HPC/Exascale Centre of Excellence in Personalised Medicine

PerMedCoE webinar series: Interrogating the effect of enzyme kinetics on metabolism using differentiable constraint-based models

Presenter: **Dr St. Elmo Wilken, Institute for Quantitative and Theoretical Biology, Heinrich Heine University, Düsseldorf (Germany)**

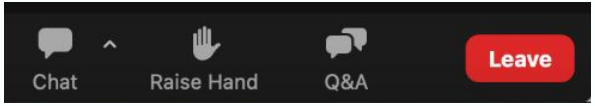
Host: **Daniel Thomas López (EMBL-EBI)**

The PerMedCoE project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement Nº951773

ABOUT THIS WEBINAR

This webinar is being recorded and will be disseminated afterwards

After the presentation we will address the questions posted by the audience using the Q&A function



PERMEDCOE MISSION

PerMedCoE is the HPC/Exascale Centre of Excellence for Personalised Medicine in Europe

The performance of current **simulation software** is still **insufficient to tackle medical problems** such as tumour evolution or patient-specific treatments.

Our goal is **to provide an efficient and sustainable entry point to the HPC/Exascale-upgraded methodology** to translate omics analysis into actionable models of cellular functions of medical relevance.

PERMEDCOE OBJECTIVES

Optimising cell-level simulation software to run in pre-exascale platforms

Integrating PerMed communities into the European HPC/Exascale ecosystem

Use cases driving the implementation of PerMedCoE solutions in HPC/Exascale environment

Building the basis for the sustainability of the PerMedCoE

Training biomedical professionals in the use of HPC/Exascale PerMedCoE tools

TODAY'S PRESENTER

St. Elmo Wilken

St. Elmo Wilken completed his undergraduate degree in Chemical Engineering at the University of Pretoria. His Ph.D. at the University of California, Santa Barbara leveraged both computational and wet lab aspects to investigate and understand the metabolism of anaerobic gut fungi. His current postdoc at the Institute of Quantitative and Theoretical Biology at the Heinrich Heine University in Düsseldorf is focused on using quantitative models to elucidate the contribution of metabolism to the stability and composition of microbial consortia.

He has partnered with PerMedCoE researchers, including Dr. Miroslav Kratochvíl, to develop a way to differentiate constraint-based models to conduct sensitivity analyses efficiently.

Interrogating the effect of enzyme kinetics on metabolism using differentiable constraint-based models

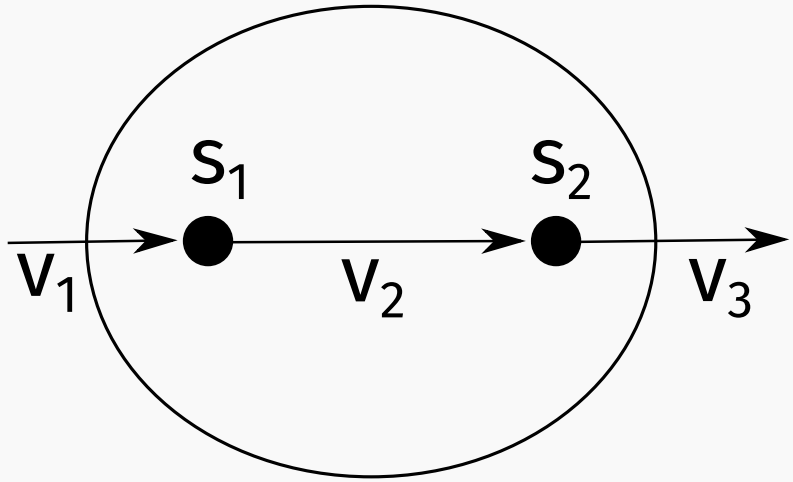
Dr. St. Elmo Wilken
Dr. Miroslav Kratochvíl
Dr. Mathieu Besançon

Heinrich Heine University, Germany
University of Luxembourg, Luxembourg
Zuse Institute, Germany

Table of contents

Modeling cellular metabolism with differential equations

Mass balance equations define an ODE metabolic model



Example of simplified metabolic model

$$\begin{aligned}\frac{\mathrm{d}\mathbf{s_1}}{\mathrm{d}t} &= v_1 - v_2 \\ \frac{\mathrm{d}\mathbf{s_2}}{\mathrm{d}t} &= v_2 - v_3\end{aligned} \iff \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \iff \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \mathbf{N} \cdot \mathbf{v}$$

Each flux is a function of metabolite concentrations, e.g.

$$v_2 = k_{\text{cat}} \cdot e \cdot \frac{s_1}{K_M + s_1}$$

Biological models have many parameters

Classic kinetic model ODE:

$$\frac{d\mathbf{s}}{dt} = \mathbf{N} \cdot \mathbf{v}$$

$$\mathbf{v} = \begin{bmatrix} k_{\text{cat},1} \cdot e_1 \cdot \frac{s_1}{K_{\text{M},1}+s_1} + \dots \\ \vdots \\ k_{\text{cat},N} \cdot e_N \cdot \frac{s_N}{K_{\text{M},N}+s_N} + \dots \end{bmatrix}$$

with $\mathbf{N}$ the stoichiometric matrix relating fluxes ($\mathbf{v}$) to metabolite concentrations ($\mathbf{s}$)

Parameters of $\mathbf{v}$ need to be measured

- Kinetic constants
  - $k_{\text{cat},i}$
  - $K_{\text{M},i}$
- Enzyme concentrations
  - e_i
- Regulatory components...
- Thermodynamic terms...

for each (i) reaction in the model

Measuring 1000s of parameters is challenging. Moreover, what effect does parameter uncertainty have on the model?

Evaluating parameter sensitivity in kinetic models

Kinetic models depend on parameters ($\mathbf{p}$):

$$\frac{d\mathbf{s}}{dt} = \mathbf{N} \cdot \mathbf{v}(\mathbf{s}(\mathbf{p}), \mathbf{p})$$

Assume steady state:

$$0 = \mathbf{N} \cdot \mathbf{v}(\mathbf{s}(\mathbf{p}), \mathbf{p})$$

Use the implicit function theorem to differentiate the model [1]:

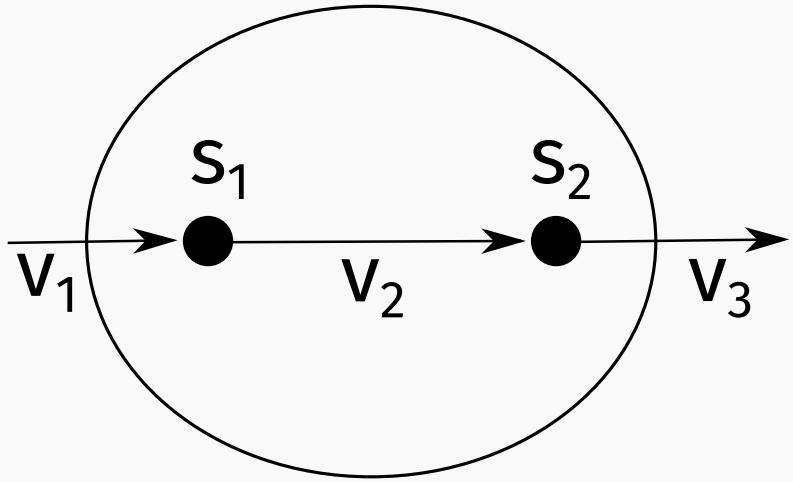
$$\implies \frac{\partial \mathbf{s}}{\partial \mathbf{p}} = \underbrace{-\left(\mathbf{N}\frac{\partial \mathbf{v}}{\partial \mathbf{s}}\right)^{-1}\mathbf{N}}_{\text{Concentration control matrix}} \frac{\partial \mathbf{v}}{\partial \mathbf{p}}$$

Metabolic control analysis (MCA) uses $\frac{\partial \mathbf{s}}{\partial \mathbf{p}}$ to evaluate the sensitivity of variables to parameters

[1] Heinrich & Schuster, Springer Science & Business Media, 2012.

Reframing a metabolic models by ignoring metabolites

Constraint-based models use fluxes as variables and assume steady state



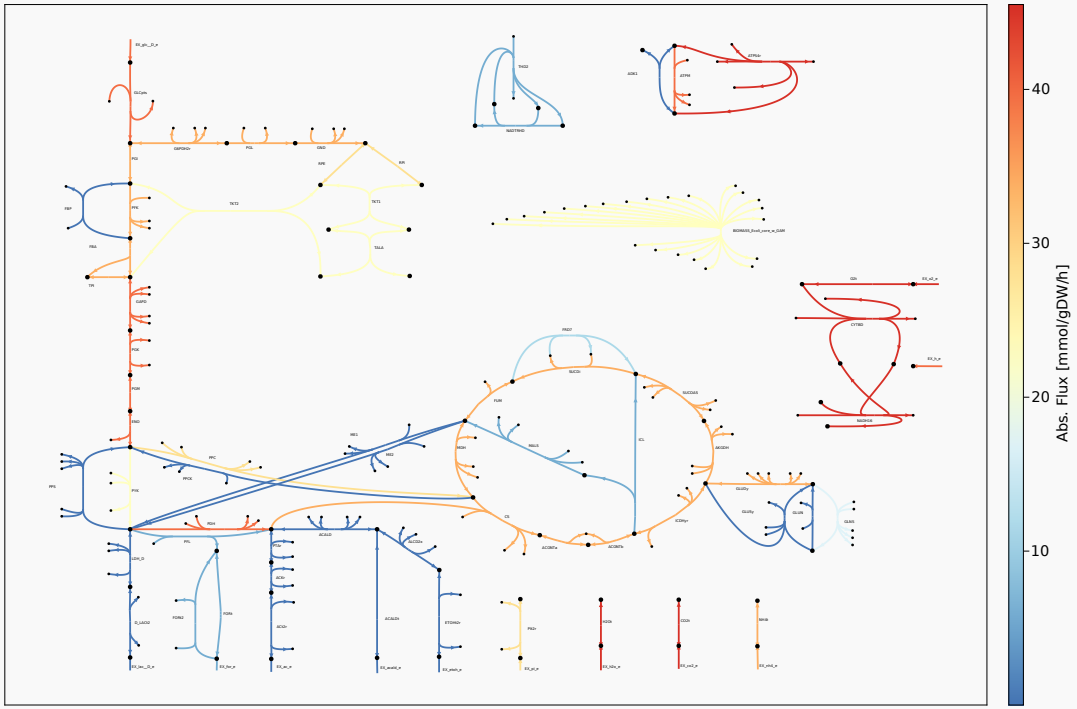
Example of simplified metabolic model

$$\frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} = \mathbf{0} = \mathbf{N} \cdot \mathbf{v}$$

Underdetermined system, need to add extra assumptions to narrow solution space

Flux balance analysis is an optimization based model

Flux balance analysis converts a metabolic model into an optimization problem



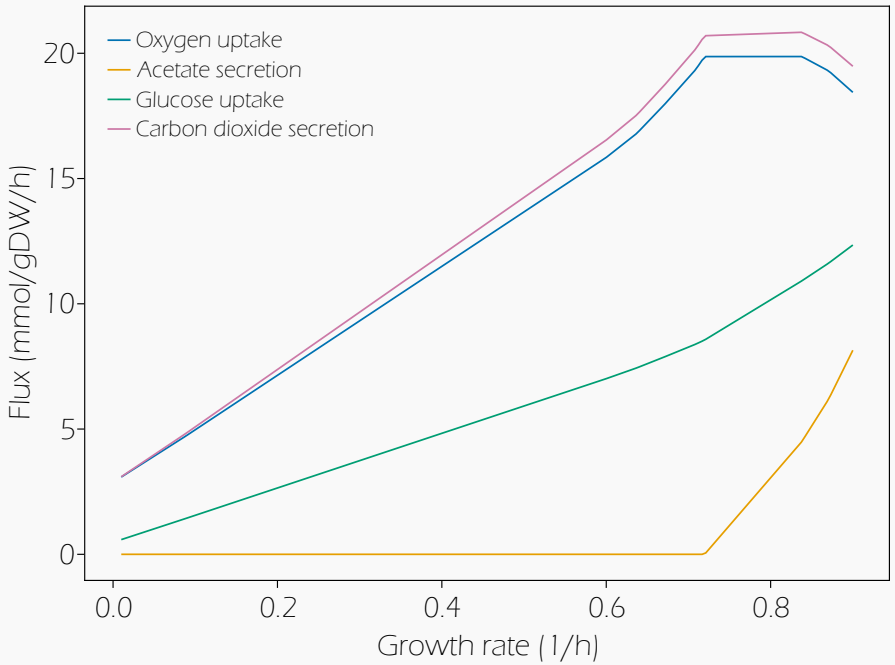
Central carbon metabolism of *E. coli*

$$\begin{aligned} \max_{\mathbf{v}} \quad & \mu(\mathbf{v}) \\ \text{s. t.} \quad & \mathbf{N} \cdot \mathbf{v} = 0 \\ & \mathbf{v}_{\text{lb}} \leq \mathbf{v} \leq \mathbf{v}_{\text{ub}} \end{aligned}$$

No kinetics and minimal experimental data are required, but additional information is simple to add

Constraint-based models reduce parameter burden

GECKO [2], MOMENT [3], etc. algorithms incorporate enzyme kinetics into flux balance analysis:



Overflow metabolism of *E. coli* can be modeled by incorporating enzyme capacity constraints

$$\begin{aligned} \max_{\mathbf{v},\mathbf{e}} \quad & \mu(\mathbf{v}) \\ \text{s. t.} \quad & \mathbf{N} \cdot \mathbf{v} = \mathbf{0} \\ & |\mathbf{v}| \leq \mathbf{k}_{\text{cat}} \cdot \mathbf{e} \\ & \sum e_i \leq E_{\text{total}} \end{aligned}$$

What is the sensitivity of the predicted fluxes and enzyme concentrations to the parameters?

[2] Sánchez, et al, Mol Syst Biol, 2017. [3] Adadi, et al, PLOS Comp Biol, 2012.

Metabolic control analysis for constraint-based models

- What is the analogue of metabolic control analysis for constraint-based models? Is there one?
- How to efficiently find:
  - $\frac{\partial v}{\partial k_{cat}}$
  - $\frac{\partial e}{\partial k_{cat}}$
  - Is there a way to avoid the finite difference based approach (flux control coefficients)?
- How can differentiable metabolic models be used to answer questions about metabolism?

Overarching question: how to differentiate through an optimization problem?

Implicit differentiation through a convex optimization problem

Flux balance analysis (FBA) type simulations are typically convex QP/LPs:

$$\left.\begin{aligned} \min_{z} \ & \frac{1}{2} z^T Q(p) z + c(p)^T z \\ \text{s. t. } & E(p) z = d(p) \\ & M(p) z \leq h(p) \end{aligned}\right\} \approx \text{FBA}(p) \text{ with e.g. } z = \begin{bmatrix} v \\ e \end{bmatrix}$$

How to find $\frac{\partial \text{FBA}}{\partial p}$?

At the optimum, z^*, the KKT conditions:

$$f(z^*(p), \nu^*(p), \lambda^*(p), p) = \begin{cases} Q(p)z^* + c(p) + E(p)^T \nu^* + M(p)^T \lambda^* & = 0 \\ E(p)z^* - d(p) & = 0 \\ \lambda^* \circ (M(p)z^* - h(p)) & = 0 \end{cases}$$

imply an implicit dependence between the primal (z^*) and dual variables (ν^*, λ^*), and the parameters (p) [4-5].

[4] Blondel, et al, Arxiv, 2021. [5] Amos, et al, P Mach Learn Res, 2017.

Implicit differentiation through a convex optimization problem

Letting $x^* = \begin{bmatrix} z^* \\ \nu^* \\ \lambda^* \end{bmatrix}$ suggests that at the optimum:

$$f(x^*(p), p) = 0$$

Hence, by implicitly differentiation [4-5], the sensitivities can be found by solving:

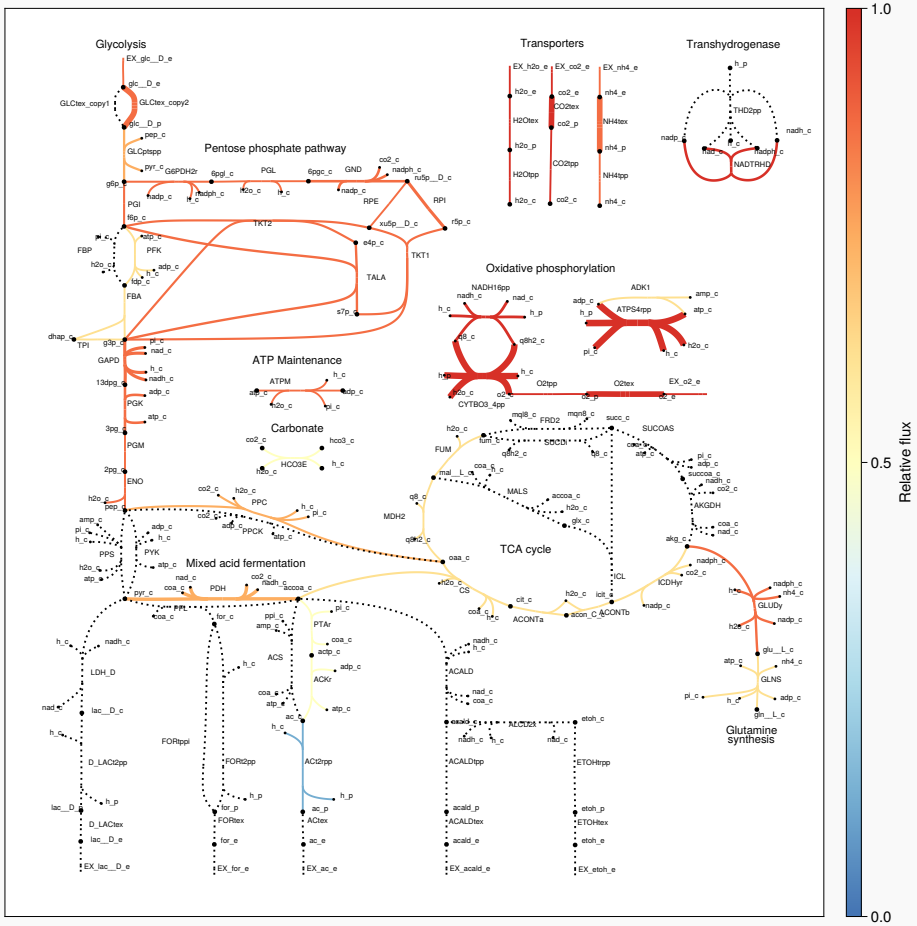
$$\frac{\partial f}{\partial x^*} \cdot \frac{\partial x^*}{\partial p} = \frac{\partial f}{\partial p}$$

$$\implies \frac{\partial x^*}{\partial p} = -\left(\frac{\partial f}{\partial x^*}\right)^{-1} \cdot \frac{\partial f}{\partial p}$$

Differentiating through an optimization problem requires finding the optimum, and thereafter solving a system of linear equations

[4] Blondel, et al, Arxiv, 2021. [5] Amos, et al, P Mach Learn Res, 2017.

Enzyme capacity constrained models require k_{cat} estimates



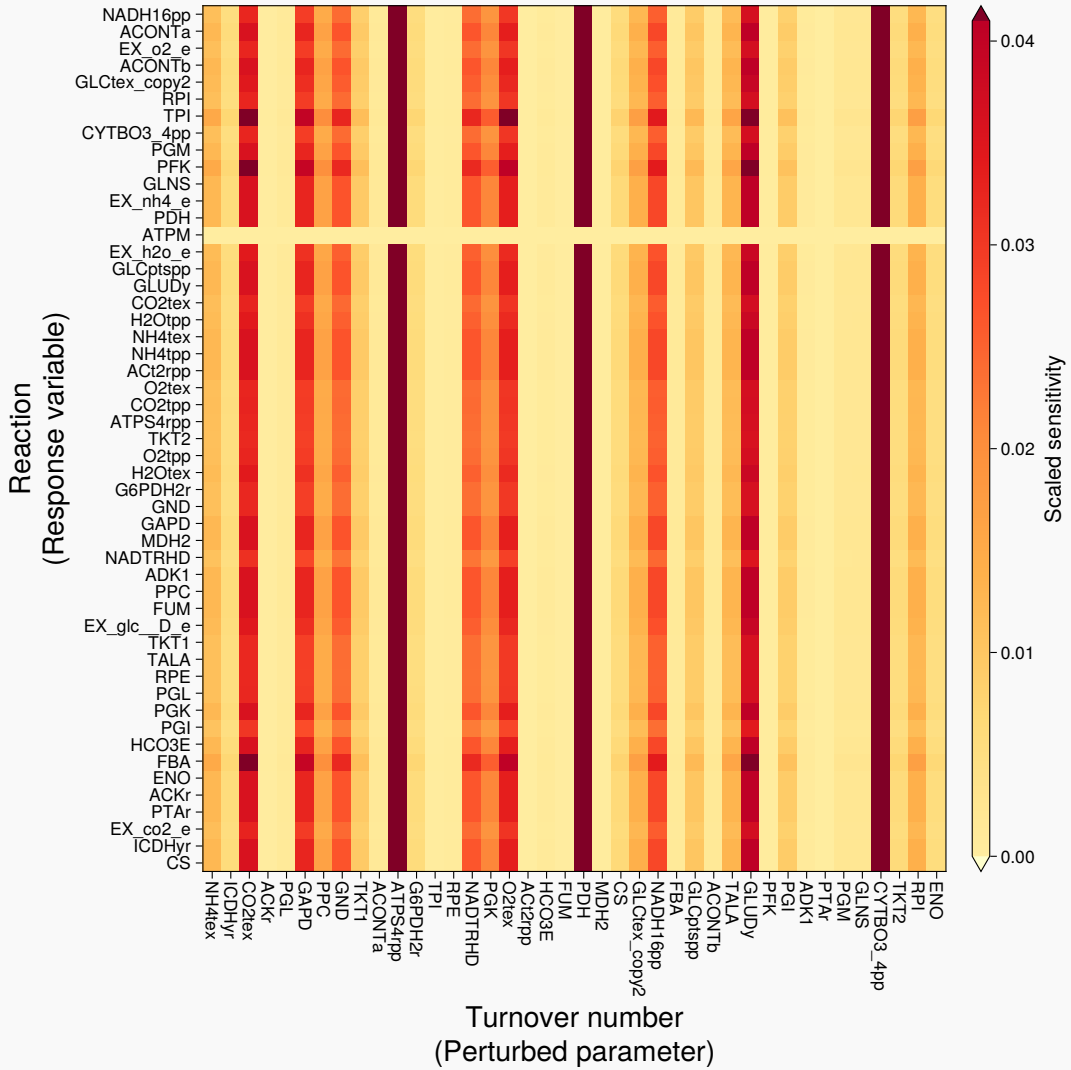
Central metabolism of *E. coli* model iML1515 [6] aerobic, glucose-fed simulation, machine learning k_{cat} estimates [7]

$$\begin{aligned} \max_{\mathbf{v},\mathbf{e}} \quad & \mu(\mathbf{v}) \\ \text{s. t.} \quad & \mathbf{N} \cdot \mathbf{v} = \mathbf{0} \\ & |\mathbf{v}| \leq \mathbf{k}_{\text{cat}} \cdot \mathbf{e} \\ & \sum e_i \leq E_{\text{total}} \end{aligned}$$

Use constraint-based MCA to find the flux control coefficients by differentiating the optimal solution

[6] Monk, et al, Nat Biotechnol, 2017. [7] Heckmann, et al, PNAS, 2020.

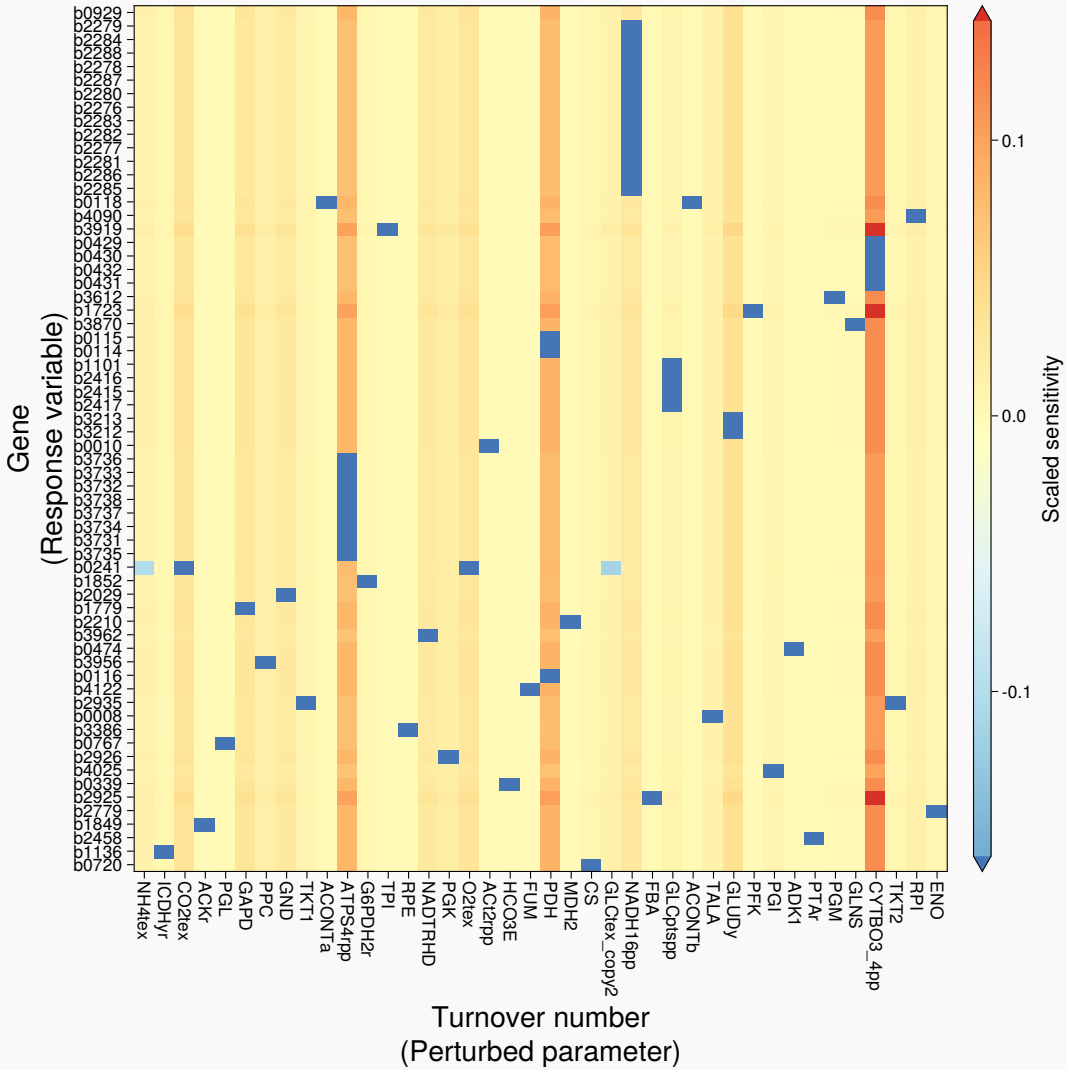
Sensitivity of fluxes to machine learning estimates of k_{cat}s



Flux sensitivities identify reaction kinetics exerting high control on solution

[8] **Wilken**, et al, Met Eng, 2022.

Sensitivity of enzyme concentrations to k_{cat}s



Enzyme sensitivities take stoichiometry into account and identifies flux controlling enzymes

[8] **Wilken**, et al, Met Eng, 2022.

Estimating k_{cat}s from data using derivatives

Increasingly quantitative absolute proteomic data, $\hat{e}$, coupled with *in vivo* flux measurements, $\hat{v}$, are available

Given estimated k_{cat}s, a **model based fit** can be calculated

$$L(k_{\text{cat}}\text{s}) = \min_{\mathbf{v},\mathbf{e}} \quad \frac{1}{I}\sum_i \left(\frac{\hat{v}_i - v_i}{\hat{v}_i}\right)^2 + \frac{1}{J}\sum_j \left(\frac{\hat{e}_j - e_j}{\hat{e}_j}\right)^2$$

$$\text{s. t.} \quad \mathbf{N}\cdot\mathbf{v} = 0$$

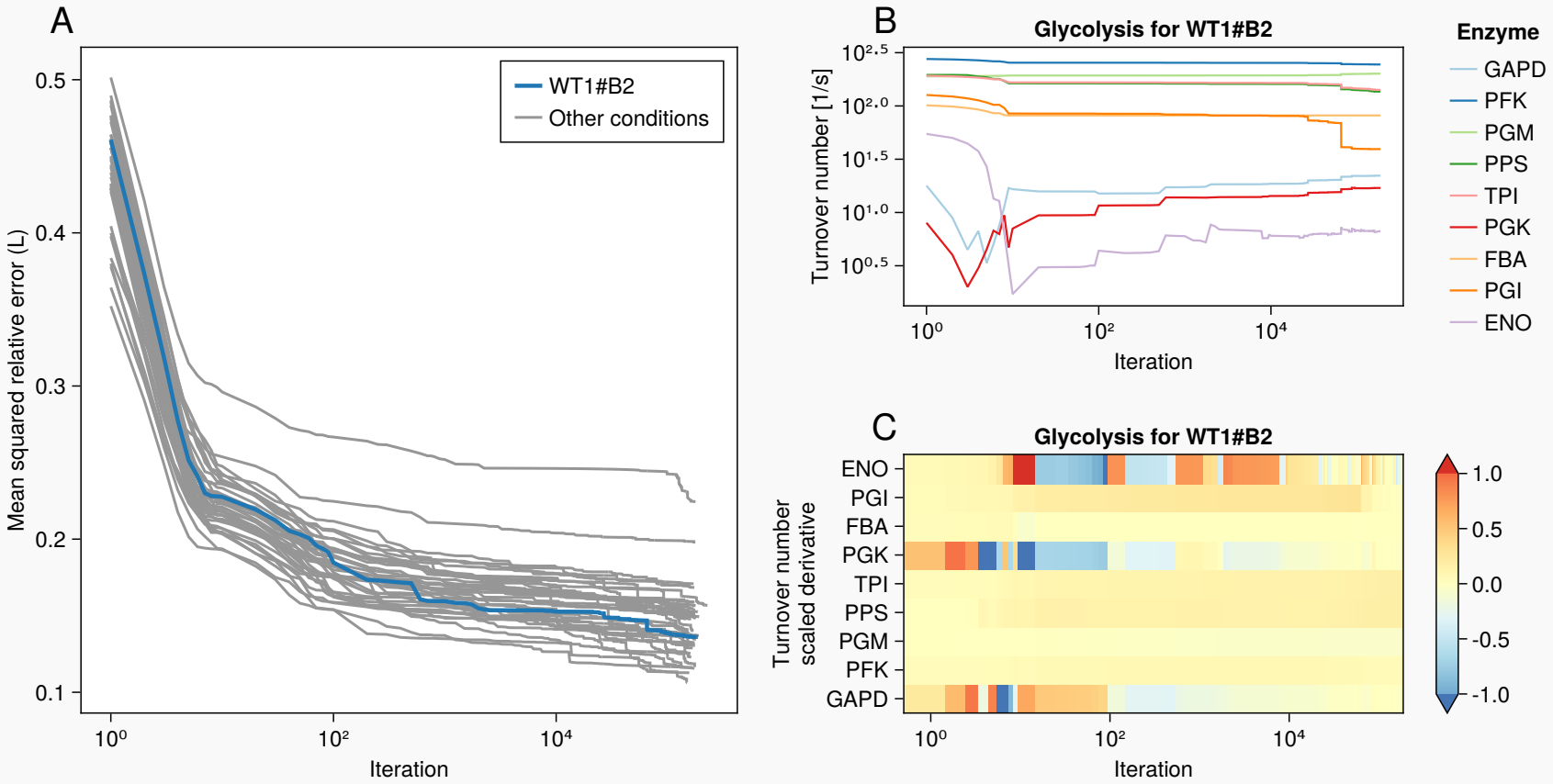
$$|v_i| \leq k_{\text{cat},i}\cdot e_i$$

$$\sum e_i \leq E_{\text{total}}$$

Using $\frac{\partial L}{\partial k_{\text{cat}}}$, gradient descent can be used to optimize the kinetic constants to fit the measured data

With derivative information, machine learning derived parameters can be tuned to better fit a metabolic model

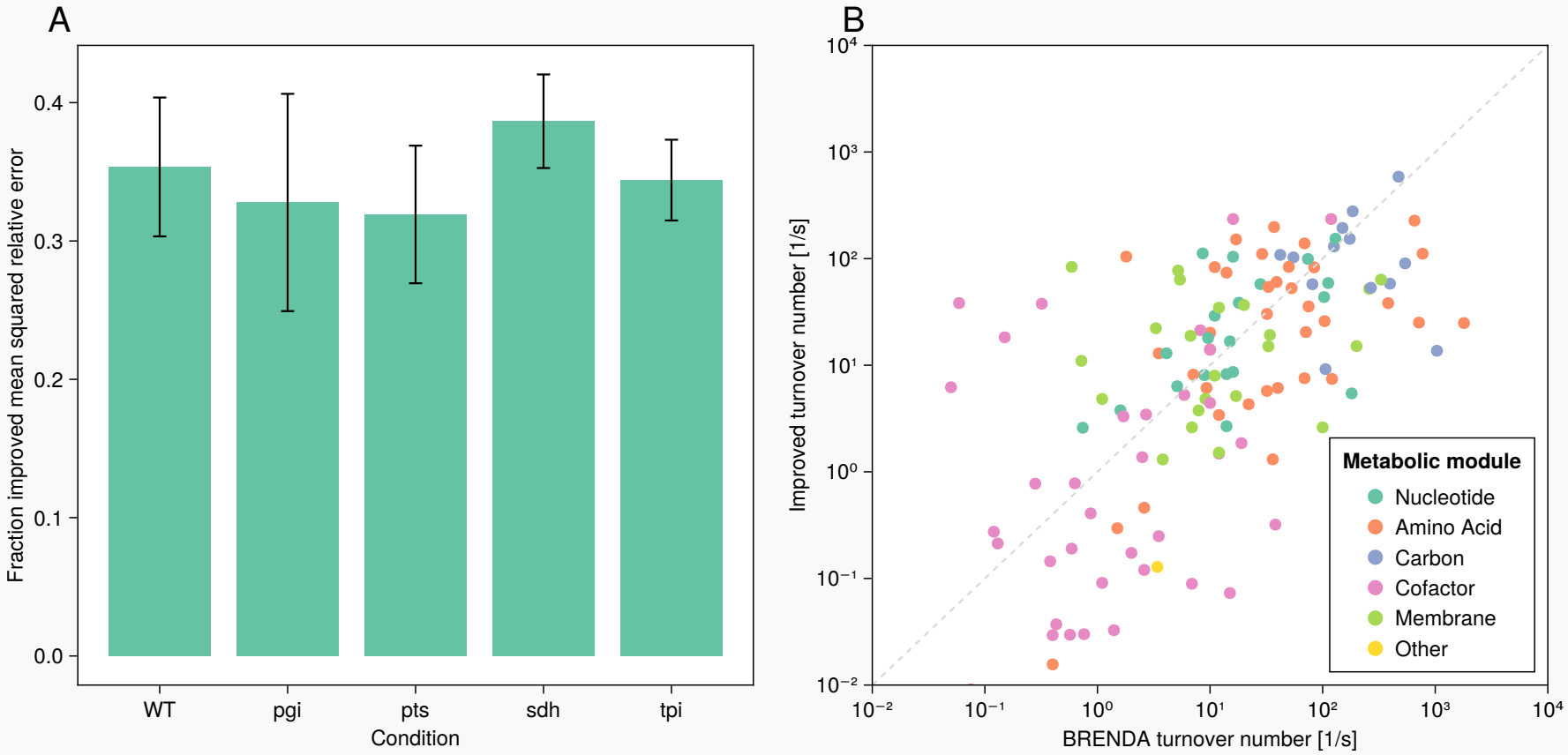
Improving ML k_{cat} estimates using gradient descent



Model-based gradient descent decreases model prediction/observation mismatch

[8] **Wilken**, et al, Met Eng, 2022.

Improving ML k_{cat} estimates using gradient descent



Model-based gradient descent using multiple datasets improves k_{cat} estimates by approximately 35%

[8] **Wilken**, et al, Met Eng, 2022.

Conclusion

Differentiable metabolic models

- Computationally efficient
- Similar mathematical foundation as classic MCA
- Forthcoming package to simplify this analysis
- Solution differentiability unlocks many gradient-based analysis techniques

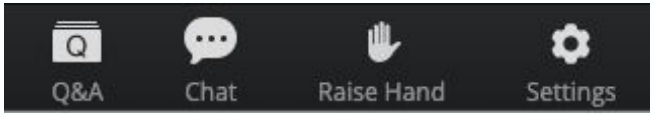
Acknowledgements

Dr. Mathieu Besançon

Dr. Miroslav Kratochvíl

QUESTIONS?

- Write your questions using the Q&A button



NEXT WEBINARS

Visit the Training section on https://permedcoe.eu/ to watch previous webinars and sign up for the upcoming sessions.

- Webinar: **Supercomputer-based modelling and simulation for advanced biomedical applications**
  - Dr Mariano Vázquez, ELEM Biotech and Barcelona Supercomputing Center (Spain)
  - Thursday 9 March 2023, 15-16 CET

- Webinar: **Development of a virtual Rheumatoid Arthritis synovial fibroblast for large-scale dynamic analysis and efficient drug-target identification**
  - Dr Anna Niarakis, University of Evry Val d'Essonne and National Institute for Research in Digital Science and Technology (INRIA)
  - Thursday 23 March 2023, 15-16 CET

Per Med CoE

HPC/Exascale Centre of Excellence in Personalised Medicine

SUMMER SCHOOL

From pathway modelling tools to cell-level simulations

Get hands-on experience using biological modelling tools and learn how high-performance computing clusters can be used to execute biomedical workflows.

25-30 June 2023

Hotel Montanyà, Seva, Catalonia

Applications open until 6 March 2023

https://permedcoe.eu/training/permedcoe-summer-school-on-cell-level-simulations/

The PerMedCoE project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement N°951773

www.linkedin.com/company/permedcoe

@permedcoe

Per Med CoE