

Abschlusskonferenz – TP7

Recipe optimization of batch distillation trajectories based on a data-driven model

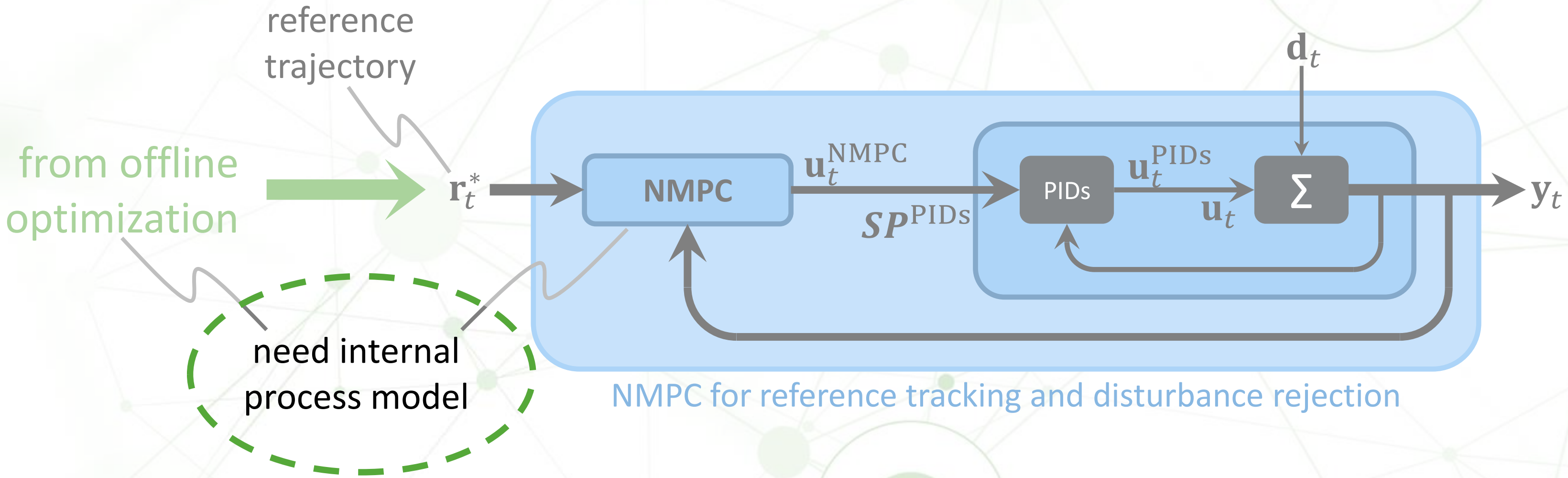
Merlin Schueler*
Corina Nentwich
Michael Kawohl

Gerardo Brand Rihm*
Erik Esche
Jens-Uwe Repke

25.05.2023

Goal

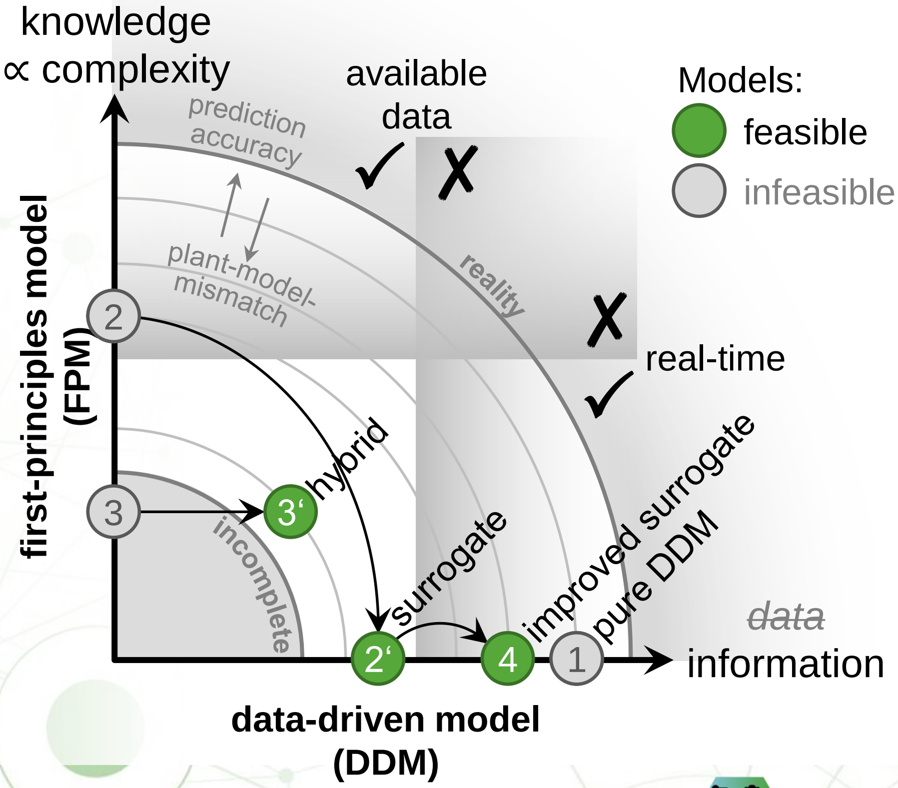
Optimal operation of batch distillation cycles in an existing plant



Why data-driven models?

When are data-driven models (DDM) justified?

1. when no first-principles model (FPM) is available, but real plant data are
   - ➢ideal case, but expensive experiments
2. when the FPM is hard to evaluate or has convergence issues (e.g., large MIMO systems)
   - ➢surrogate DDM for real-time
3. as part of a hybrid (grey-box) model
4. when the FPM is inaccurate, but also little real-plant data are available
   - ➢transfer learning (TL)/domain adaptation



Case study: Batch distillation

Based on real plant at EVONIK Leading Beyond Chemistry

- mixture: methanol + water
- packed vacuum column, fully automated

Simulation (first-principles)

- **pressure-driven**
- **whole batch distillation cycles**
  - from cold and empty startup (inertized at 1 bar)
  - to shutdown (cold and inertized)
- flowsheet simulator: Aspen Plus dynamics

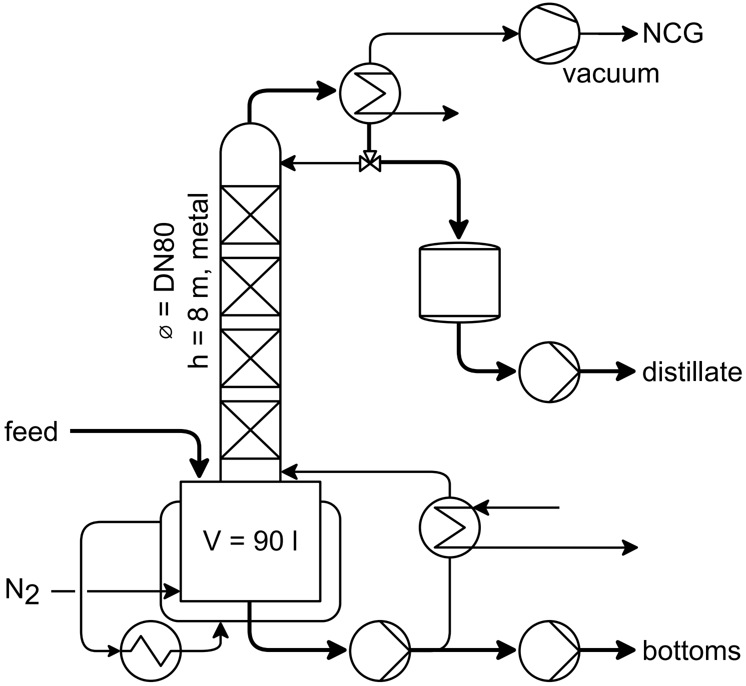
FPM

Data-driven model (DDM)

- large MIMO system
- RNN (2 LSTM layers) in Tensorflow/Keras

dynamic long-term dependencies

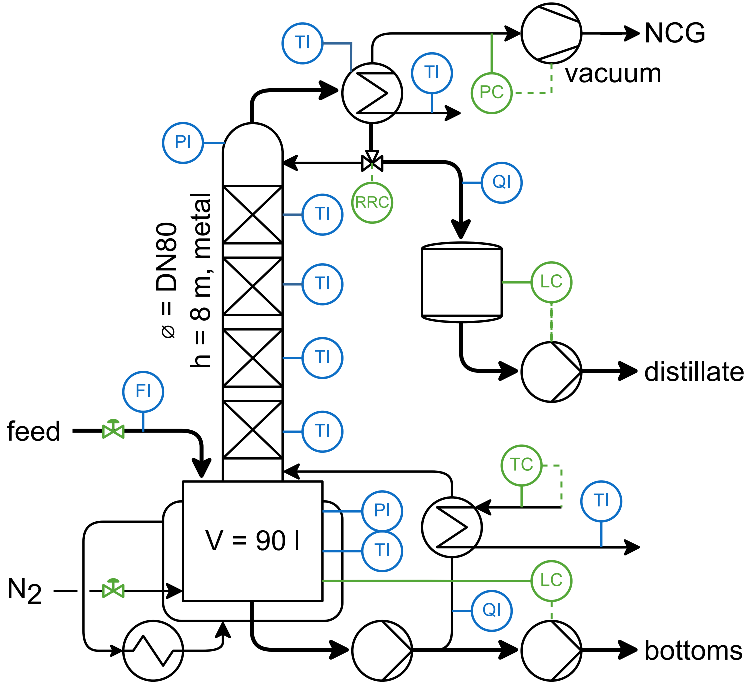
TensorFlow Keras DDM



Case study: Batch distillation

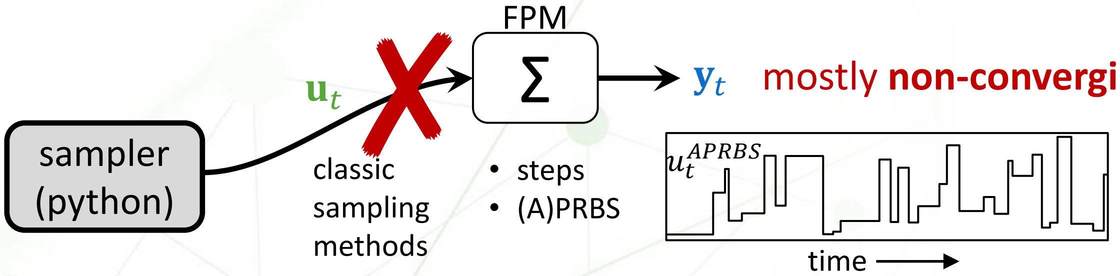
MIMO system

inputs / controls	#
valve positions	2
heating medium temperature	1
reflux ratio	1
controller setpoints (SP)	2
controller modes (AUTO/MAN)	3
total manipulated variables (MV) n_u	**9**



outputs / states	#
temperatures	6
mass flows	6
distillate mass fractions	3
pressures	3
heating/cooling duties	2
liquid levels	4
cumulative mass flows	2
total process variables (PV) n_y	**26**

Dynamic sampling



mostly **non-converging** simulations

...and many others

Input signal (time seires)

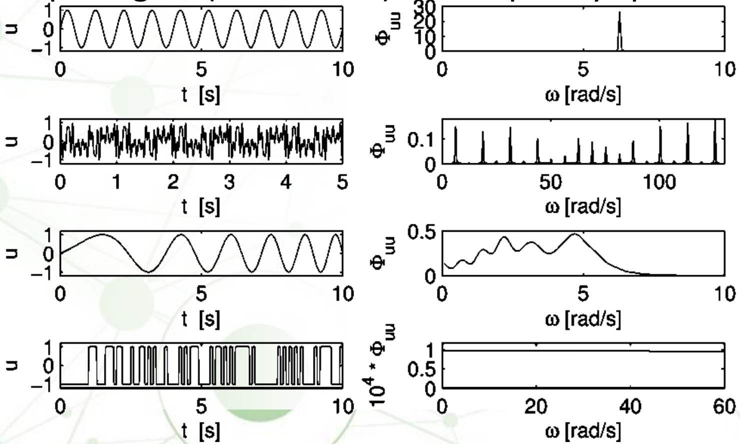
Frequency spectra

sine

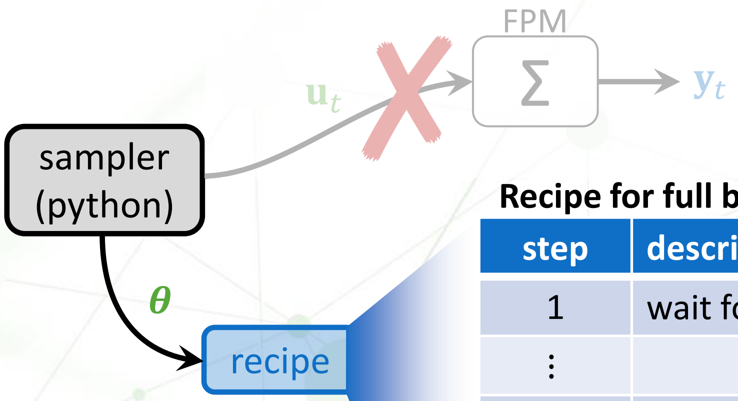
Multiple sine with random phase

Chirp

Pseudorandom binary sequence (PRBS)



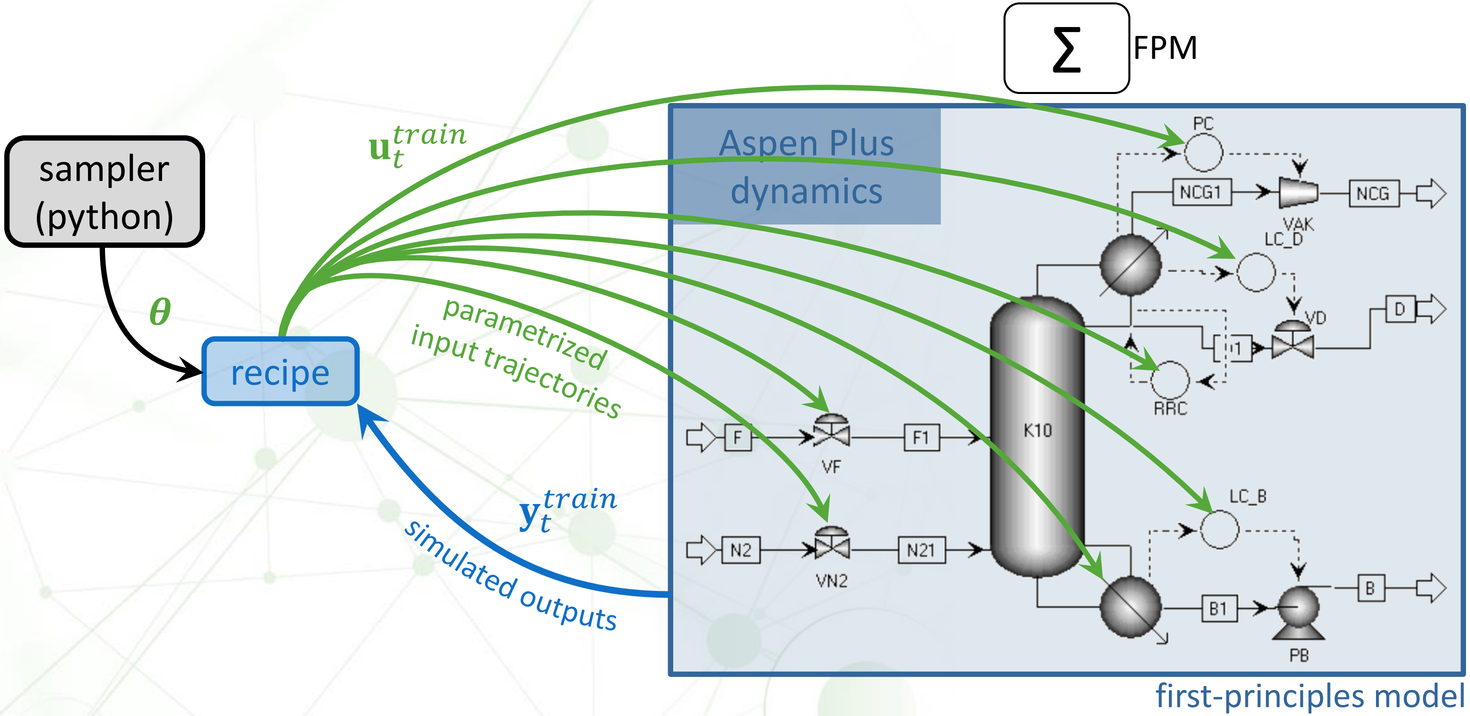
Recipe sampling



Recipe for full batch distillation cycle

step	description
1	wait for time θ_1 with cold and empty system
⋮	⋮
6	heat up reboiler ramping up $T^{Reb}_{med,in}$ to θ_7 during a period of θ_8
⋮	⋮
11	wait until achieving a cumulative distillate composition of θ_{14}
⋮	⋮

Recipe sampling



Dynamic sampling

sampler (python)

$\boldsymbol{\theta}^{real}$ → recipe → $\mathbf{u}_t^{real}$ → Real plant → $\mathbf{y}_t^{real}$

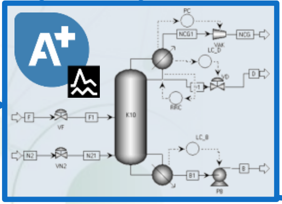
$\boldsymbol{\theta}^{sim}$ → recipe → $\mathbf{u}_t^{sim}$ → First-principles model → $\mathbf{y}_t^{sim}$

Real plant

- packed column
- 39 equilibrium stages (estimated)
- unknown parameters:
  - × heat transfer coefficients
  - × geometries
  - × feed compositions

RELATED
(but different)

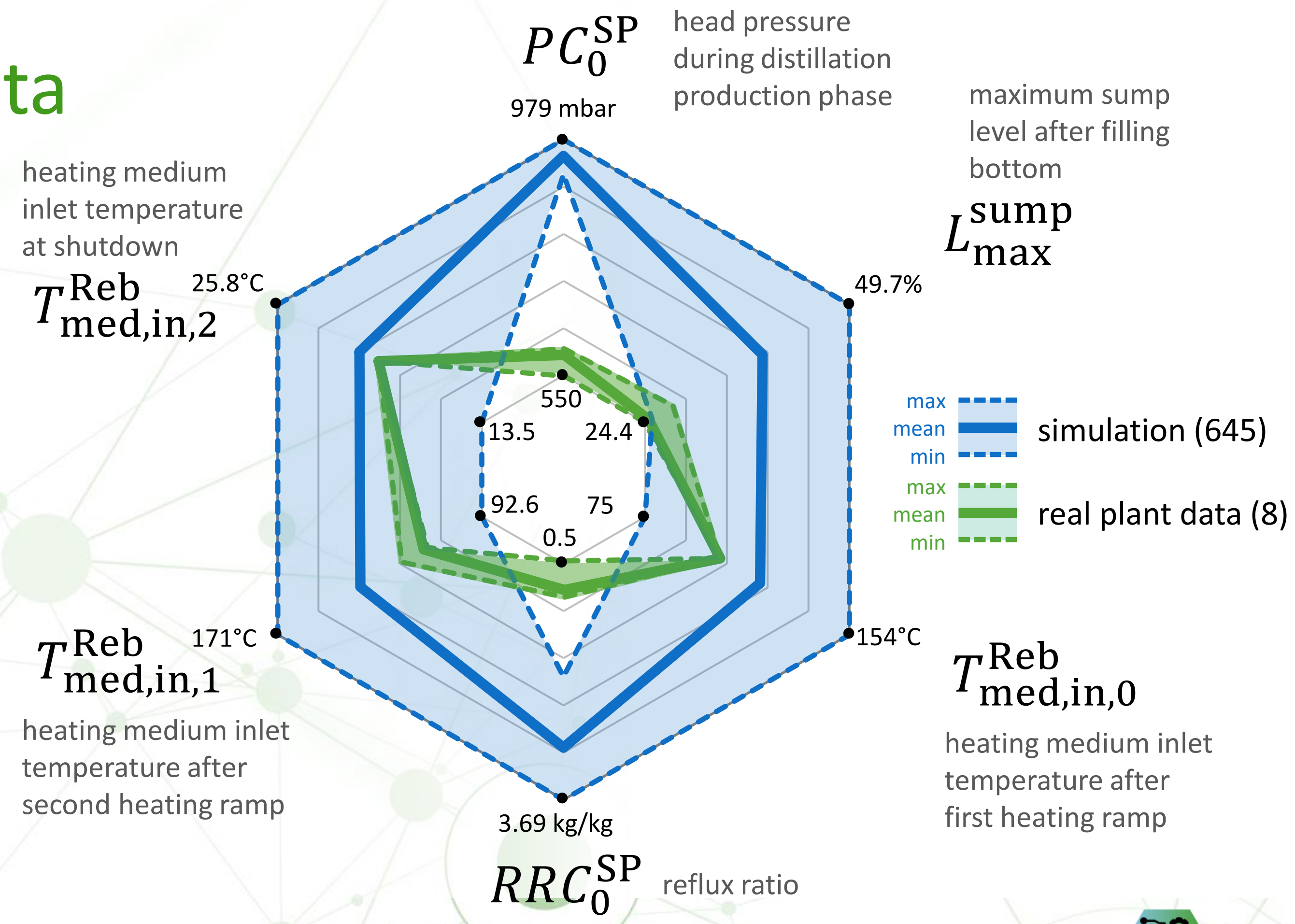
First-principles model



- sieve trays
- 10 equilibrium stages
- “guessed” parameters
- only accepts ‘sramp’s

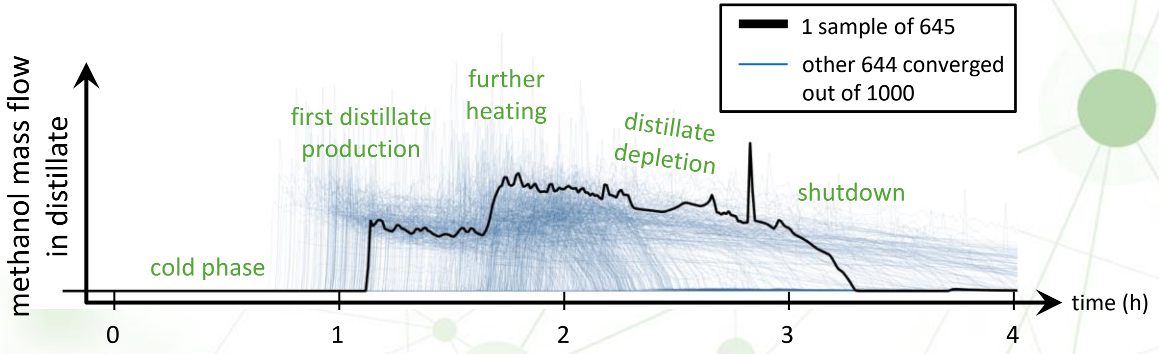
Available data

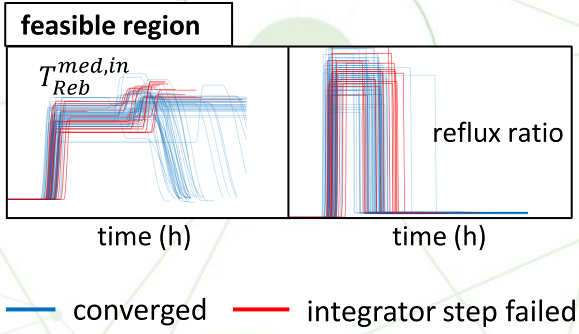
6 common (sim, real) recipe parameters were identified



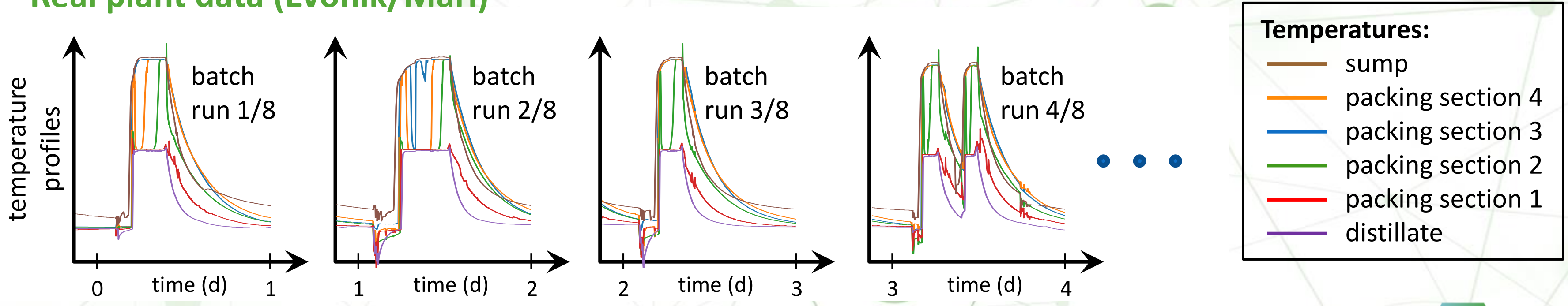
Datasets

Simulated data from first-principles model (Aspen Plus Dynamics)

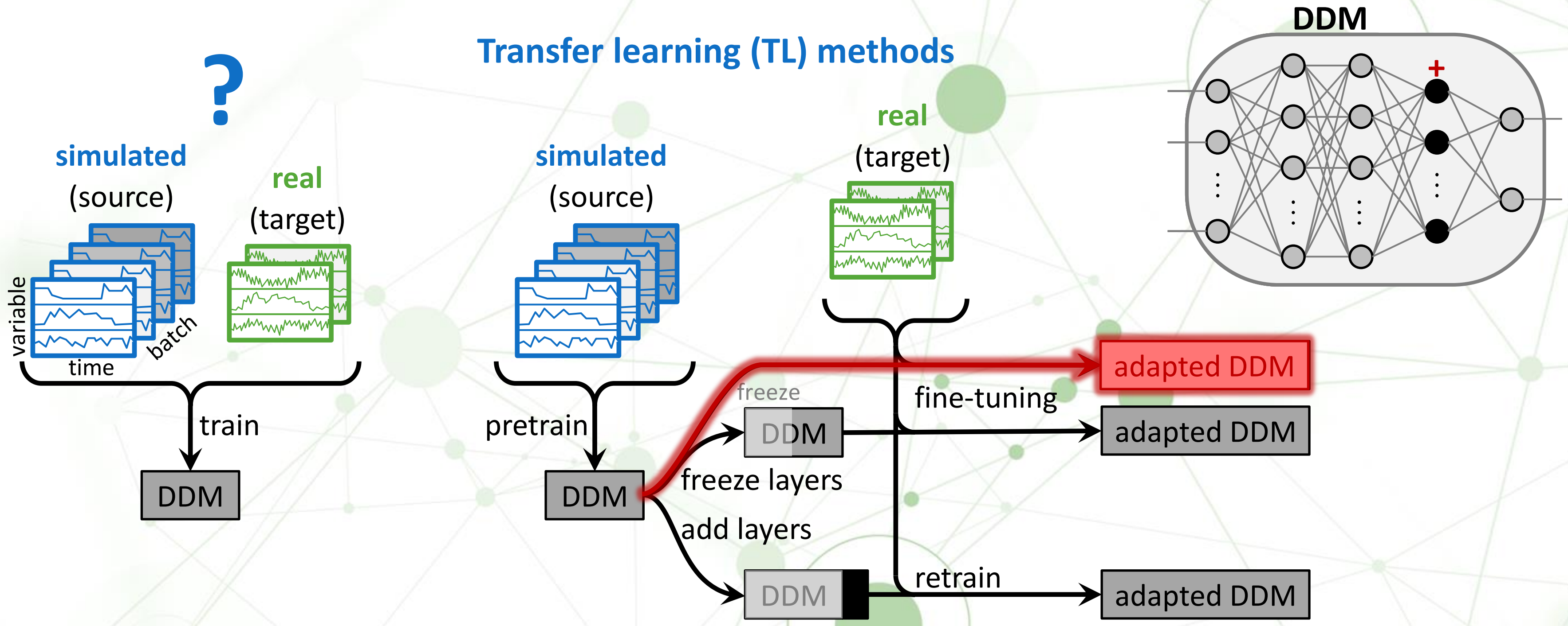




Real plant data (Evonik/Marl)

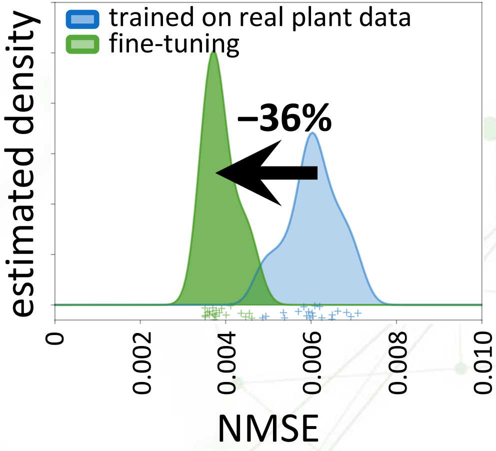


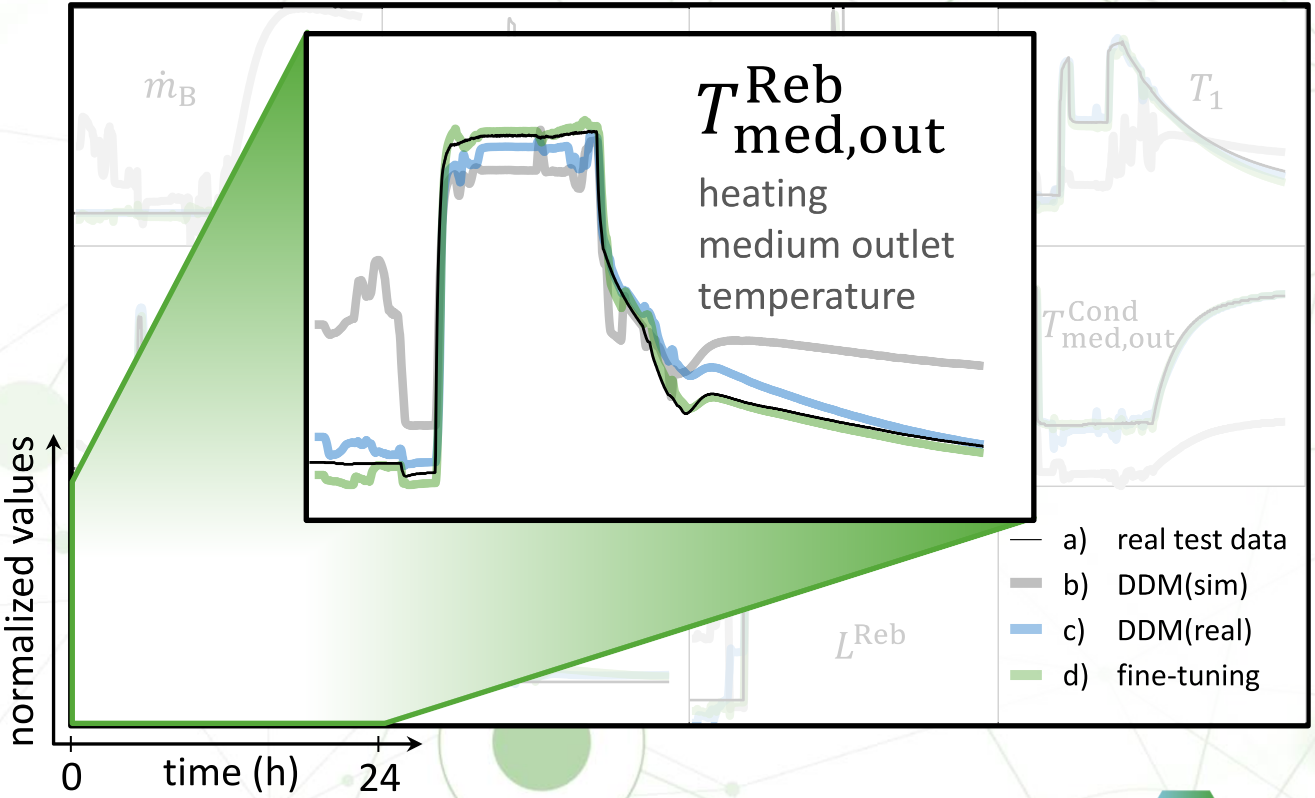
How to unify both data sources into one model?



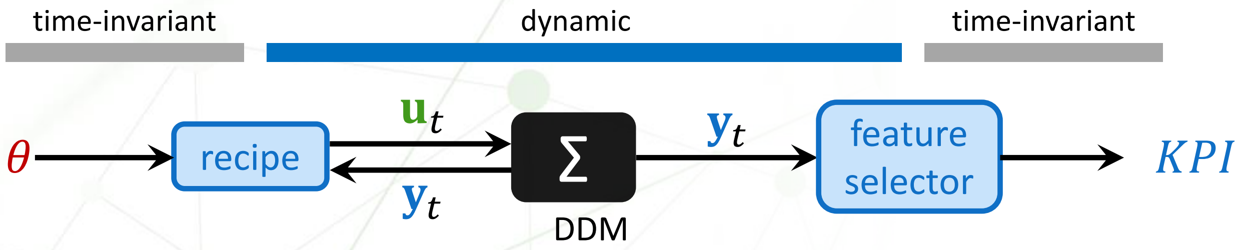
Test predictions

Predictions over real plant test data





Offline optimization options



Optimal control problem (OCP)

$\mathbf{u}_t \xrightarrow{\text{RNN}} \mathbf{y}_t$

$$\boxed{\min_{\{\mathbf{u}_t\}} f(\mathbf{u}_t)}$$

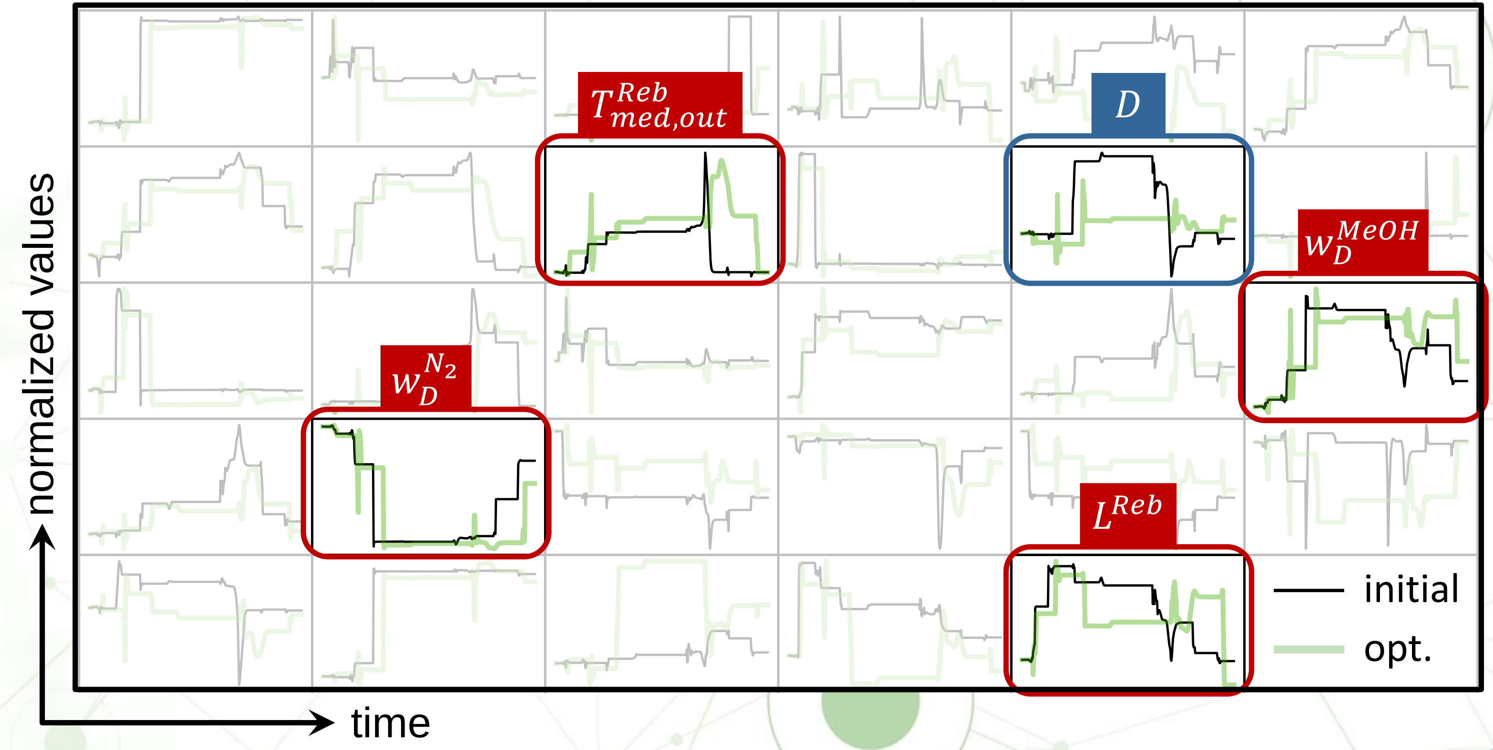
Nonlinear optimization problem (NLP)

$\theta \xrightarrow{\text{multilayer perceptron (MLP)}} KPI$

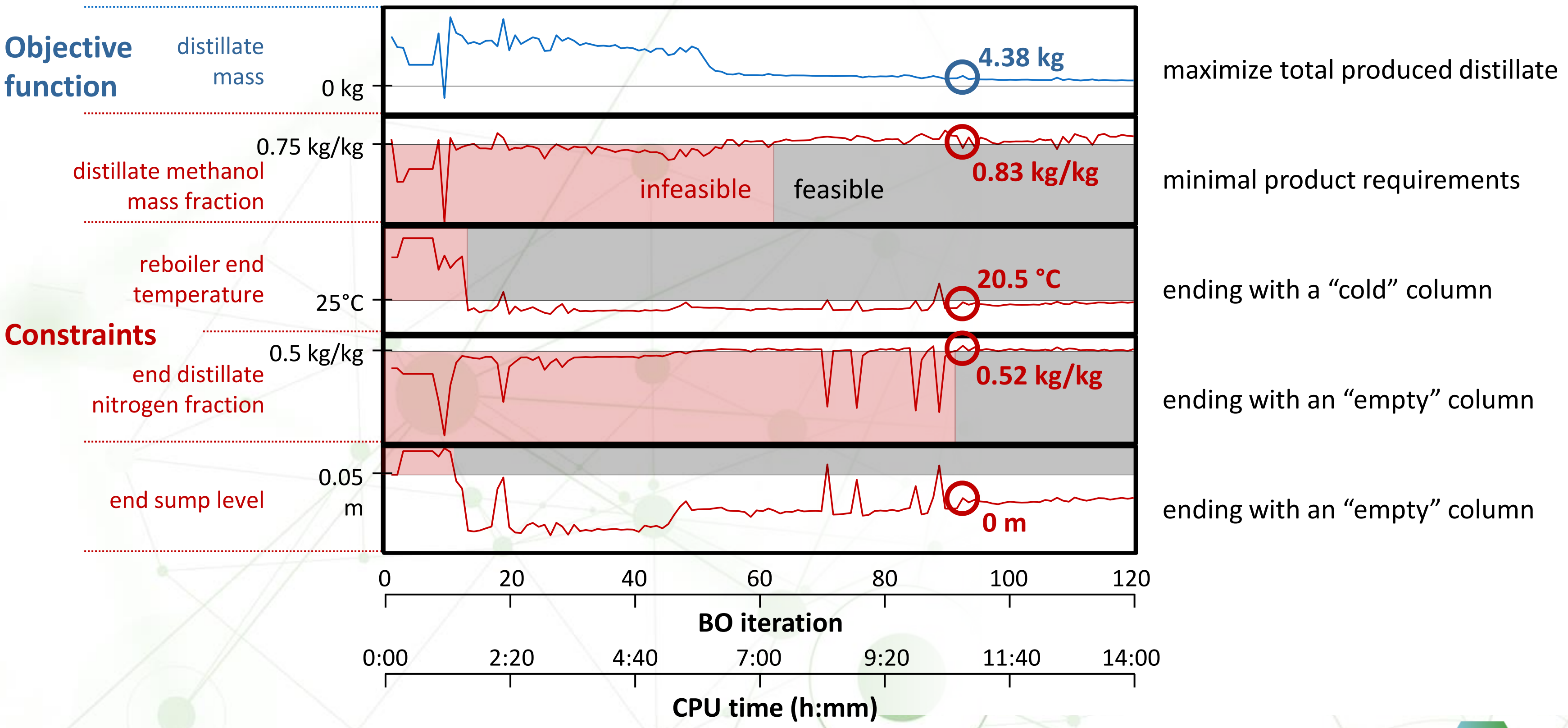
$$\boxed{\min_{\{\theta\}} f\big(\mathbf{u}_t(\theta)\big)}$$

Optimized trajectories

States $\hat{\mathbf{y}}_t\left(\mathbf{u}_t(\theta^*)\right)$

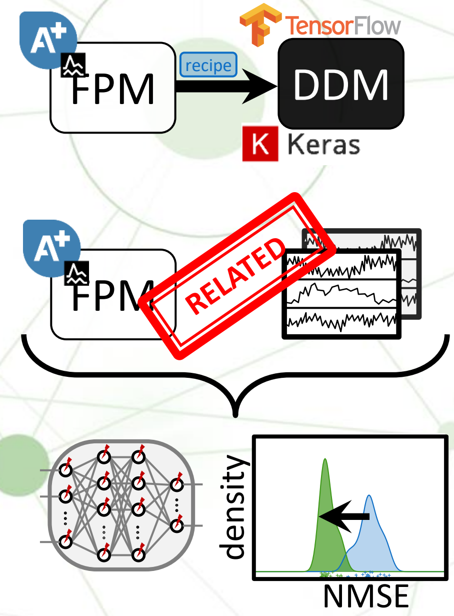


Optimization results



Conclusions

- **Main findings**
  - a **surrogate dynamic data-driven model** of a batch distillation could be trained on **simulated data** from **recipe** sampling
  - **scarce real plant data** and a related **first-principles** model can be unified via **transfer learning** (TL)
  - an **offline trajectory** for a simulated batch distillation column was **optimized**
- **Outlook**
  - **validation** of offline trajectories in the **real plant**
  - use the data-driven model for **NMPC**



$$\min_{\{\theta\}} f\left(\mathbf{u}_t(\theta)\right)$$

Acknowledgements: This research has been supported by the project "KI-Inkubator-Labore in der Prozessindustrie - KEEN", funded by the Bundesministerium für Wirtschaft und Energie (**BMWK**) under grant number **01MK20014T**.