On the identifiability of kernels for Population Balance Modelling

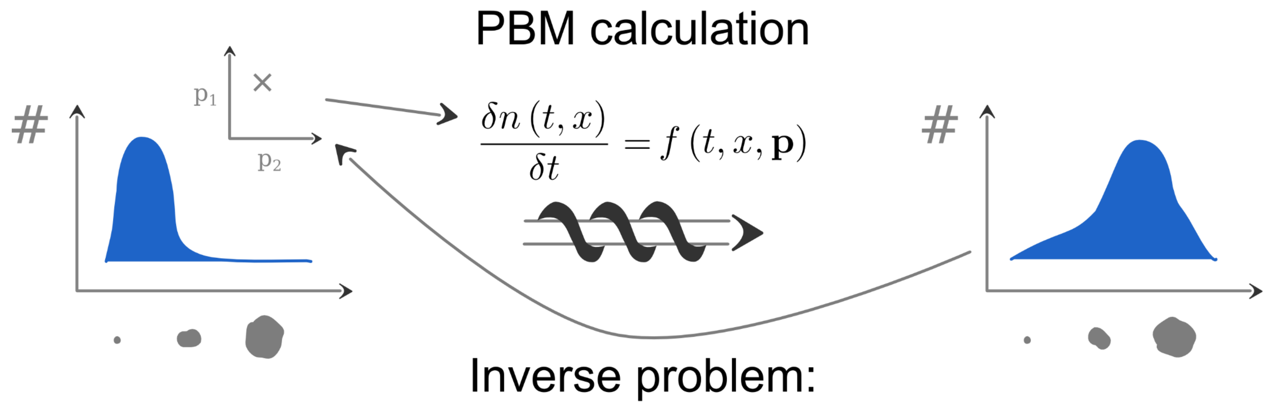
Daan Van Hauwermeiren [1,2], Thomas De Beer[2], and Ingmar Nopens[1]

[1]BIOMATH, Department of Data Analysis and Mathematical Modelling, Faculty of Bioscience Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium
[2]Laboratory of Pharmaceutical Process Analytical Technology, Department of Pharmaceutical Analysis, Ghent University, Ottergemsesteenweg 460, 9000 Gent, Belgium

GHENT UNIVERSITY

1. Why identifiability?

Forward problem:
PBM calculation



Inverse problem:
is **p** unique?

- Identifiability is the inverse modelling problem:

$$f(t,x,\mathbf{p}) = f(t,x,\mathbf{q}) \implies \mathbf{p} = \mathbf{q}$$

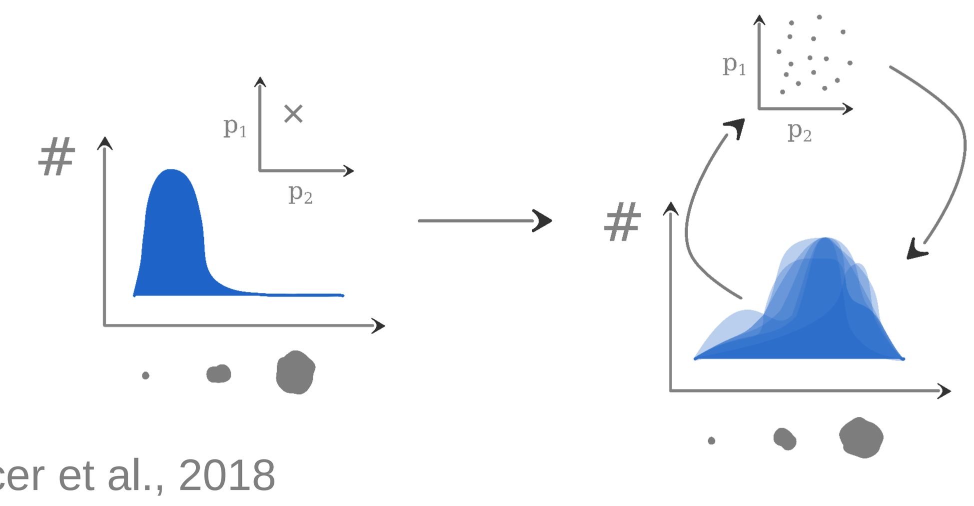
- Important when parameters are linked to physical mechanisms (e.g. machine/operational settings)
- Important for broader application of model

2. How is it calculated for PBM?

- Analytical methods: not applicable
- Most suitable method: repeated numerical parameter estimations *

$$\hat{\mathbf{p}} = \underset{\mathbf{p}}{\operatorname{argmin}}\ \mathrm{RMSE}(y, f(t,x,\mathbf{p}))$$

- Practical identifiability: synthetic data with relative Gaussian noise



* reference: Tuncer et al., 2018

3. Focus of this poster

$$\frac{\delta n(t,x)}{\delta t} = \frac{1}{2}\int_0^x \beta(t, x-\varepsilon, \varepsilon)\, n(t, x-\varepsilon)\, n(t,\varepsilon)\, d\varepsilon$$
$$- n(t,x)\int_0^\infty \beta(t,x,\varepsilon)\, n(t,\varepsilon)\, d\varepsilon$$
$$+ \int_x^\infty b(t,x,\varepsilon)\, S(t,\varepsilon)\, n(t,\varepsilon)\, d\varepsilon - S(t,x)\, n(t,x)$$

- Implementation in Python using Cell Average Technique
- This poster: pure aggregation process, two kernels are assessed:

$$\beta = \beta_0 \cdot (x-\varepsilon)^{\frac{1}{3}} \cdot \varepsilon^{\frac{1}{3}}$$

$$\beta = \beta_0 \cdot (x-\varepsilon)^{\frac{1}{3}} \cdot \varepsilon^{\frac{1}{3}} \cdot (1 + 0.5(step - 1)$$
$$(1 + tanh(R - \sqrt{(x-\varepsilon)^2 + \varepsilon^2})))$$

4. Results (part 1)

- Analysis of optimization results: average relative error of estimated parameters:

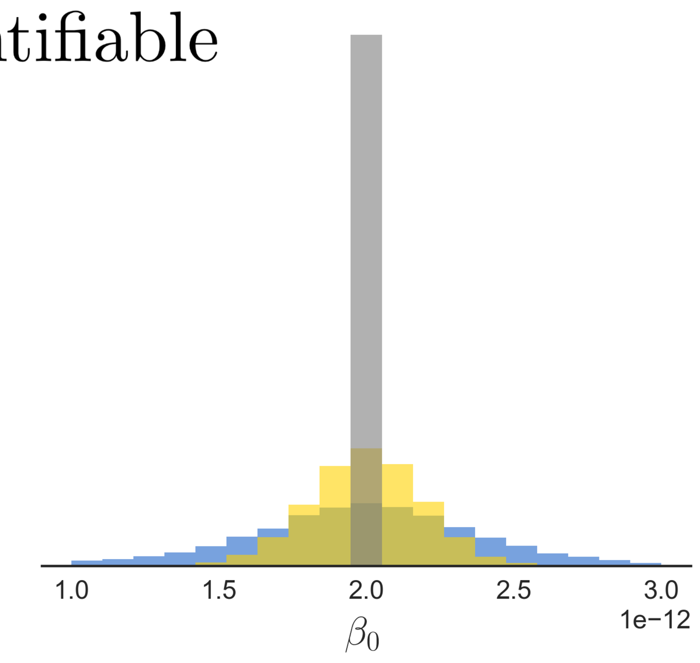
$$\mathrm{ARE}(\hat{p}^k) = 100\% \frac{1}{M}\sum_{j=1}^{M}\frac{|p^k - \hat{p}^k_j|}{p^k}$$

$$\mathrm{ARE} < \sigma_{noise} \implies \text{identifiable}$$

- Product kernel: structurally and practically identifiable

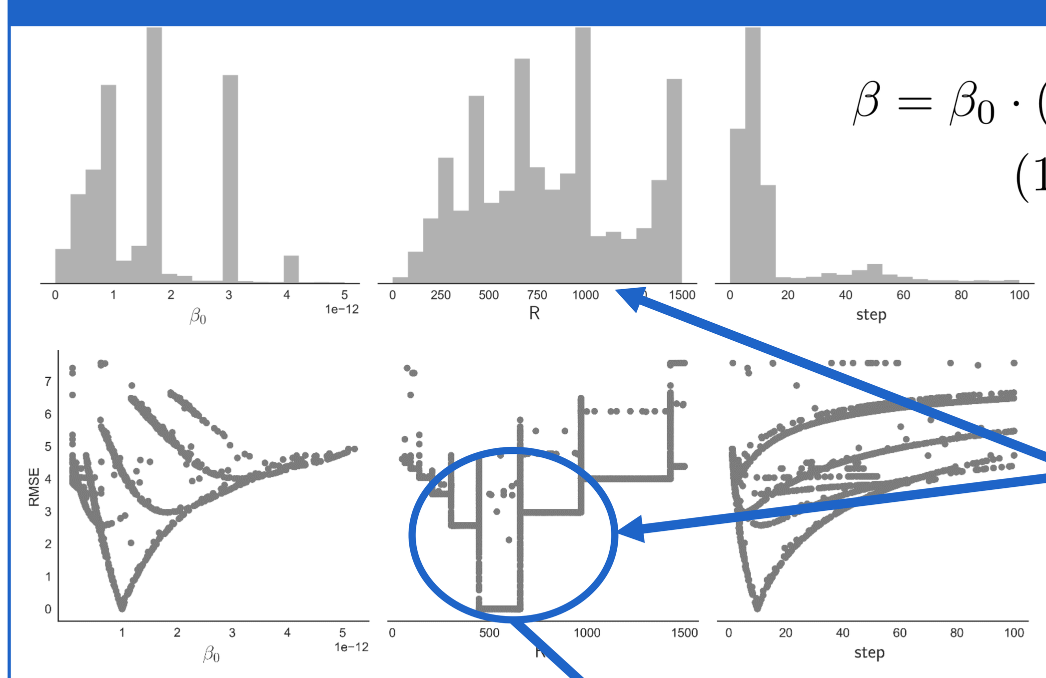
$$\beta = \beta_0 \cdot (x-\varepsilon)^{\frac{1}{3}} \cdot \varepsilon^{\frac{1}{3}}$$

structural
practical (25% relative error)
practical (50% relative error)



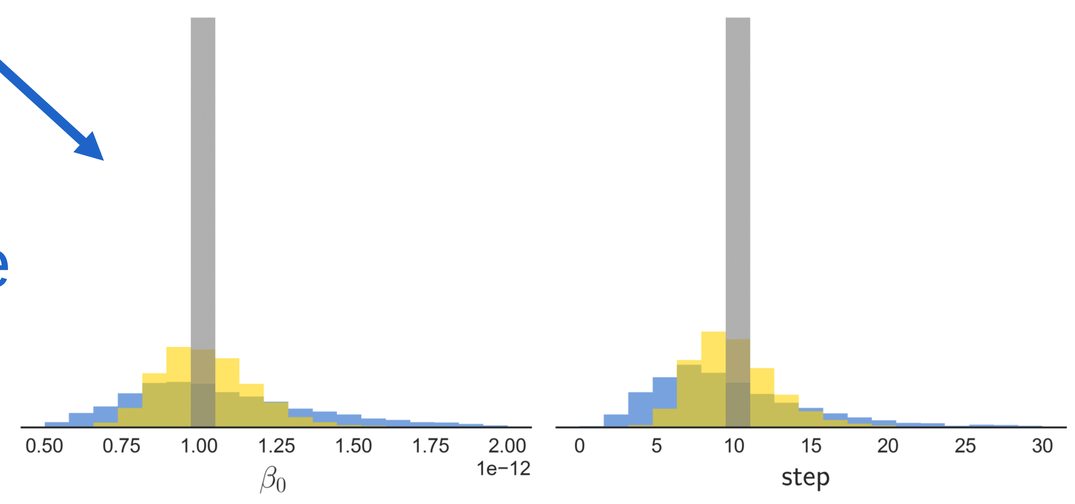
5. Results (part 2)

$$\beta = \beta_0 \cdot (x-\varepsilon)^{\frac{1}{3}} \cdot \varepsilon^{\frac{1}{3}} \cdot (1 + 0.5(step - 1)$$
$$(1 + tanh(R - \sqrt{(x-\varepsilon)^2 + \varepsilon^2})))$$



Identifiability issues in model!

Fixing **R** solves identifiability issue



6. Conclusion and prospects

- The used method is elegant, widely applicable, and yields good results
- Analysis for new kernels can be performed in a short timeframe
- It is possible to assess and improve kernel structures

- Future work will include more kernels, and combinations of mechanisms: aggregation and breakage
- Other identifiability techniques can be explored (e.g. profile likelihood)