



HAL
open science

cMFA for multi-omics data integration in microbial community models

Sthyve Junior Tatho Djeanou, Simon Labarthe, Valentina Baldazzi

► **To cite this version:**

Sthyve Junior Tatho Djeanou, Simon Labarthe, Valentina Baldazzi. cMFA for multi-omics data integration in microbial community models. 25e Forum des jeunes mathématiciennes et mathématiciens 2025, Nov 2025, Bordeaux, France. <hal-05469230>

HAL Id: hal-05469230

<https://hal.science/hal-05469230v1>

Submitted on 21 Jan 2026

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 CC BY 4.0 - Attribution - International License

➤ cMFA for multi-omics data integration in microbial community models.

Sthyve Tatho, Simon Labarthe, Valentina Baldazzi

PhD Student
27 novembre 2025

FJMM 2025

26-28 nov. 2025 Talence

Sthyve Tatho

Univ. Bordeaux, INRAE, BIOGECO

Univ. Bordeaux, Inria-Inrae, Pléiade

sthyve.tatho@inria.fr



STHYVE TATHO



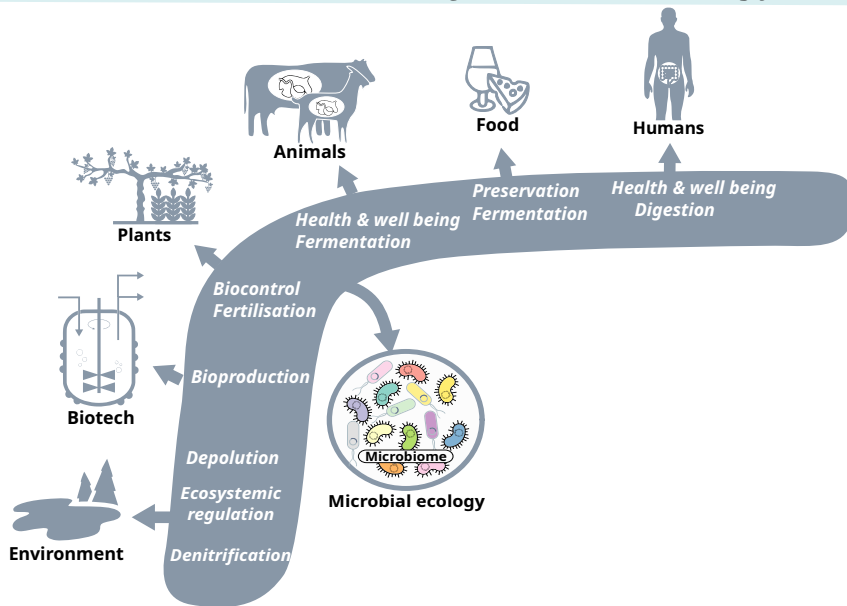
INRAE

Inria

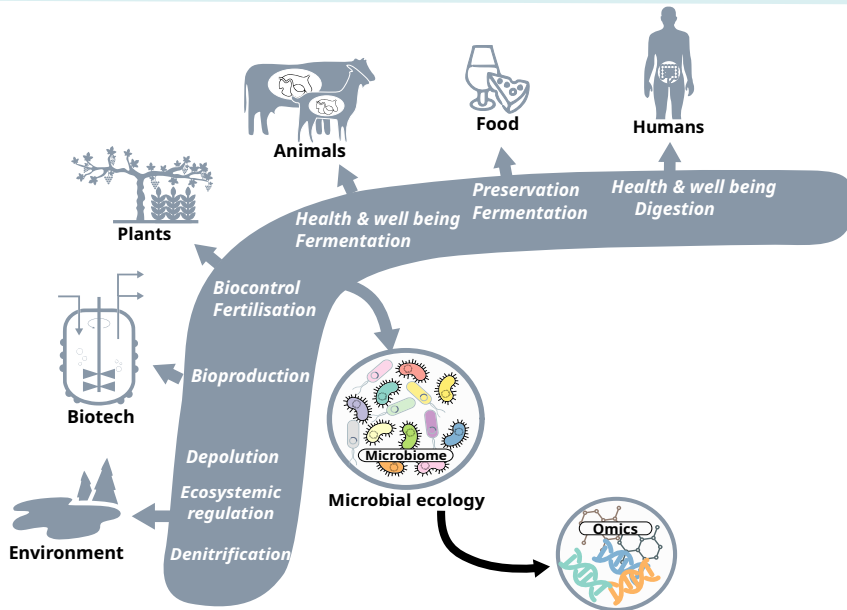
université
de BORDEAUX

How understanding microbial community functioning ?

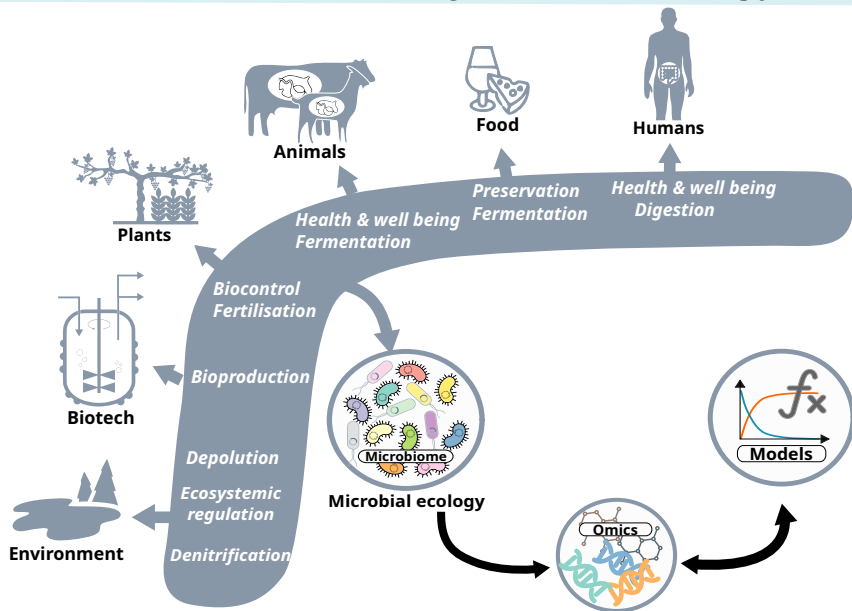
Context : How Understanding microbial community functioning?



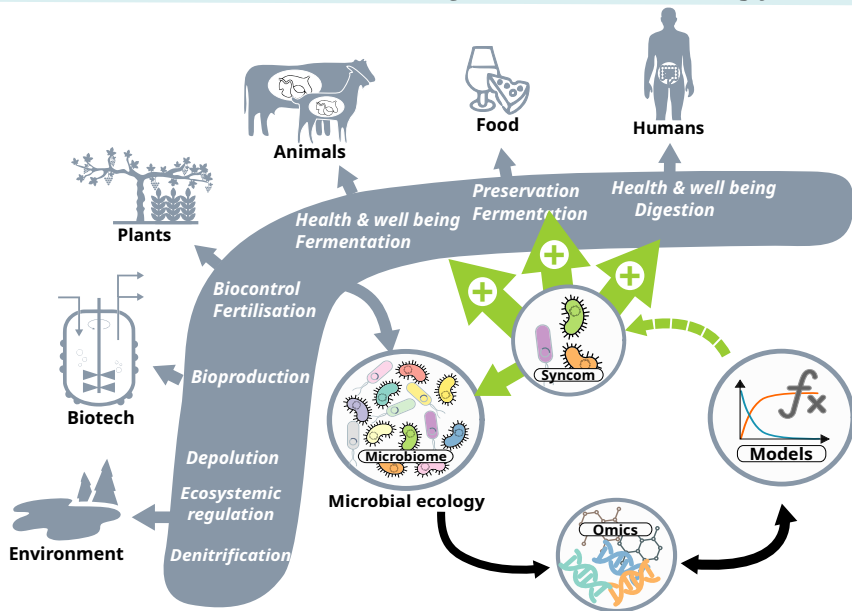
Context : How Understanding microbial community functioning?



Context : How Understanding microbial community functioning?



Context : How Understanding microbial community functioning?



Context : How Understanding microbial community functioning?

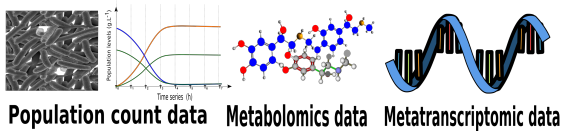
Context :

- ▶ Developing a mathematical inference method based on metabolic networks.
- ▶ Inferring metabolic fluxes by integrating multi-omics data.

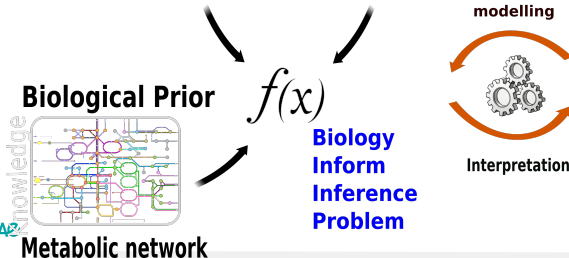
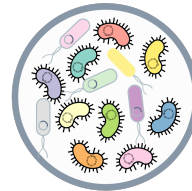
Context : How Understanding microbial community functioning ?

Context :

- Developing a mathematical inference method based on metabolic networks.
- Inferring metabolic fluxes by integrating multi-omics data.



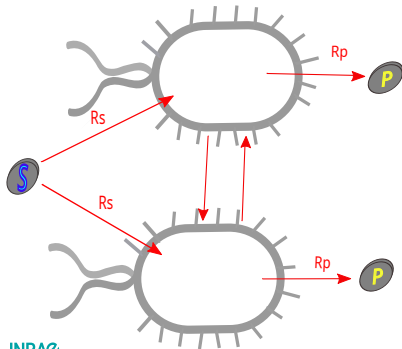
Microbial Community



Biology-informed inference from multi-omics data

Objectives :

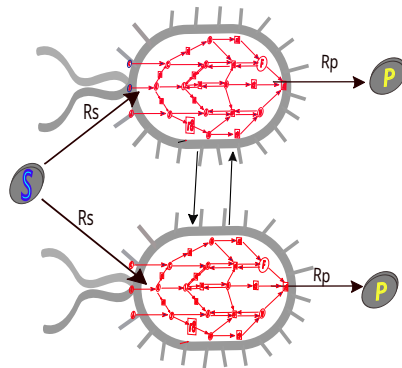
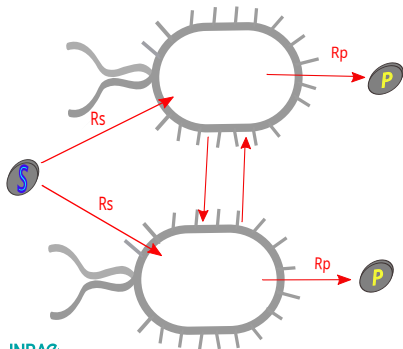
- Estimating the contribution of each microorganism to the community.



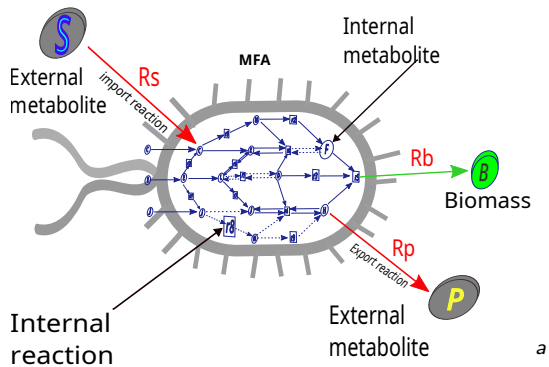
Biology-informed inference from multi-omics data

Objectives :

- Estimating the contribution of each microorganism to the community.
- Reconstructing internal fluxes of individual micro-organisms.

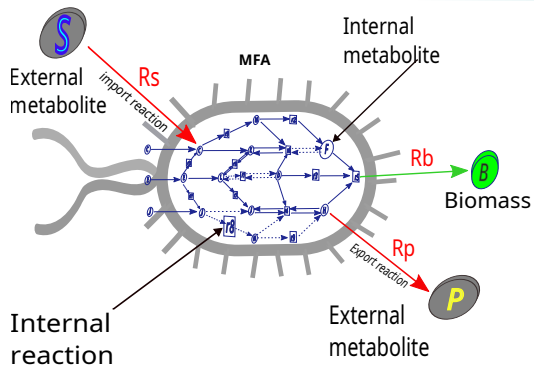


cMFA for community models



a. Gerken Starepravo, M. et al. (2022). *Digital Chemical Engineering*

cMFA for community models

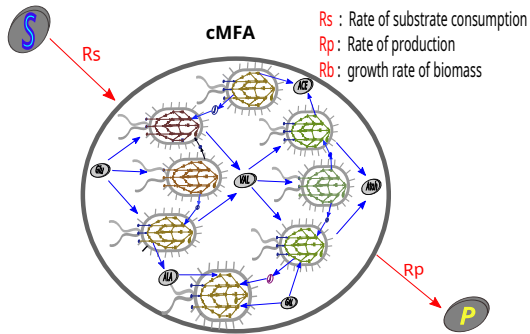
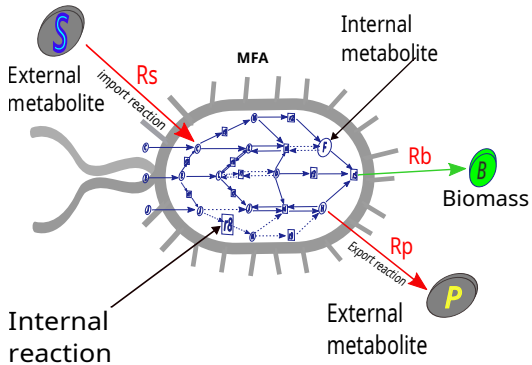


$N_b \approx 1, N_m \approx 3$

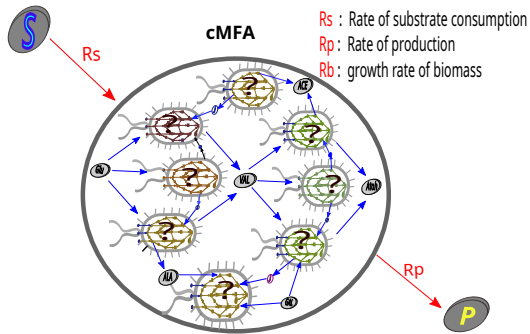
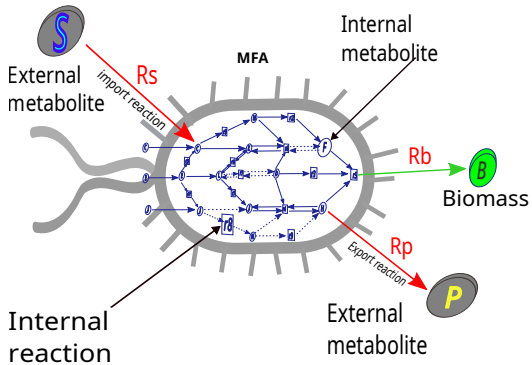
$N_r \approx 2500$ High-dimensional

$N_r \gg N_m$ Highly-underdetermined

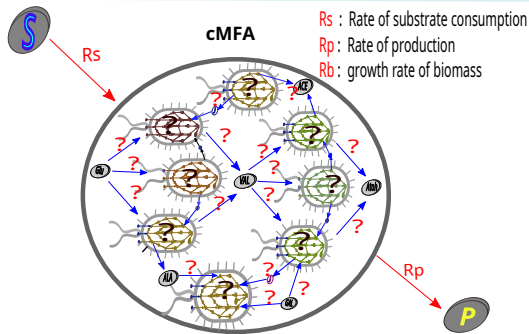
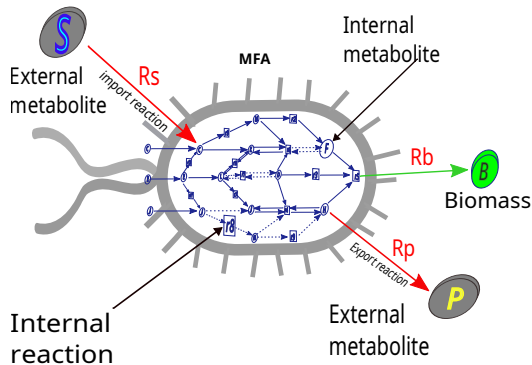
cMFA for community models



cMFA for community models



cMFA for community models

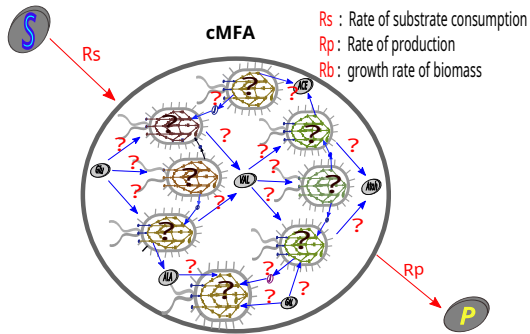
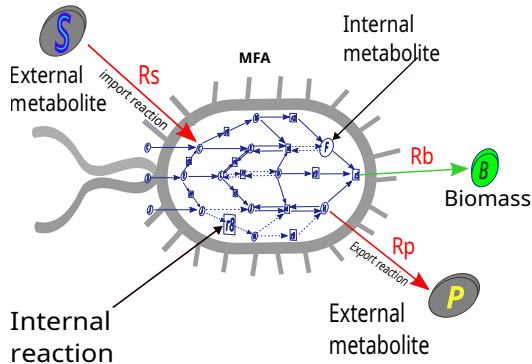


$$N_m \approx 3 * N_b$$

$$N_r \approx 2500 * N_b \text{ High-dimensional}$$

$$N_r \gg N_m \text{ Highly-underdetermined}$$

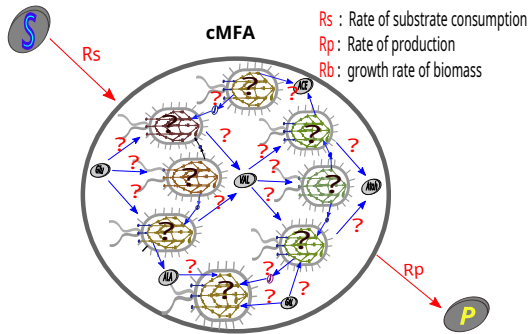
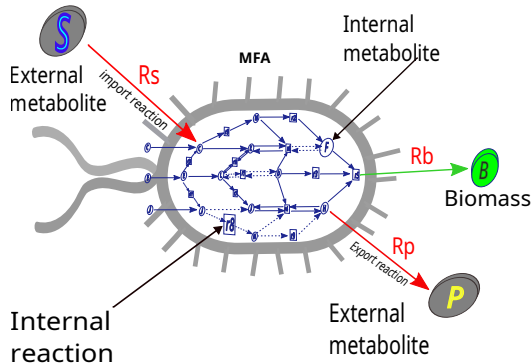
cMFA for community models



$$\hat{\nu}(t) := \arg \min_{m \in \mathcal{M}} \sum_{m \in \mathcal{M}} \left\| \overbrace{R_{m,t}}^{\text{Data rate}} - \sum_{b \in \mathcal{B}_m} \overbrace{\chi_t^{(b)} \nu_{m,t}^{(b)}}^{\text{Model}} \right\|_2^2 + \underbrace{\lambda \mathcal{R}(\nu)}_{\text{Regularization term}}$$

Set of (Metabolites & Micro organisms)

cMFA for community models



R_s : Rate of substrate consumption
 R_p : Rate of production
 R_b : growth rate of biomass

$$\hat{\nu}(t) := \arg \min_{m \in \mathcal{M}} \sum_{m \in \mathcal{M}} \left\| R_{m,t} - \sum_{b \in \mathcal{B}_m} \overset{\text{Biomass}}{\chi}_t^{(b)} \nu_{m,t}^{(b)} \right\|_2^2 + \underbrace{\lambda \mathcal{R}(\nu)}_{\text{Regularization term}}$$

\mathcal{M} : Set of (Metabolites & Micro organisms)
 \mathcal{B}_m : Biomass

biological Information

stochiometry matrix

$$\begin{cases} S^{(b)} \cdot \nu^{(b)} = 0, \\ C_{min}^{(b)} \leq \nu^{(b)} \leq C_{max}^{(b)}. \end{cases}$$

Fluxes

Meta-transcriptomic data inclusion

$$\hat{\nu}(t) := \arg \min_{\nu_t} \sum_{m \in \mathcal{M}} \|R_{m,t} - \sum_{b \in \mathcal{B}_m} \chi_t^{(b)} \nu_{m,t}^{(b)}\|_2^2 + \lambda \mathcal{R}(\mathcal{W}\nu) \quad (1)$$

subject to : $\begin{cases} S^{(b)} \cdot \nu^{(b)} = 0 \\ C_{min}^{(b)} \leq \nu^{(b)} \leq C_{max}^{(b)}, \end{cases}$ for $b = 1, \dots, N_b$

Inclusion by Penalization

e.g : $C_t^{(b)} \mapsto f(C_r^b(t)) := (\max(C_r^b) + \min(C_r^b) - C_r^b(t)) = \mathcal{W}_t^{(b)}$

with $C_r^b = \left| \frac{T_r^b}{\max T_r^b} \right|$ $r \in \mathcal{R}_b$ following^a

$\mathcal{W}_t^{(b)}(r)$: the weight associated with reaction r , which is small if the reaction is activated and high if it is deactivated.

f : is the mapping function : transcript of reaction r to weight

\mathbf{T} : transcript level

a. Jenior, M.L. et al. (2020). *PLoS Comput. Biol.*, **16**, e1007099.

Benchmark for Robustness of the method : Different steps

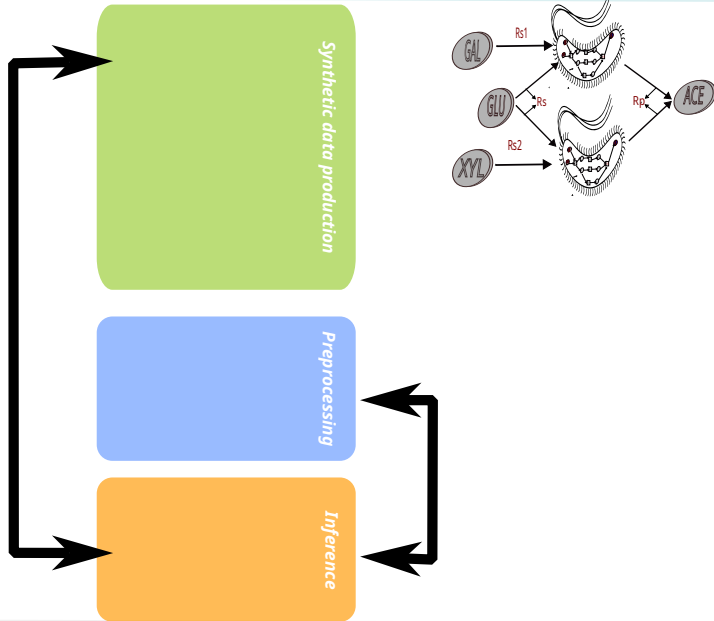
▣ Simulated data.

1. Explore different types of interactions.
2. Benchmark noise levels
3. Test different community sizes (2–56)

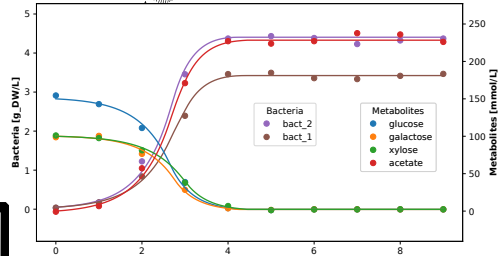
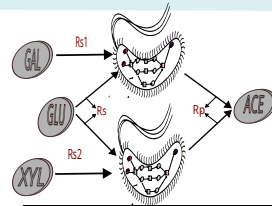
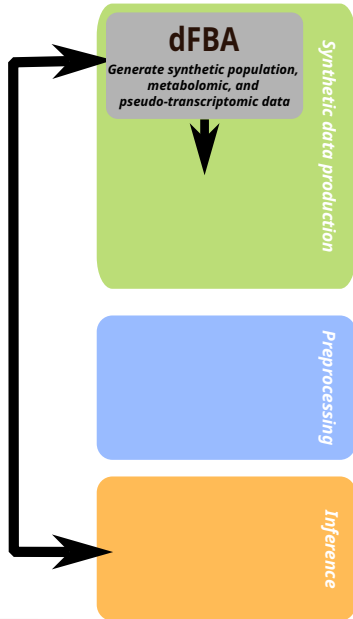
▣ Real data.

1. TANGO data

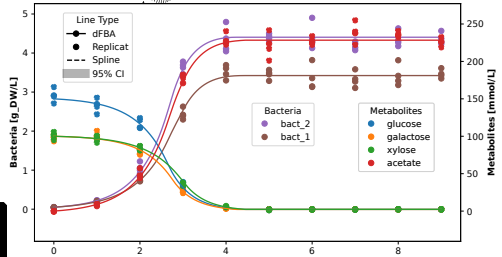
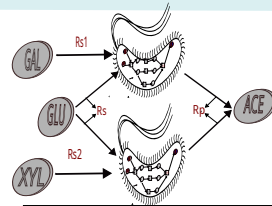
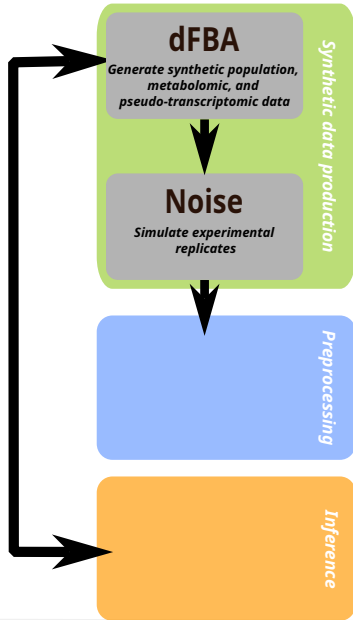
Benchmark method



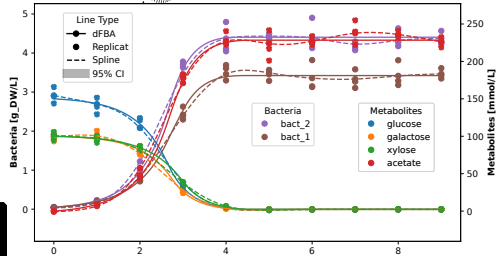
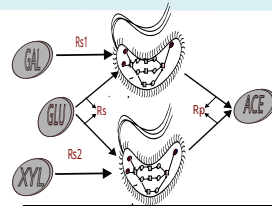
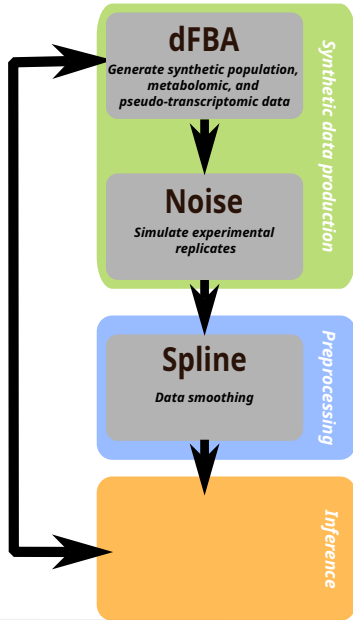
Benchmark method



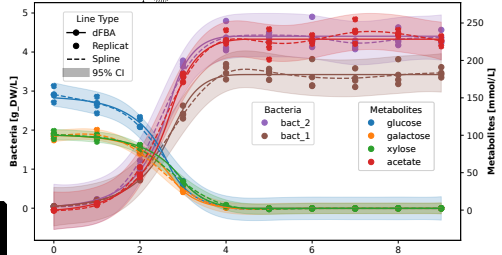
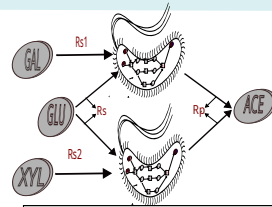
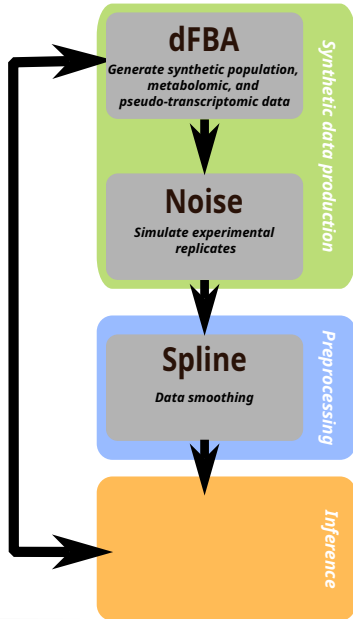
Benchmark method



Benchmark method

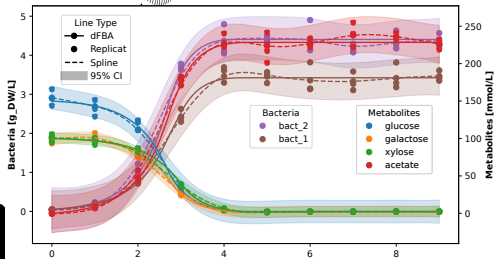
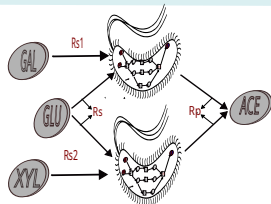
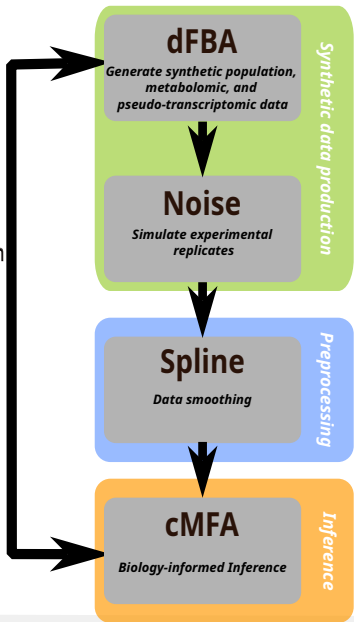


Benchmark method



Benchmark method

Comparison between intracellular rates in dFBA and cMFA

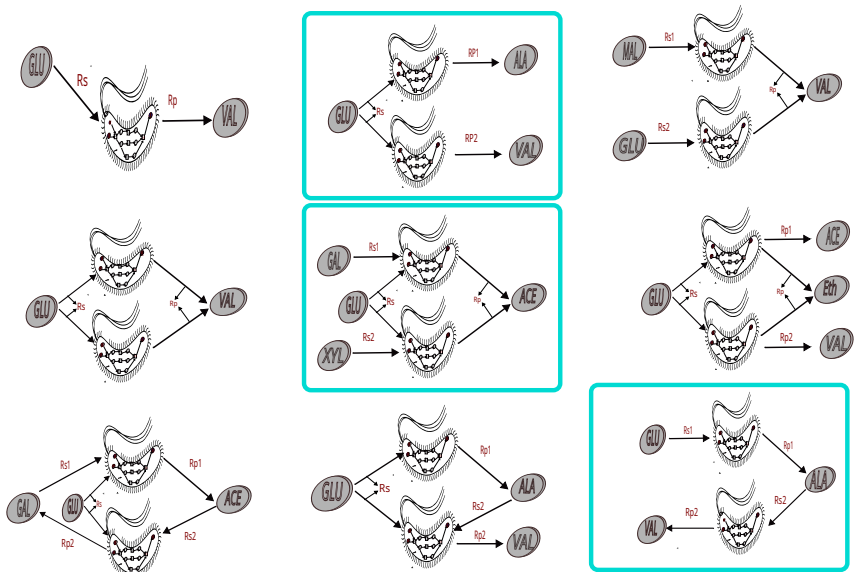


Comparison between extracellular rates in spline and cMFA



Simulated data

1. Different Interactions



1. Three Biological Cases of Two-Species Communities

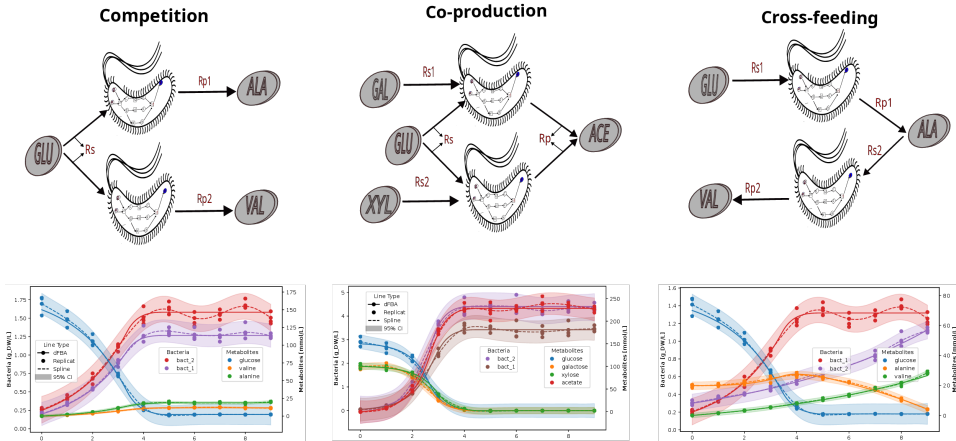
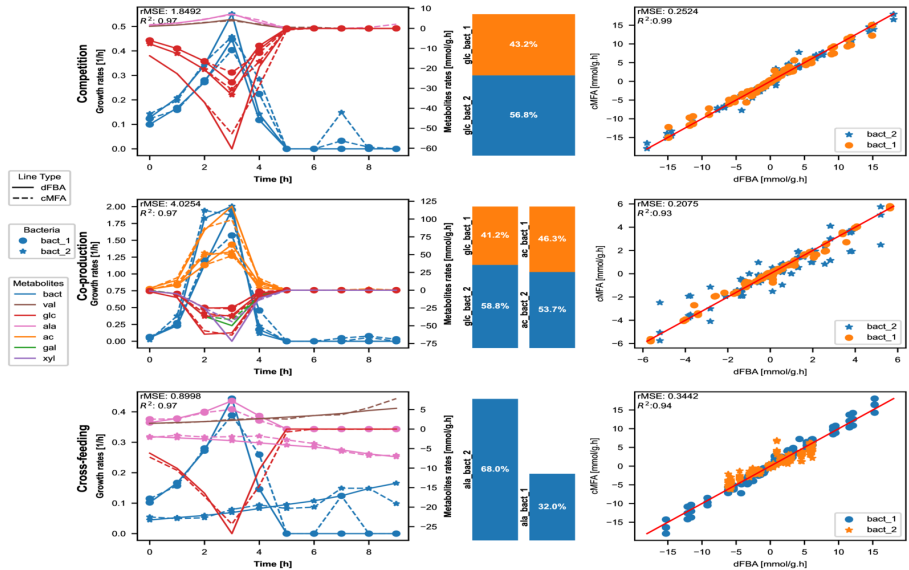


Figure – Synthetic data for cMFA benchmark. the different cases of increasingly complex ecological interactions are sketched



1. Three Biological Cases of Two-Species Communities



2. Data noise for replication

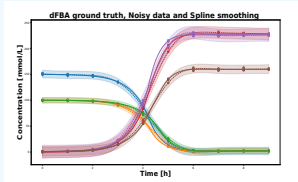
Noise percent in data

$$D_{\text{final}}(t) = \max(0, D_{\text{dFBA}}(t) \times (1 + \epsilon(t))), \text{ with } \epsilon(t) \sim \mathcal{N}(0, \sigma^2)$$

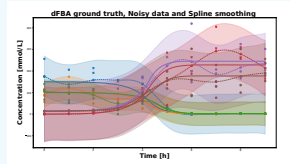
$\sigma = q\sigma_b$, where $\sigma_b = 1$

$q \in \{[1\%, 5\%, 10\%, 20\%, 30\%, 40\%, 50\%, 60\%, 70\%, 80\%, 90\%, 100\%]\}$ are used for the data noise benchmark

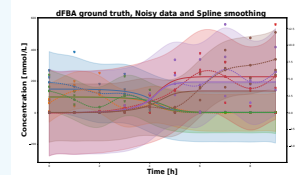
For each noise level, we replicated the whole benchmark pipeline (i.e. data noising, pre-processing with spline-smoothing and cMFA inference) $n = 7$ times,



1%

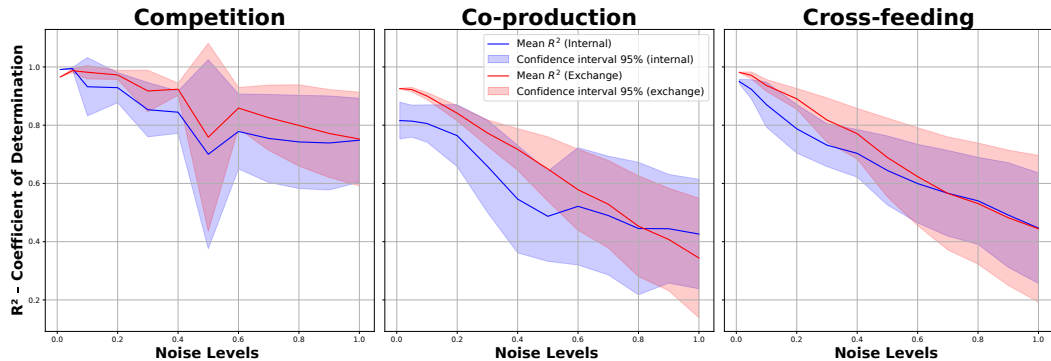


50%

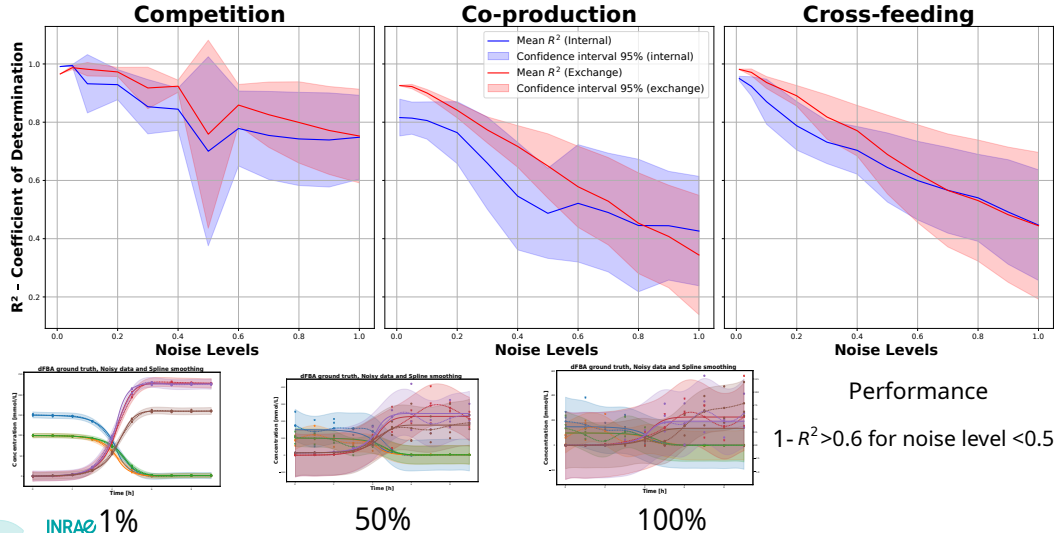


100%

2. Data noise for replication



2. Data noise for replication



3. Large Community Benchmark

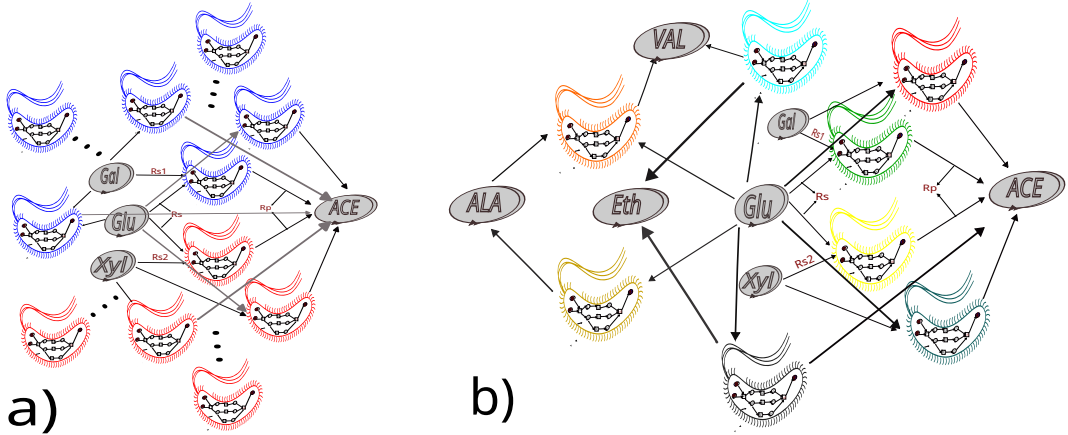


Figure – Large size community of : a) Variance Same strain, b) Variance Different strain

3. Large Community Benchmark

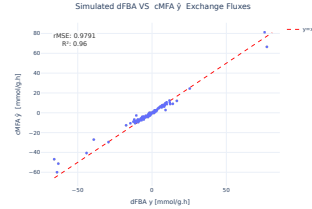
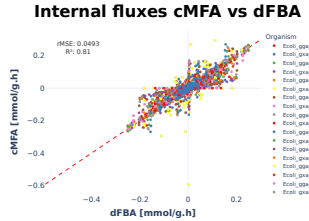
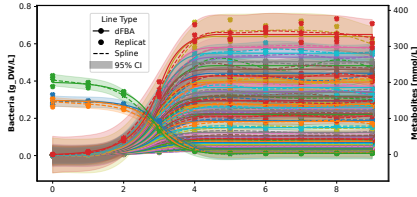


Figure – Variance within same strain group : 56 organisms (2 guild of 28).

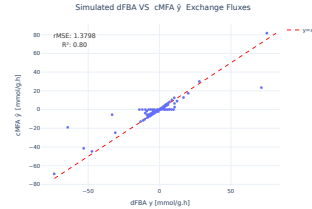
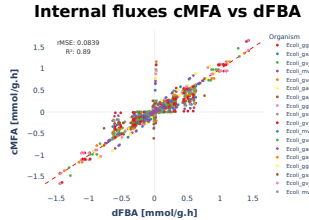
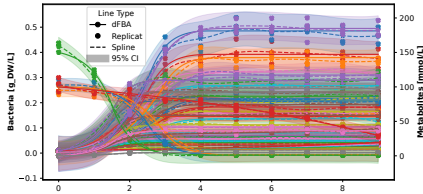


Figure – Variance across different strain guild : 49 organisms (7 guild of 7).
25e Forum des jeunes mathématiciennes et mathématiciens 26-28 nov. 2025 Talence



Real data

Community of three microorganisms

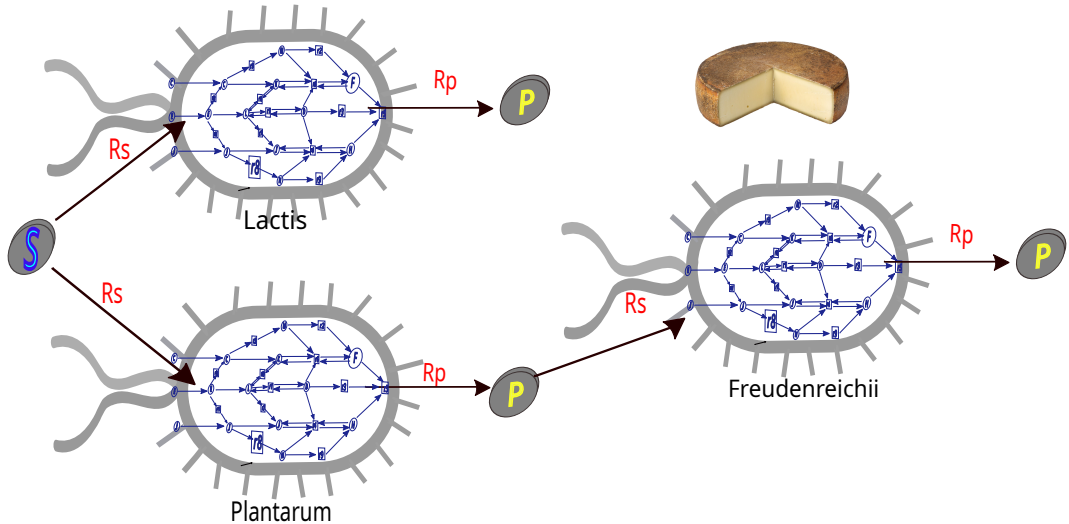
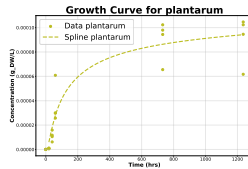
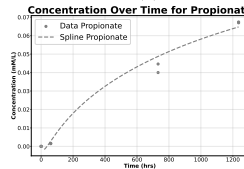
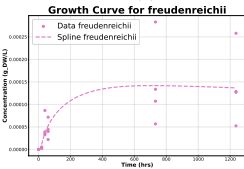
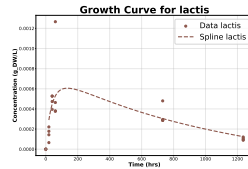
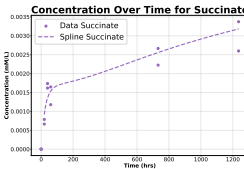
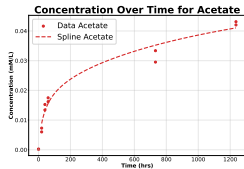
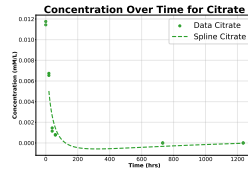
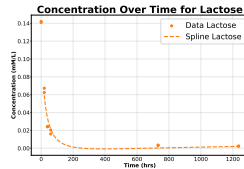
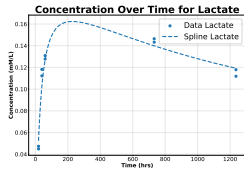
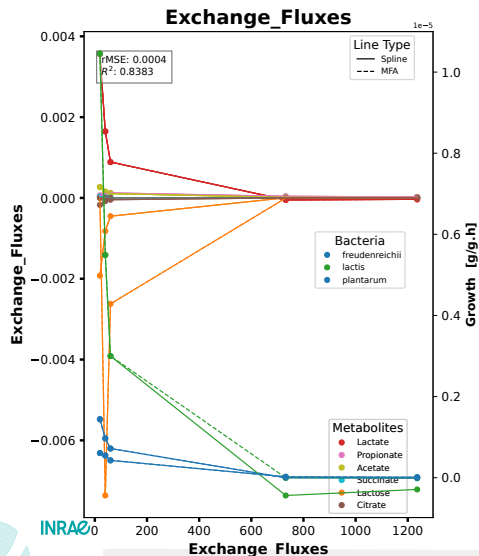


Figure 1 Community of three microorganisms involved in cheese fermentation and maturation

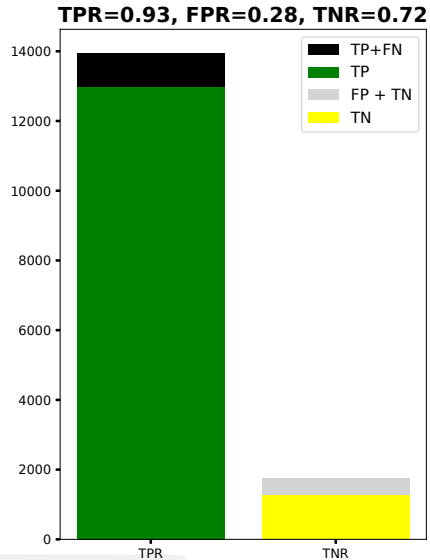
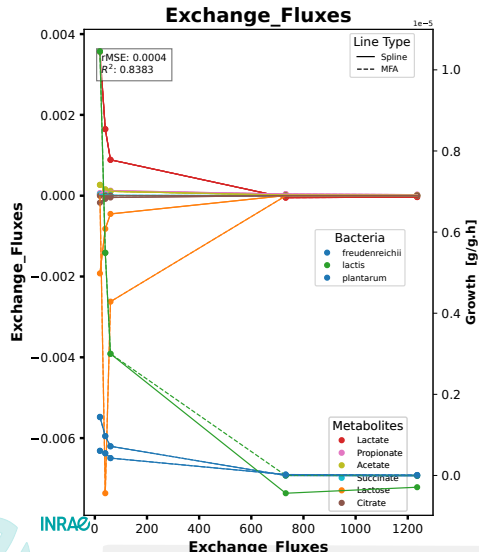
Smoothing spline of metabolites and organisms data



Results of inference



Results of inference



Conclusion

We present a Statistical inference method informed by biology : *cMFA*

Work Completed

- ➡ Explore different interaction types
- ➡ Benchmark noise levels
- ➡ Handle incomplete metatranscriptomic data
- ➡ Vary metabolic import rates
- ➡ Test different community sizes (2–56)
- ➡ Missing metabolites

Advantages

- ➡ Robust to noise
- ➡ Robust to variations in metabolic import rates
- ➡ Robust for exchange rate inference with both high- and low-quality metatranscriptomic data.
- ➡ Scalable to large communities

Disadvantages

- ➡ Internal flux inference is less robust when metatranscriptomic data is of lower quality.










Application on real data

- ➡ Synthetic communities produced by colleagues at the Institut de Recherche en Horticulture et Semences (IRHS) in Angers
- ➡ 48 microorganisms of Vitae Project

Software deployment

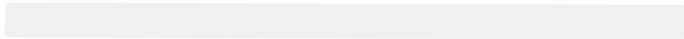
- ➡ Facilitation of utilisation by making on package

References

-  Brunner et al. (2023). *PLoS Comput. Biol.*
-  Gerken Starepravo, M. et al. (2022). *Digital Chemical Engineering*, **2**.
-  Lecomte, A. et al. (2023). *Metabolic Engineering*.
-  Mahadevan, R. et al. (2002). *Biophysical Journal*, **83**.
-  Pansu, M. (2004). *Courbes B-splines*.
-  Pande, A. et al. (2013). *The ISME Journal*, **8**(5).
-  Jenior, M.L. et al. (2020). *PLoS Comput. Biol.*, **16**, e1007099. .



INRAE



Preprocessing : improving solver convergence

Vectorization Renormalization and Dimensionless

$$\sigma_{1_t}^{-1} = \frac{1}{R_t}, \quad \sigma_{2_t} = \frac{R_t}{B_t}$$

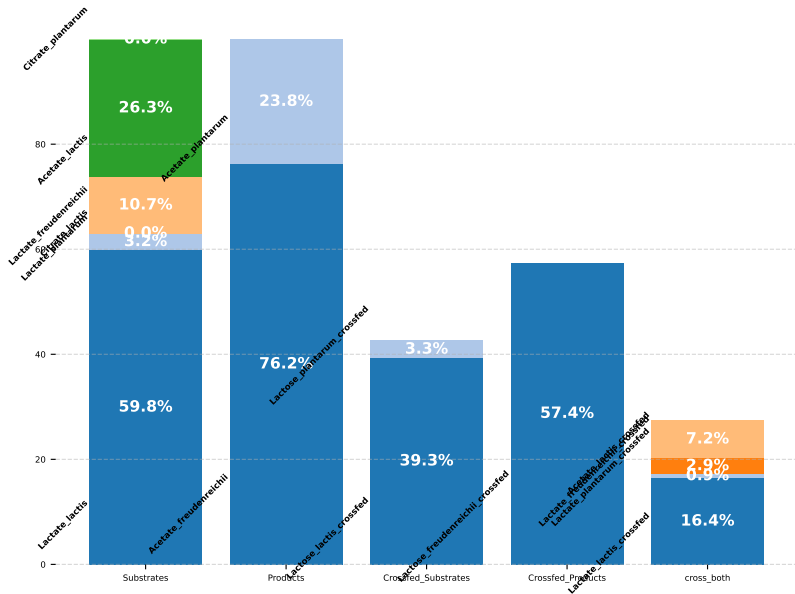
so that after renormalization, the dimensionless problem becomes

$$\hat{\nu}_t = \arg \min_{\tilde{\nu}_t} \|\tilde{R}_t - \tilde{B}_t \tilde{\nu}_t\|_2^2 + \lambda \|w_t \tilde{\nu}_t\|_1$$

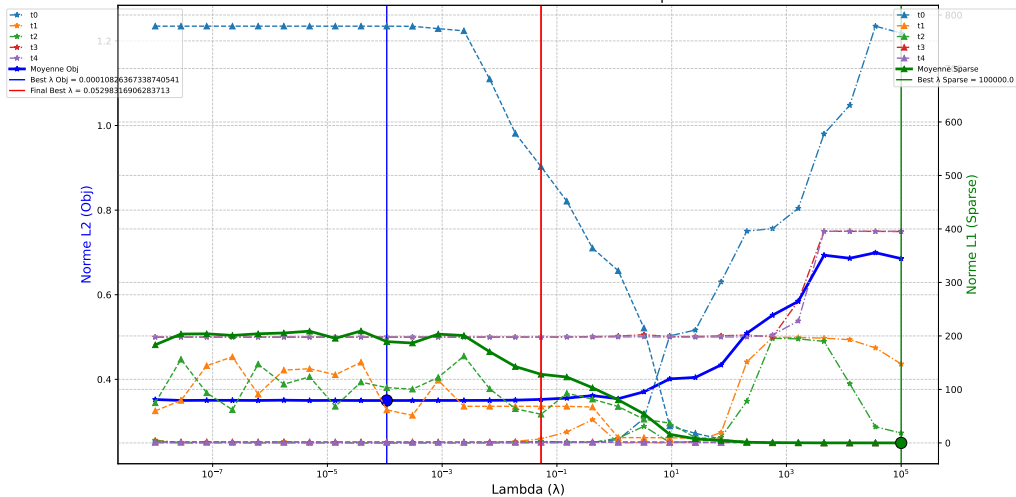
subject to :
$$\begin{cases} \tilde{S} \tilde{\nu}_t = 0 \\ \tilde{\nu}_{\min} \leq \tilde{\nu}_t \leq \tilde{\nu}_{\max} \end{cases} \quad (2)$$

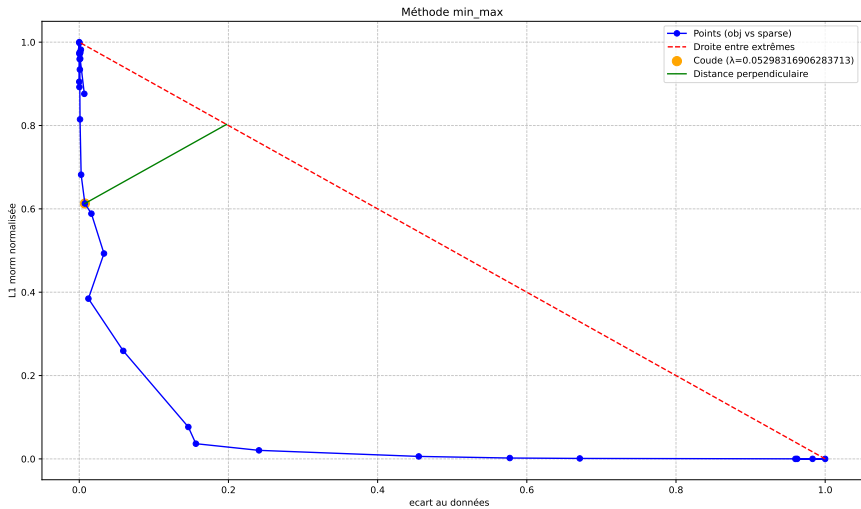
$$\tilde{R}_t = \sigma_{1_t}^{-1} R_t, \quad \tilde{B}_t = \sigma_{1_t}^{-1} B_t \sigma_{2_t}, \quad \tilde{\nu}_t = \sigma_{2_t}^{-1} \nu_t$$
$$\tilde{S} = S \sigma_{2_t}, \quad \tilde{\nu}_{\min} = \sigma_{2_t}^{-1} \nu_{\min}, \quad \tilde{\nu}_{\max} = \sigma_{2_t}^{-1} \nu_{\max}, \quad B_t \nu_t = \sum_{b \in \mathcal{B}_m} \chi_t \nu_t$$

This renormalization harmonizes the scales reactions fluxes (exchange & internal)

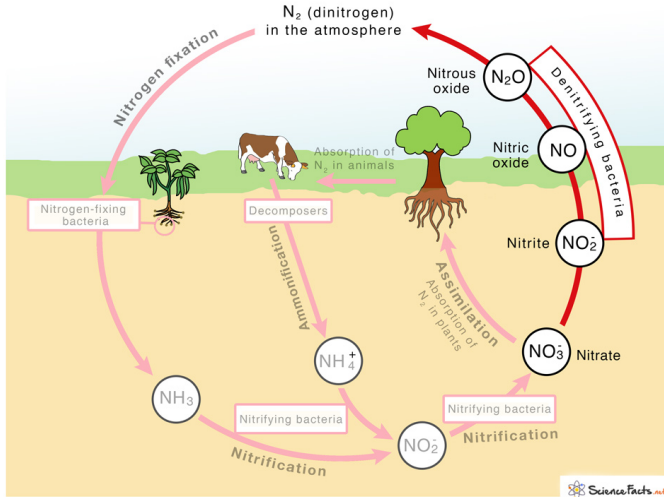


Grid Search : distribution des normes L2 par λ

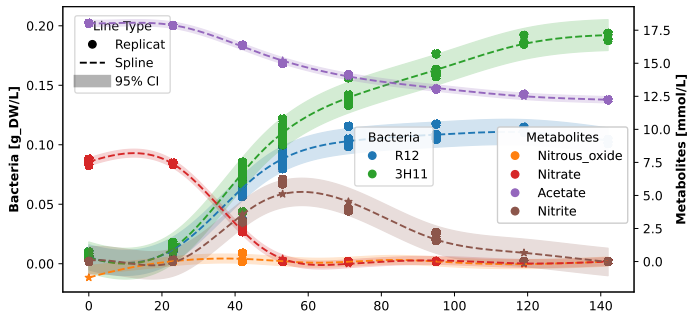




Denitrification



Concentrations



3. Incomplete meta-transcriptomic data

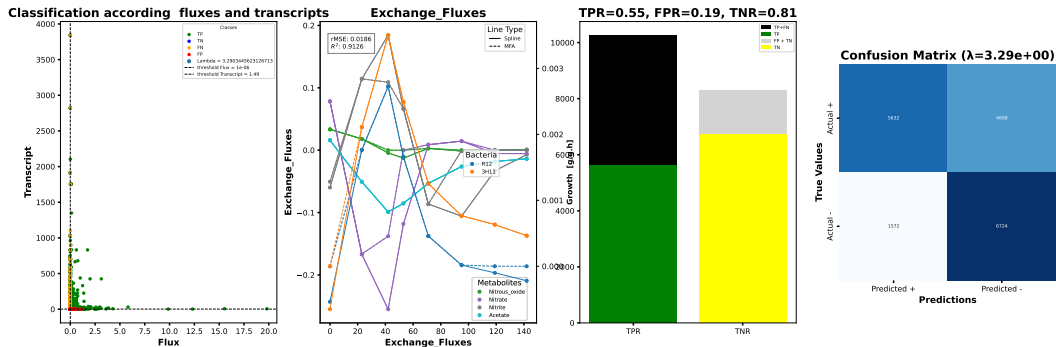
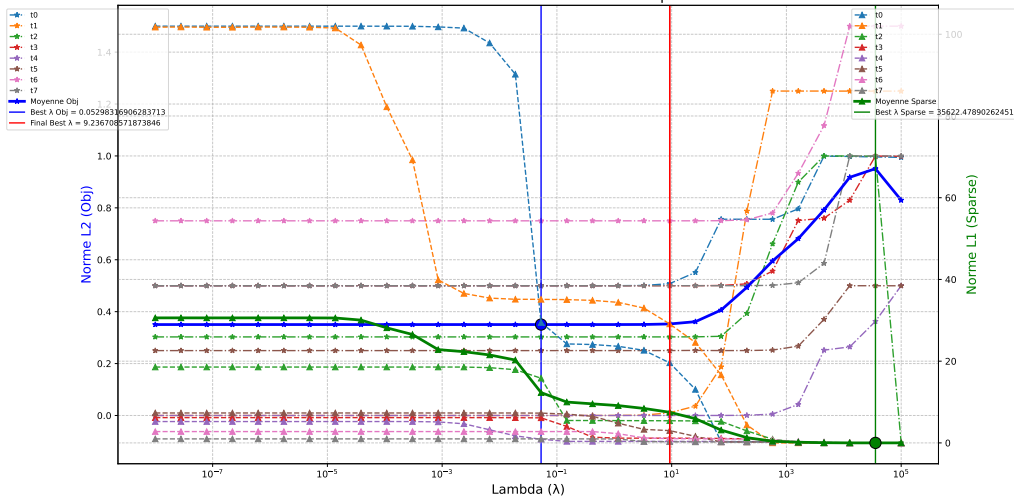
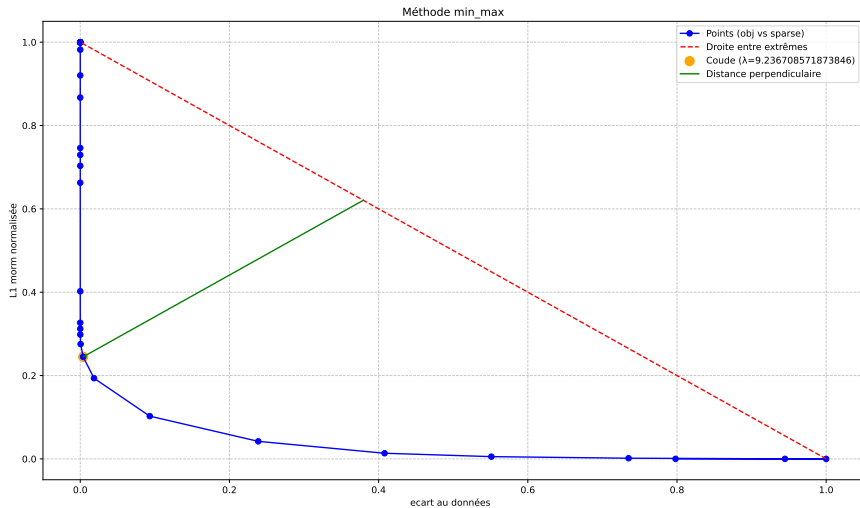


Figure – Incomplete meta-transcriptomic information

Grid Search : distribution des normes L2 par λ





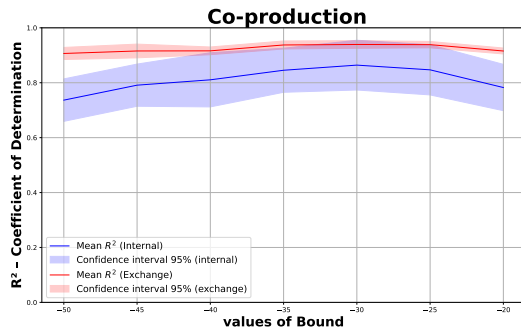
Benchmark for Robustness of the method : metabolic import rates

Intrinsic Flux

$$C_{min}^{(b)} = \max \left(\nu_{intr}, -\frac{m}{\Delta_t \cdot \sum_{b \in \mathcal{B}_m} \chi_b} \right)$$

$$\nu_{intr} \in [-50, -45, -40, -35, -30, -25, -20]$$

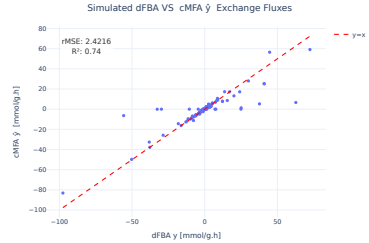
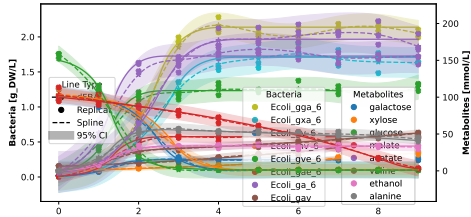
reference value $\nu_{intr} = -30$



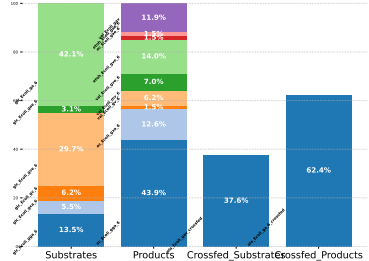
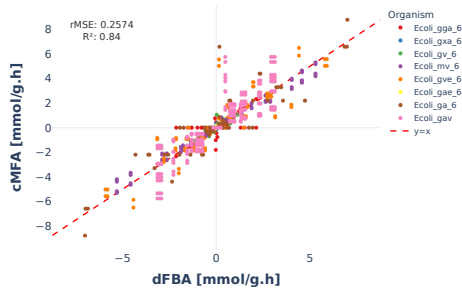
Performances

- ➡ Good estimation of exchange fluxes
- ➡ Good estimation of internal fluxes

Large Community : Different strain



Internal fluxes cMFA vs dFBA



Meta-transcriptomic data inclusion

$$\hat{\nu}(t) := \arg \min_{\nu_t} \sum_{m \in \mathcal{M}} \|R_{m,t} - \sum_{b \in \mathcal{B}_m} \mathcal{X}_t^{(b)} \nu_{m,t}^{(b)}\|_2^2 + \lambda \mathcal{R}(\nu)$$

(3)

$$\text{subject to : } \begin{cases} S^{(b)} \cdot \nu^{(b)} = 0 \\ C_{min}^{(b)} \leq \nu^{(b)} \leq C_{max}^{(b)}, \end{cases} \quad \text{for } b = 1, \dots, N_b$$

Constraint

$$\mathcal{T}_t^{(b)}(r) = \mathbb{1}_{\mathcal{A}_t^{(b)}(r)}$$

$$C_{min}^{(b)} := \bar{C}_{min}^{(b)} \mathcal{T}_t^{(b)} \leq \nu_t^{(b)} \leq C_{max}^{(b)} := \bar{C}_{max}^{(b)} \mathcal{T}_t^{(b)}$$

$\mathcal{A}_t^{(b)}$: Activate of bacteria b at time t.

$\bar{C}_{min}^{(b)}$ and $\bar{C}_{max}^{(b)}$ are nominal constraints for internal reactions.

$\mathcal{T}_t^{(b)}$ is a boolean vector indicating if the expression of the corresponding reaction

Penalization

$$\mathcal{T}_t^{(b)}(r) := f(\mathbf{T}_t^{(b)}(r))$$

$$\mathcal{R}(\nu) := \mathcal{R}(\mathcal{T}\nu)$$

$\mathcal{T}_t^{(b)}(r)$: the linear coefficient of ponderation associated with reaction r , which is small if the reaction is activated and high if it is deactivated.

f : is the mapping function : transcript of reaction r to linear coefficient

\mathbf{T} : transcript level

Synthetic Metatranscriptomic Data Generation

Constraint

$$\mathcal{T}_t^{(b)}(r) = \mathbb{1}_{\mathcal{A}_t^{(b)}}(r) \text{ where } \mathcal{A}_t^{(b)} = \{r \in \mathbf{R} \mid |\nu_r^b(t)| > 10^{-15}\}$$

$\nu_r^b(t) :=$ value of the reaction flux r for bacteria b

Penalization

Let M be mappings from flux to transcript and C mapping of transcript to a coefficient.

$\nu_r^b(t)$ fluxes values in the model in response to gene expression levels.

$$\mathcal{T}_t^{(b)}(r) := f(\nu_r^b(t))$$

where f is defined as the composition of two functions, C , M .

$$f = C \circ M \text{ where } C, M : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

$$T_r := M(\nu_r^b) = \frac{1}{1 + e^{-r|\nu_r^b|}} - \frac{1}{2}$$

$$C = (\max(C_s) - \min(C_s) - C_s)$$

$$C_s = \frac{T_r}{\max T_b} \quad r \in \mathcal{R}_b$$

T_r : transcript level



Preprocessing and Rescaling

Penalization + Transcriptomic data

$$\hat{\nu}(t) := \arg \min_{\nu_t} \sum_{m \in \mathcal{M}} \|R_{m,t} - \sum_{b \in \mathcal{B}_m} \mathcal{X}_t^{(b)} \nu_{m,t}^{(b)}\|_2^2 + \lambda \mathcal{R}(\mathcal{T}\nu) \quad (4)$$

subject to : $\begin{cases} \mathcal{S}^{(b)} \cdot \nu^{(b)} = 0 \\ C_{min}^{(b)} \leq \nu^{(b)} \leq C_{max}^{(b)}, \end{cases}$ for $b = 1, \dots, N_b$

Preprocessing and Rescaling

Penalization + Transcriptomic data

$$\hat{\nu}(t) := \arg \min_{\nu_t} \sum_{m \in \mathcal{M}} \|R_{m,t} - \sum_{b \in \mathcal{B}_m} \mathcal{X}_t^{(b)} \nu_{m,t}^{(b)}\|_2^2 + \lambda \mathcal{R}(\mathcal{T}\nu)$$

(4)

$$\text{subject to : } \begin{cases} \mathcal{S}^{(b)} \cdot \nu^{(b)} & = 0 \\ C_{\min}^{(b)} \leq \nu^{(b)} \leq C_{\max}^{(b)}, \end{cases} \quad \text{for } b = 1, \dots, N_b$$

Vectorization

$$\forall t \geq 0, \quad \hat{\nu}(t) := \arg \min_{\nu_t} \|R_t - B_t \nu_t\|_2^2 + \lambda \mathcal{R}(\mathcal{T}\nu_t)$$

(5)

$$\text{subject to : } \begin{cases} \mathcal{S} \cdot \nu_t & = 0 \\ \nu_{\min} \leq \nu_t \leq \nu_{\max} \end{cases}$$

with

$$\nu_{\min} = C_{\min} \mathcal{T}_t, \quad \nu_{\max} = C_{\max} \mathcal{T}_t, \quad B_t \nu_t = \sum_{b \in \mathcal{B}_m} \mathcal{X}_t^{(b)} \nu_t^{(b)}$$

Penalization

Let B be adding bias function in generated flux where f is defined as the composition of three functions, C, M, B .

$f = C \circ M \circ B$ where $C, M : \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$T_r := M(\nu_r^b) = \frac{1}{1 + e^{-r|\nu_r^b|}} - \frac{1}{2}$$

$$C = (\max(C_s) - \min(C_s) - C_s)$$

$$C_s = \frac{T_r}{\max T_b} \quad r \in \mathcal{R}_b$$

T_r : transcript level



Ryan J. Tibshirani, Carnegie Mellon University : The Lasso Problem and Uniqueness

B-splines are piecewise polynomials defined as linear combinations of spline basis elements. These spline base elements are themselves linear combinations of the usual monomials $1, x, x^2, \dots, x^k$. Let $S : [a, b] \rightarrow \mathbb{R}$ a spline,
 $a = t_0 \leq t_1 \dots \leq t_{j-1} \leq t_j \dots \leq t_n = b$

$$S_k(x) = \sum_{j=0}^{n-1} B_{j,k,t}(x) \cdot P_j, \quad x \in [t_j, t_{j+1}]$$

P_j form a polygon called the control polygon,

$B_{j,k,t}$ are B-spline basis functions of degree k and nodes $t = (t_0, \dots, t_n)$.

$$B_{j,0}(x) := \begin{cases} 1 & \text{si } t_j \leq x < t_{j+1} \\ 0 & \text{else} \end{cases}$$

Moving a control point only locally modifies the shape of the curve.

$$B_{j,k}(x) := \frac{x - t_j}{t_{j+k} - t_j} B_{j,k-1}(x) + \frac{t_{j+k+1} - x}{t_{j+k+1} - t_{j+1}} B_{j+1,k-1}(x)$$

Conclusion

We present a Statistical inference method informed by biology : *cMFA*

Implementation

- Python
- Cobrapy
- PYGAM
- OSQP

Perspectives

- Incomplete information in Network
- Real data

Context : How Understanding microbial community functioning?

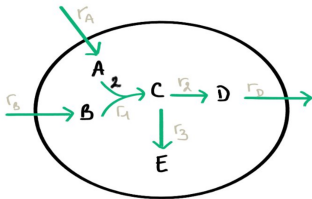
Context :

- ▶ Developing a mathematical inference method based on metabolic networks.
- ▶ Inferring metabolic fluxes by integrating multi-omics data.

Context : How Understanding microbial community functioning?

Context :

- ▶▶▶ Developing a mathematical inference method based on metabolic networks.
- ▶▶▶ Inferring metabolic fluxes by integrating multi-omics data.



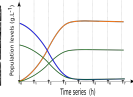
$$S = \begin{matrix} & r_A & r_B & r_1 & r_2 & r_3 & r_D \\ \begin{matrix} A \\ B \\ C \\ D \\ E \end{matrix} & \begin{pmatrix} 1 & 0 & -2 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

$$Sv = 0$$

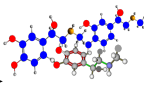
Context : How Understanding microbial community functioning?

Context :

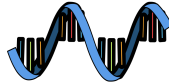
- Developing a mathematical inference method based on metabolic networks.
- Inferring metabolic fluxes by integrating multi-omics data.



Population count data

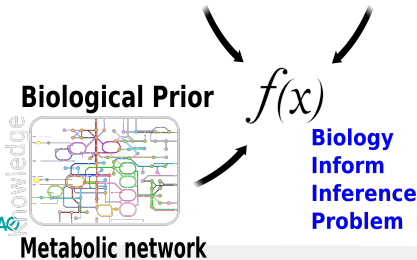
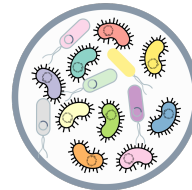


Metabolomics data



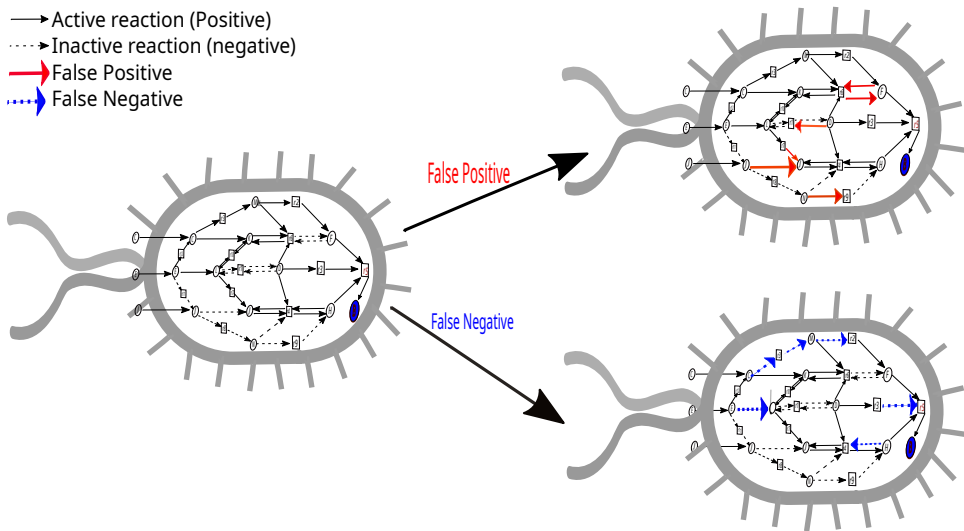
Metatranscriptomic data

Microbial Community

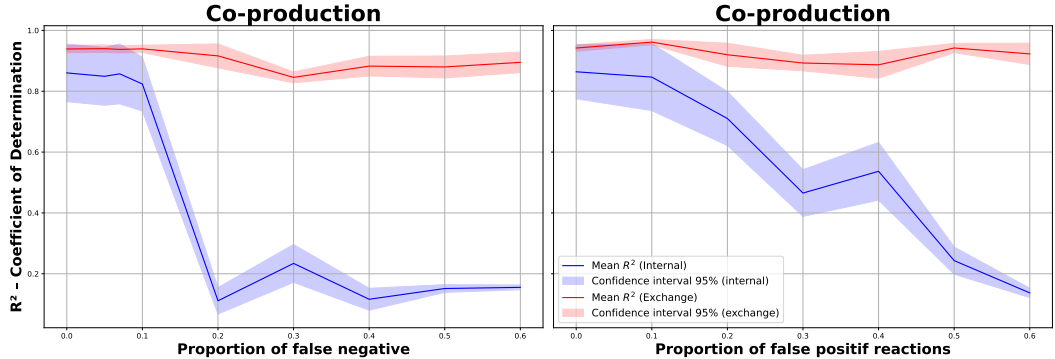


3. Incomplete meta-transcriptomic data

- ▶ Active reaction (Positive)
- ⋯▶ Inactive reaction (negative)
- ▶ False Positive
- ▶ False Negative



3. Incomplete meta-transcriptomic data



Performances

- ➡ Good estimation of exchange fluxes
- ➡ Good estimation of internal fluxes when proportion < 10 %

INRAE

Difficulty

- ➡ Sensitivity to the quality of metatranscriptomic data



Context : Complexity of natural microbial communities

Growth Stimulation (Phytohormone Production)

Fertilisation (nitrogen fixation)

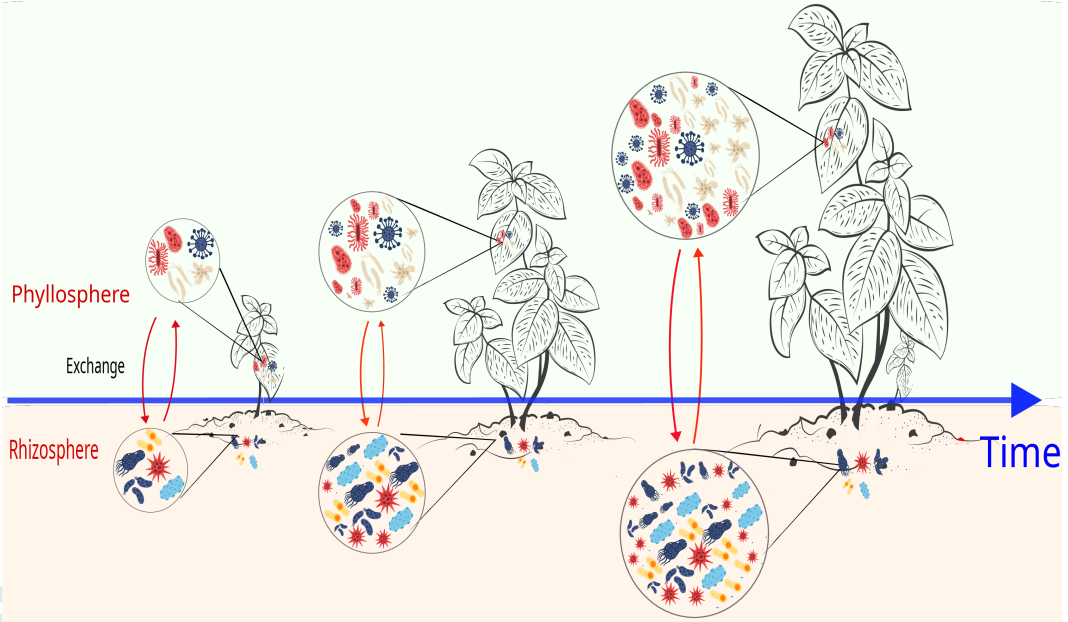
Protection (Pathogen resistance)



Context : Complexity of natural microbial communities

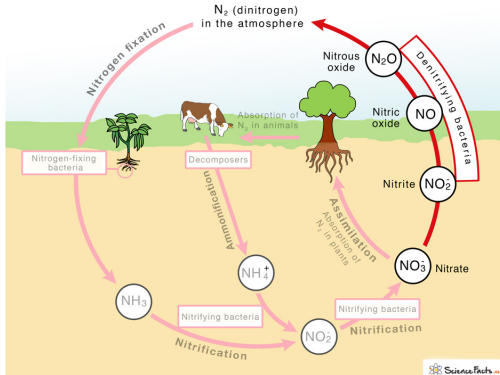


Context : Complexity of natural microbial communities



Context : Complexity of natural microbial communities

Denitrification







Method : Consumer-resource , Constraint Based


Consumer-resource

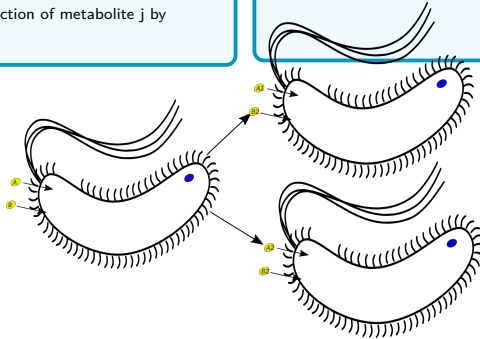
$$\partial_t b_i = \nu_i(M, b) b_i$$

$$\partial_t M_j = \sum_i \nu_{i,j}(M, b) b_i$$

-  b_i Concentration of microorganisms i
-  M_j Concentration of metabolites j
-  ν_i Growth rate of the microorganism i
-  $\nu_{i,j}$ Rate of consumption or production of metabolite j by microorganism i

Advantages and disadvantages

-  Integrates population count and metabolomic data







Method : Consumer-resource , Constraint Based



Consumer-resource

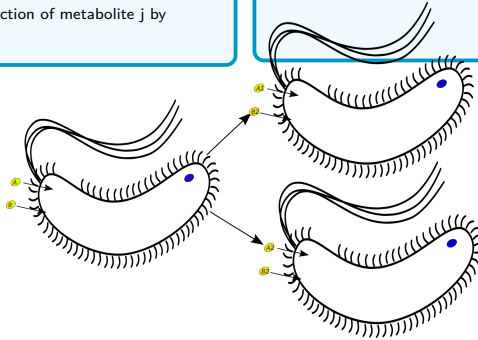
$$\partial_t b_i = \nu_i(M, b) b_i$$

$$\partial_t M_j = \sum_i \nu_{i,j}(M, b) b_i$$

-  b_i Concentration of microorganisms i
-  M_j Concentration of metabolites j
-  ν_i Growth rate of the microorganism i
-  $\nu_{i,j}$ Rate of consumption or production of metabolite j by microorganism i

Advantages and disadvantages

-  Integrates population count and metabolomic data
-  Low calculation costs







Method : Consumer-resource , Constraint Based





Consumer-resource

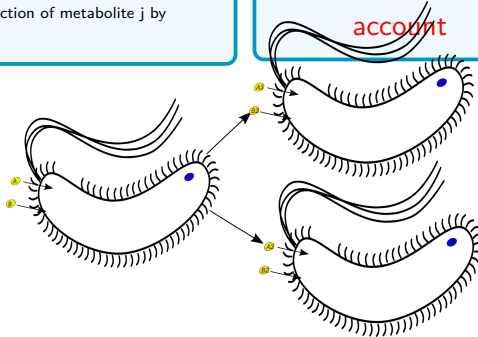
$$\partial_t b_i = \nu_i(M, b) b_i$$

$$\partial_t M_j = \sum_i \nu_{i,j}(M, b) b_i$$

-  b_i Concentration of microorganisms i
-  M_j Concentration of metabolites j
-  ν_i Growth rate of the microorganism i
-  $\nu_{i,j}$ Rate of consumption or production of metabolite j by microorganism i

Advantages and disadvantages

-  Integrates population count and metabolomic data
-  Low calculation costs
-   Does not take metatranscriptomic data into account



Dynamic Flux Balance Analysis

dFBA

$$\partial_t b_i = B \hat{v}_i(M, b) b_i$$

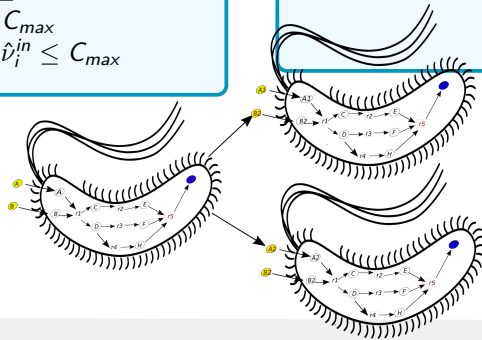
$$\partial_t M_j = \sum_{i \in \mathcal{C}} \hat{v}_{i,j}(M, b) b_i$$

$$\hat{v}_i = \arg \max B v_i$$

$$\text{s.c.} \left\{ \begin{array}{l} S_i \hat{v}_i = 0 \\ C_M \leq \hat{v}_i^s \leq 0 \\ 0 \leq \hat{v}_i^p \leq C_{\max} \\ -C_{\max} \leq \hat{v}_i^{\text{in}} \leq C_{\max} \end{array} \right.$$

Advantages and disadvantages

- Integrates population count, metabolomic data and metatranscriptomic



Dynamic Flux Balance Analysis

dFBA

$$\partial_t b_i = B \hat{\nu}_i(M, b) b_i$$

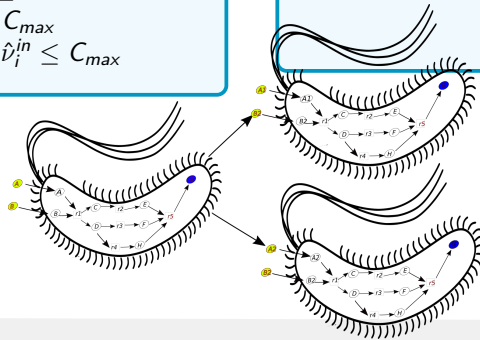
$$\partial_t M_j = \sum_{i \in \mathcal{C}} \hat{\nu}_{i,j}(M, b) b_i$$

$$\hat{\nu}_i = \arg \max B \nu_i$$

$$\text{s.c.} \left\{ \begin{array}{l} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{\max} \\ -C_{\max} \leq \hat{\nu}_i^{\text{in}} \leq C_{\max} \end{array} \right.$$

Advantages and disadvantages

- Integrates population count, metabolomic data and metatranscriptomic
- ⚠ Very high computing resources**



Dynamic Flux Balance Analysis

dFBA

$$\partial_t b_i = B \hat{\nu}_i(M, b) b_i$$

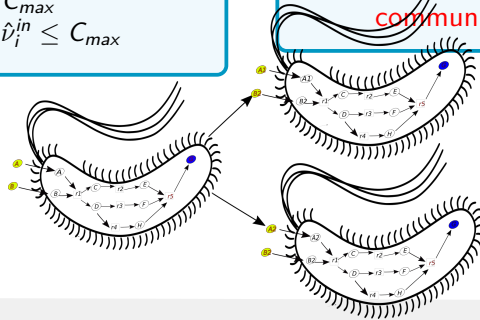
$$\partial_t M_j = \sum_{i \in \mathcal{C}} \hat{\nu}_{i,j}(M, b) b_i$$

$$\hat{\nu}_i = \arg \max B \nu_i$$

$$\text{s.t.} \left\{ \begin{array}{l} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{max} \\ -C_{max} \leq \hat{\nu}_i^{in} \leq C_{max} \end{array} \right.$$

Advantages and disadvantages

- Integrates population count, metabolomic data and metatranscriptomic
- ⚠ Very high computing resources
- ⚠ Does not adapt to large communities



Methods : Metabolic Flux Analysis


MFA

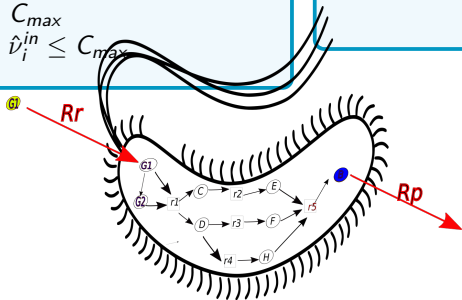
Find $\hat{\nu} = (\nu^{\hat{1}}, \dots, \nu^{\hat{N}_t})$ such that

$$\hat{\nu} = \arg \min \frac{1}{N_t} \sum_{t=1}^{N_t} \|R_t - \mathcal{X}_t \nu_t\|^2 + \lambda \mathcal{R}(\nu)$$

$$\text{s.c.} \left\{ \begin{array}{l} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{max} \\ -C_{max} \leq \hat{\nu}_i^{in} \leq C_{max} \end{array} \right.$$

Advantages and Limitations

-  Precisely quantifies metabolite fluxes in a metabolic network



Methods : Metabolic Flux Analysis




MFA

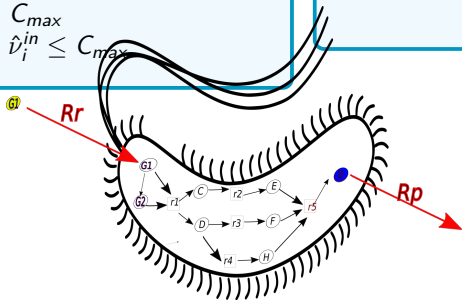
Find $\hat{\nu} = (\nu^{(\hat{1})}, \dots, \nu^{(\hat{N}_t)})$ such that

$$\hat{\nu} = \arg \min \frac{1}{N_t} \sum_{t=1}^{N_t} \|R_t - \mathcal{X}_t \nu_t\|^2 + \lambda \mathcal{R}(\nu)$$

$$s.c \left\{ \begin{array}{l} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{max} \\ -C_{max} \leq \hat{\nu}_i^{in} \leq C_{max} \end{array} \right.$$

Advantages and Limitations

-  Precisely quantifies metabolite fluxes in a metabolic network
-   No application to community models



Methods : Metabolic Flux Analysis






MFA

Find $\hat{\nu} = (\nu^{\hat{1}}, \dots, \nu^{\hat{N}_t})$ such that

$$\hat{\nu} = \arg \min \frac{1}{N_t} \sum_{t=1}^{N_t} \|R_t - \mathcal{X}_t \nu_t\|^2 + \lambda \mathcal{R}(\nu)$$

$$\text{s.c.} \left\{ \begin{array}{l} S_i \hat{\nu}_i = 0 \\ C_M \leq \hat{\nu}_i^s \leq 0 \\ 0 \leq \hat{\nu}_i^p \leq C_{max} \\ -C_{max} \leq \hat{\nu}_i^{in} \leq C_{max} \end{array} \right.$$

Advantages and Limitations

-  Precisely quantifies metabolite fluxes in a metabolic network
-   No application to community models
-   No application with metatranscriptomic data

