Minimizing multi-pathogen risk and costs through control measures

Subhasish Basak[1,2], Julien Bect[2], Laurent Guillier[1], Fanny Tenenhaus-Aziza[3], Janushan Christy[4] & Emmanuel Vazquez[2]

ArtiSaneFood: Biopreservation and Risk Modelling Approaches
May 25, 2023 - Bragança, Portugal

1. Agence Nationale de Sécurité Sanitaire (ANSES), Maison-Alfort, France
2. Université Paris-Saclay, CNRS, CentraleSupelec, L2S, Gif-sur-yvette, France
3. Centre national interprofessionnel de l'économie laitière (CNIEL), Paris, France
4. Centre technique d'expertise agroalimentaire (ACTALIA), La-Roche-sur-foron, France

This work is part of the ArtiSaneFood project (grant number : ANR-18-PRIM-0015) which is part of the PRIMA program supported by the European Union.

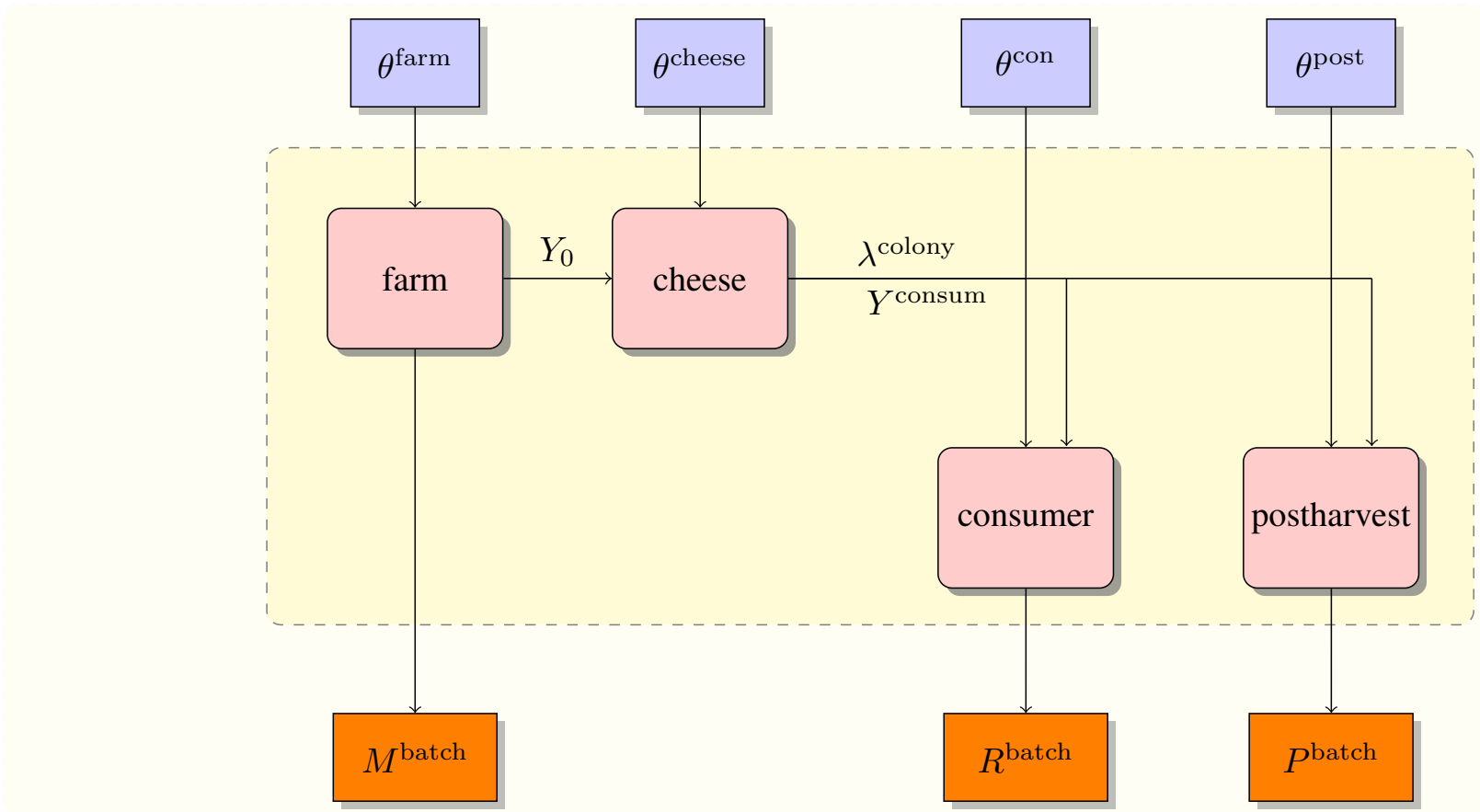
Motivation & application

- Aim : Control risk of illness from pathogens in raw milk cheese
  - STEC → Haemolytic Uremic Syndrome (HUS)

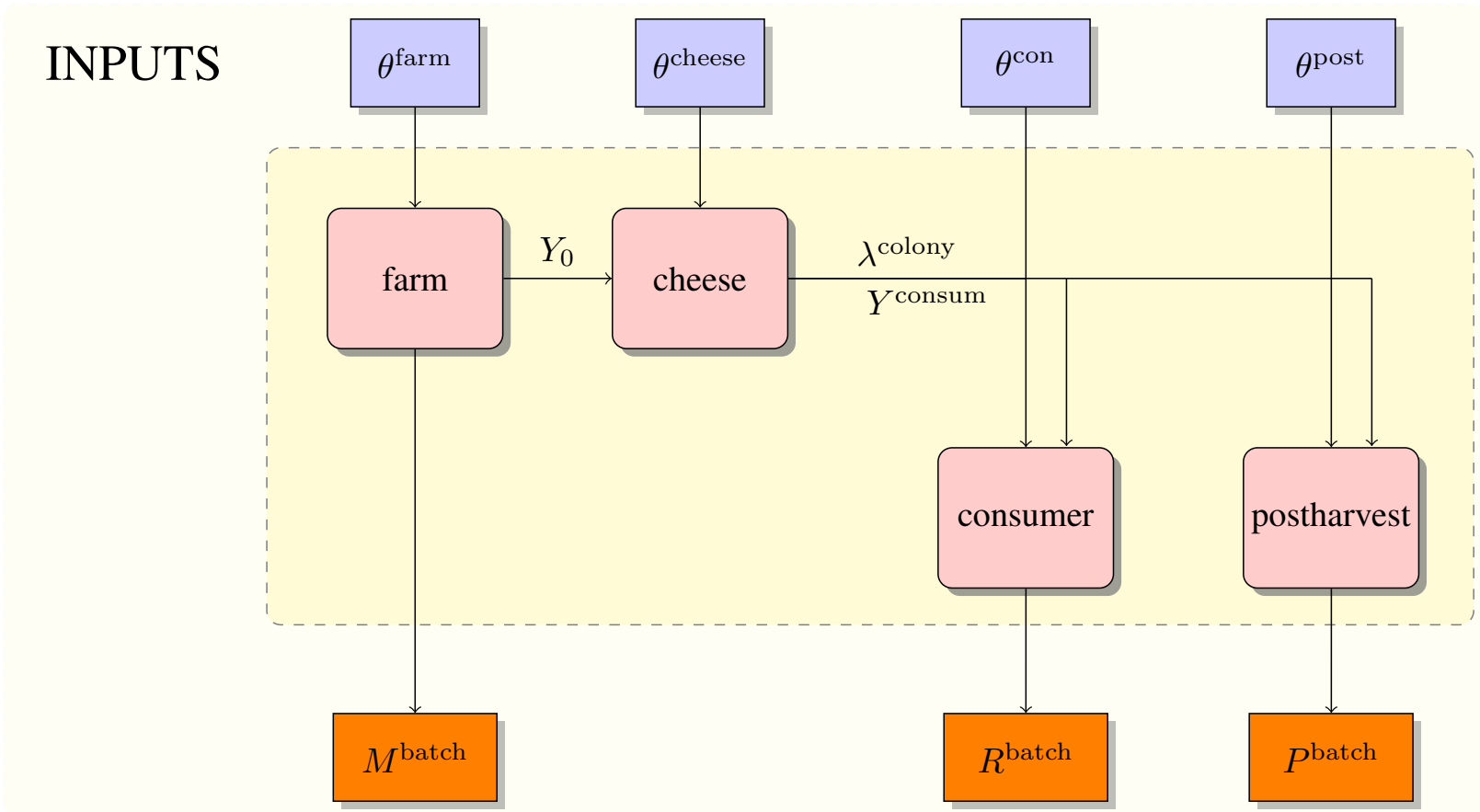
    STEC & MPS STEC
  - Salmonella → Salmonellosis

    High & Low Virulent Salmonella
- Quantitaive Risk Assesement (QRA)
  - Based on *Perrin et al. (2014)* and *Basak et al. (in prep.)*
- Optimization : Find intervention parameters that minimize the objectives
  - Relative risk of illness & intervention cost (€)

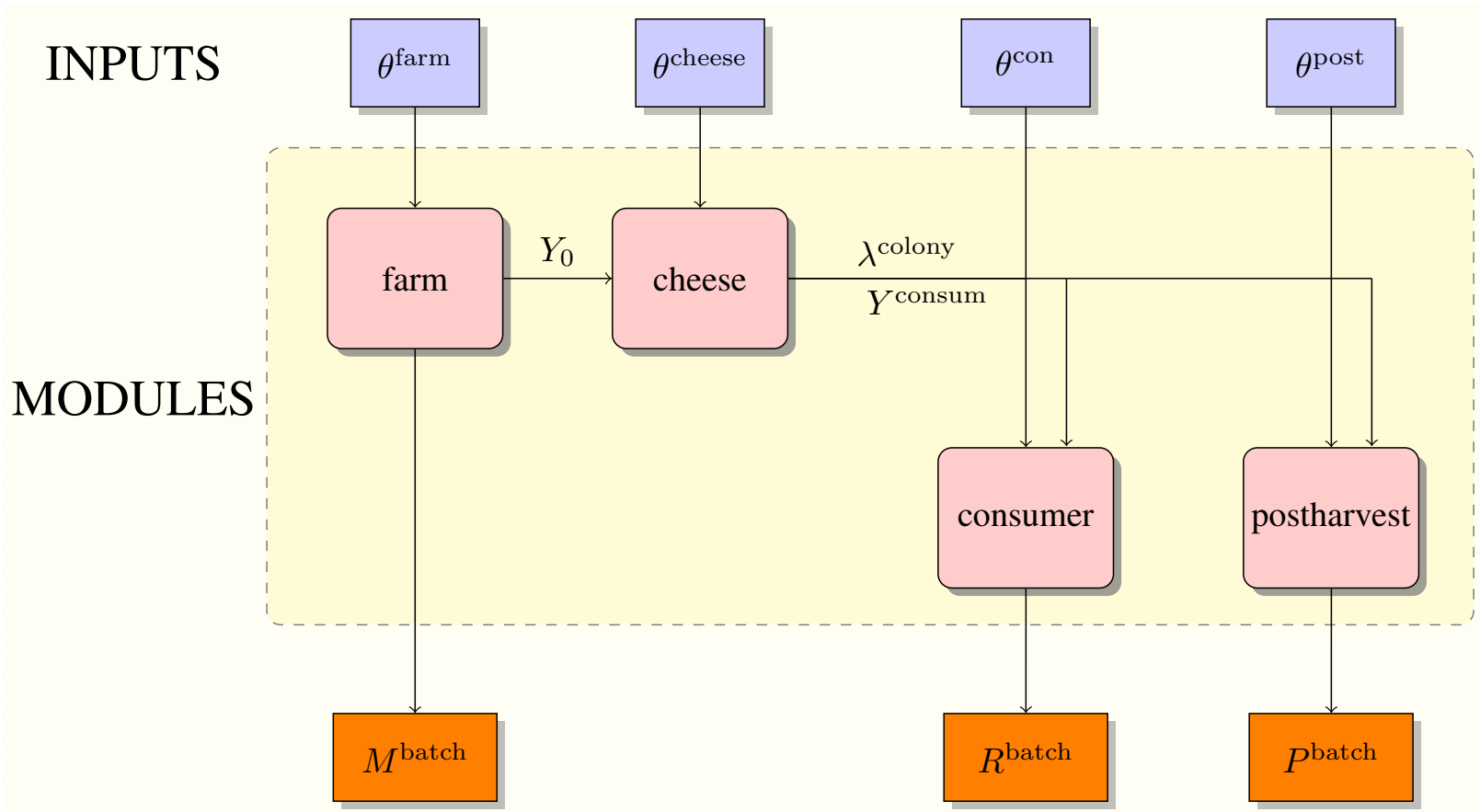
QRA model – Raw milk cheese



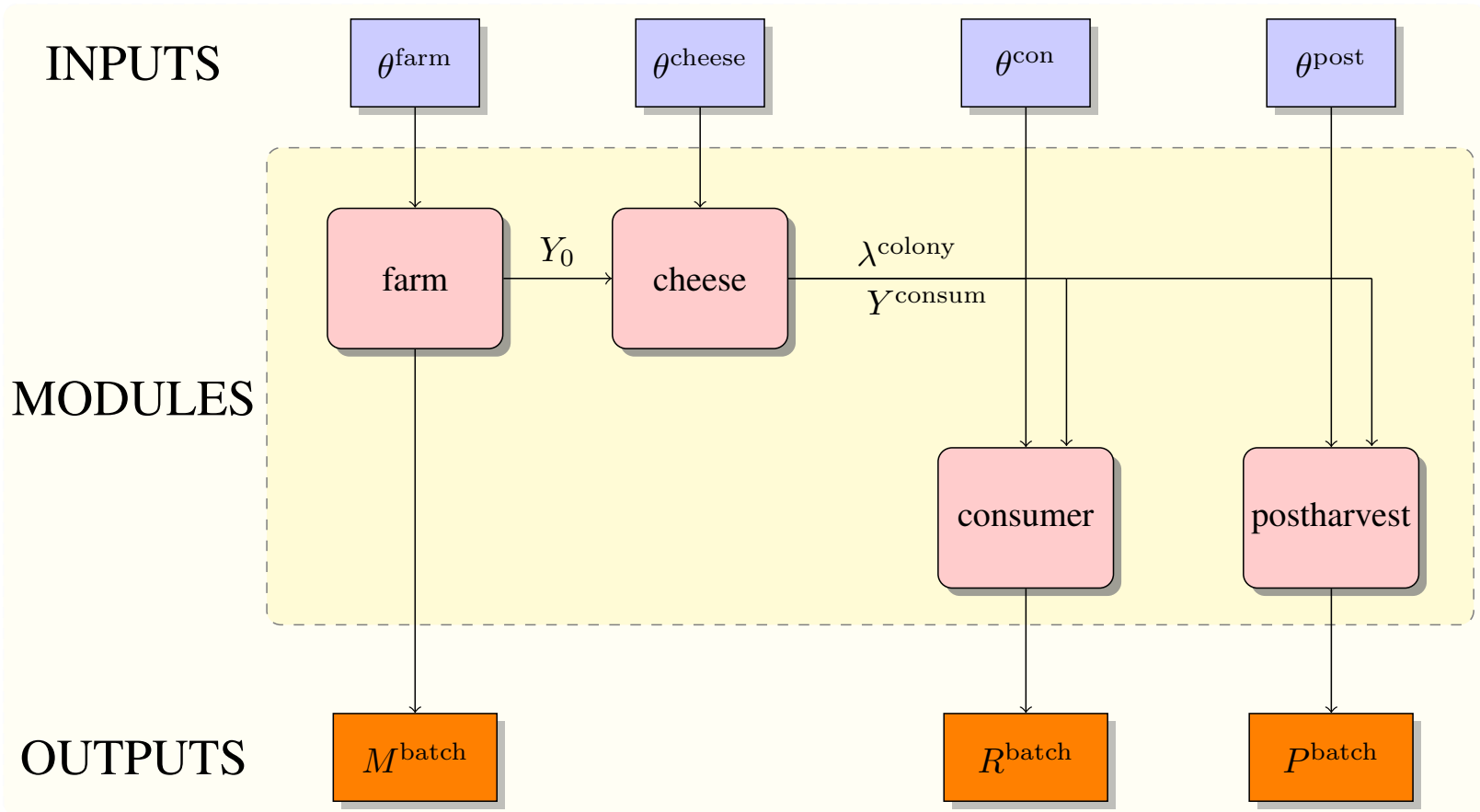
QRA model – Raw milk cheese



QRA model – Raw milk cheese



QRA model – Raw milk cheese



Farm module : Milk collection

- Milk sorting: Reject farm if E.coli conc. $(Y_{\text{milk}}^{\text{Ecoli}}) >$ threshold

  $Y_{\text{milk}}^{\text{Ecoli}} \sim$ Lognormal with parameters based on *CNIEL + ACTALIA* data

- Concentrations are computed indirectly due to limit of detection

$$Y_{\text{milk}}^{\text{X}} = Y_{\text{milk}}^{\text{Ecoli}} \cdot (Y_{\text{feces}}^{\text{X}} / Y_{\text{feces}}^{\text{Ecoli}})$$

  - $\text{X} \in \{\text{STEC}, \text{Salmonella}\}$
  - $Y_{\text{feces}}^{\text{Salmo}} \sim$ Lognormal, $Y_{\text{feces}}^{\text{STEC}} \sim$ Weibull, $Y_{\text{feces}}^{\text{Ecoli}} \sim$ Lognormal

- Prevalence rates and distribution parameters:
  - *(Perrin et al., 2014)* & *(Bonifait et al., 2021)*

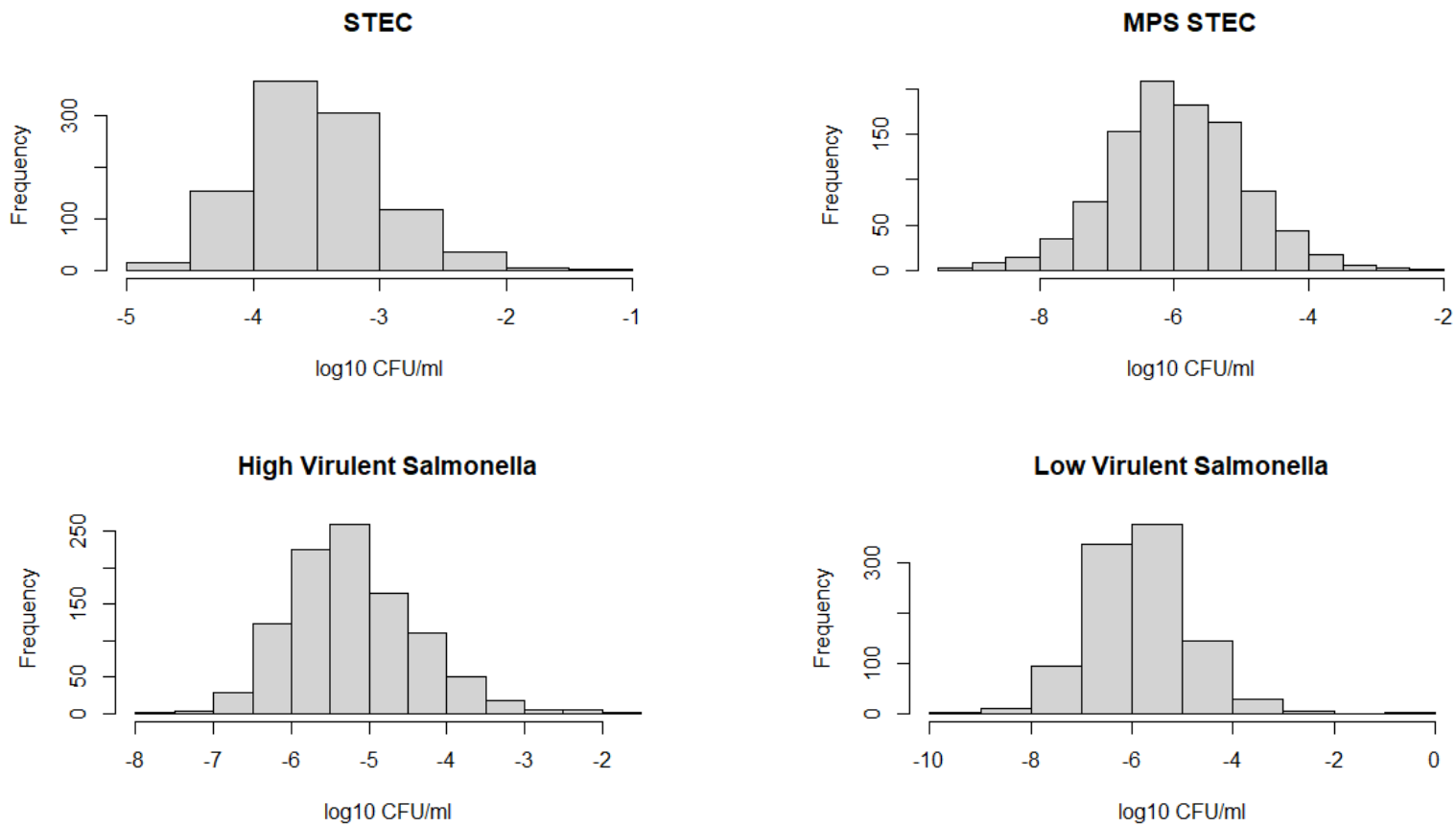


Figure 1: Simulated concentration of bacterias in milk in production

Farm module outputs

Cheese module

- Growth ↗ phase

$$\frac{dy}{dt} = \mu_{\mathrm{X}}^{\max}(t) \cdot y(t) \cdot (1 - \frac{y(t)}{y^{\max}})$$

Milk storage → Molding → *Colony formation* → Draining → Salting

- Decline ↘ phase

$$\text{STEC} : Y_{\mathrm{STEC}}^{\mathrm{consume}} = Y_{\mathrm{STEC}}^{\mathrm{salting}} \cdot 10^{-\rho \cdot t}$$

$$\text{Salmonella} : Y_{\mathrm{Salmo}}^{\mathrm{consume}} = Y_{\mathrm{Salmo}}^{\mathrm{salting}} \cdot 10^{-(t/\delta)^{p}}$$

Ripening → Cheese storage → Consumption

→ Decline rates based on *ACTALIA* report

Cheese module : Pre-molding stages

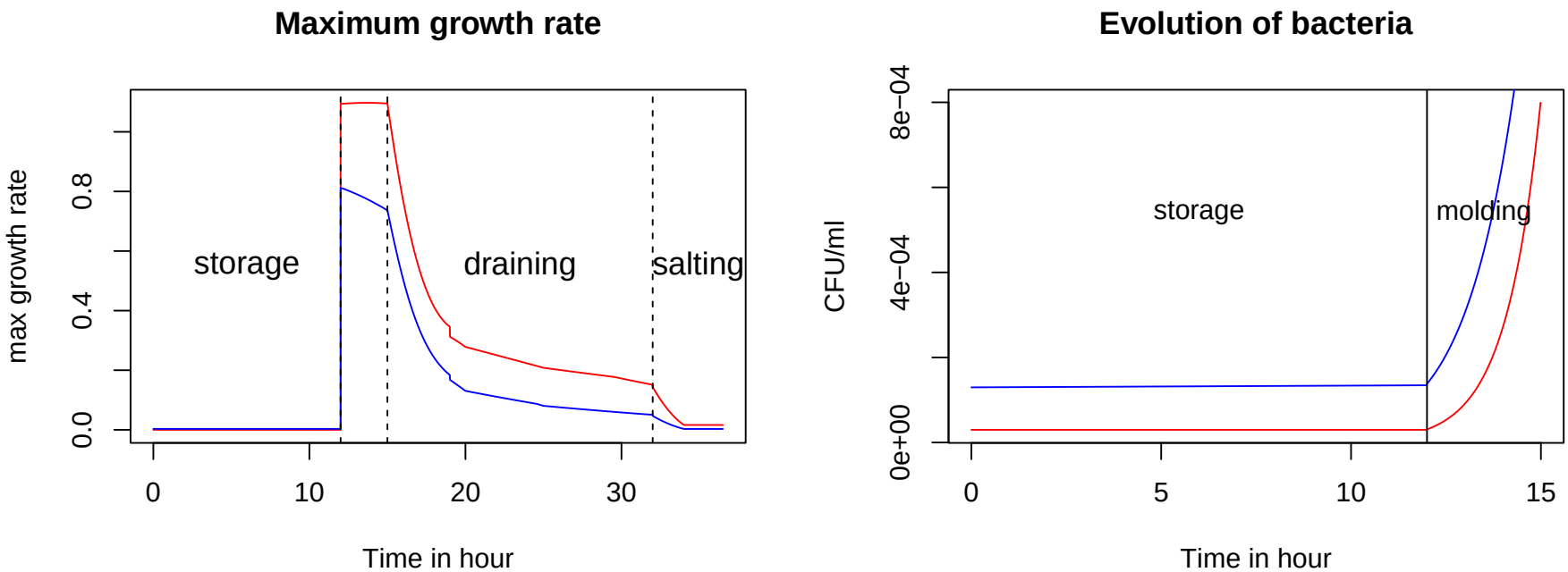


Figure 2: Growth rate $\mu^{\max}(t)$ and evolution $y(t)$ of STEC and Salmonella

$\rightarrow$ Number of colonies: $N_{\mathrm{X}}^{\mathrm{colony}} \sim \mathrm{Poisson}(Y_{\mathrm{X}}^{\mathrm{molding}} \cdot c)$

Cheese module : Evolution of colonies

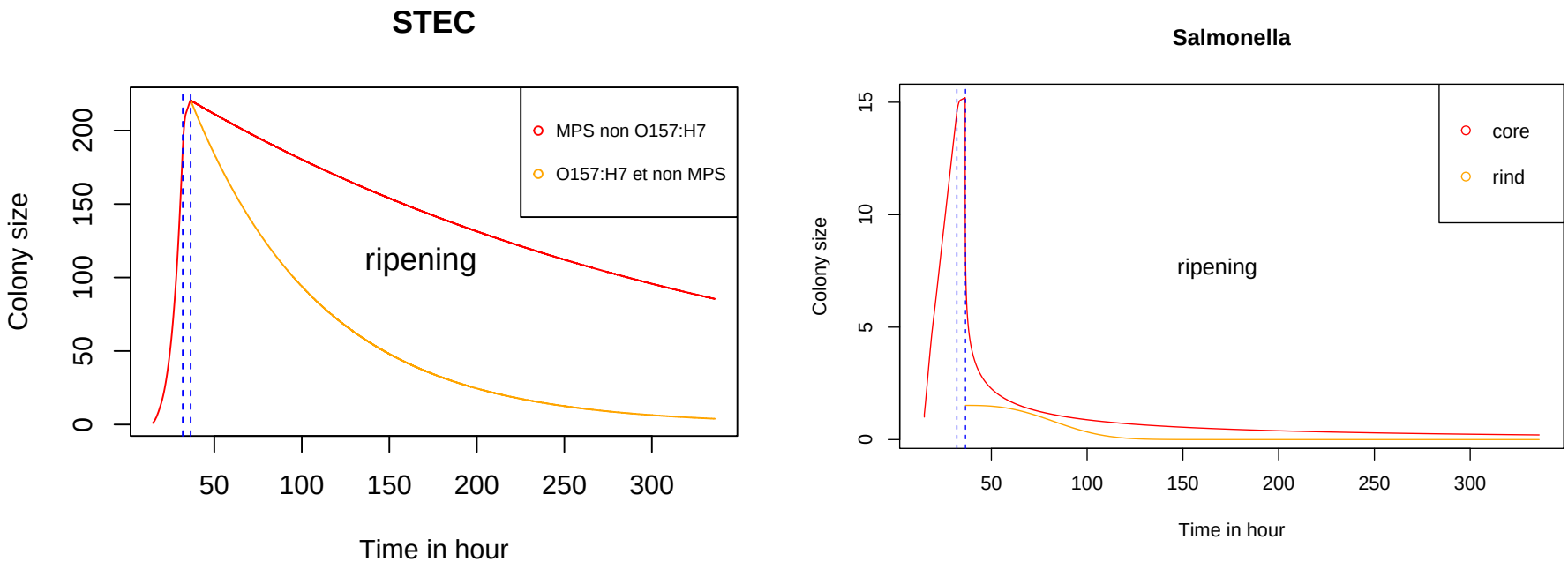


Figure 3: Evolution of colonies of different strains of STEC and Salmonella

$\rightarrow$ Size of colonies: $Y_{\mathrm{X}}^{\mathrm{colony}} \sim \mathrm{Lognormal}$ with median $Y_{\mathrm{X}}^{\mathrm{consume}}$

Consumer module

- Dose in cheese serving:

$$\Gamma_{\mathrm{X}} = \sum_{s \in \mathrm{strains}} N_s^{\mathrm{colony}} \cdot Y_s^{\mathrm{colony}}$$

- Dose-response model:

$$P_{\mathrm{STEC}}^{\mathrm{illness}} = 1 - (1 - r_{\mathrm{age}})^{\Gamma_{\mathrm{STEC}}} \textit{ (Perrin et al., 2014)}$$

$$P_{\mathrm{Salmo}}^{\mathrm{illness}} = 1 - (1 + \frac{\Gamma_{\mathrm{Salmo}}}{\beta})^{-\alpha} \textit{ (Strickland et al., 2023)}$$

- Batch risk:

$$R_{\mathrm{X}}^{\mathrm{batch}} = \sum_{\mathrm{age}=1}^{15} g(\mathrm{age}) \cdot \mathbb{E}_{\Gamma_{\mathrm{X}}}[P_{\mathrm{X}}^{\mathrm{illness}}]$$

$\rightarrow$ Effect of Salmonella is independent of consumer age

Intervention steps

- Milk testing: Farm module
  - Proportion of farm milk tested $\rightarrow p_{\text{milk}}$
  - Threshold of E.coli test $\rightarrow l_{\text{milk}}$

  $\rightarrow$ Computes M^{batch} milk rejected (in liters)
- Cheese testing: Postharvest module
  - Proportion of cheese batch tested $\rightarrow p_{\text{cheese}}$
  - Number of cheese samples tested $\rightarrow n_{\text{sample}}$

  $\rightarrow$ Computes $P_{\text{X}}^{\text{batch}}$ Probability of rejecting the batch

$$P_{\text{X}}^{\text{batch}} = 1 - (1 - P[\Gamma_{\text{X}} > 0])^{n_{\text{sample}}}$$

Quantities of interest (QoI)

- QoIs are computed using $R_{\mathrm{X}}^{\mathrm{batch}}$, $P_{\mathrm{X}}^{\mathrm{batch}}$ and M^{batch}

$$R_{\mathrm{X}}^{\mathrm{illness}} = \frac{\mathbb{E}[R_{\mathrm{X}}^{\mathrm{batch}} \cdot (1 - P_{\mathrm{X}}^{\mathrm{batch}} \cdot p^{\mathrm{cheese}})]}{\mathbb{E}[1 - P_{\mathrm{X}}^{\mathrm{batch}} \cdot p^{\mathrm{cheese}}]}$$

$$C^{\mathrm{avg}} = (c_1 + c_2 \cdot \mathbb{E}[M^{\mathrm{batch}}]) + (c_3 + c_4 \cdot \mathbb{E}[P_{\mathrm{X}}^{\mathrm{batch}} \cdot p^{\mathrm{cheese}}])$$

- Several batches are simulated to estimate

  $\rightarrow$ c_i's denote cost values *(COPIL ArtiSaneFood, Caen, Normandie, 2022)*

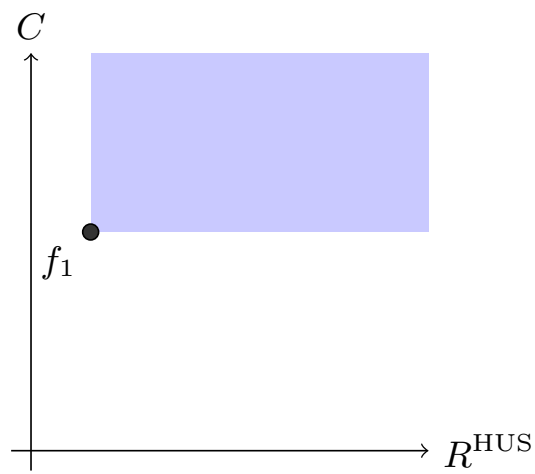
- DALY: Disability-adjusted life year
  - Batch risk for HUS and Salmonellosis are computed
  - Overall risk was assessed by combining the burden of disease metrics

Multi-objective optimization $\rightarrow f = (R^{\mathrm{HUS}}, C)$

- Consider an input: $x_1 = (p_{\mathrm{milk}},\, l_{\mathrm{milk}},\, p_{\mathrm{cheese}},\, n_{\mathrm{sample}})$

$$p_{\mathrm{milk}} = 30\%,\; l_{\mathrm{milk}} = 50\,\mathrm{CFU},\; p_{\mathrm{cheese}} = 50\%,\; n_{\mathrm{sample}} = 5\,\mathrm{units}$$

- Corresponding actual output: $f_1 = (R_1^{\mathrm{HUS}} = 2.2, C_1 = 1000\,\mathrm{EUR})$

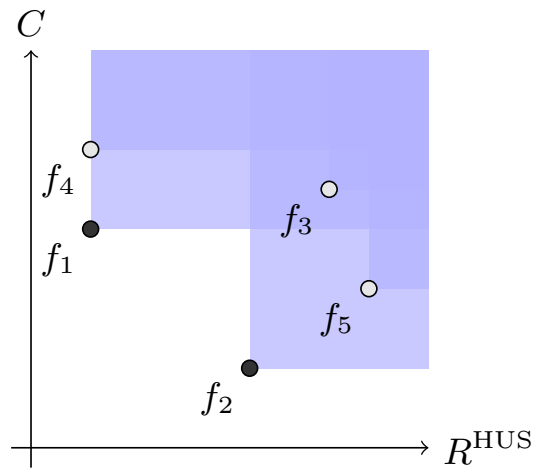


Dominated area by f_1 in objective space

Pareto optimal solutions

- Example: Inputs $\{x_i\}$ and outputs $\{f_i\}$, for $i = 1, \ldots, 5$

Minimizing two conflicting functions: $\min_x f(x)$



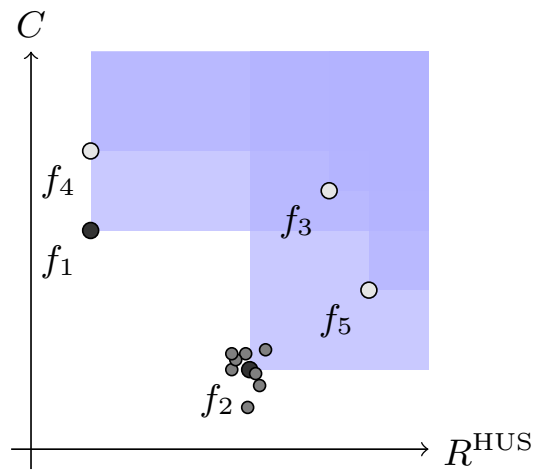
f_3, f_4 & f_5 dominated by f_1 and f_2

- Goal: Estimate the Pareto set $\mathcal{P} = \{x_1, x_2\}$ and Pareto front $\mathcal{F} = \{f_1, f_2\}$

Multi-objective stochastic optimization

- Simulator: Inputs $\{x_i\}$ and outputs $\{z_i = f_i + \text{noise}\}$, for $i = 1, \ldots, 5$

Minimizing two expensive & noisy functions: $\min_x f(x)$

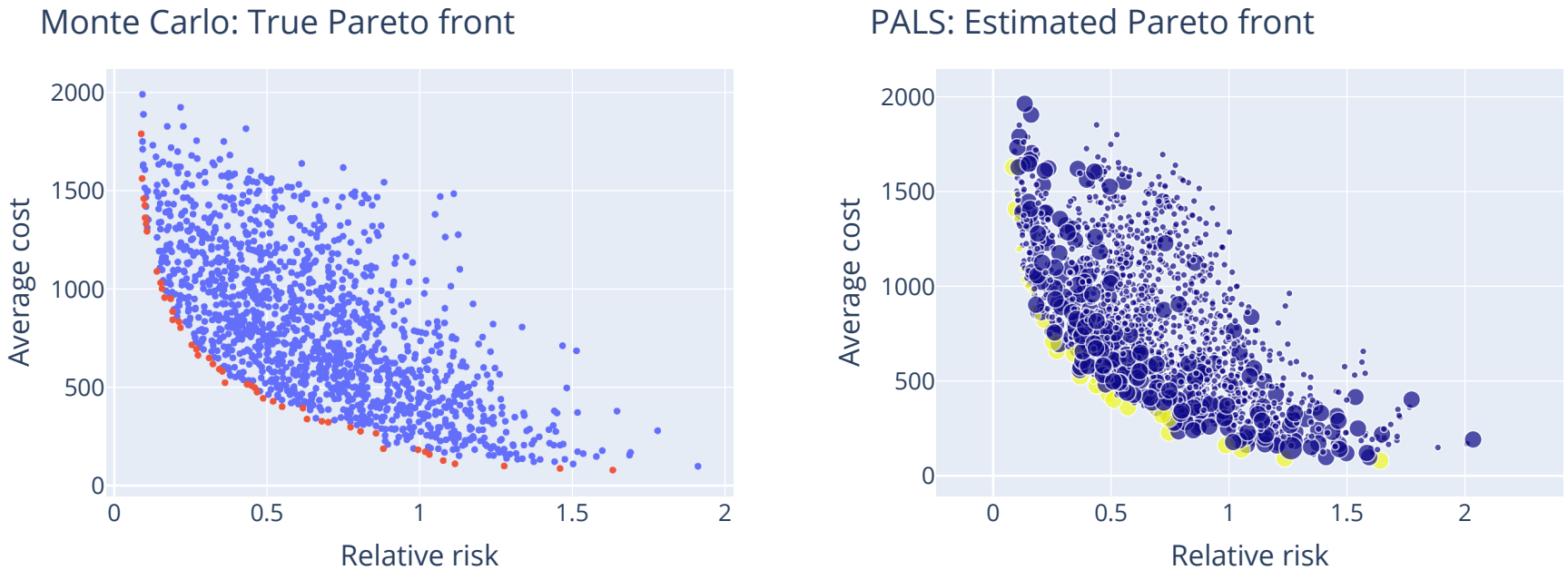


f_3, f_4 & f_5 dominated by f_1 and f_2

Naive approach Use Monte Carlo simulations → computationally expensive

PALS (Barracosa et al., 2021) and (Basak et al., 2022)

- Input points: $(p^i_{\text{milk}}, l^i_{\text{milk}}, p^i_{\text{cheese}}, n^i_{\text{sample}})_{i=1,2,\ldots,1500}$



(a) Samples ALL $\underline{1500 \text{ points}}$ $\times 5000$ (b) Samples only $\underline{300 \text{ points}}$ $\times 200$ (size)

- PALS has $\sim 4\%$ misclassification rate in estimating Pareto optimal points

Conclusions & Perspectives

- QRA model for STEC and Salmonella
- Bayesian Optimization for finding the optimal intervention parameters

→ Open sourcing (Github + FSKX) multipathogen QRA model

→ Optimization on DALY metrics

→ Propose a 3-pathogen model with Listeria

Thank you for your attention!

COPIL ArtiSaneFood 2022, Caen, Normandie

References

F. Perrin, F. Tenenhaus-Aziza, V. Michel, S. Miszczycha, N. Bel, and M. Sanaa. Quantitative risk assessment of haemolytic and uremic syndrome linked to O157:H7 and non-O157:H7 shiga-toxin producing escherichia coli strains in raw milk soft cheeses. Risk Analysis, 2014.

S. Basak, J. Christy, L. Guillier, F. Audiat-Perrin, M. Sanaa, F. Tenenhaus-Aziza, J. Bect, and E. Vazquez. Quantitative risk assessment of haemolytic and uremic syndrome (hus) from consumption of raw milk soft cheese. in prep.

L. Bonifait, A. Thépault, L. Baugé, S. Rouxel, F. Le Gall, and M. Chemaly. Occurrence of salmonella in the cattle production in france. Microorganisms, 9(4), 2021. ISSN 2076-2607. doi: 10.3390/microorganisms9040872. URL https://www.mdpi.com/2076-2607/9/4/872.

J.C. Augustin, V. Zuliani, M. Cornu, and L. Guillier. Growth rate and growth probability of listeria monocytogenes in dairy, meat and seafood products in suboptimal conditions. Journal of Applied Microbiology, 99:1019–1042, 2005. doi: 10.1111/j.1365-2672.2005.02710.x.

A.J Strickland, F Sampedro, and C.W Hedberg. Quantitative risk assessment of salmonella in ground beef products and the resulting impact of risk mitigation strategies on public health. Journal of Food Protection, 86(6):100093, 2023. ISSN 0362-028X. doi: https://doi.org/10.1016/j.jfp.2023.100093. URL https://www.sciencedirect.com/science/article/pii/S0362028X23067650.

B. Barracosa, J. Bect, H. Dutrieux Baraffe, J. Morin, J. Fournel, and E. Vazquez. Extension of the pareto active learning method to multi-objective optimization for stochastic simulators. Virtual Conference originally scheduled in Fort Worth, Texas, United States, 2021.

Subhasish Basak, Julien Bect, Laurent L. Guillier, Fanny Tenenhaus-Aziza, Janushan Christy, and Emmanuel Vazquez. Bayesian multi-objective optimization for quantitative risk assessment in microbiology. MASCOT-NUM 2022, June 2022. URL https://hal.science/hal-03715857. Poster.

A Bayesian Optimization (BO) framework

Expensive evaluations + Noisy observations

Algorithm 1 Using a Gaussian process regression (GPR) model ξ on f

Sample f at n_0 points ▷ initialization step
while budget > 0 **do**
  Update : GPR posterior ξ_n
  Compute : acquisition function $J_n(x)$
  Next point : $x_{n+1} = \arg\max_{x \in \mathbb{X}} J_n(x)$
  Sample : f at x_{n+1}
end while
Estimate $\widehat{\mathcal{P}}$ and $\widehat{\mathcal{F}}$ with GPR posterior mean ▷ Prediction step

Weighted Mean Square Error (W-MSE)

- PALS *(Barracosa et al., 2021)* + extension *(Basak et al., 2022)*
- MSE is used as a measure of uncertainty
- The new sample X_{n+1} corresponds to highest uncertain region of $\mathbb{X}$

$$X_{n+1} = \arg\max_{x \in \mathbb{X}} \left(w_n(x) \cdot \sum_{j=1}^{q} \frac{\sigma_{j,n}^2(x)}{R_{j,n}^2} \right)$$

- $R_{j,n}$ is a normalizing constant for $j = 1, 2, \ldots, q$ -th objective
- $\sigma_{j,n}^2(x)$ is the GP posterior variance at $x \in \mathbb{X}$
- Non-zero weights are given to "potentially Pareto optimal" points