



POLITECNICO
MILANO 1863

DEPARTMENT
OF MATHEMATICS



fdaPDE
Physics-Informed Statistical Learning

Kinetic Models by Physics-Informed Statistical Learning Methods

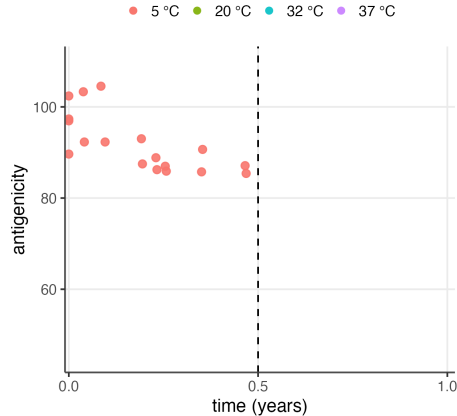
NCS 2026 - Padua, Italy, September 23-25, 2026

Marco Galliani¹ Bernard G. Francq² Laura M. Sangalli¹

¹MOX, Department of Mathematics, Politecnico di Milano, Milan, Italy.

²CMC Statistical Modeling, GSK, Rixensart, Belgium.

CQA: antigenicity



STABILITY STUDIES

■ Quantity of Interest:

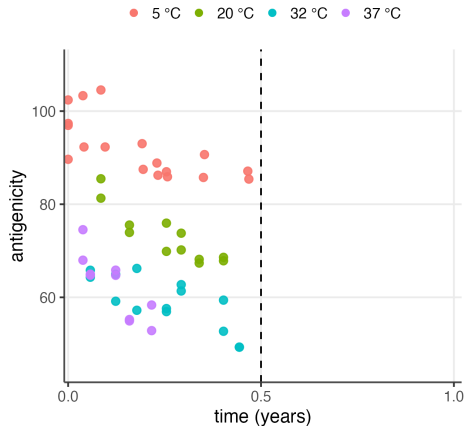
Critical Quality Attribute (CQA), proxy of product effectiveness, e.g.,

- ▶ concentration of Active Pharmaceutical Ingredient
- ▶ vaccine potency

■ Goal:

- ▶ estimate CQA beyond the data time-window
- ▶ determine product **shelf-life** with confidence

CQA: antigenicity



ACCELERATED STABILITY STUDIES

■ Stressed conditions:

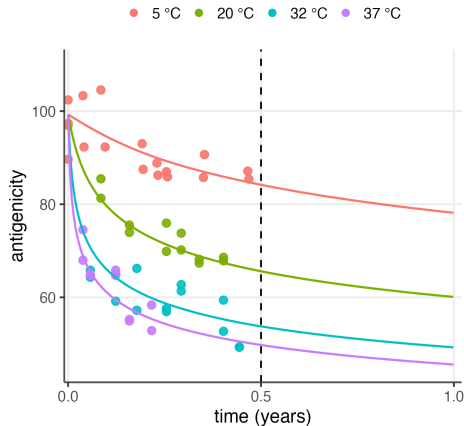
Accelerate product degradation

- ▶ higher storage temperatures
- ▶ higher humidity

■ Goal:

- ▶ estimate CQA beyond the data time-window
- ▶ determine product **shelf-life** with confidence
- ▶ Reduce time window necessary for prediction

CQA: antigenicity



ACCELERATED STABILITY STUDIES

■ Stressed conditions:

Accelerate product degradation

- ▶ higher storage temperatures
- ▶ higher humidity

■ Goal:

- ▶ estimate CQA beyond the data time-window
- ▶ determine product **shelf-life** with confidence
- ▶ Reduce the time window necessary for prediction

TIME

Šesták-Berggren model

$\alpha(t)$: reacted fraction of CQA at time t

$$\frac{d\alpha}{dt} = \alpha^m (1 - \alpha)^n, \quad \alpha(0) = 0$$

m : autocatalytic term, n : reaction order

TEMPERATURE

Arrhenius equation

$k(T)$: reaction rate at temperature T

$$k(T; A, E_a) = A \exp \left\{ -\frac{E_a}{RT} \right\}$$

E_a : activation energy, A : frequency of collisions

Final decay model

$Y(t)$: remaining qty of CQA at time t , $Y(t) = (C_0 - L_0)(1 - \alpha(t)) + L_0 \in [L_0, C_0]$

$$\begin{cases} \frac{dY}{dt} = -\frac{k(T; A, E_a)}{(C_0 - L_0)^{m+n-1}} (C_0 - Y)^m (Y - L_0)^n \\ Y(0) = y_0 \end{cases}$$

Model parameters: $\theta = \{E_a, A, m, n, C_0, L_0\}$

Nonlinear Least Squares:

- Difficult optimisation landscape
- Model misspecification

$$J(\boldsymbol{\theta}) = \sum_{j=1}^k (y_j - y(t_j))^2 \quad \text{s.t.} \quad \frac{dy}{dt} = f_{\boldsymbol{\theta}}(t, y)$$

Penalised Regression Framework

$$J(y, \boldsymbol{\theta}) = \sum_{j=1}^k (y_j - y(t_j))^2 + \lambda \int \left\| \frac{dy}{dt} - f_{\boldsymbol{\theta}}(t, y) \right\|^2 dt$$

Kinetic model as a soft penalty,
 λ governs the trade-off:

- $\lambda \rightarrow 0$: completely data-driven estimate
- $\lambda \rightarrow \infty$: completely model-based estimate

FORWARD PROBLEM

Define $u = \frac{dy}{dt} - f_\theta(t, y)$, model misfit.

Optimal Control Problem (OCP) formulation:

$$\min_u J(u, \theta) = \sum_{j=1}^k (y_j - y(t_j))^2 + \lambda \int \|u\|^2 dt$$

$$\text{s.t. } \frac{dy}{dt} = f_\theta(t, y) + u$$

Optimality system

$$u \in L^2([0, T]), y \in H^1([0, T])$$

$$\begin{cases} \frac{dy}{dt} = f(t, y) + u \\ -\frac{dp}{dt} = \frac{df}{dy}(t, y)^\top p + 2 \sum_j (y(t_j) - y_j) \delta(t - t_j) \\ \nabla_u J = 2\lambda u - p = 0 \end{cases}$$

INVERSE PROBLEM

The problem reads as

$$\min_{\theta} \underbrace{\left\{ \sum_{j=1}^k (y_j - y_{\theta}^*(t_j))^2 + \lambda \int \|u_{\theta}^*\|^2 dt \right\}}_{H(\theta)}$$

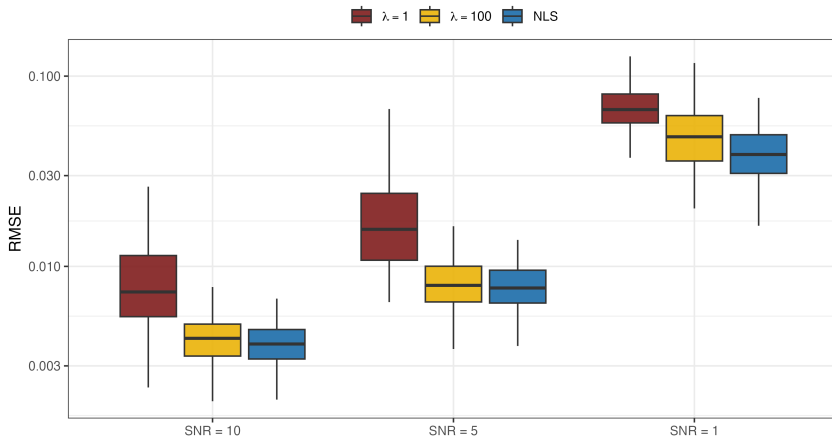
where $u_{\theta}^*, y_{\theta}^*$ solves the forward problem for the corresponding θ . To compute the gradient

$$\frac{dH}{d\theta} = \frac{\partial J}{\partial y} \frac{dy}{d\theta} + \frac{\partial J}{\partial u} \frac{du}{d\theta}$$

where $\frac{dy}{d\theta}, \frac{du}{d\theta}$ obtained by differentiating forward problem's optimality system w.r.t. θ

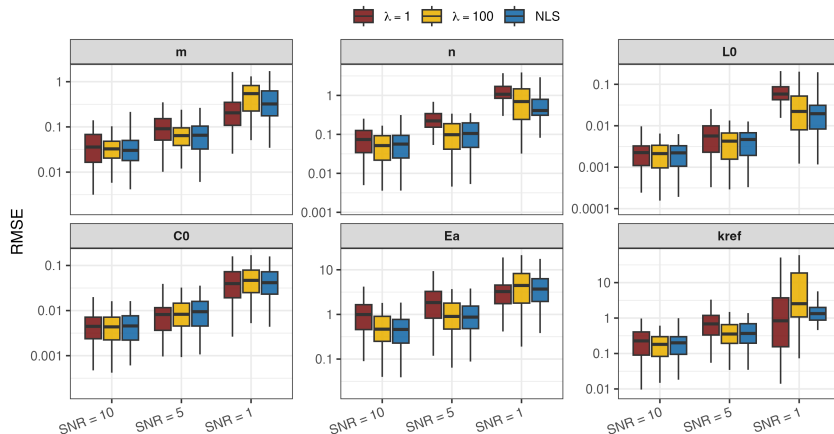
Šesták-Berggren model: $\frac{dY}{dt} = -\frac{k(T;A,E_a)}{(C_0-L_0)^{m+n-1}}(C_0-Y)^m(Y-L_0)^n$, $\theta = \{E_a, A, m, n, C_0, L_0\}$

Trajectory RMSE against the true signal



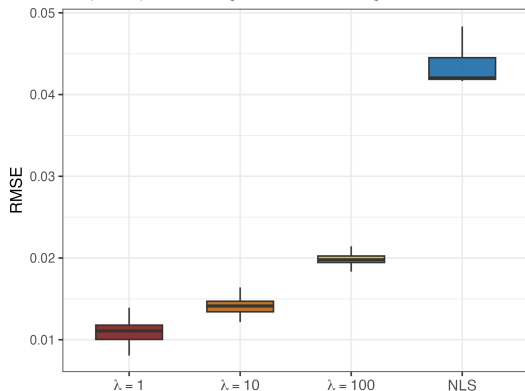
Šesták-Berggren model: $\frac{dY}{dt} = -\frac{k(T;A,E_a)}{(C_0-L_0)^{m+n-1}}(C_0-Y)^m(Y-L_0)^n$, $\theta = \{E_a, A, m, n, C_0, L_0\}$

Parameter error against the true values

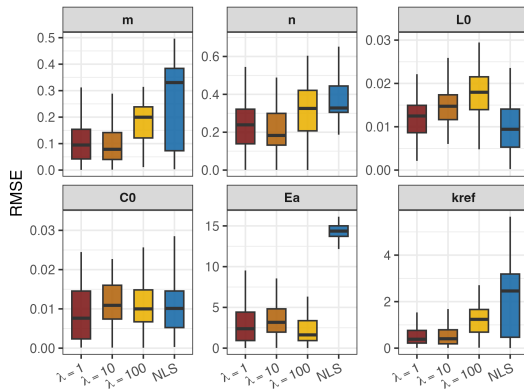


Perturbed SB model: $\frac{dY}{dt} = -\frac{k(T;A,E_a)}{(C_0-L_0)^{m+n-1}}(C_0-Y)^m(Y-L_0)^n + \sin(t)$ [to generate data],
 $\theta = \{E_a, A, m, n, C_0, L_0\}$

Trajectory RMSE against the true signal



Parameter error against the true values



ACHIEVEMENTS

- A **robust** method to estimate (y, θ) jointly
 - ▶ robust to complex optimization landscapes
 - ▶ robust to model misspecification
-

OPEN PROBLEMS

- Uncertainty quantification for $(\hat{y}, \hat{\theta})$
- λ selection (GCV)
- Assess extrapolation properties

References

-  Quentin Clairon and Nicolas J.-B. Brunel.
Optimal control and additive perturbations help in estimating ill-posed and uncertain dynamical systems.
Journal of the American Statistical Association, 113(523):1195–1209, 2018.
-  Bernard G. Francq.
AccelStab: Accelerated Stability Kinetic Modelling, 2025.
-  James O. Ramsay and Giles Hooker.
Dynamic Data Analysis: Modeling Data with Differential Equations.
Springer New York, 2017.
-  James O. Ramsay, Giles Hooker, David Campbell, and Jiguo Cao.
Parameter estimation for differential equations: A generalized smoothing approach.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69(5):741–796, 2007.



Discretized optimality system:

$$\begin{cases} y_{t+1} = \Phi_t(y_t, u_t), & y(0) = y_0 \\ p_t = \frac{\partial \Phi_t^\top}{\partial y} p_{t+1} - 2 \mathbb{I}_{\text{obs}}(y_t - y_{\text{obs},t}), & p_m := 0 \\ 2\lambda W_t u_t - \frac{\partial \Phi_t^\top}{\partial u} p_{t+1} = 0 \end{cases}$$

Integration scheme: an s -stage Runge–Kutta method with tableau (A, b, c) ,

$$\begin{aligned} \Phi_t(y, u) &= y + h_t \sum_{i=1}^s b_i k_i, \\ k_i &= f\left(t_t + c_i h_t, y + h_t \sum_{j=1}^s a_{ij} k_j\right) + u_i, \end{aligned}$$

Collocation methods: on each interval the scheme carries a polynomial of degree s , determined by its stage derivatives $k_{t,i}$ and evaluable at any $\theta \in [0, 1]$,

$$y(t_t + \theta h_t) = y_t + h_t \sum_{i=1}^s \beta_i(\theta) k_{t,i}, \quad u(t_t + \theta h_t) = \sum_{i=1}^s \ell_i(\theta) u_{t,i},$$

where $\{\ell_i\}$ is the **Lagrange basis** on the nodes $\{c_i\}$ and β_i its primitive,

$$\ell_i(\theta) = \prod_{j \neq i} \frac{\theta - c_j}{c_i - c_j}, \quad \beta_i(\theta) = \int_0^\theta \ell_i(\tau) d\tau, \quad \beta_i(c_j) = a_{ji}, \quad \beta_i(1) = b_i.$$