Development and utilization of a dynamic gray-box model for a fermentation process of spore production

Joschka Winz [a], Supasuda Assawajaruwan [b],
Uwe Piechottka [b], Sebastian Engell [a]

[a]: TU Dortmund university, Process dynamics and operations group

[b]: Evonik Operations GmbH

23.05.2023

Supported by:

Federal Ministry for Economic Affairs and Climate Action

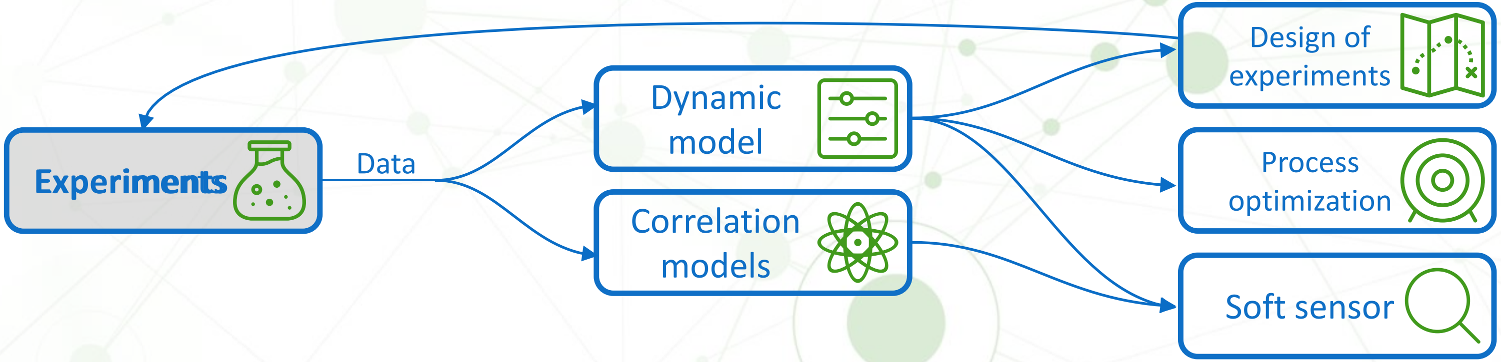
on the basis of a decision by the German Bundestag

EVONIK Leading Beyond Chemistry

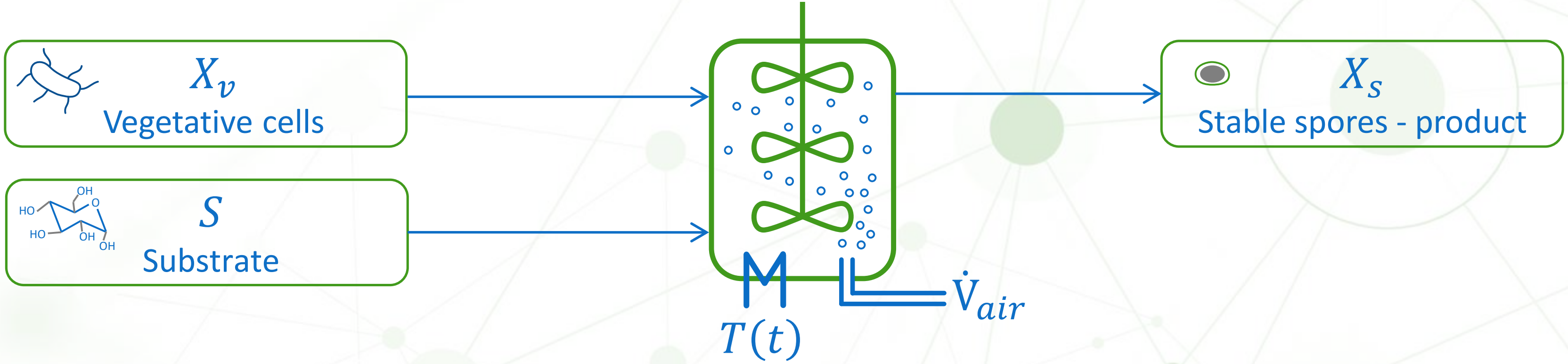
Use case fermentation of *Bacillus Subtilis*

- Fermentation of the sporulating bacterium *Bacillus Subtilis*
  - Long batch time
  - Unknown optimal operation
  - Time- and labor-intensive analytical processes to control the quality and quantity of the product

- **Objective 1: Process optimization -** Optimize the batch time while maintaining a sufficient product yield
- **Objective 2: Soft sensor -** Develop a state estimator to gain online insights into the fermenter state

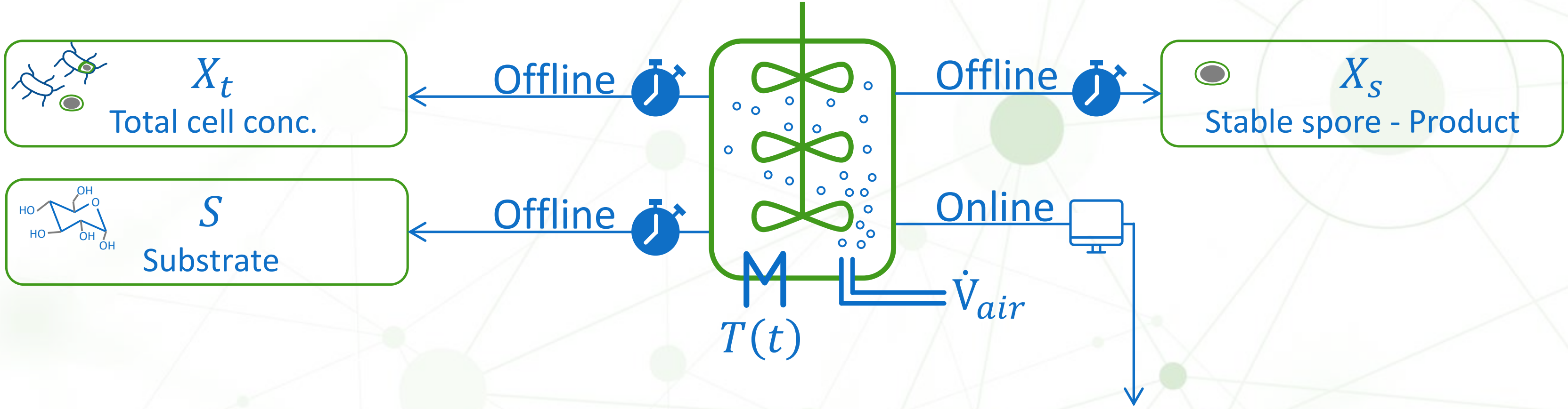


Sporulation of *Bacillus Subtilis*



- Process optimization: Find the optimal temperature trajectory to minimize the batch time

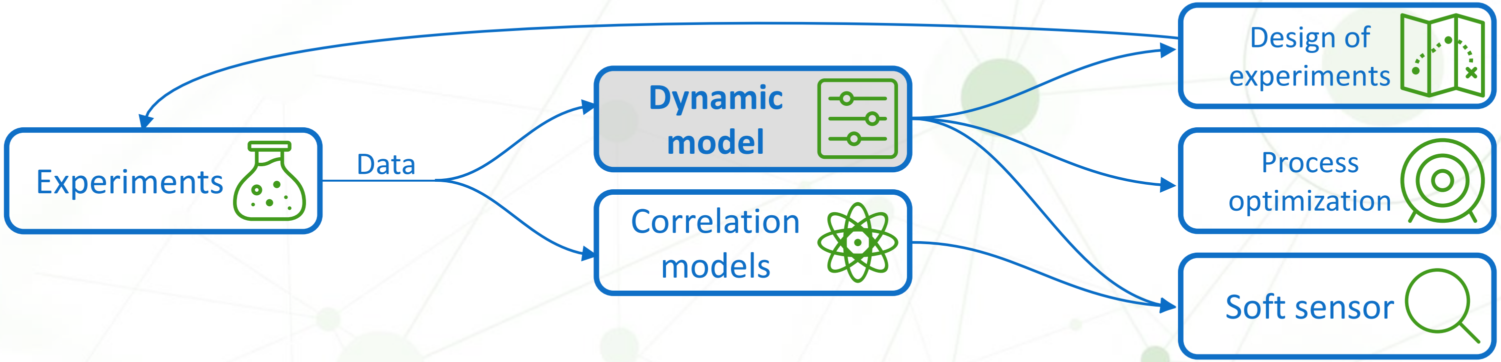
Sporulation of *Bacillus Subtilis*



- Process optimization: Find the optimal temperature trajectory to minimize the batch time
- Soft sensor: Use the online measurements for real time monitoring of X_t, X_s and S

Online measurements

- Turbidity
- CO_2 / O_2 in off gas
- pH value
- Capacitance (dielectric spectroscopy)



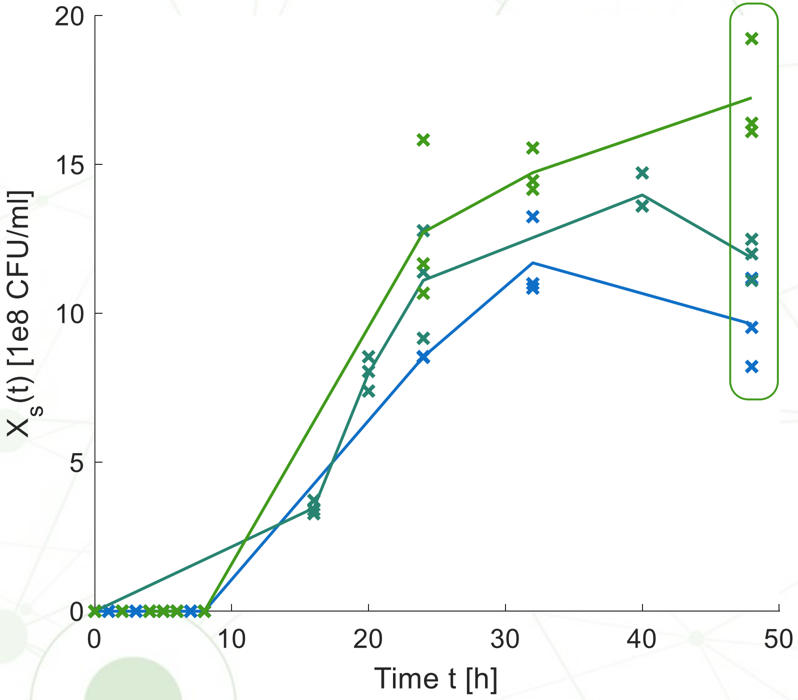
Dynamic gray-box modelling

Dynamic gray-box modelling

- Knowledge about the mass and energy balances and state variables available
- But e.g. kinetic expressions unknown
- Database from the fermentation experiments
  - ≈600 data points
  - Large measurement uncertainty
  - Significant process variability

→ **Dynamic gray-box model with embedded ML-submodels**

How to identify the model structure and parameters without a lot of trial-and-error?

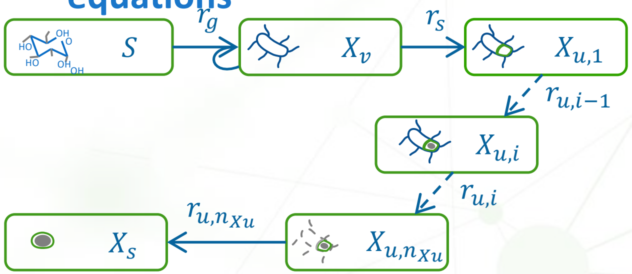


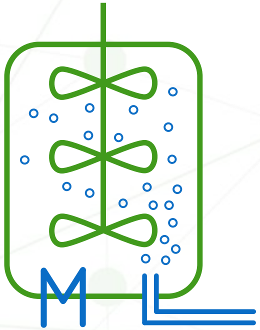
Problem decomposition

[1-3]

1

1. **Set up the first principles model equations**





2

2. **Specify the embedded variables that are described by the unknown submodels**

$$\frac{dX_v}{dt} = \boxed{\boldsymbol{\mu}} \cdot X_v - \frac{n_{Xu}}{\boxed{\boldsymbol{\tau}}} \cdot \frac{\boxed{\boldsymbol{\eta}} \cdot X_v \cdot K_s}{K_s + S}$$

$\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\eta}$: embedded variables, which are described by an ML-submodel

[1] J. Winz, S. Engell, *A methodology for gray-box modeling of nonlinear ODE systems*, in: L. Montastruc, S. Negny (Eds.), Computer Aided Chemical Engineering, Elsevier, (2022): pp. 1483–1488.
[2] J. Winz, S. Engell, *Reliable nonlinear dynamic gray-box modeling by regularized training data estimation and sensitivity analysis*, IFAC-PapersOnLine. 55 (2022) 86–93.
[3] J. Winz, S. Assawajaruwan, S. Engell, *Development of a Dynamic Gray-Box Model of a Fermentation Process for Spore Production*, Chemie Ingenieur Technik, in Press. (2023).

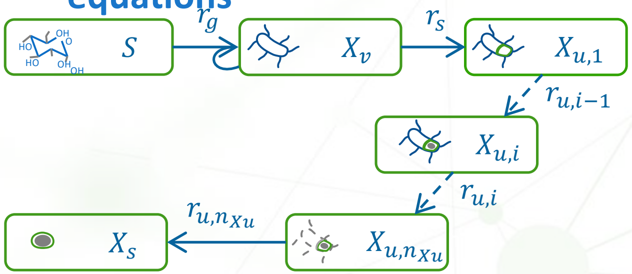
KE:3N

Problem decomposition

[1-3]

1

1. Set up the first principles model equations



2

2. Specify the embedded variables described by the unknown submodels

$$\frac{dX_v}{dt} = \boxed{\mu} \cdot X_v - \frac{n_{Xu}}{\boxed{\tau}} \cdot \frac{\boxed{\eta} \cdot X_v \cdot K_s}{K_s + S}$$

3

3. Estimate a training set for the embedded variables

Training set:

$$\begin{bmatrix} \vdots & \vdots \\ X_v(t_k) & T(t_k) \\ \vdots & \vdots \end{bmatrix} \qquad \begin{bmatrix} \vdots \\ \mu(t_k) \\ \vdots \end{bmatrix}$$

inputs outputs

What values should $\boldsymbol{\mu}$,$\boldsymbol{\tau}$,$\boldsymbol{\eta}$ assume to describe the experimental data?

4

4. Use the estimated training set for input determination and model selection

Training set $\xrightarrow{\text{ML toolbox}}$ ML Model structure + Θ_{ML}

5

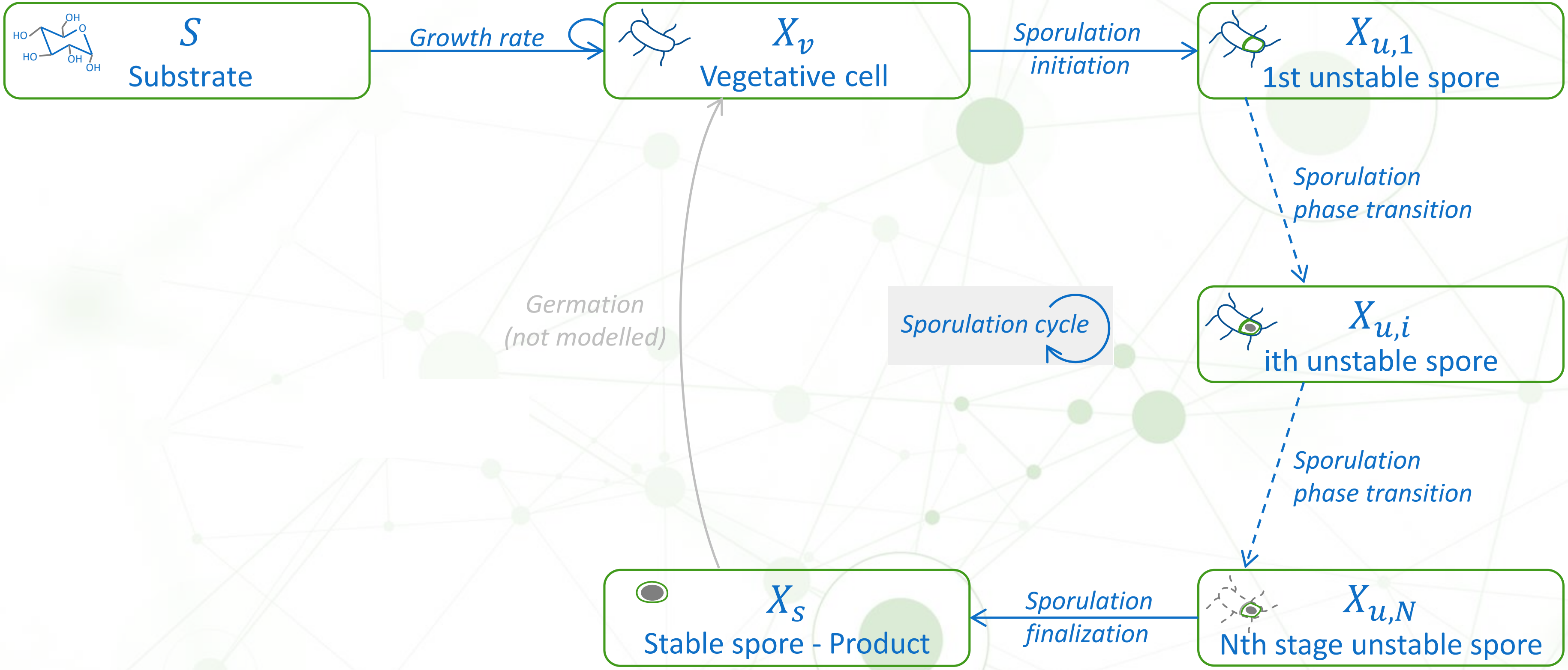
5. Full dynamic parameter estimation with the previously trained ML-model parameters as initial values

$$\frac{dX_v}{dt} = \boxed{\boldsymbol{\mu_\theta(X_v, T)}} X_v - \frac{n_{Xu}}{\boxed{\boldsymbol{\tau_\theta(T)}}} \cdot \frac{\boxed{\boldsymbol{\eta_\theta(T)}} X_v K_s}{K_s + S}$$

[1] J. Winz, S. Engell, *A methodology for gray-box modeling of nonlinear ODE systems*, in: L. Montastruc, S. Negny (Eds.), Computer Aided Chemical Engineering, Elsevier, (2022): pp. 1483–1488.
[2] J. Winz, S. Engell, *Reliable nonlinear dynamic gray-box modeling by regularized training data estimation and sensitivity analysis*, IFAC-PapersOnLine. 55 (2022) 86–93.
[3] J. Winz, S. Assawajaruwan, S. Engell, *Development of a Dynamic Gray-Box Model of a Fermentation Process for Spore Production*, Chemie Ingenieur Technik, in Press. (2023).

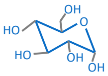
KEEN

Model structure



[1] J. Winz, S. Assawajaruwan, S. Engell, *Development of a Dynamic Gray-Box Model of a Fermentation Process for Spore Production*, Chemie Ingenieur Technik, in Press. (2023).

Model structure



S
Substrate

Growth rate

X_v
Vegetative cell

Sporulation initiation

$X_{u,1}$
1st unstable spore

Sporulation phase transition

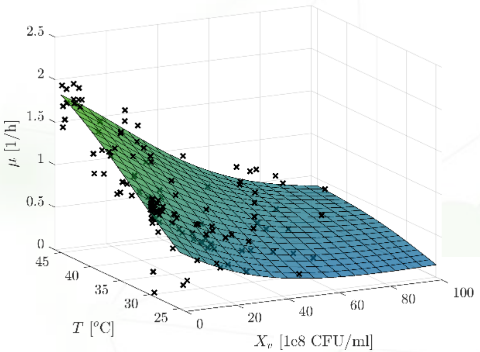
$X_{u,i}$
ith unstable spore

Sporulation phase transition

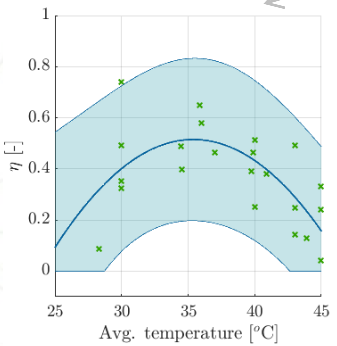
$X_{u,N}$
Nth stage unstable spore

Sporulation finalization

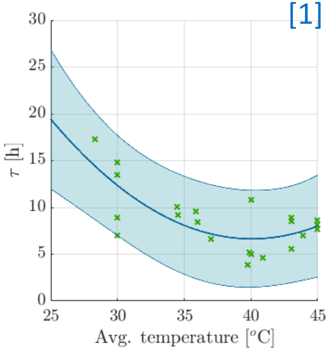
X_s
Stable spore - Product



Specific growth rate



Sporulation efficiency

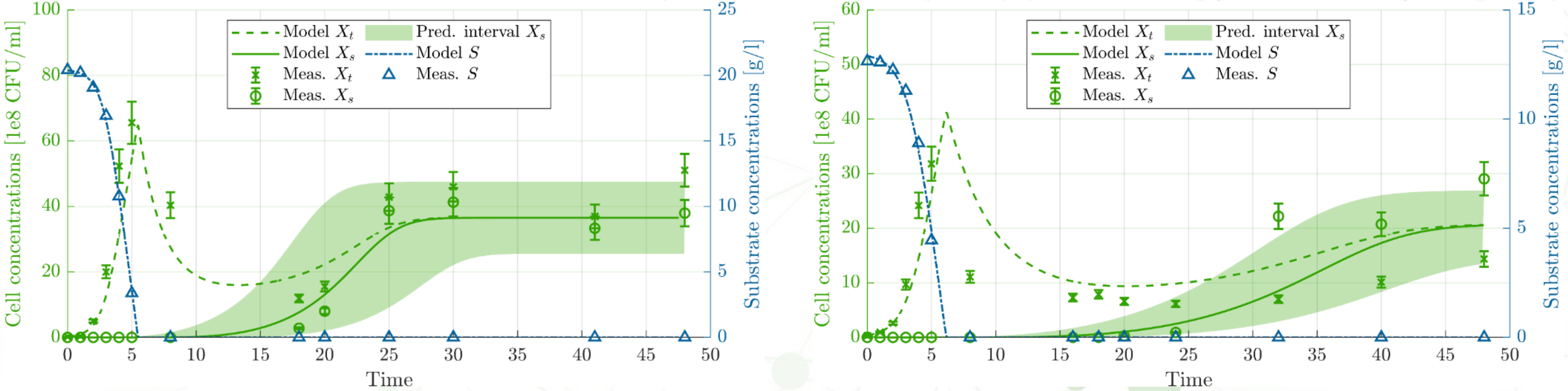


Sporulation delay time

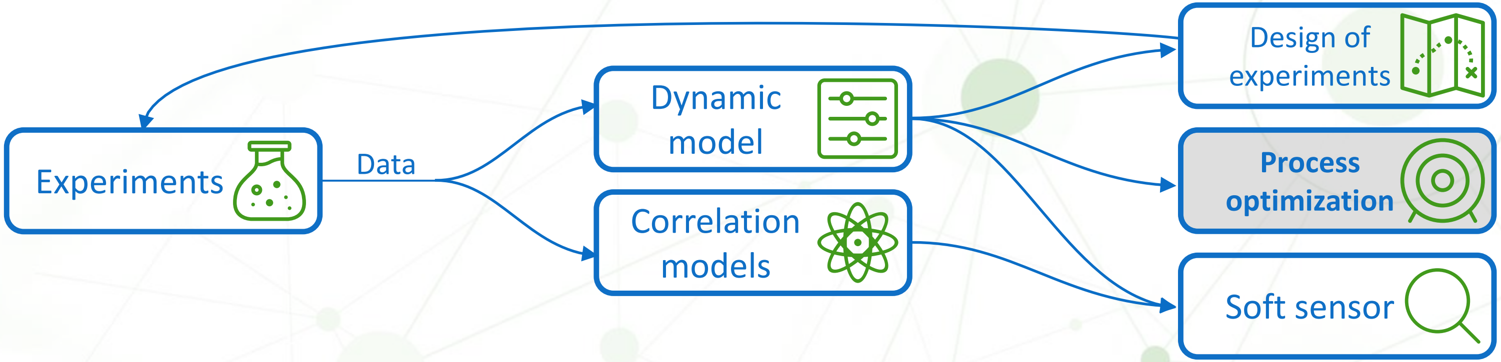
[1] J. Winz, S. Assawajaruwan, S. Engell, Development of a Dynamic Gray-Box Model of a Fermentation Process for Spore Production, Chemie Ingenieur Technik, in Press. (2023).

KEEN

Dynamic gray-box model predictions



- Accurate predictions for all state variables
- Enables model-based applications
  - Model-based optimal design of experiments
  - Process optimization
  - Soft sensor

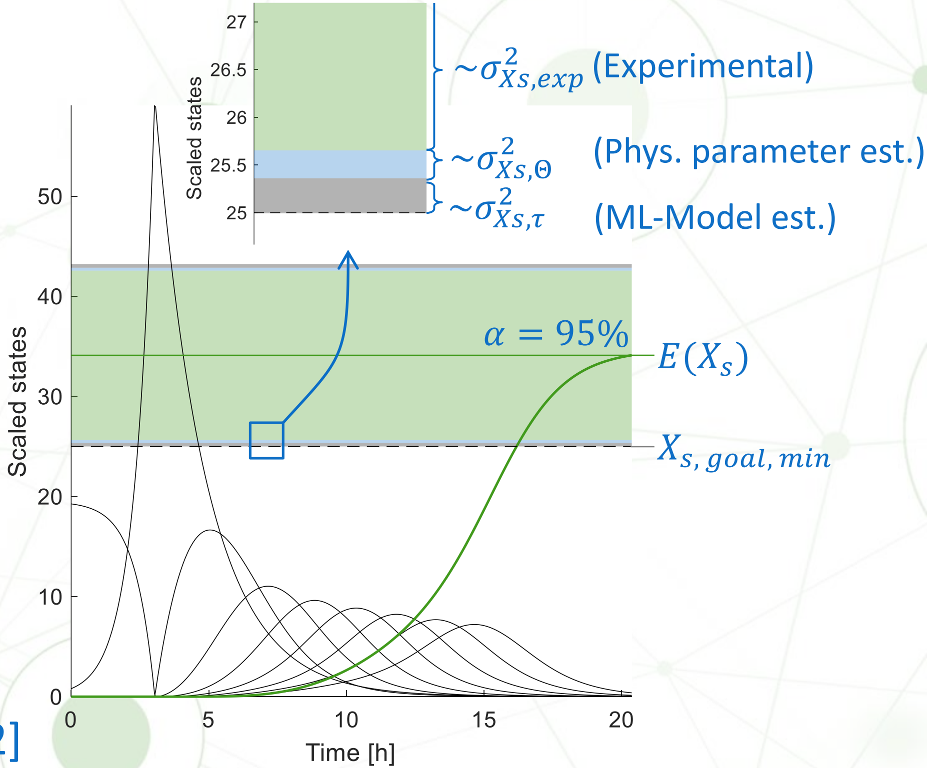


Process optimization methodology

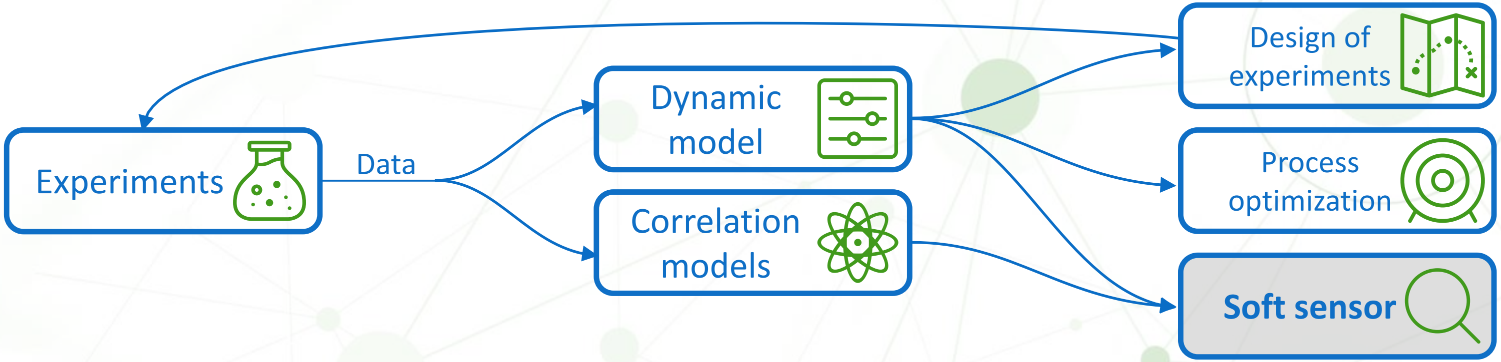
Chance constrained optimization

- Optimization goal: Minimize the batch time, while **statistically guaranteeing** a minimum harvest of spores
  - Chance constrained optimization [1]
- Challenge: Multiple different sources of uncertainty

$$\sigma^2_{Xs} = \sigma^2_{Xs,exp} + \sigma^2_{Xs,\Theta} + \underbrace{\sigma^2_{Xs,ML}}_{\text{Uncertainty quantified using the jackknife variance [2]}}$$



[1] Li, Pu, Harvey Arellano-Garcia, and Günter Wozny. "Chance constrained programming approach to process optimization under uncertainty." Computers & chemical engineering 32.1-2 (2008): 25-45.
[2] R.G. Miller, The jackknife - a review, Biometrika. 61 (1974) 1–15.

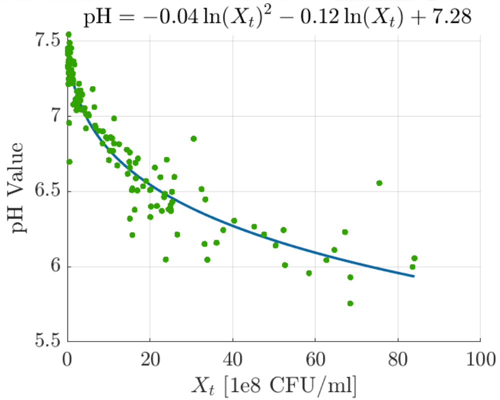


Soft sensor development

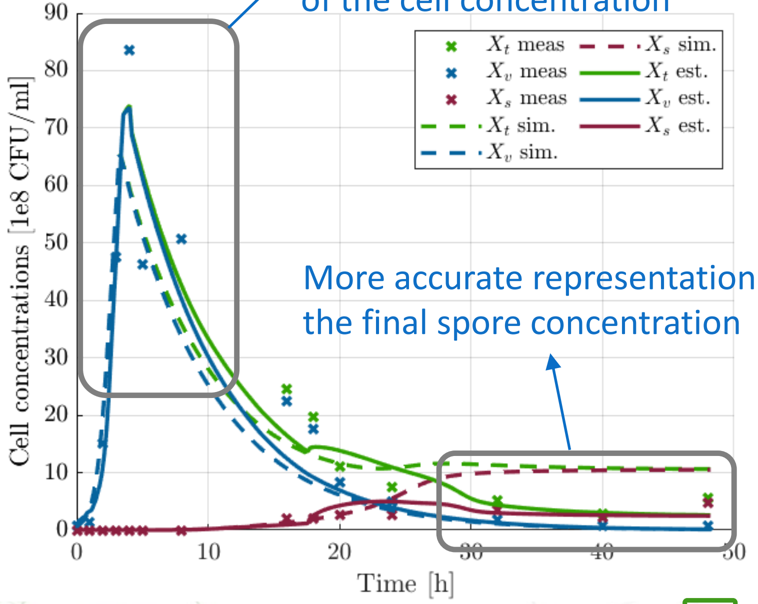
Phase adaptive state estimator

- Different phases during the batch: growth, sporulation & stable phase
- A simple correlation model is developed for each phase

→ Unscented Kalman Filter



Correction for higher values of the cell concentration



More accurate representation of the final spore concentration

→ **Implemented at the real plant**

Summary and further work

Summary

- Development of a dynamic gray-box model for a complex process
- Utilization of chance constrained optimization
- Implementation of the phase adaptive state estimator on the real plant

Outlook

- Description of the model structure uncertainty
- Use of multivariate correlations
- Benchmarking of EKF vs. UKF vs. Particle Filter