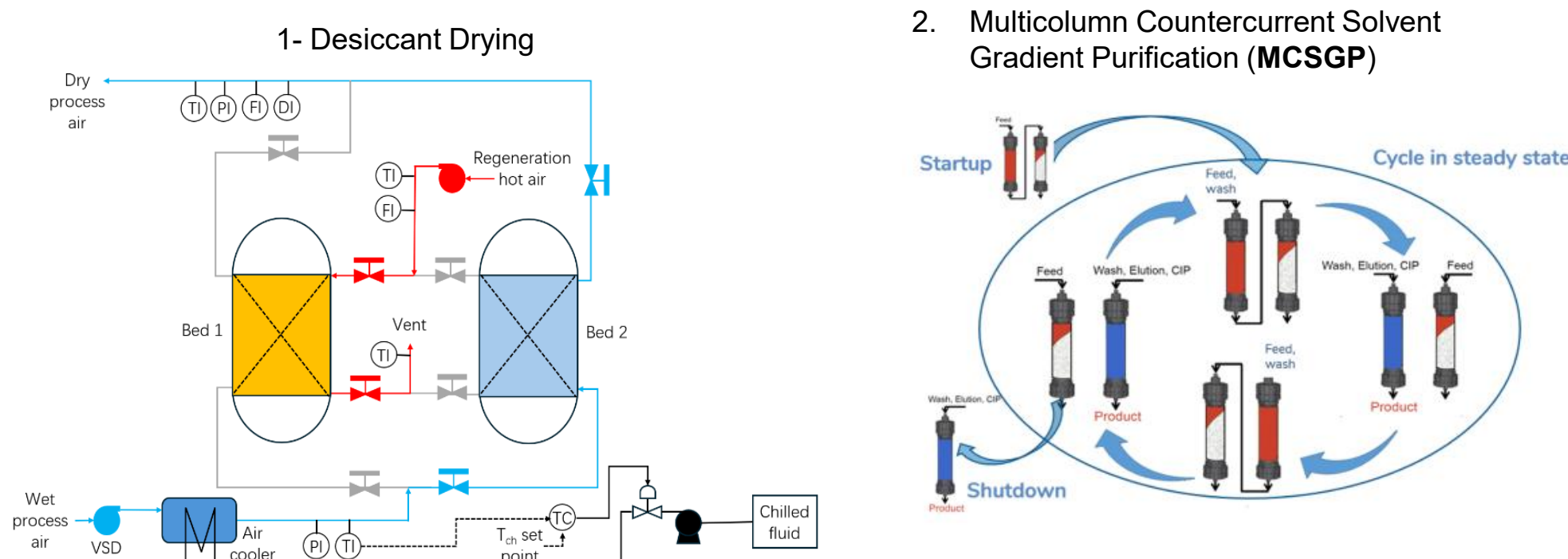


Introduction

- Several chemical processes run in **cycles** (dryers, PSA, continuous chromatography)
- Operators care about **when to switch** or **when cyclic steady state (CSS) is reached**
- **MF-kNN** reuses historical cycles to give **one-shot forecasts/ time-to-event** estimates
- Here we illustrate two applications:



MF-kNN Framework and Methodology

- MF-kNN extends traditional k-nearest neighbours (k-NN) by incorporating multiple process variables for enhanced time series forecasting

$$\mathbf{Y}_M(N) = \begin{bmatrix} y_{M1}(N) \\ y_{M2}(N) \\ \vdots \\ y_{Mp}(N) \end{bmatrix} = \begin{bmatrix} y_1(1) & y_1(2) & \dots & y_1(N) \\ y_2(1) & y_2(2) & \dots & y_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ y_p(1) & y_p(2) & \dots & y_p(N) \end{bmatrix}$$

- **Univariate k-NN** looks at one tag only

→ picks neighbours that look right locally but mismatch overall process behaviour

- **MF-kNN** uses multiple process tags

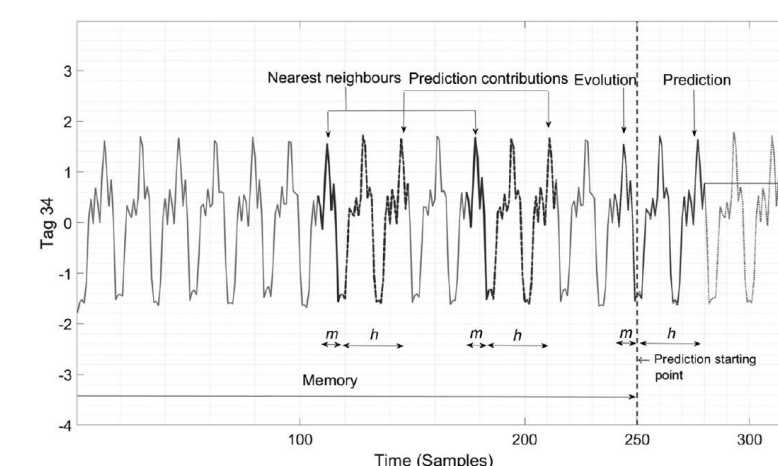
→ selects neighbours consistent with the full process behaviour

- Forecast & decision metrics:

- The **multi-feature distance** determines the nearest neighbours

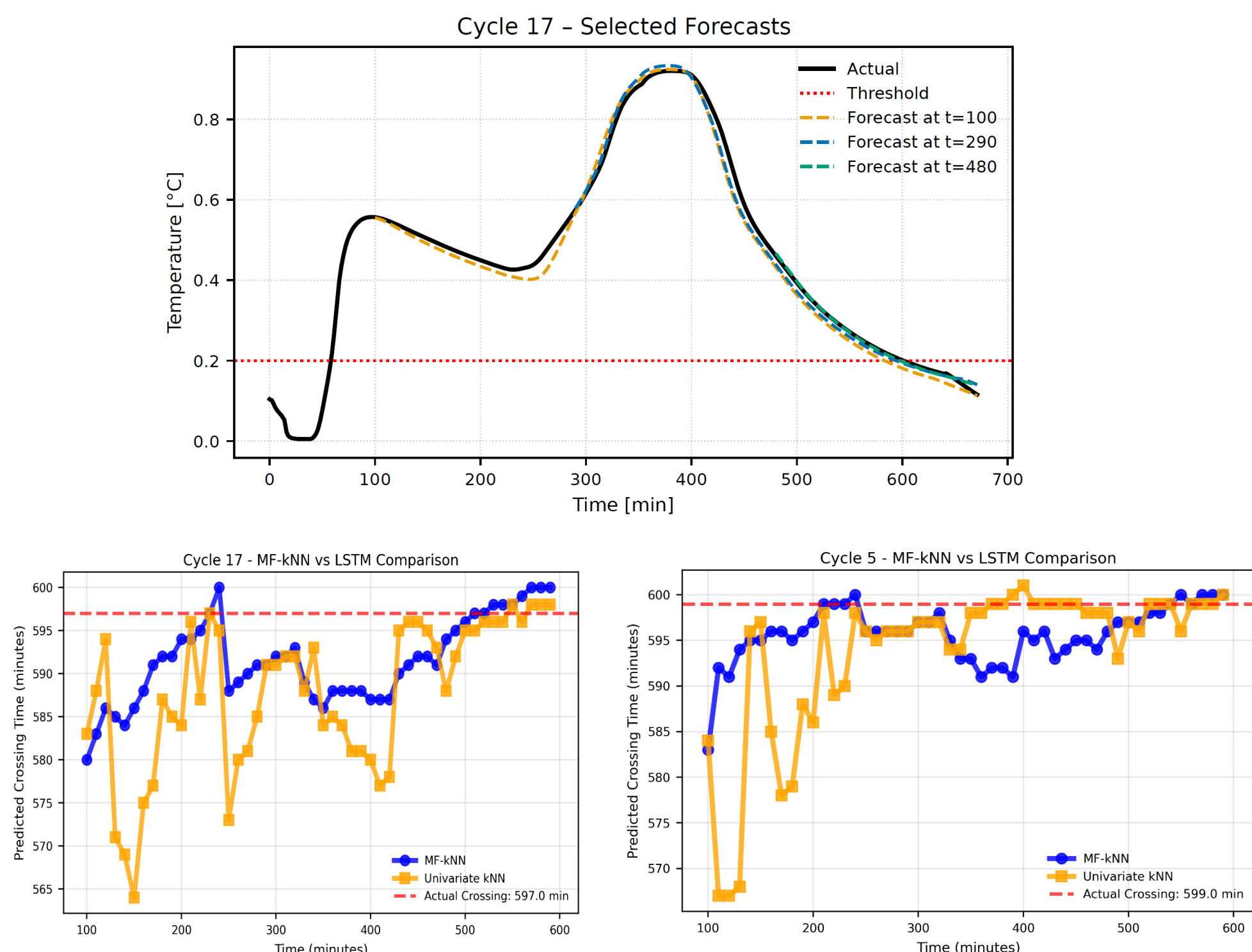
- Once selected, **only the target variable's future trajectories** are combined to produce the **forecast**

$$D(y_E, y_r) = \sqrt{\sum_{j=1}^p \sum_{i=1}^{m_j} w_{f,j} \left(y_E^{(j)}(i) - y_r^{(j)}(i) \right)^2}$$



Case Study 1 – Desiccant Drying

- Industrial **desiccant dehumidification** with alternating drying / regeneration cycles
- Decision: **when will ROT drop below 40 °C?** → time-to-threshold (TTT) problem
- Data: 52 regeneration cycles with multiple temperature and flow measurements
- MF-kNN uses multi-feature segments (RIT, ROT, etc.) and cluster-based memory
- Compare against: **univariate k-NN, ARMA, Theta, LSTM** (all one-shot)



- MF-kNN achieves the lowest threshold MAE
- **LSTM** requires training, tuning, and large datasets while MF-kNN achieves better performance *training-free*
- **Theta and ARMA struggle** because they assume linear/stationary structure, which cyclic processes violate

Model Variant	Relative Speedup
Univariate kNN	1.00x
MF-kNN	~5.00x

Model	Threshold MAE (%)
MF-kNN	2.94
LSTM	4.58
ARMA	4.64
Theta	15.45

- Neighbour search is a bottleneck for online deployment

- Vectorization and early pruning removed most of the overhead.

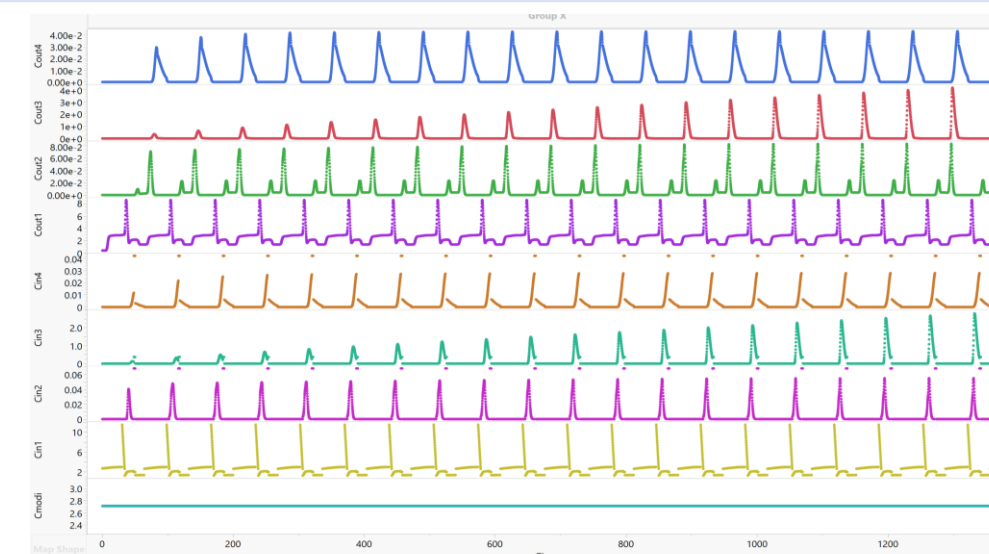
- **Result:** ~80% faster runtime with no loss in accuracy

Case Study 2 – Bioprocess MCSGP

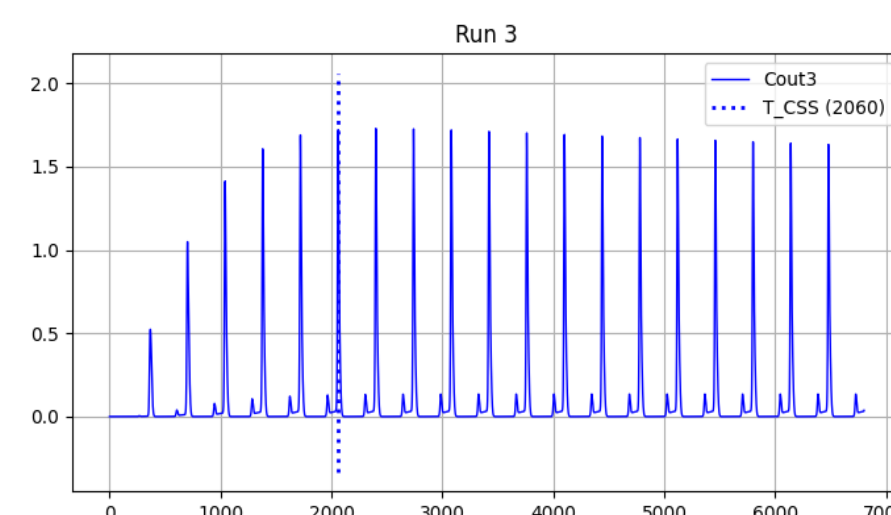
- **MCSGP**: multicolumn counter-current solvent-gradient purification; inherently cyclic.
- Need to know:

- Does the run reach **CSS**? (classification)
- When is **CSS** reached? (TCSS forecasting)

- CSS detection uses **peak amplitude differences** between consecutive cycles.



$$T_{CSS} \text{ is reached when: } x_{peak_i} - x_{peak_{i+1}} < \epsilon$$



The objective function combines profile accuracy and CSS correctness:

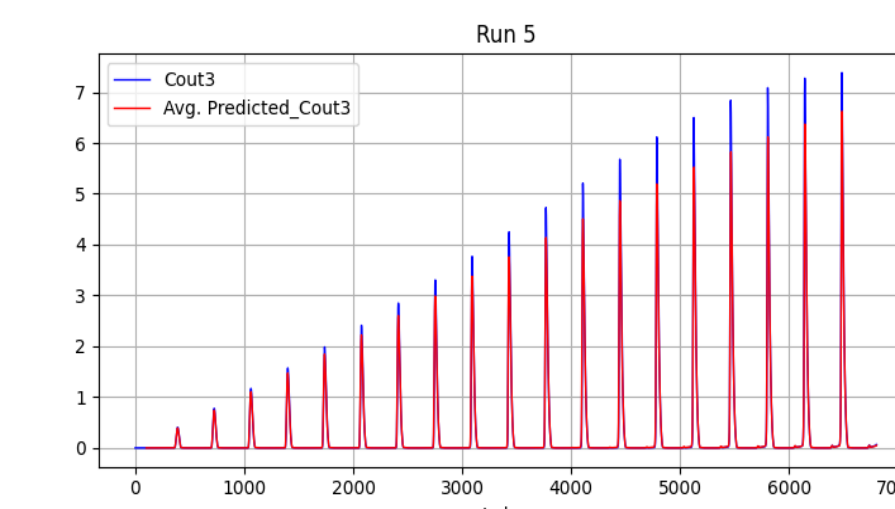
$$Obj_1 = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$Obj_2 = \begin{cases} \frac{1}{M} \sum_{i=1}^M |T_{CSS,i} - \hat{T}_{CSS,i}|, & \text{if CSS exists in both} \\ 0, & \text{if neither has CSS} \\ \text{penalty}, & \text{if incorrect CSS} \end{cases}$$

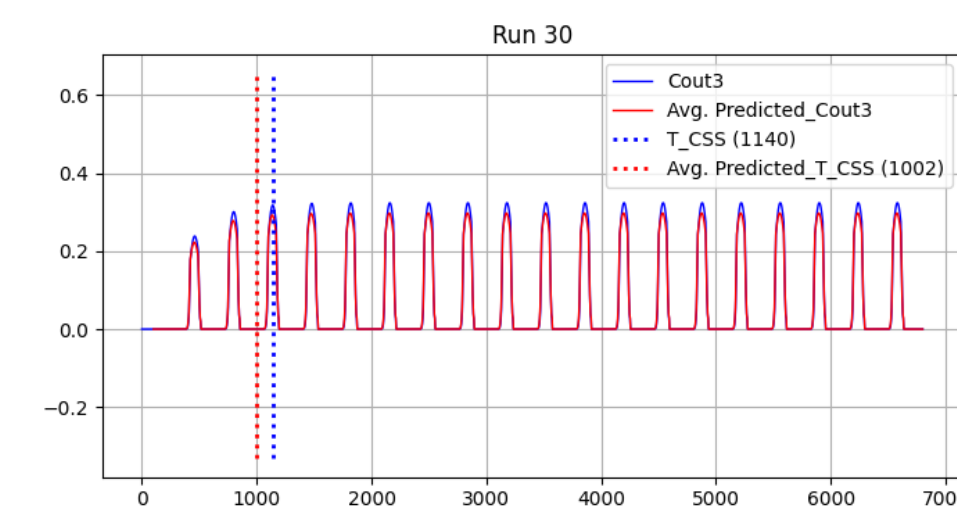
$$Obj_{overall} = (1 - \beta) \times Obj_1 + \beta \times Obj_2$$

- ❖ From the first ~100 timesteps, MF-kNN:

- ❑ Forecasts the **target concentration profile**,
- ❑ Classifies CSS vs non-CSS (neighbour majority),
- ❑ Forecasts TCSS when CSS is classified



- Correct classification



- Correct classification, and
- Accurate T_{CSS} forecast

Conclusions & Outlook

- **MF-kNN provides a unified, interpretable forecasting tool** for cyclic chemical and biochemical processes
- In **desiccant drying**, it delivers **accurate time-to-threshold (TTT)** predictions from early-cycle data
- In **MCSGP**, it correctly predicts **CSS vs non-CSS** and provides **time-to-CSS (TCSS)** forecasts using only the first ~100 timesteps
- **Multi-feature distance matching** enables more reliable neighbour selection than univariate k-NN
- **Future work:** integrate MF-kNN forecasts into **supervisory control / switching logic** for fully automated cycle management

Deployment Advantages:

- ✓ No retraining
- ✓ Fast neighbour search
- ✓ TTT from early-cycle data
- ✓ Robust for cyclic industrial units

[1] Algoufily, Y., Borghesan, F., & Mercangöz, M. (2025). Predictive Modelling of Desiccant Drying Processes Using Multi-Feature k-Nearest Neighbours Algorithm. IFAC PapersOn-Line, 59(6), 205-210. <https://doi.org/10.1016/j.ifacol.2025.07.146>

[2] Algoufily, Y., Borghesan, F., Vanhoutte, D., Mercangöz, M., 2025. Multi-Feature k-Nearest Neighbors for Time-Series Forecasting in Cyclic Processes: Algorithmic Enhancements and Performance Analysis. Submitted to *Computers & Chemical Engineering* (under review).

[3] Borghesan, F., Chioua, M., & Thornhill, N. F. "Forecasting of process disturbances using k-nearest neighbours, with an application in process control." *Computers & Chemical Engineering*, 128, 188–200 (2019).