An anticipatory approach to optimal experimental design for model discrimination

Brecht MR Donckels[▲,▼], Dirk JW De Pauw[▼], Jo Maertens[▲], Bernard De Baets[▼] and Peter A Vanrolleghem[△,▲]

▲ BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Belgium
▼ KERMIT, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Belgium
△ model*EAU*, Département de Génie Civil, Université Laval, Quebec, Canada

UNIVERSITEIT GENT

Introduction

Several mathematical models seem to give a reasonable description of the same process, but which model is the **best** one? Therefore, **new experiments** have to be performed in order to discriminate among them.

HOW TO FIND THESE EXPERIMENTS ?

OPTIMAL EXPERIMENTAL DESIGN FOR MODEL DISCRIMINATION

Mathematical models

The **mathematical models** used in our research are sets of (nonlinear) differential equations

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{\xi}(t), \boldsymbol{\theta}, t); \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \\ \hat{\boldsymbol{y}}(t) &= \boldsymbol{g}(\boldsymbol{x}(t)) \end{aligned}$$

where $\boldsymbol{x}(t)$ are the state variables; $\boldsymbol{\xi}(t)$ represents the experimental degrees of freedom; $\boldsymbol{\theta}$ are the model parameters; and $\hat{\boldsymbol{y}}(t)$ are the measured variables.

Optimal experimental design

The **optimal experiment** ($\xi^\star$) is found after optimization of a well-defined objective function within the design space Ξ

$$\xi^\star = \arg\max_{\xi \in \Xi} T(\xi)$$

The **objective function** to discriminate between model i and j (T_{ij}) should maximize the difference between the model predictions ($\Delta\hat{\boldsymbol{y}}_{ij}$) and should take into account the uncertainty on the measurements and the model predictions, as illustrated in the figure below.

$$T_{ij}(\xi) = \sum_{k=1}^{n_{sp}} \Delta\hat{\boldsymbol{y}}'_{ijk} \cdot \boldsymbol{\psi}_{ij}^{-1} \cdot \Delta\hat{\boldsymbol{y}}_{ijk}$$

sampling times → n_{sp}

$$\text{with} \quad \Delta\hat{\boldsymbol{y}}_{ijk} = \hat{\boldsymbol{y}}_i\left(\xi, \hat{\boldsymbol{\theta}}_i, t_k\right) - \hat{\boldsymbol{y}}_j\left(\xi, \hat{\boldsymbol{\theta}}_j, t_k\right)$$

estimated parameters → $\hat{\boldsymbol{\theta}}_j$

$$\boldsymbol{\psi}_{ij} = \boldsymbol{\Sigma} + \boldsymbol{\Omega}_i + \boldsymbol{\Sigma} + \boldsymbol{\Omega}_j$$

uncertainty on measurements → $\boldsymbol{\Sigma}$; uncertainty on model predictions → $\boldsymbol{\Omega}_j$; both together: uncertainty on predicted outcome of the new experiment

$$\text{and} \quad \boldsymbol{\Omega} = S_{\hat{\theta}}^{\hat{y}} \cdot \underbrace{\left(S_{\hat{\theta}}^{\hat{y}\,\prime} \cdot \boldsymbol{\Sigma} \cdot S_{\hat{\theta}}^{\hat{y}}\right)^{-1}}_{} \cdot S_{\hat{\theta}}^{\hat{y}\,\prime}$$

$$\boldsymbol{\Phi} = \mathsf{FIM}^{-1}$$

uncertainty on parameter estimates ← $\boldsymbol{\Phi}$; parameter sensitivities → $S_{\hat{\theta}}^{\hat{y}\,\prime}$

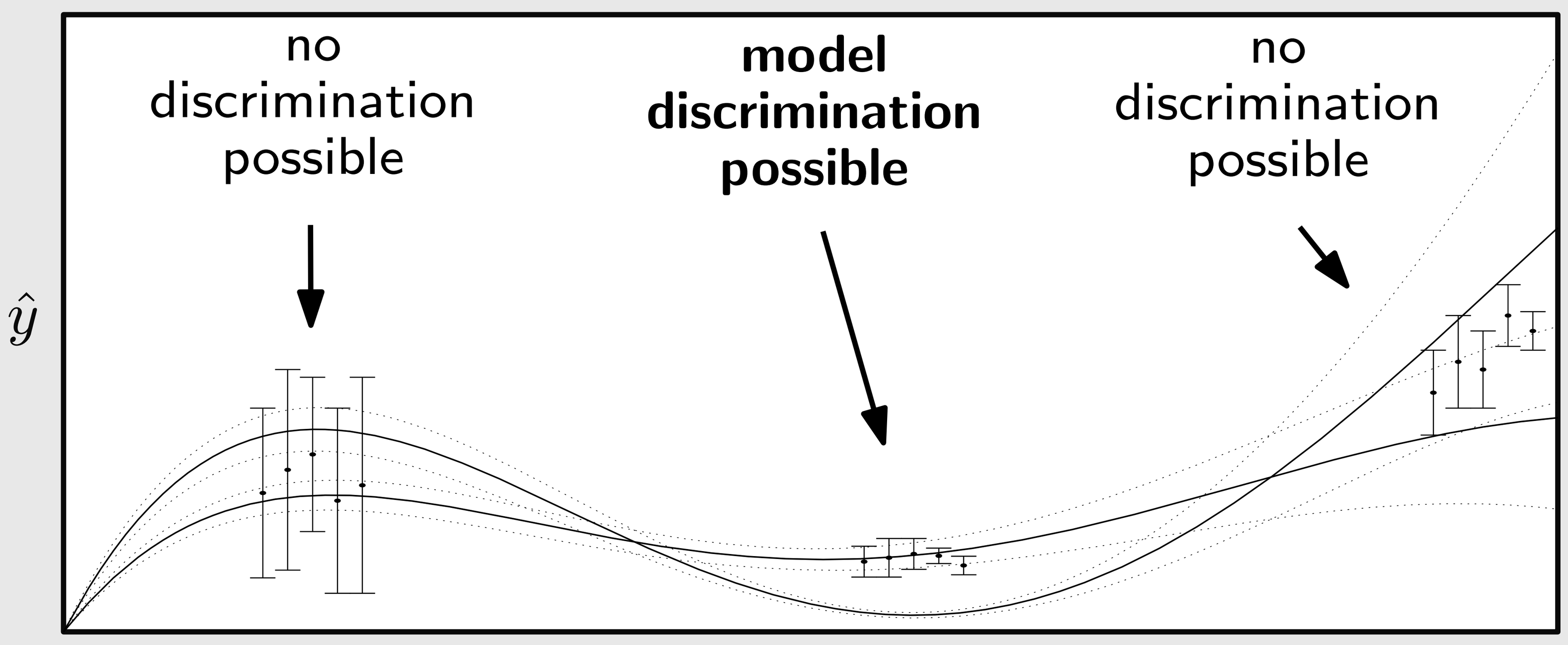

Figure: Illustrating the importance of uncertainty on the measurements and on the model predictions in the context of model discrimination

Current approach

$\boldsymbol{\Phi}$ is estimated from information collected in the n_e already performed experiments

$$\boldsymbol{\Phi}^{-1} = \sum_{k=1}^{n_e} \mathsf{FIM}(\boldsymbol{\xi}_k)$$

→ Information that will be collected is **ignored** in the estimation of the model prediction uncertainty $\boldsymbol{\Omega}$

Proposed anticipatory approach

$\boldsymbol{\Phi}$ is estimated from information collected in already performed experiments **and** from the designed experiment

$$\boldsymbol{\Phi}^{-1} = \sum_{k=1}^{n_e} \mathsf{FIM}(\boldsymbol{\xi}_k) + \mathsf{FIM}(\boldsymbol{\xi}_{n_e+1})$$

→ Taking into account the information that **will be collected** leads to a better estimation of the model prediction uncertainty $\boldsymbol{\Omega}$ that will determine the dicriminatory power of the data sets

→ Experiments could be obtained where parts of the collected data serves to **decrease the model prediction uncertainty** in regions where the difference in the model predictions is big, but would not have been exploited by the current approach because of the high model prediction uncertainty.

INCREASED DISCRIMINATORY POTENTIAL

For further information on this work, do not hesitate to contact me at **brecht.donckels@ugent.be**

This work was done in the framework of the MEMORE project. MEMORE is an IWT-funded project with partners from universities of Ghent, Brussels and Delft (NL). For further information, contact gino.baart@ugent.be.