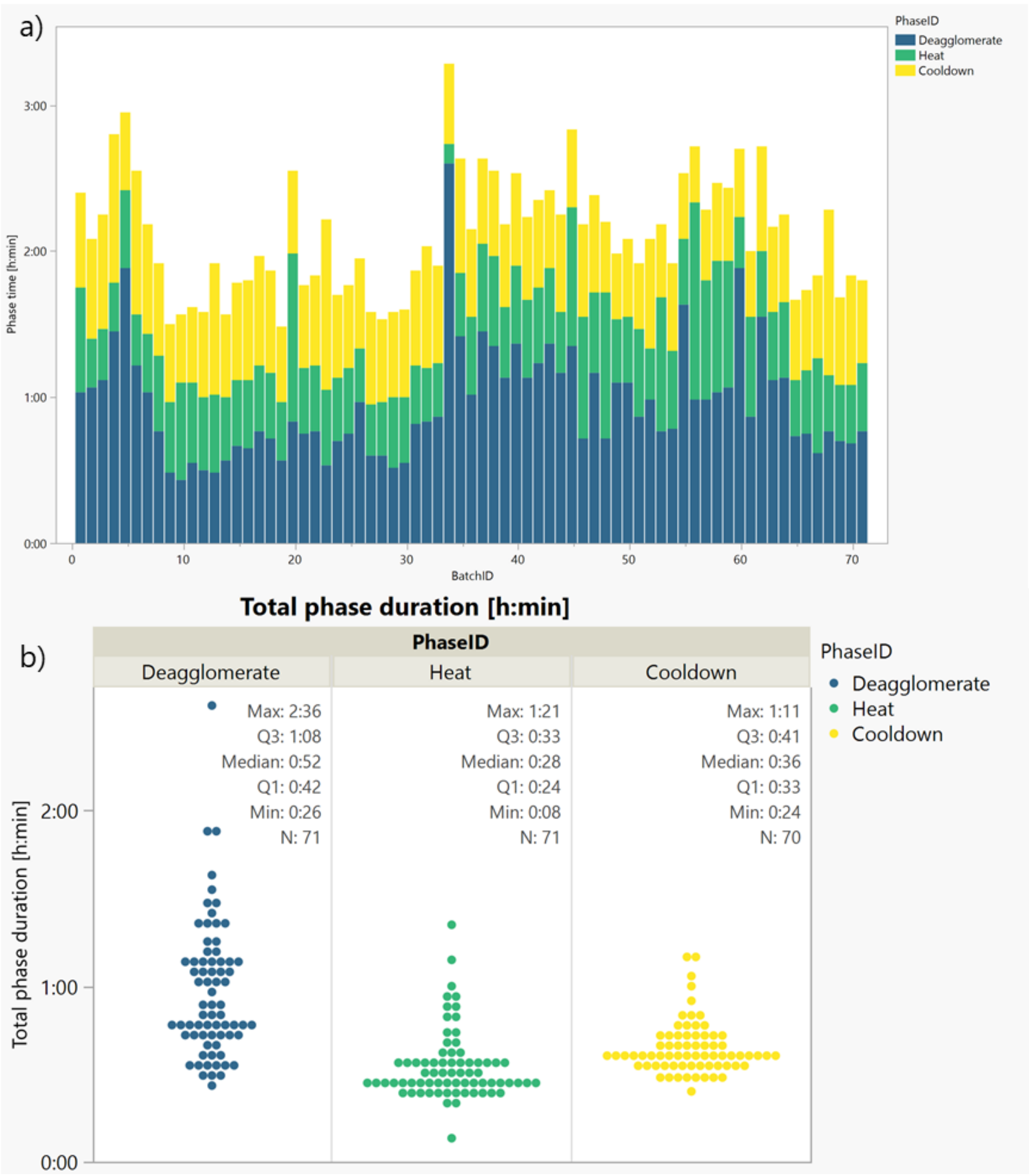


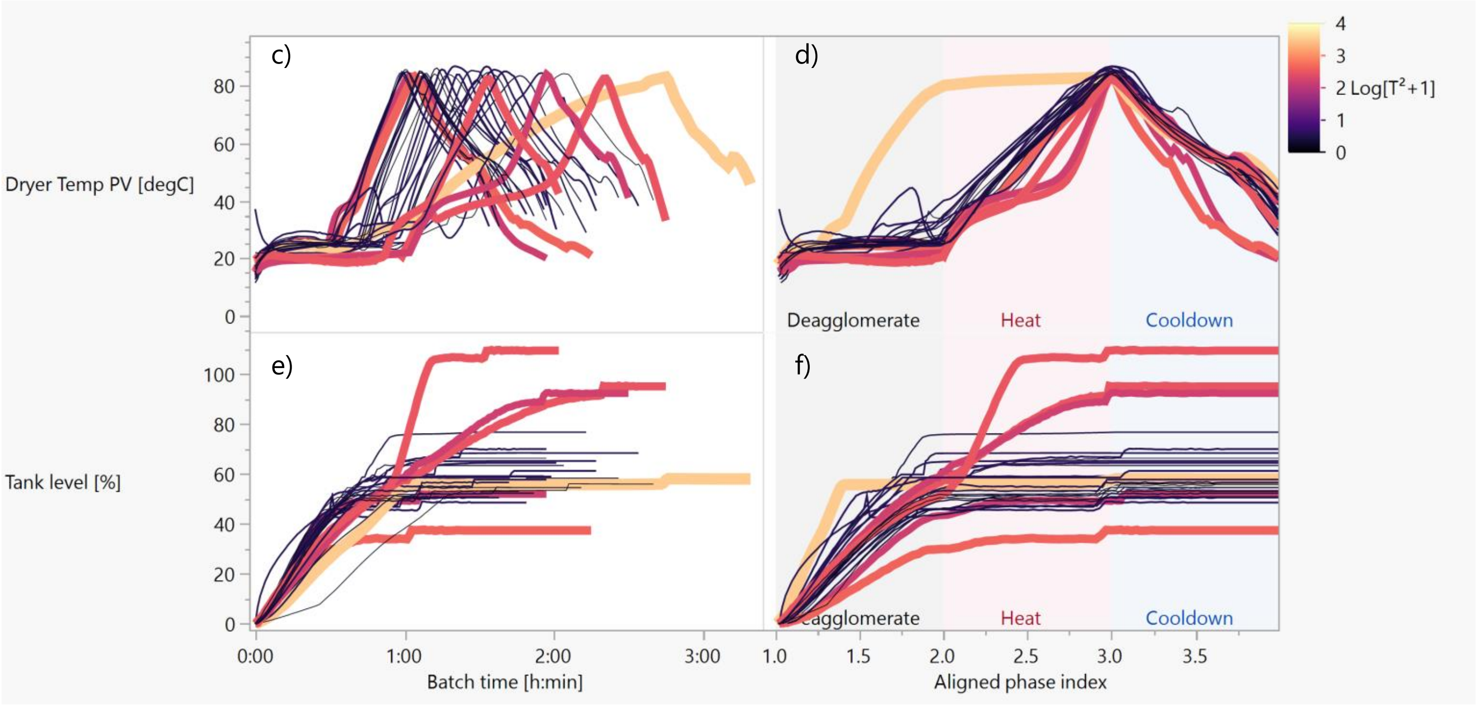


Contrary to summary statistics, Functional PCA can capture and decompose the whole trajectories seen in batch processes. Batch curves can be decomposed using the average trajectory and a combination of characteristic and independent trajectories (called shape functions). Similar to PCA, these are ordered by the amount variability they are capturing.

Source:
Industrial Data Science for Batch Manufacturing Processes
arXiv:2209.09660

Anomaly detection for batch processes





The duration of a batch drying process shows variability (a, b) due to various reasons (e.g., humidity of feeding material). This makes it difficult to monitor, since the different batch runs show different temperature and volume trajectories (c, e). Aligning batches helps identifying common profiles regardless of their duration (d, f).

Source:
Industrial Data Science for Batch Manufacturing Processes
arXiv:2209.09660

Thank you!

Benjamin Katz
Philippe Neyraval

SOLVAY

Mattia Vallerio

Carlos Perez-Galvan

Antonio
Del Rio Chanona

Francisco J.
Navarro-Brull

Sargent Centre for Process Systems Engineering

Imperial College London