



**HAL**  
open science

## Solving regression problems using bacterial metabolism

Paul Ahavi, Bastien Mollet, Antoine Cornuéjols, Evelyne Lutton, Alberto Tonda, Audrey Le Gouellec, Jean-Loup Faulon

### ► To cite this version:

Paul Ahavi, Bastien Mollet, Antoine Cornuéjols, Evelyne Lutton, Alberto Tonda, et al.. Solving regression problems using bacterial metabolism. 2023 Living Machines @ Work day, Oct 2023, Gif-sur-Yvette, France. . hal-04312498

**HAL Id: hal-04312498**

**<https://hal.science/hal-04312498>**

Submitted on 28 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Evaluation of information processing capacity of bacterial metabolism through regression problem solving

BioRetroSynth and Ekinocs teams

Ahavi Paul<sup>1\*</sup>, Bastien Mollet<sup>2\*</sup>, Antoine Cornuejols<sup>2</sup>, Evelyne Lutton<sup>2</sup> and Alberto Tonda<sup>2\*\*</sup>, Audrey Le Gouellec<sup>3\*\*</sup>, Jean-Loup Faulon<sup>1\*\*</sup>

<sup>1</sup>Micalis Institute (UMR 1319), INRAE - Paris-Saclay University - 78352 Jouy-en-Josas, France

<sup>2</sup>UMR MIA, AgroParisTech, INRAE - Paris-Saclay University - 91120 Palaiseau, France

<sup>3</sup>TIMC (UMR 5525 CNRS UGA) - CHU Grenoble Alpes - 38000 Grenoble, France

\* co-author

\*\* co-last author

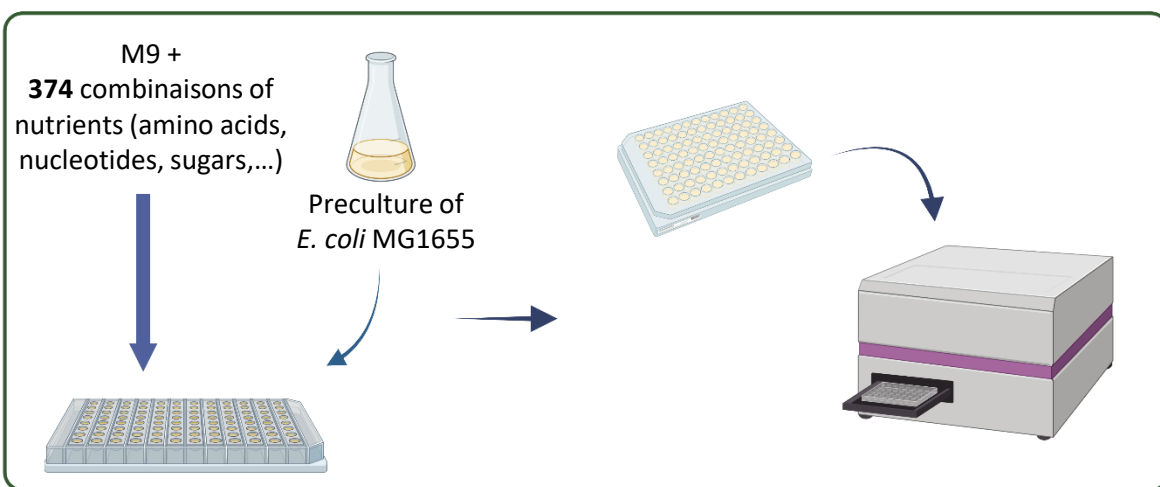


## Context

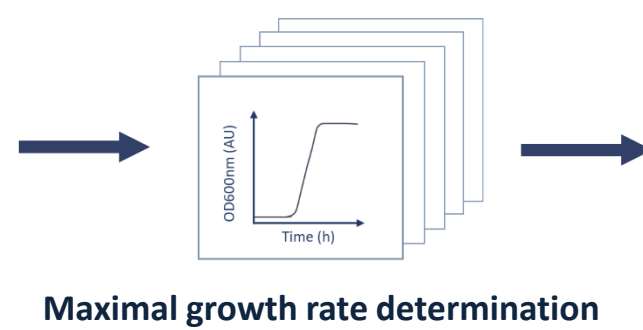
- Bacteria can **sense chemical variations** in their growth medium and **adapt their metabolism** accordingly<sup>[1]</sup>.
- Bacterial metabolism can therefore be compared to an **algorithm** taking as inputs the media composition and yielding as outputs metabolic fluxes describing a metabolic phenotype<sup>[2]</sup>.
- Our lab has focused on the perspective of metabolism as an information processing unit and previously developed a hybrid modelling method, the AMN for **Artificial Metabolic Networks**<sup>[3]</sup> (see QRcode), which is able to **predict the maximal growth rate** and other metabolic fluxes from a **medium composition**.

➤ **At present, we would like to go one step further and assess the capacity of *E. coli*'s metabolism by solving regression problems of increasing complexity.**

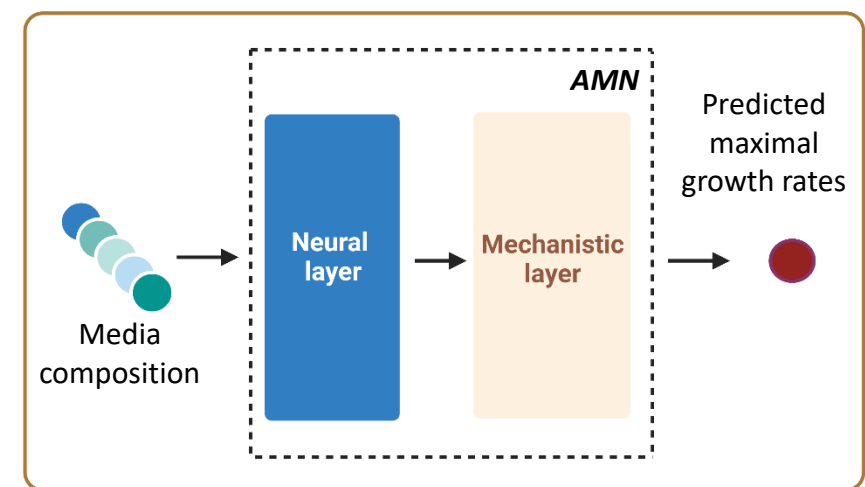
## Construction of a metabolic model of *E. coli*



Acquisition of growth kinetics of *E. coli* MG1655 in 374 different media



Maximal growth rate determination



AMN training

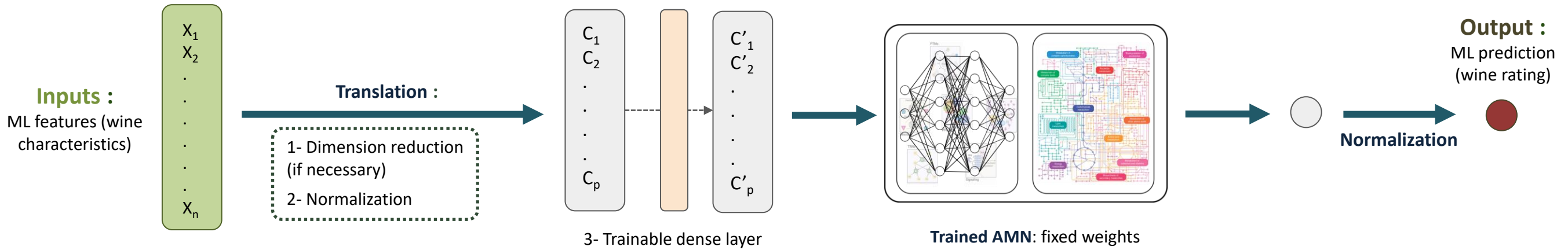
## Problem-solving

Example : White wine quality dataset (score from 1 to 10)

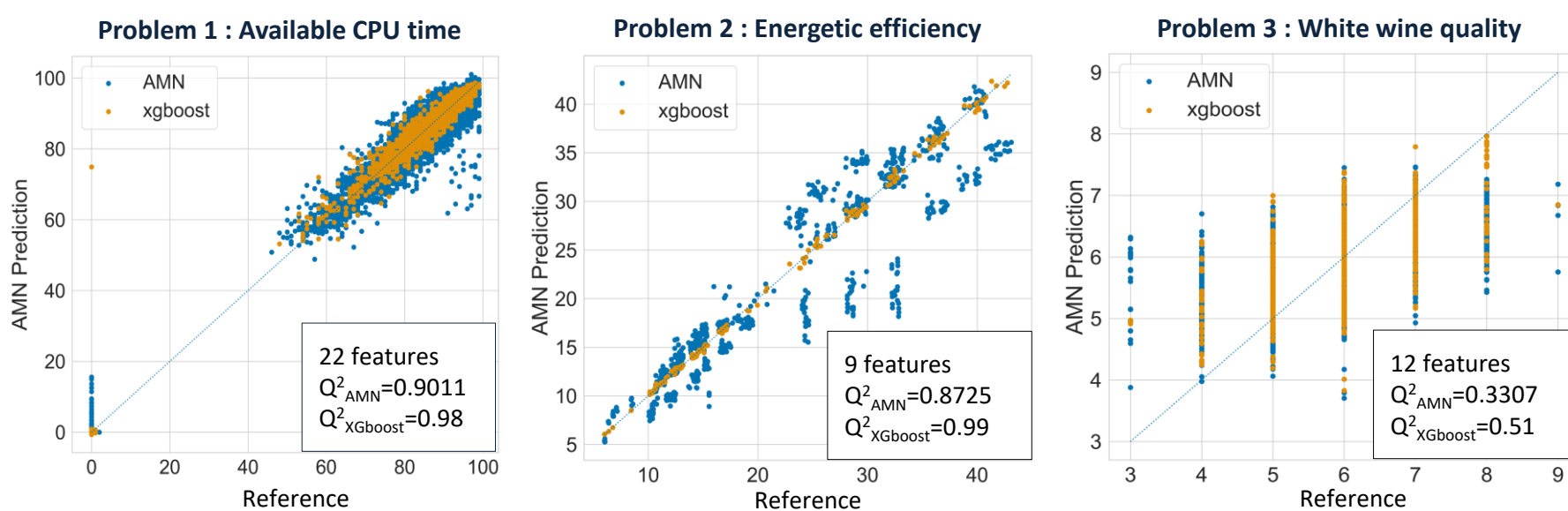
Densité	pH	Sulfates
0.9956	3.13	45

Glucose	Pi	Ala
0.9956	0.43	0.25

$\mu_{max}$	Wine rating
1,12	0,72 (7.2/10)



## Results



The AMN model obtains satisfactory performance on regression problem, but it is less effective than XGBoost, the state of the art

Open questions for the next steps:

- What is the role of the trainable dense layer ?
- What is the influence of feature pairing ?

## Perspectives

- Experimental validation** : An independent test set will be converted into growth media for bacteria after training the dense layer. The maximal growth rates will be experimentally verified.
- Kinetic modelling** : The AMN is a hybrid model based on an artificial neural network (ANN) and flux-balance analysis (FBA), the latter being a stoichiometric model. Using a kinetic model will increase the number of outputs and may enable to solve more complex problems
- Computing more metabolic fluxes** : Same principle, by increasing the number of outputs, the model may be able to solve more complex problems.
- Performance on classification problems**

### References

- [1] Zampieri, Mattia, Manuel Hörl, Florian Hotz, Nicola F. Müller, et Uwe Sauer. « Regulatory Mechanisms Underlying Coordination of Amino Acid and Glucose Catabolism in *Escherichia coli* ». *Nature Communications* 10, n° 1 (26 juillet 2019): 3354. <https://doi.org/10.1038/s41467-019-11331-5>.
- [2] Lahoz-Beltra, Rafael, Jorge Navarro, et Pedro C. Marijuán. « Bacterial computing: a form of natural computing and its applications ». *Frontiers in Microbiology* 5 (25 mars 2014). <https://doi.org/10.3389/fmicb.2014.00101>.
- [3] Faure, Léon, Bastien Mollet, Wolfram Liebermeister, et Jean-Loup Faulon. « A Neural-Mechanistic Hybrid Approach Improving the Predictive Power of Genome-Scale Metabolic Models ». Preprint. *Systems Biology*, 11 janvier 2022. <https://doi.org/10.1101/2022.01.09.475487>.