Hochschule Ostwestfalen-Lippe
University of Applied Sciences

BIO PROCESS AND SYSTEMS TUHH ENGINEERING

Produktion von Biopharmazeutika – Ein MATLAB©-Tool zur Simulation und Optimierung von Zellvermehrungsverfahren

MATLAB EXPO 2017, München

Tanja Hernández Rodríguez

Kirsten S. Koopmann, Jonas Klaes, Simon Kern, Oscar Platas Barradas, Ralf Pörtner , Björn Frahm

Überblick

- Einführung in den biopharmazeutischen Anwendungsbereich
- Entwickeltes Software-Tool zur Optimierung: Struktur, Modellierung & Mathematik, wesentliche Elemente, Realisierung in MATLAB
- Anwendungen und Einsatzfelder des Tools
- Ein Blick auf experimentelle Anwendungen/Verifizierungen
- Fazit & Ausblick

Biopharmazeutika & die Herausforderungen bei ihrer Herstellung



Wozu dienen Biopharmazeutika?


Krebs
Diabetes
Multiple
Sklerose
Thrombose
Blutererkrankung

Relevanz / Bedarf



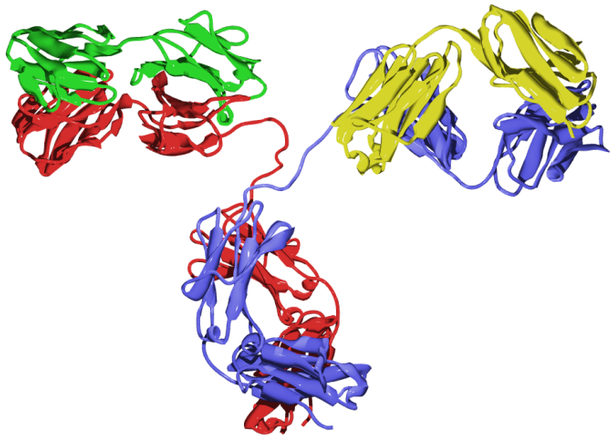
Mehr als 300 biopharmazeutische Produkte weltweit

Jahresumsatz: ca. 200 Milliarden US$ weltweit

→ Weltbevölkerung wächst
→ Zahl der älteren Menschen steigt

Quelle Bild: Pixabay

Was sind Biopharmazeutika?



Biologische Wirkstoffe, i.d.R. komplexe Proteine

- Bis zu 1000 x größer als chemisch hergestellte Wirkstoffe
- Besitzen eine sehr komplexe Struktur
- Sie benötigen lebende Zellen (meist Säugetierzellen), in denen sie produziert werden. In Bakterien, Pilzen und Hefen ist das nicht möglich (anders als bei Antibiotika).

Komplexe Proteine, wie:

- Monoklonale Antikörper (z.B. für die Krebsbehandlung)
- tPA gegen Thrombose
- Gerinnungsfaktoren (z.B. zur Behandlung der Bluterkrankheit)

Quelle Bild: commons.wikimedia

Biopharmazeutische Produktion



Produktions-Bioreaktoren, Quelle: COTTER

Viele Schritte unter stetiger Kontrolle notwendig

Sicherstellung, dass die 3-dimensionale Proteinstruktur präzise gebildet wird.

Hohe Anforderungen an die Produktqualität

(schon kleine Abweichungen in der komplexen Proteinstruktur können bei den Patienten zu Immunreaktionen führen)

Viele Faktoren können die Produktqualität beeinflussen/beeinträchtigen (Kultivierungsparameter, Zellzustand,…)

Von der mL-Ampulle bis in den Produktionsreaktor (2.000 - 20.000 L)



Produktions-Bioreaktoren



Bioreaktor

Schüttelkolben



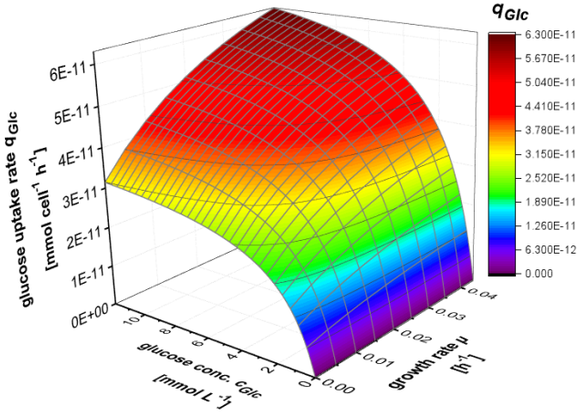
Quellen: Kern et al. (links), Sartorius Stedim (Mitte), Cotter (rechts)

Ziel: Hohe Produktausbeute bei hoher Produktqualität

Zellvermehrungsverfahren (Seed-Trains) zeit- und kostenintensiv

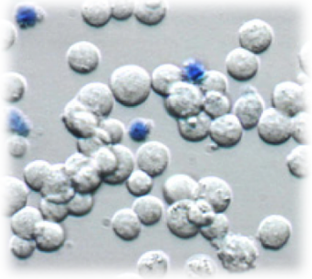
Lebende Zellen verhalten sich individuell

Komplexer Zellmetabolismus



Optimierungsmethoden erforderlich

→ IT- und modellgestützte Methoden



Bsp. für Zellen, Quelle: ProBioGen

Zellwachstum und Proteinproduktion sind von Kultivierungsbedingungen und Zellperformance abhängig

Zellvermehrungsverfahren hat Einfluss auf die Erfolgsquote der Wirkstoffproduktion

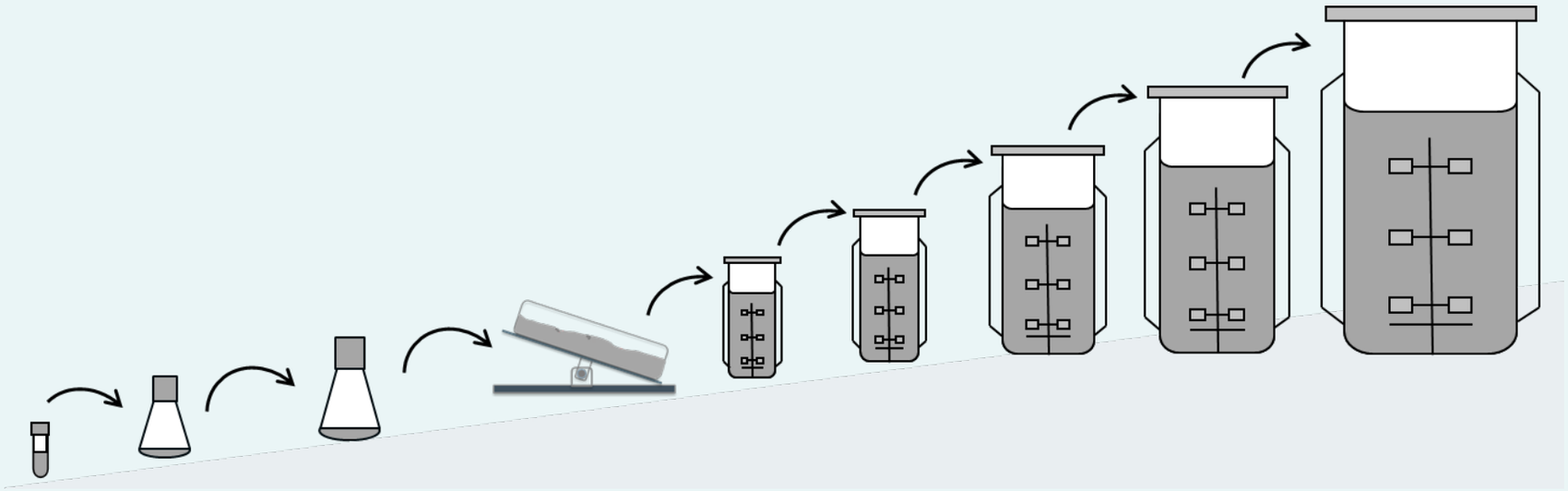
Entwickeltes Software-Tool zur Optimierung: Struktur, Modellierung & Mathematik, wesentliche Elemente, Realisierung in MATLAB

Struktur

Anpassung an Zelllinie

Kultivierungs-daten Zelllinie → **Modell**

Seed-Train



Mathematisches Modell - Beispiel

Bilanzgleichungen

Bio-Phase

$$\frac{dX_v}{dt} = \left(\mu - \mu_d - \frac{F_{Glc} + F_{Gln} + F_{Medium}}{V}\right) \cdot X_v$$

$$\frac{dX_t}{dt} = \mu \cdot X_v - \frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot X_t - K_{Lys} \cdot (X_t - X_v)$$

Flüssig-Phase

$$\frac{dc_{Glc}}{dt} = \frac{F_{Glc}}{V} \cdot c_{Glc,F} + \frac{F_{Medium}}{V} \cdot c_{Glc,Medium} - \frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{Glc} - q_{Glc} \cdot X_v$$

$$\frac{dc_{Gln}}{dt} = \frac{F_{Gln}}{V} \cdot c_{Gln,F} + \frac{F_{Medium}}{V} \cdot c_{Gln,Medium} - \frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{Gln} - q_{Gln} \cdot X_v$$

$$\frac{dc_{LS}}{dt} = \frac{F_{Medium}}{V} \cdot c_{LS,Medium} - \frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{LS} - q_{LS} \cdot X_v$$

$$\frac{dc_{Lac}}{dt} = -\frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{Lac} + q_{Lac} \cdot X_v$$

$$\frac{dc_{Amm}}{dt} = -\frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{Amm} + q_{Amm} \cdot X_v$$

$$\frac{dc_{MAb}}{dt} = -\frac{F_{Glc} + F_{Gln} + F_{Medium}}{V} \cdot c_{MAb} + q_{MAb} \cdot X_v$$

$$\frac{dV}{dt} = F_{Glc} + F_{Gln} + F_{Medium}$$

Kinetiken

Zell-Wachstum/Absterben

$$\mu = \mu_{\mathbf{max}} \cdot \frac{c_{Glc}}{c_{Glc} + \mathbf{K_{S,Glc}}} \cdot \frac{c_{Gln}}{c_{Gln} + \mathbf{K_{S,Gln}}} \cdot \frac{c_{LS}}{c_{LS} + \mathbf{K_{S,LS}}}$$

$$\mu_d = \mu_{\mathbf{d,min}}$$

Substrataufnahme / Metabolitenproduktion

$$q_{Glc} = \mathbf{q_{Glc,max}} \cdot \frac{c_{Glc}}{c_{Glc} + \mathbf{k_{Glc}}} \cdot \left(\frac{\mu}{\mu + \mu_{\mathbf{max}}} + 0.5\right)$$

$$q_{Gln} = K_{p,Gln} \cdot \mathbf{q_{Gln,max}} \cdot \frac{c_{Gln}}{c_{Gln} + \mathbf{k_{Gln}}}$$

$$K_{p,Gln} = 1, \quad \text{if} \quad c_{Glc} > \mathbf{c_{cGlc,thr}}$$

$$K_{p,Gln} = 1 + \mathbf{a} \cdot (\mathbf{c_{Glc,thr}} - c_{Glc}), \quad \text{if} \quad c_{Glc} \leq \mathbf{c_{Glc,thr}}$$

$$q_{LS} = \mathbf{q_{LS,max}} \cdot \frac{c_{LS}}{c_{LS} + \mathbf{k_{LS}}}$$

$$q_{Lac} = \mathbf{Y_{Lac/Glc}} \cdot q_{Glc} - \mathbf{q_{Lac,uptake}} \frac{\mu_{\mathbf{max}} - \mu}{\mu_{\mathbf{max}}}$$

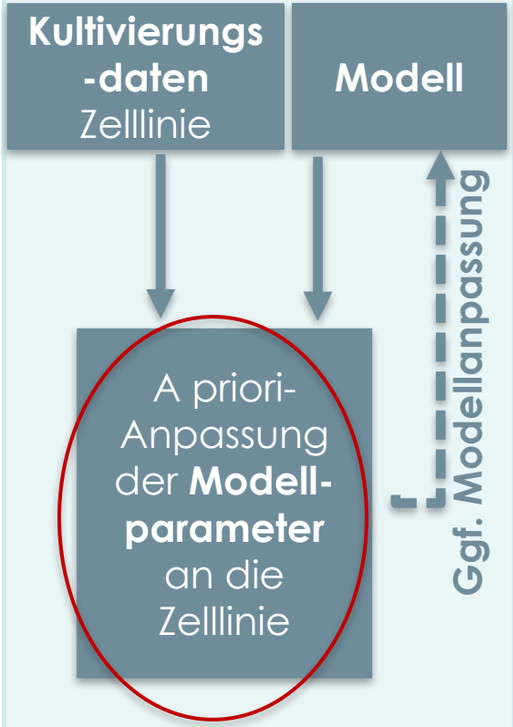
$$q_{Amm} = \mathbf{Y_{Amm/Gln}} \cdot q_{Gln}$$

$$q_{MAb} = \mathbf{q_{MAb,max}}$$

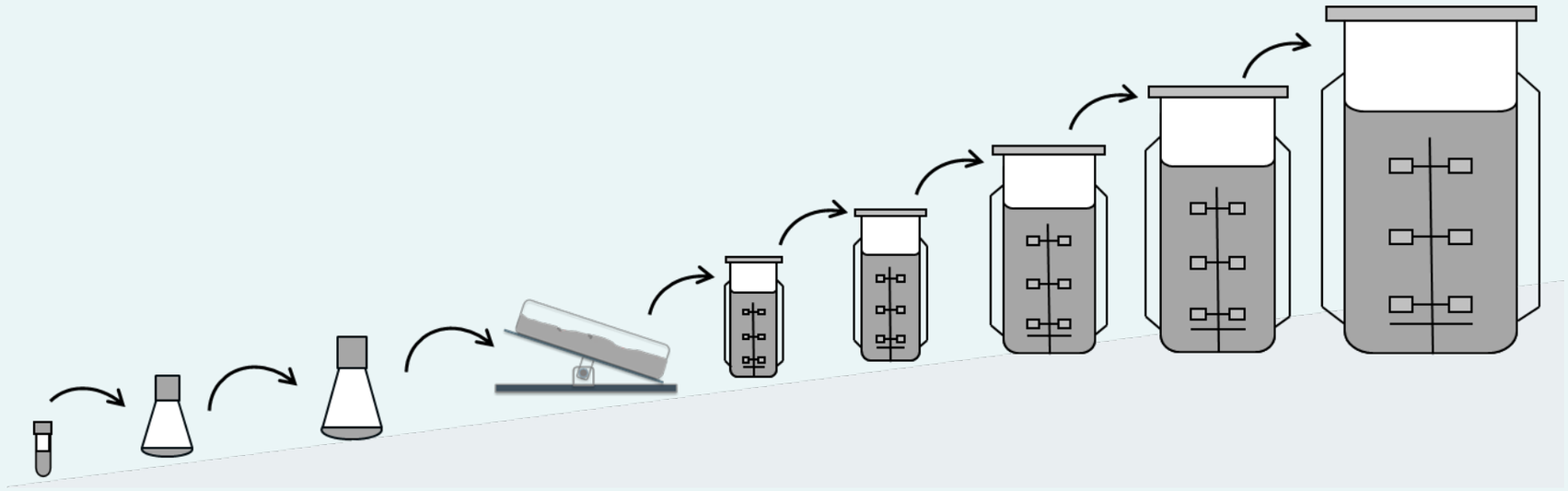
Modellparameter rot hervorgehoben

Struktur

Anpassung an Zelllinie



Seed-Train



Parameterschätzung – Mathematische Formulierung

Parametervektor p bestimmen:
- Zu schätzende Modellparameter
- Konzentrationsstartwerte (Messungenauigkeiten)

Anfangswerte für die in p enthaltenen Parameter

Randbedingungen, d.h. Ober- und Untergrenzen für die in p enthaltenen Parameter

Kultivierungsdaten

Nichtlineares Optimierungsproblem mit Nebenbedingungen

Gütefunktional
(= Summe der gewichteten Fehlerquadrate)

$$\min_{p \in \mathbf{R}^n} \left(J(p) = \sum_{k=1}^{M} \sum_{i=1}^{N_k} w_k \left(y_{k,mess}(t_i) - y_{k,sim}(t_i, p) \right)^2 \right)$$

unter der Nebenbedingung: $lb \leq p \leq ub$

Minimierung der Fehlerquadrate mittels Nelder-Mead-Simplex-Algorithmus

$$w_k = \frac{WF_k}{m_k^2}$$

$$m_k = \max_i y_{k,mess}(t_i)$$



Geschätzte Modellparameter und Startwerte

$p = (p_1, \dots, p_n)$: Parametervektor
n: Anzahl der zu schätzenden Modellparameter
M: Anzahl der Messgrößen
N_k: Anzahl der Messungen
w_k: Gewichtungen
WF_k: Gewichtungsfaktor für die k-te Messgröße
$y_{k,mess}$: k-ter Messwert
$y_{k,simu}$: k-te mittels ODE-Löser berechnete Näherungslösung, basierend auf dem aktuellen Parametersatz
$lb = (lb_1, \dots, lb_n)$: untere Parametergrenzen (*lower bounds*)
$ub = (ub_1, \dots, ub_n)$: obere Parametergrenzen (*upper bounds*)

Lösen des Optimierungsproblems

Nichtlineares Optimierungsproblem mit Nebenbedingungen



Hier verwendet wurde der **File Exchange Select** file: `fminsearchbnd`

Dieser Algorithmus transformiert das beschränkte Optimierungsproblem in ein freies Optimierungsproblem und wendet anschließend den Löser `fminsearch` aus der **Optimization Toolbox** an.

`fminsearch` verwendet den ableitungsfreien **Nelder-Mead-Simplex-Algorithmus**.

Falls einige Modellparameter bereits festgesetzt werden können, fließen nur noch die übrigen unbekannten Modellparameter in die Parameteridentifikation ein.

Struktur

Anpassung an Zelllinie

Seed-Train

Kultivierungs-daten Zelllinie

Modell

Konzentrations-Startwerte (Beginn des Seed-Trains)

A priori-Anpassung der **Modell-parameter** an die Zelllinie

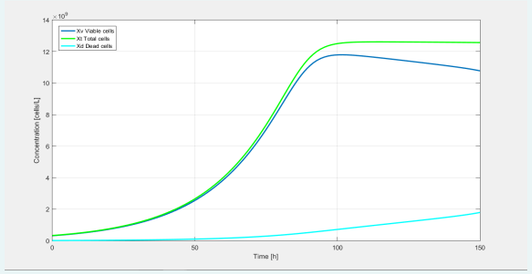
Ggf. Modellanpassung

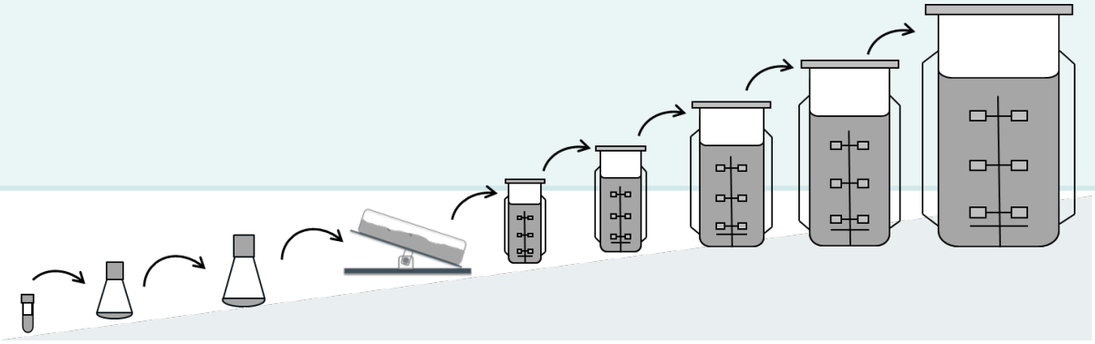
Vorhersage der zeitlichen Verläufe pro Scale

Modell-parameter

Lösen der Differentialgleichungen mittels ODE-Löser: `ode23`

Vorhersage der Konzentrationsverläufe





Struktur

Anpassung an Zelllinie

Seed-Train

Kultivierungs-daten Zelllinie

Modell

Konzentrations-Startwerte (Beginn des Seed-Trains)

Passagier-Strategie

Seed-Train-**Behältnisse & Einstellungen**

Medium-konzentra-tionen

A priori-Anpassung der **Modell-parameter** an die Zelllinie

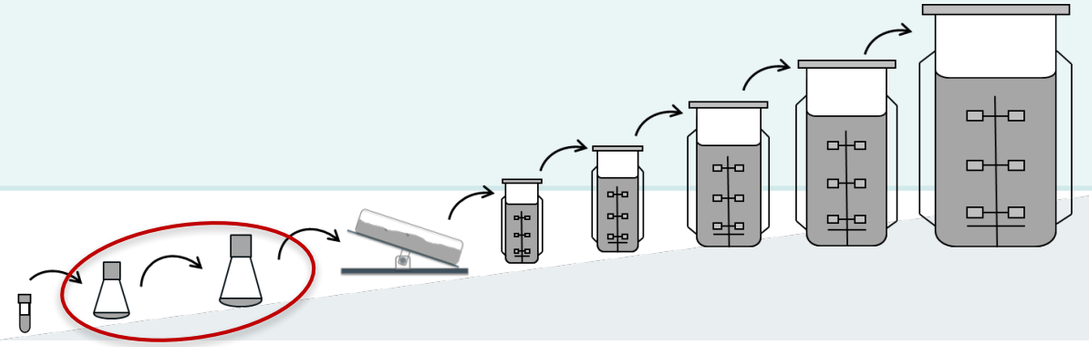
Ggf. Modellanpassung

Vorhersage der zeitlichen Verläufe pro Scale

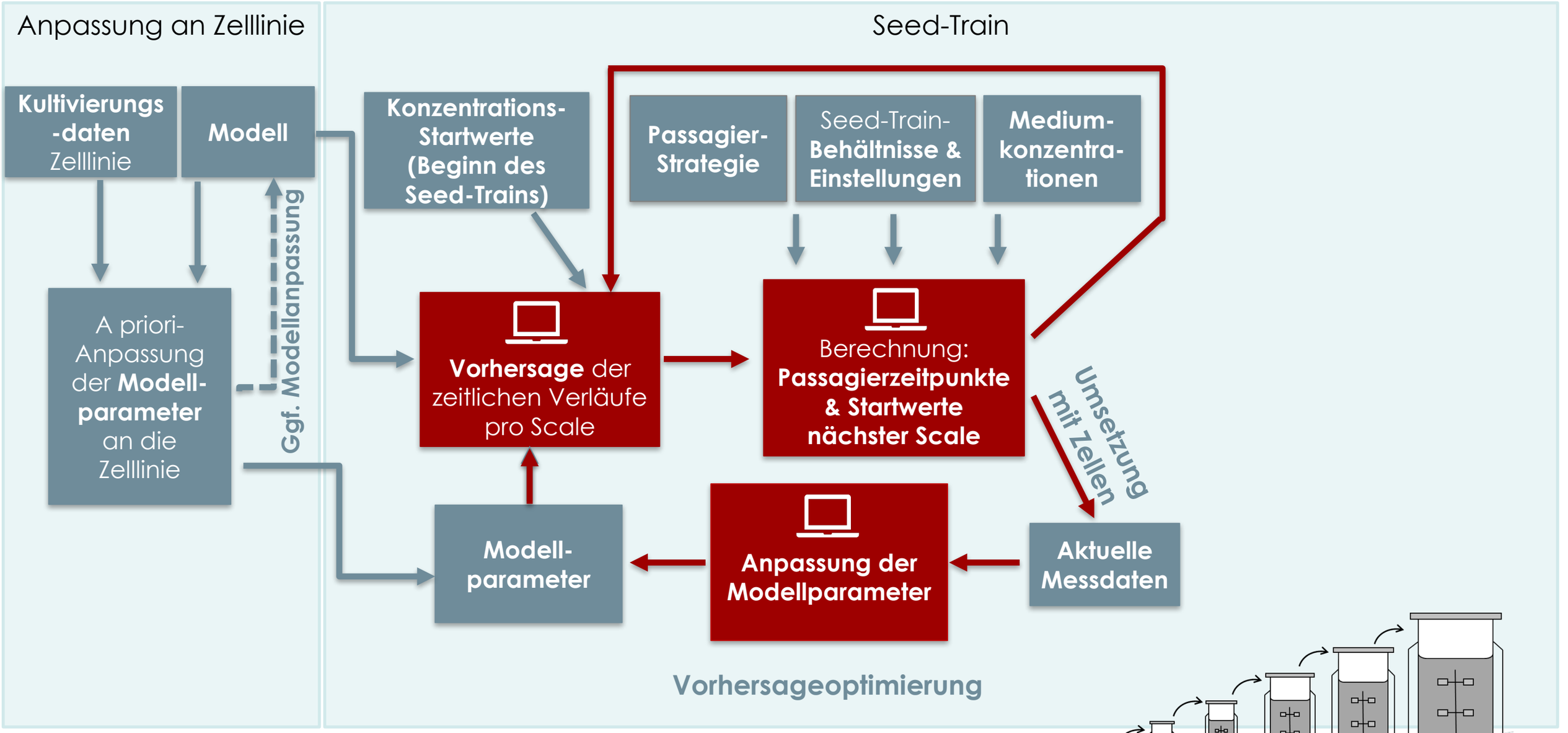
Berechnung: **Passagierzeitpunkte & Startwerte nächster Scale**

Modell-parameter

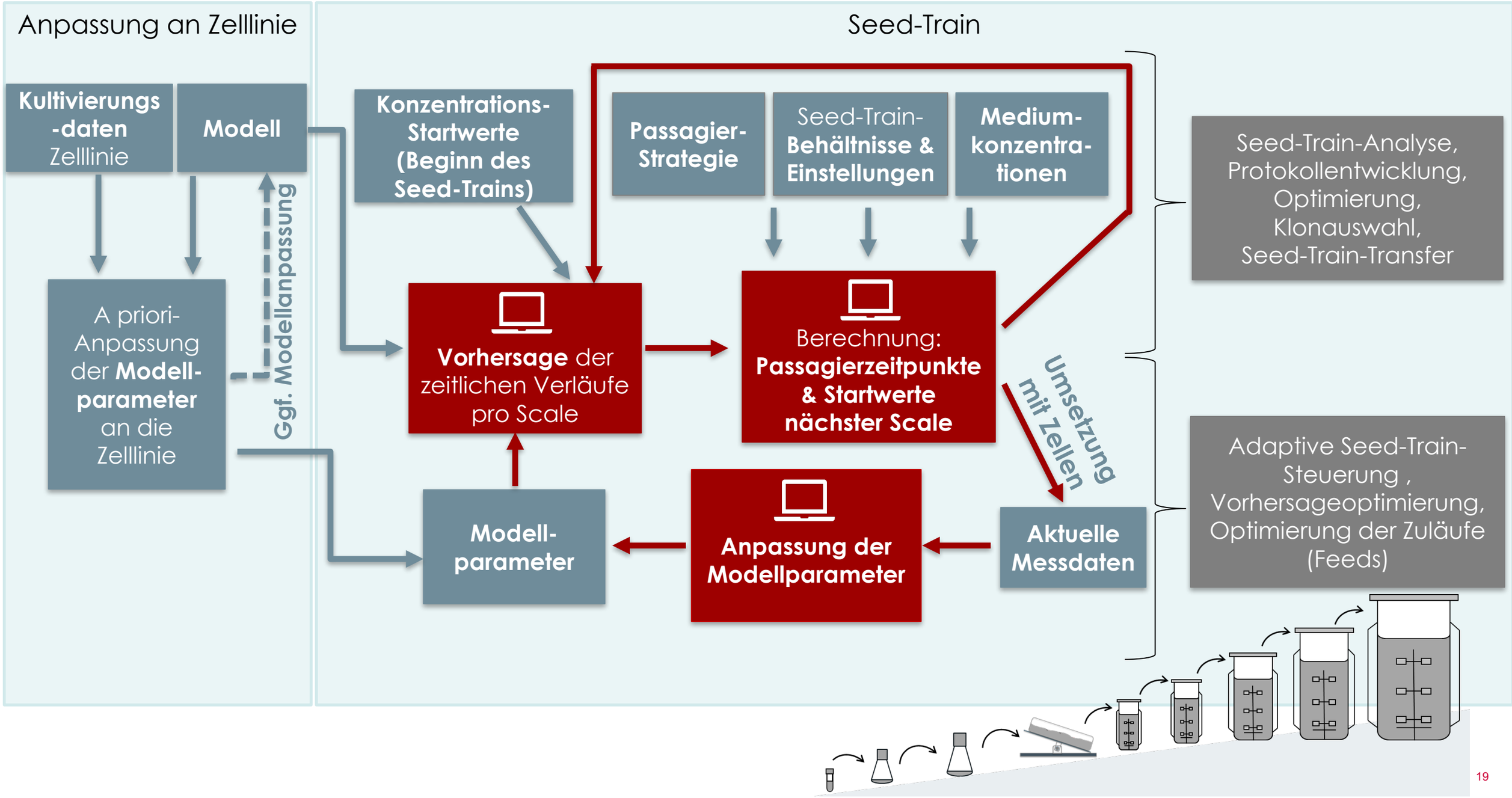
Übergang / Transfer in den nächsten Scale



Struktur


Anpassung an Zelllinie
Seed-Train
Kultivierungs-daten Zelllinie
Modell
Konzentrations-Startwerte (Beginn des Seed-Trains)
Passagier-Strategie
Seed-Train-Behältnisse & Einstellungen
Medium-konzentra-tionen
A priori-Anpassung der Modell-parameter an die Zelllinie
Ggf. Modellanpassung
Vorhersage der zeitlichen Verläufe pro Scale
Berechnung: Passagierzeitpunkte & Startwerte nächster Scale
Umsetzung mit Zellen
Modell-parameter
Anpassung der Modellparameter
Aktuelle Messdaten
Vorhersageoptimierung

Struktur


Anpassung an Zelllinie
Seed-Train
Kultivierungs-daten Zelllinie
Modell
Konzentrations-Startwerte (Beginn des Seed-Trains)
Passagier-Strategie
Seed-Train-Behältnisse & Einstellungen
Medium-konzentra-tionen
Seed-Train-Analyse, Protokollentwicklung, Optimierung, Klonauswahl, Seed-Train-Transfer
A priori-Anpassung der Modell-parameter an die Zelllinie
Ggf. Modellanpassung
Vorhersage der zeitlichen Verläufe pro Scale
Berechnung: Passagierzeitpunkte & Startwerte nächster Scale
Umsetzung mit Zellen
Adaptive Seed-Train-Steuerung , Vorhersageoptimierung, Optimierung der Zuläufe (Feeds)
Modell-parameter
Anpassung der Modellparameter
Aktuelle Messdaten



Einsatzfelder und Anwendungen des Tools

Entwicklung & Produktion (Biopharmazeutische Industrie)

Einsatzfelder

Forschung (Hochschulen & Universitäten)

Optimierung eines bestehenden Seed-Trains hinsichtlich Zellausbeute (→ Zeit- und Kostenoptimierung)

Anwendungs-Möglichkeiten

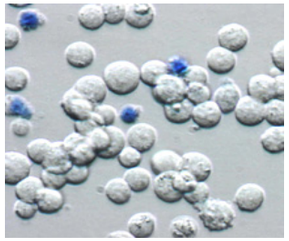
Planung eines Seed-Trains für ein anderes Produkt oder eine andere Produktionsstätte



Online-Entscheidungshilfe bzgl. der nächsten Prozessschritte mittels Vorhersage,
z.B. Zeitpunkt für den Transfer von einem Scale in den nächsten

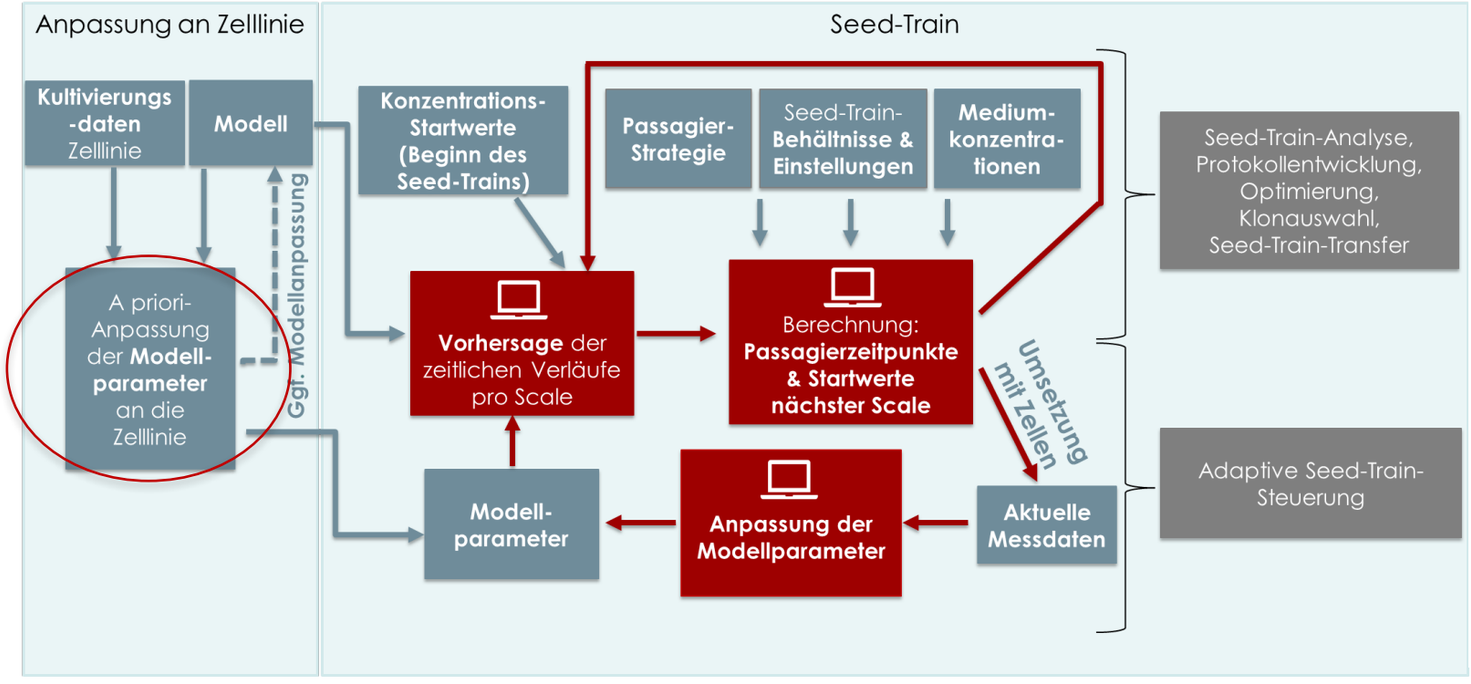
Ein Blick auf Anwendungen / Verifizierungen

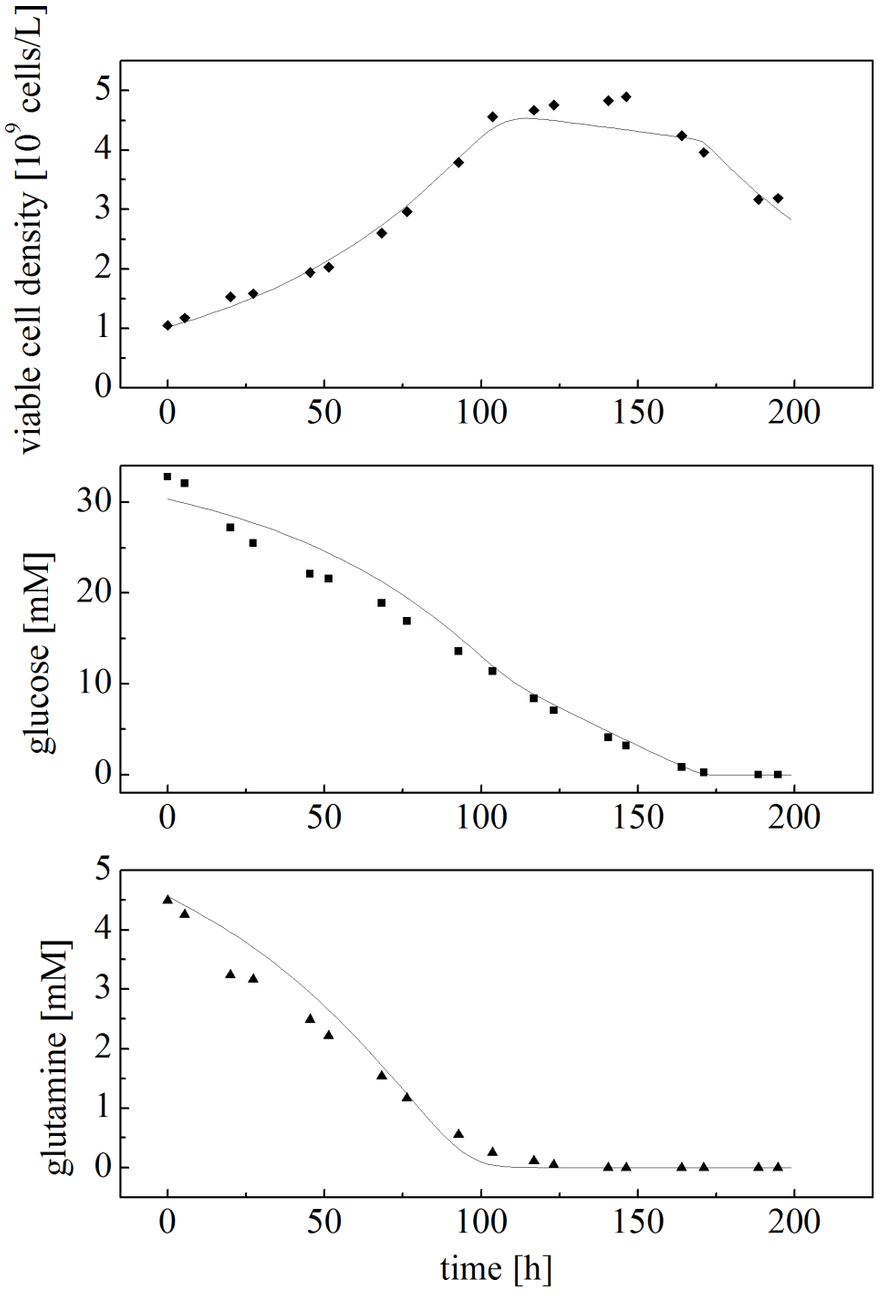
Parameterschätzung der Modellparameter einer **AGE1.HN$_{AAT}$ –Zelllinie** basierend auf zwei Datensätzen



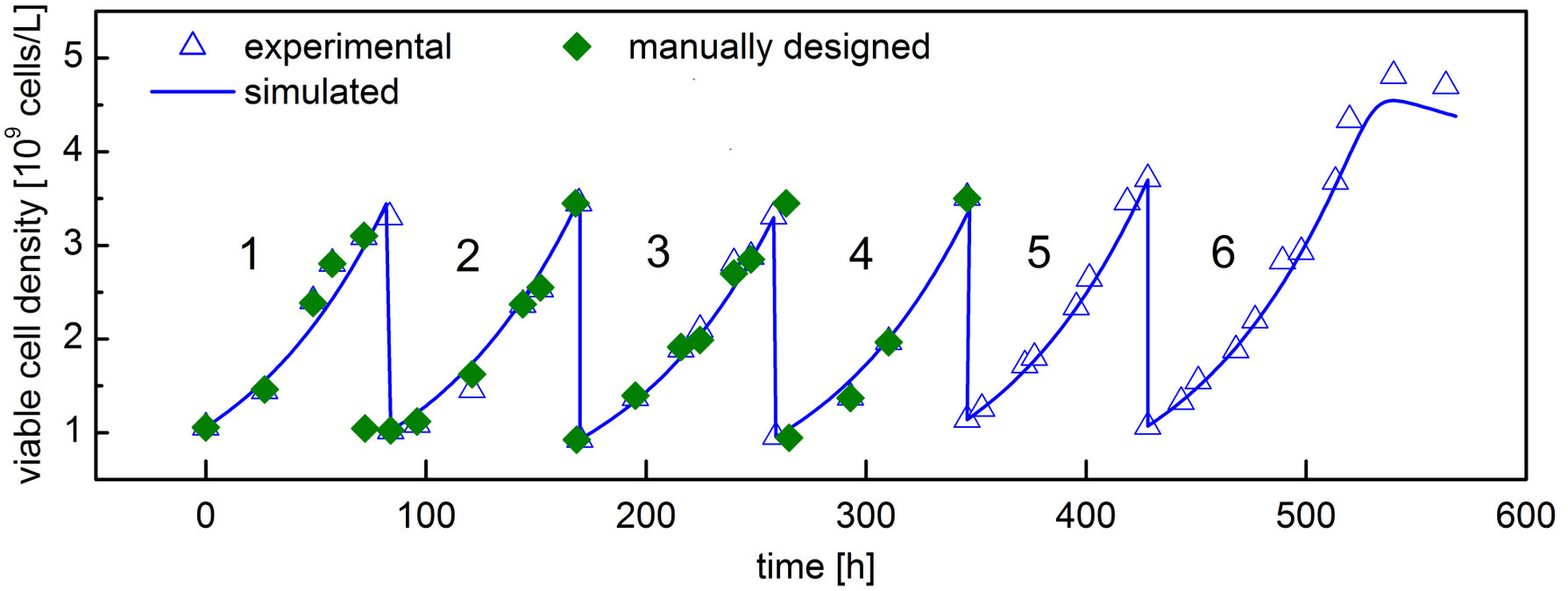
Quelle: ProBioGen, SysLogics

- Menschliche Zelllinie (ProBioGen AG, Berlin, Germany)
- Produziert das therapeutische Protein α-1-Antitrypsin





Verifizierung Vorhersage - Zelllinie AGE1.HN$_{AAT}$ und Vergleich zur einer manuellen Optimierung



Seed-Train über 6 Scales:
1. culture tube (0.01 L)
2. shake flask (0.035 L)
3. shake flask (0.13 L),
4. Vario 1000 (0.35 L)
5. VSF 2000 (1 L)
6. Labfors 5 Cell (2.5 L)

Simulierter Verlauf

Experimentelle Daten der Durchführung

Experimentelle Daten eines manuell entworfenen Verlaufs

Gute Seed-Train-Auslegung

Weitere Anwendungen

Seed-Train-Vorhersage CHO-K1-Zelllinie

Modellierung industrieller Seed-Trains bis in den großskaligen Produktionsmaßstab

Sensitivitätsanalyse

Fazit

Erfolgreiche Erstellung eines umfangreichen Optimierungs-Tools

Ausblick

- Implementierung zusätzlicher Funktionen
- (z.B. Einbinden von Unsicherheit → Vorhersagebereiche, Evaluation der Modellgüte,…)
- Weitere Untersuchungen
- Weiterer industrieller Einsatz

Hochschule Ostwestfalen-Lippe
University of Applied Sciences

Vielen Dank für Ihre Aufmerksamkeit!

M. Sc. Tanja Hernández Rodríguez, Prof. Dr.-Ing. Björn Frahm
FB4 – Life Science Technologies

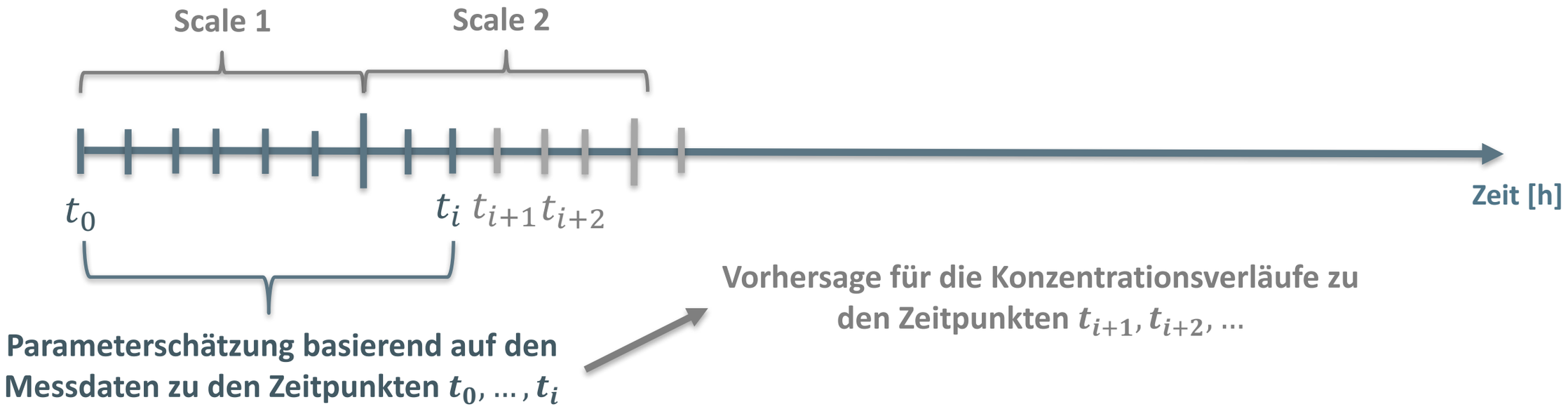
Telefon +49 5261 – 702 5399 (Tanja Hernández Rodríguez), +49 5261 – 702 5662 (Björn Frahm)
Telefax +49 5261 – 702 85662
tanja.hernandez@hs-owl.de, bjoern.frahm@hs-owl.de
www.hs-owl.de/fb4

Back up: Sensitivitätsanalyse

- Modellparameter einer CHO-K1-Zelllinie (zur Verfügung gestellt von Prof. Noll, Universität Bielefeld)
- Variation der Startwerte (I), der wichtigen Mediumkonzentrationen (II) und der zellspezifischen Modellparameter (III) um +/-5 % .
- Der Einfluss auf den Passagierzeitpunkt wurde mit Hilfe eines vollfaktoriellen Versuchsplans (*one factor at a time (OFAT)*) untersucht.
- Passagierzeitpunkt: Mittelung aus Zeitpunkt der optimalen Raum-Zeit-Ausbeute (RZA) und Zeitpunkt, bei dem die scheinbare Wachstumsrate auf 90 % der maximalen scheinbaren Wachstumsrate abgefallen ist.

	Deviation 0 %	Deviation < 1 %	Deviation 1 - 5 %	Deviation 5 - 7 %
(I) Seed train starting conditions	c_{Glc} (95 % / 105 %)	c_{Gln} (95 % / 105 %)	X_v, X_t (95 % / 105 %)	
	c_{Lac} (95 % / 105 %)			
(II) Important medium concentrations	$c_{Lac,medium}$ (95 % / 105 %)	$c_{Gln,medium}$ (95 % / 105 %)		
	$c_{Glc,medium}$ (95 % / 105 %)			
(III) Cell line parameters for model	k_{Glc} (95 % / 105 %)	$K_{S,Gln}$ (95 % / 105 %)	$q_{Gln,max}$ (95 % / 105 %)	μ_{max} (95 % / 105 %)
	$\mu_{d,max}$ (95 % / 105 %)	$\mu_{d,min}$ (95 % / 105 %)	k_{Gln} (95 % / 105 %)	
	K_{Lys} (95 % / 105 %)			
	$K_{S,Glc}$ (95 % / 105 %)			
	$q_{Lac,uptake,max}$ (95 % / 105 %)			
	$Y_{Lac/Glc}$ (95 % / 105 %)			
	$q_{Glc,max}$ (95 % / 105 %)			
	$Y_{Amm/Gln}$ (95 % / 105 %)			

Back up: Wesentliche Elemente – Optimierte Vorhersage



Sequentielle Vorhersagen: Nach jeder neuen Messung erneute Parameterschätzung
→ Verbesserung der Vorhersage
→ Bewegliches Zeitfenster „Moving Horizon“
→ Entscheidungshilfe bzgl. der nächsten Prozessschritte