Minimizing the risk of foodborne illness and analytical costs using a QMRA model for raw milk cheeses.

Subhasish Basak[1,2], Janushan Christy[4], Laurent Guillier[1], Fréderique Audiat-Perrin[1], Moez Sanaa[5], Fanny Tenenhaus-Aziza[3], Julien Bect[2] & Emmanuel Vazquez[2]

ICPMF 12

June 14, 2023 - Sapporo, Japan

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
5. World Health Organization (WHO), Geneva, Switzerland

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

- Aim : Control bacteria contamination in raw milk cheese
    1. Pathogenic *E. coli* (MPS-STEC) → Haemolytic Uremic Syndrome
    2. *Salmonella* → Salmonellosis
    3. *Listeria monocytogenes* → Listeriosis
- Control measures:

Farm milk testing

Test *E. coli* in farm milk

→ test frequency p_{milk}

→ threshold limit l_{milk}

Chesse batch testing

Test for cheese contamination

→ test frequency p_{cheese}

→ sample units n_{sample}

Optimal choice of parameters?

- Parameters: $\{p_{\text{milk}},\ l_{\text{milk}},\ p_{\text{cheese}},\ n_{\text{sample}}\}$
- Minimize both objectives:

Risk of illness

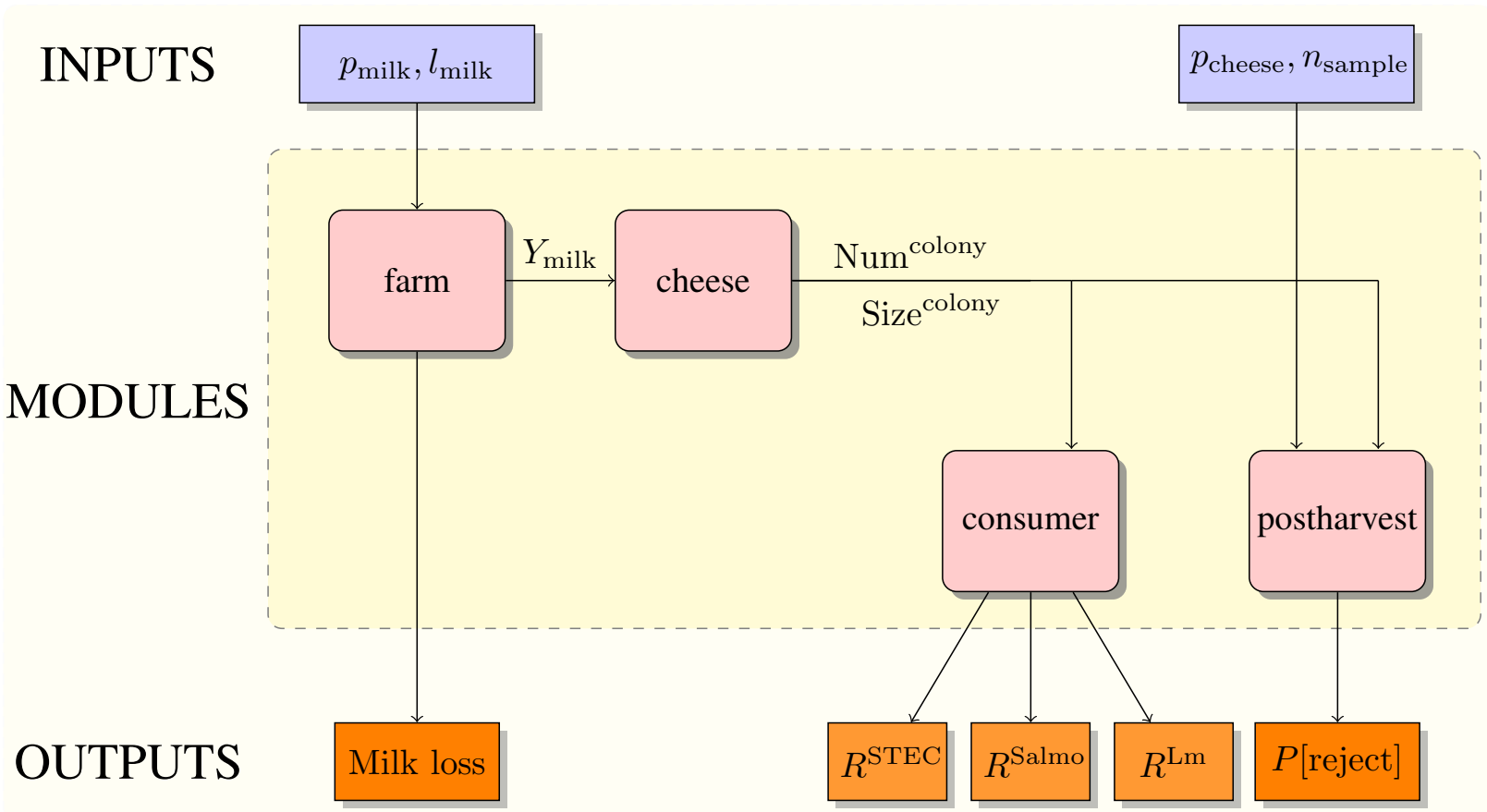
conflicting ↑ ↓ trade-off

Cost of intervention

1. Quantitative Risk Assesement *Perrin et al. (2014)*, *Basak et al. (in prep.)*
   - QRA model estimating risk and cost
2. Optimization algorithm *Barracosa et al. (2021)*, *Basak et al. (2022)*
   - Find best trade-off among the objectives & optimize parameters

Quantitative Risk Assessement – Raw milk cheese

Multipathogen QRA model



Farm module : Milk collection

- Milk testing: Rejecting farms with high *E. coli* contamination

$$Y_{\text{milk}}^{\text{Ecoli}} > l_{\text{milk}} \quad \text{CFU}$$

- Simulate concentration in milk
  - Indirect approach : limit of detection for $\text{x} \in \{\text{STEC},\ \text{Salmonella}\}$

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

$$Y_{\text{feces}}^{\text{Salmo}},\ Y_{\text{milk}}^{\text{Ecoli}},\ Y_{\text{feces}}^{\text{Ecoli}} \sim \text{Lognormal},\ Y_{\text{feces}}^{\text{STEC}} \sim \text{Weibull}$$

  - Direct approach : $Y_{\text{milk}}^{\text{Lm}} \sim \text{Lognormal}$

- Parameters: *Perrin et al. (2014)*, *Bonifait et al. (2021)* & *ACTALIA*

Cheese module: Maximum growth rate $\mu_{\mathrm{x}}^{\max}(t)$

- $\mu_{\mathrm{x}}^{\max}(t)$ is defined by a secondary cardinal model *(Augustin et al., 2005)*

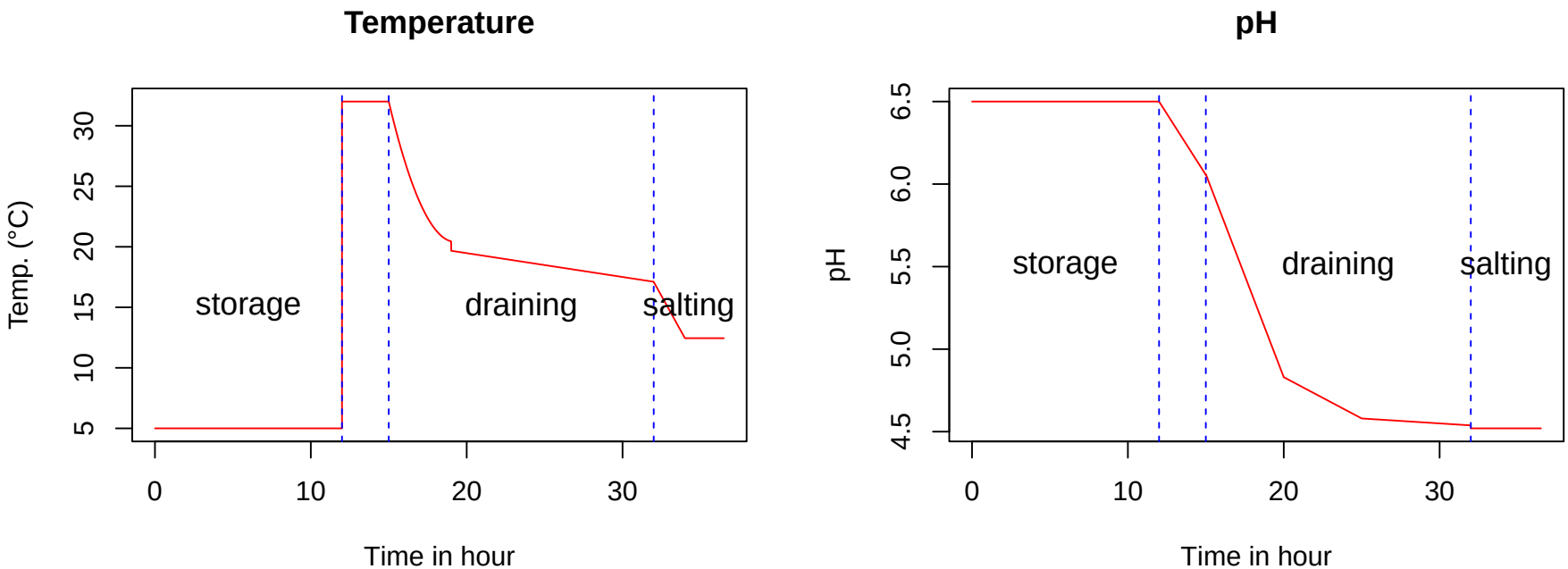


Figure 1: Physico-chemical params. for $\mu_{\mathrm{x}}^{\max}(t)$, $\mathrm{x} \in \{\mathrm{STEC},\ \mathrm{Salmo.},\ \mathrm{List.}\}$

Cheese module: Bacteria evolution

- 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

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

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

$$\text{Listeria} : \text{No decline} + \text{Second growth phase}$$

Ripening → Cheese storage → Consumption

- Parameters: *Perrin et al. (2014)*, *ACTALIA* challenge tests data

Cheese module: Pre-molding steps

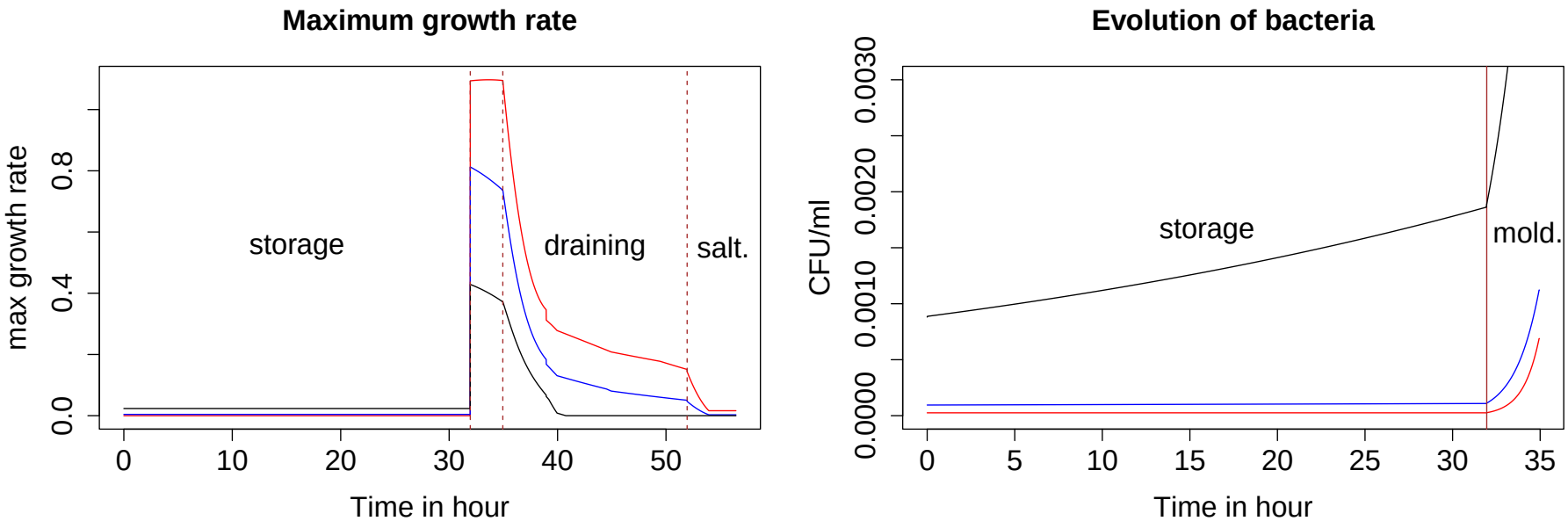


Figure 2: Growth rate $\mu_{\mathrm{x}}^{\max}(t)$ and evolution for STEC, Salmonella & Listeria

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

Cheese module: Evolution of colonies

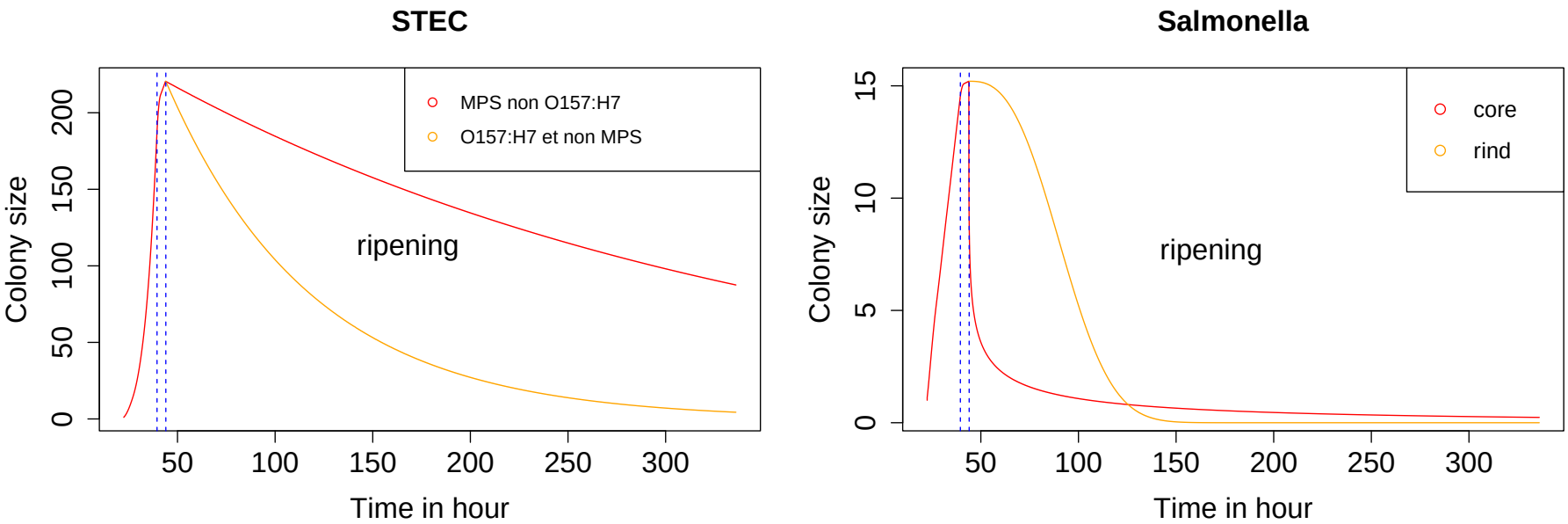


Figure 3: Evolution of colonies of different strains of STEC and Salmonella

$\rightarrow$ Size of colonies: $Y_{\mathrm{x}}^{\mathrm{colony}}$ based on $Y_{\mathrm{x}}^{\mathrm{consume}}$

Consumer module

- Dose in cheese serving: $\Gamma_{\mathrm{x}} = \sum_{s \in \mathrm{strains}} N_{\mathrm{x,s}}^{\mathrm{colony}} \cdot Y_{\mathrm{x,s}}^{\mathrm{colony}}$
- Dose-response model:

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

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

$P_{\mathrm{Listeria}}^{\mathrm{illness}} \rightarrow$ EFSA model *(Ricci et al., 2018)* + JEMRA *(Cadavez et al.)*

- 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

Quantities of interest (QoI)

- Outputs corresponding to one simulated batch
  - Consumer module → Risk of illness $\underline{R_{\mathrm{x}}^{\mathrm{batch}}}$
  - Cheese testing → Prob. of batch rejection $\underline{P^{\mathrm{batch}}}$
  - Milk testing → Milk loss (in liters) $\underline{M^{\mathrm{batch}}}$
- Several batches are simulated to estimate the QoIs

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

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

→ c_i's denote cost values *(COPIL ArtiSaneFood, Caen, Normandie, 2022)*

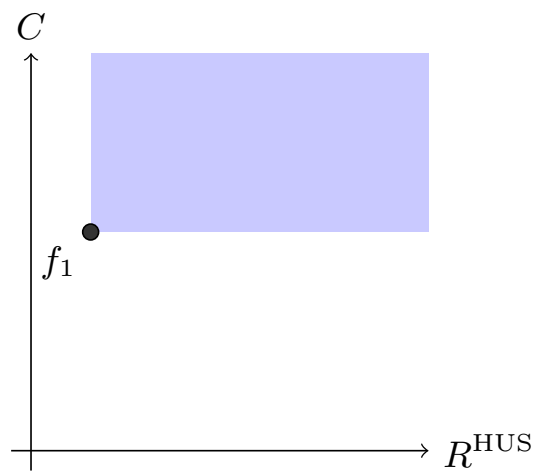
Multiobjective optimization of QRA simulator

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

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

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

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

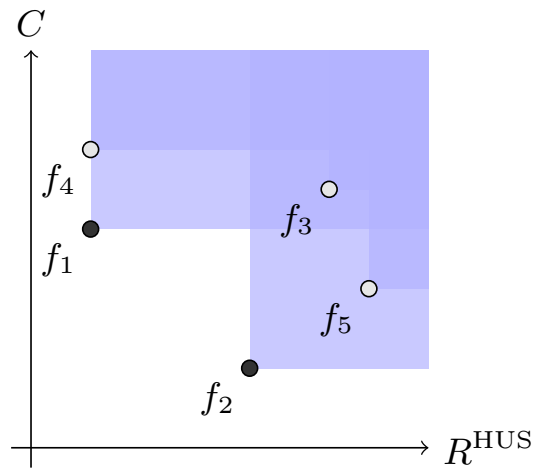


Dominated area by f_1 in objective space

Pareto optimal solutions

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

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



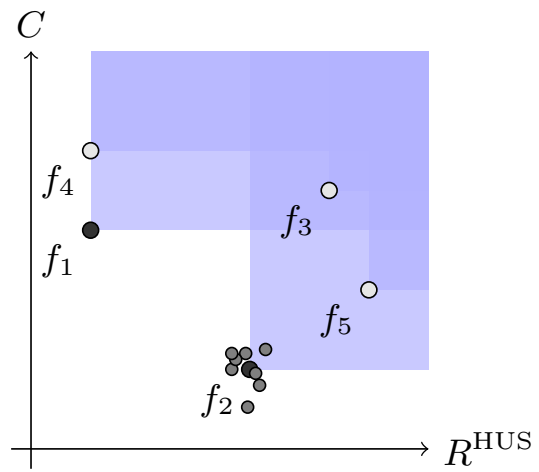
f_3, f_4 & f_5 dominated by f_1 and f_2

- 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 $i = 1, \ldots, 5$

QRA simulator produces noisy outputs



f_3, f_4 & f_5 dominated by f_1 and f_2

Naive approach Use Monte Carlo simulations → computationally expensive

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

Choice of acquisition function

- PALS *(Barracosa et al., 2021)* + extension *(Basak et al., 2022)*
- Weighted Mean Squared Error → measure of uncertainty
- The new sample X_{n+1} corresponds to highest uncertain region of $\mathbb{X}$

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

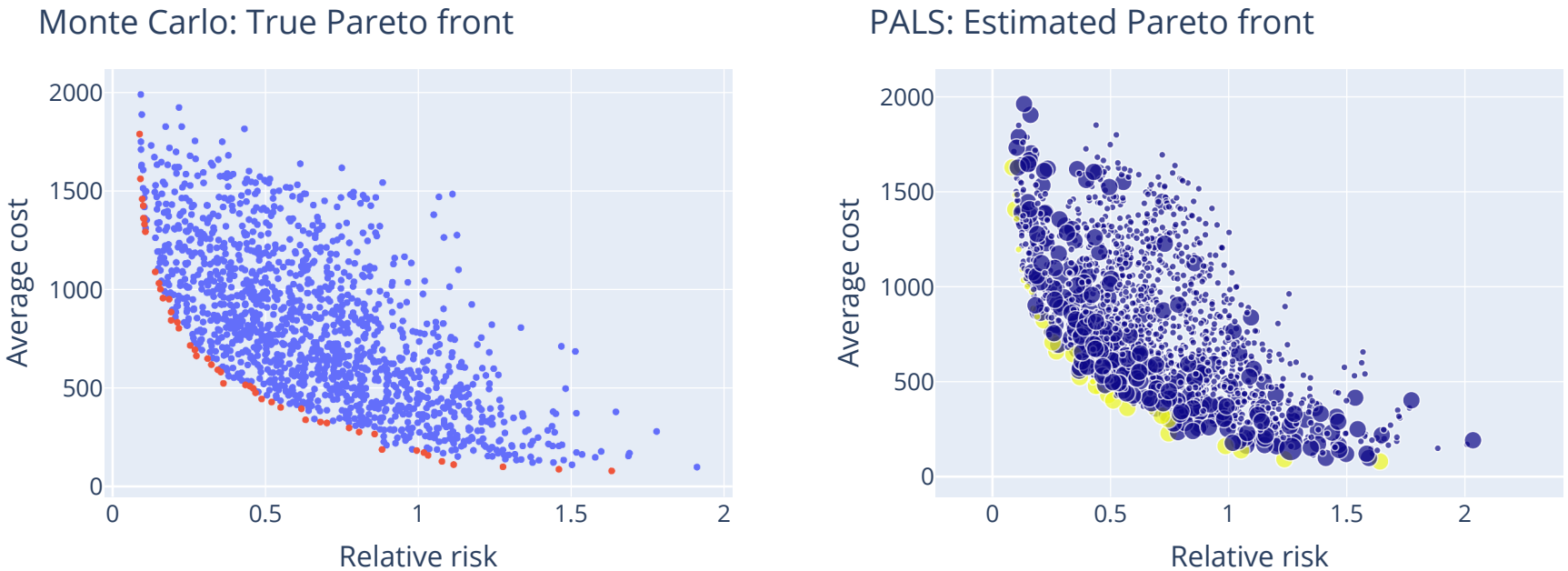
→ $R_{j,n}$ is a normalizing constant for $j = 1, 2, \ldots, q$ -th objective

→ $\sigma_{j,n}^2(x)$ is the GP posterior variance at $x \in \mathbb{X}$

- Non-zero weights are given to "potentially Pareto optimal" points

PALS w/ quantiles (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

Contributions of this work

- Multipathogen QRA model
  - STEC, Salmonella and Listeria monocytogenes
- Implementation of the model in R + FSKX
- Bayesian optimization algorithm
  - To optimize noisy and costly simulators
  - Multiple conflicting objectives
  - Finding Pareto optimal parmeters in a limited budget

$\rightarrow$ Perspectives: Optimization on DALY metrics

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.

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.

S. Basak, J. Bect, L. Guillier, F. Tenenhaus-Aziza, J. Christy, and E. Vazquez. Bayesian multi-objective optimization for quantitative risk assessment in microbiology. MASCOT-NUM 2022, June 2022. URL `https://hal.science/hal-03715857`. Poster.

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.

A. Ricci, A. Allende, D. Bolton, M. Chemaly, R. Davies, P. S. Fernández Escámez, R. Girones, L. Herman, K. Koutsoumanis, B. Nørrung, L. Robertson, G. Ru, M. Sanaa, M. Simmons, P. Skandamis, E. Snary, N. Speybroeck, B. Ter Kuile, J. Threlfall, H. Wahlström, J. Takkinen, M. Wagner, D. Arcella, M. T. Da Silva Felicio, M. Georgiadis, W. Messens, R. Lindqvist, and EFSA Panel on Biological Hazards (BIOHAZ). Listeria monocytogenes contamination of ready-to-eat foods and the risk for human health in the eu. EFSA Journal, 16(1):e05134, 2018. doi: https://doi.org/10.2903/j.efsa.2018.5134. URL https://efsa.onlinelibrary.wiley.com/doi/abs/10.2903/j.efsa.2018.5134.

V. Cadavez, R. Pouillot, L. Guillier, U. Gonzales-Barron, and M. Sanaa. JEMRA QMRA models for Listeria monocytogenes. URL https://github.com/vcadavez/JEMRA/. R package version 0.2.0.