Contents

- **Quantitative Risk Assessment (QRA)**
  1. Farm-to-Fork continnum
  2. Modules and outputs
- **Intervention Strategies**
  1. Parameters and outputs
  2. Cost of intervention
- **The Optimization problem**
  1. Quantities of interest
  2. Control parameters
- **Pareto Optimal solution**
  1. Estimation of the Pareto set
  2. Monte-Carlo vs PALS

Quantitative Risk Assessment (QRA)

- Model by Perrin et al (2014)

- Simulates <u>one cheese batch</u>

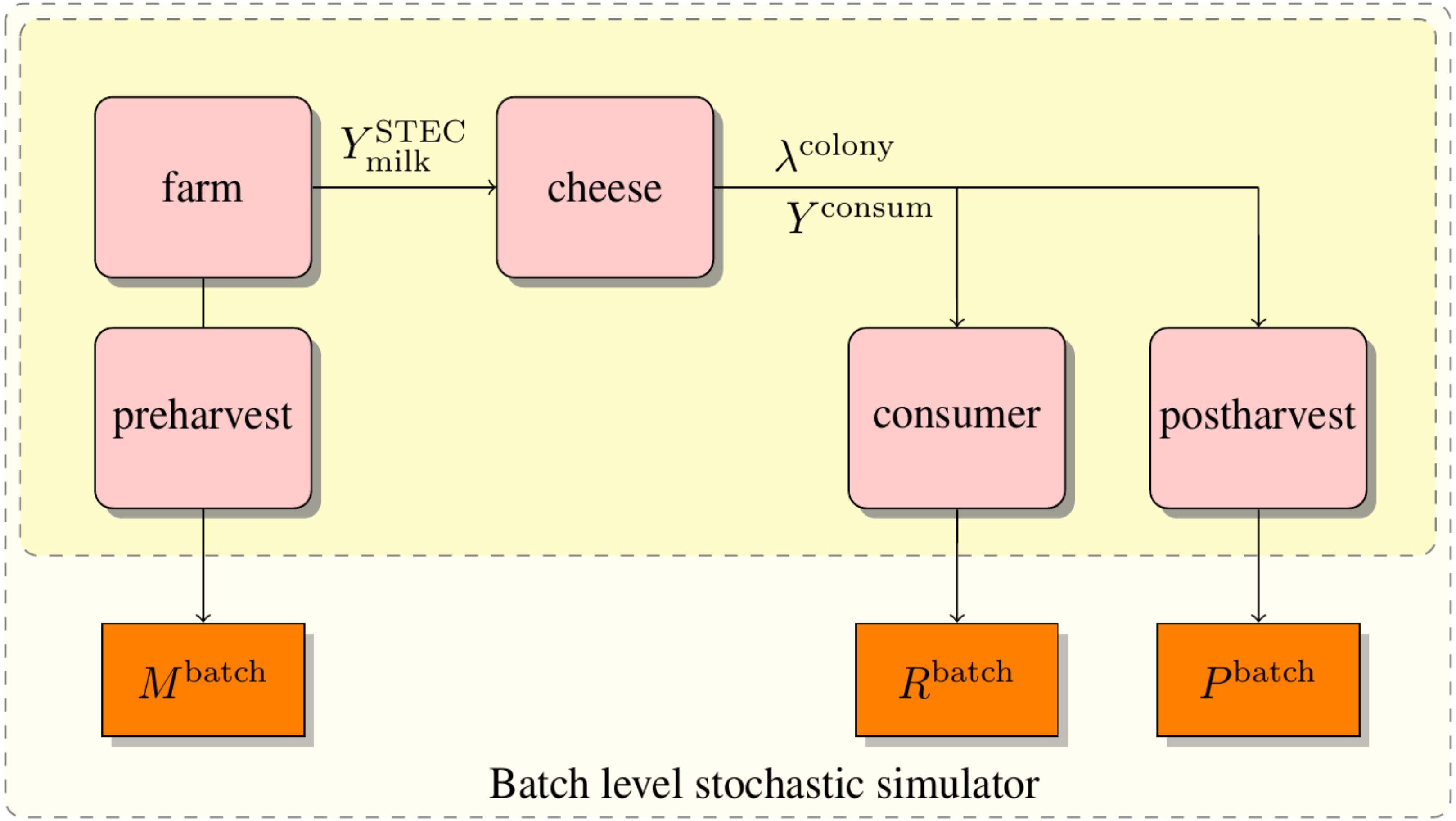
~20 000 cheese (250 g)

- Outputs

Milk loss (in Liters)

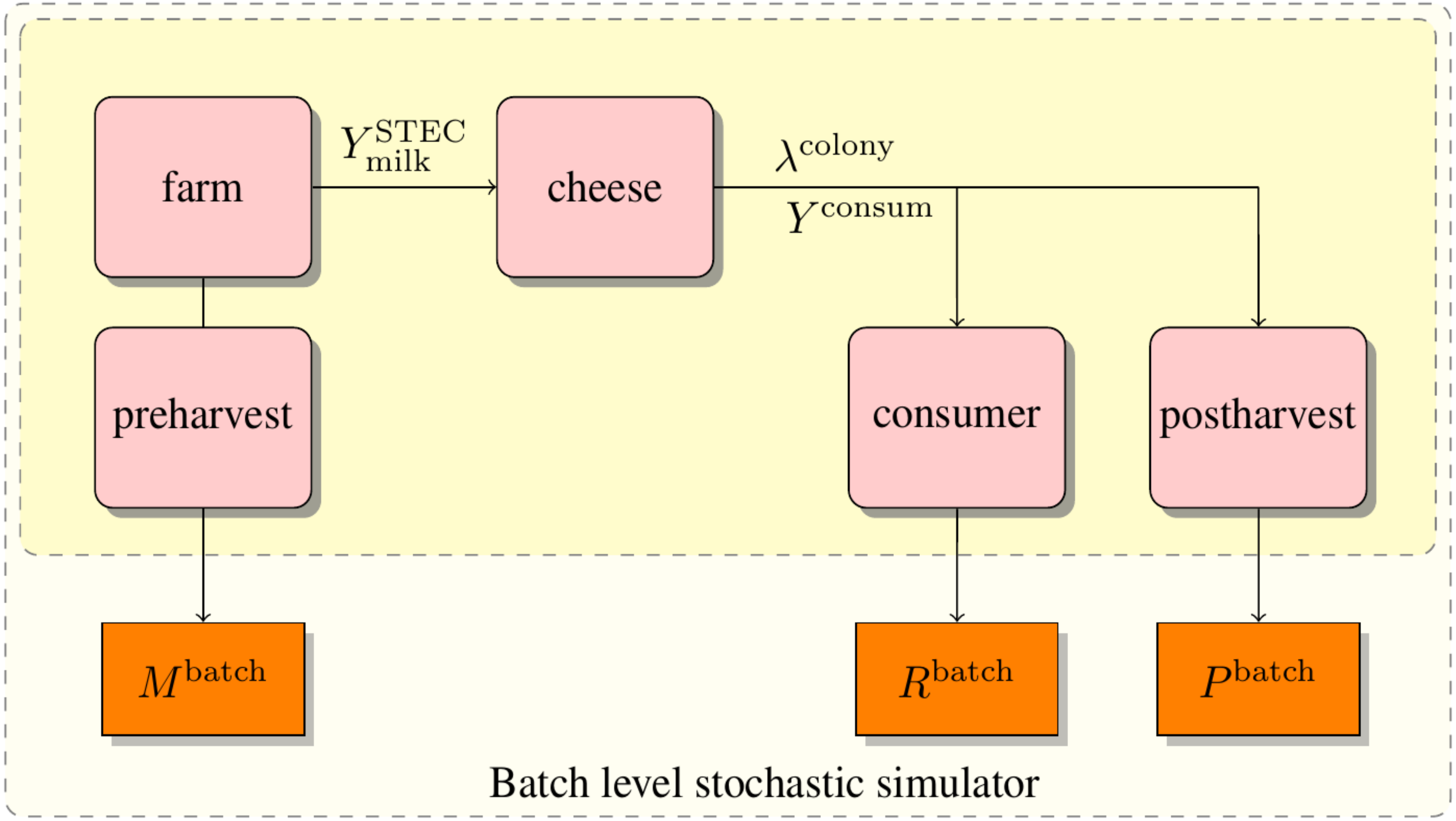
Risk of HUS from MPS-STEC

Probability of batch rejection



Quantitative Risk Assessment (QRA)

- Farm module
- Preharvest Step
- Cheese module
- Postharvest Step
- Consumer module



Quantitative Risk Assessment (QRA)

- Farm module

Assumption:

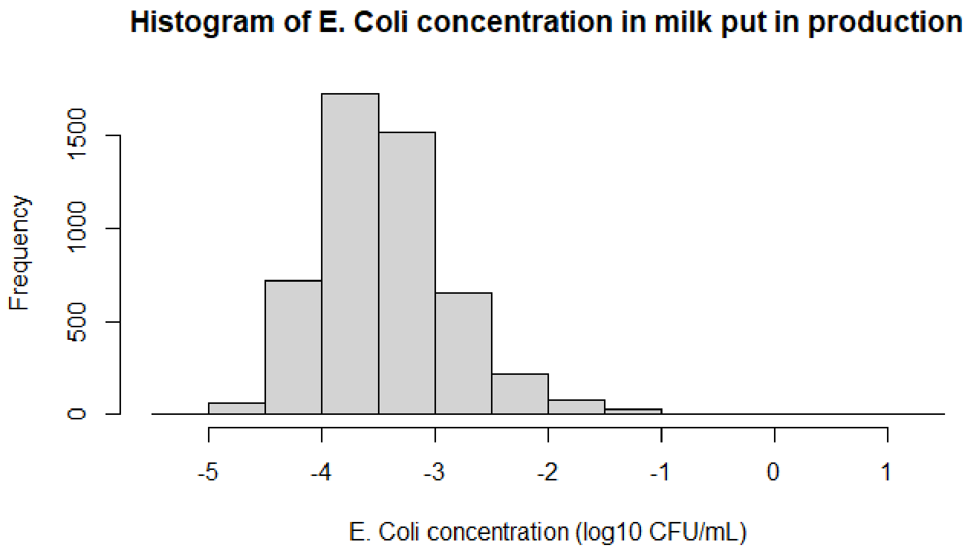
E. coli and STEC follows the same fecal route!

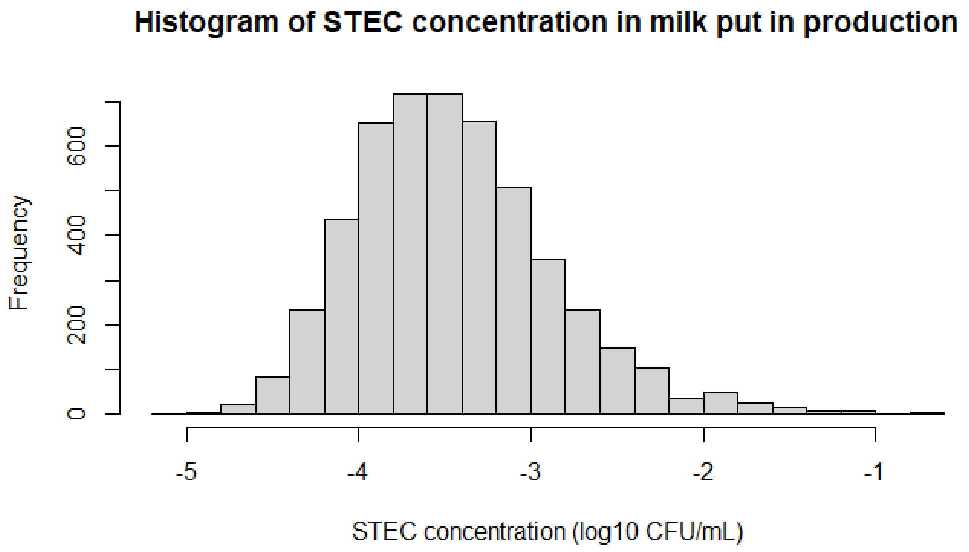
Output:

Concentration of STEC (**y0**) (CFU/ml) in milk used for cheese making

Plot:

Distribution of STEC (both MPS and Non-MPS) and *E. coli*





Quantitative Risk Assessment (QRA)

- Preharvest Step
  (inside farm module)

Parameters:

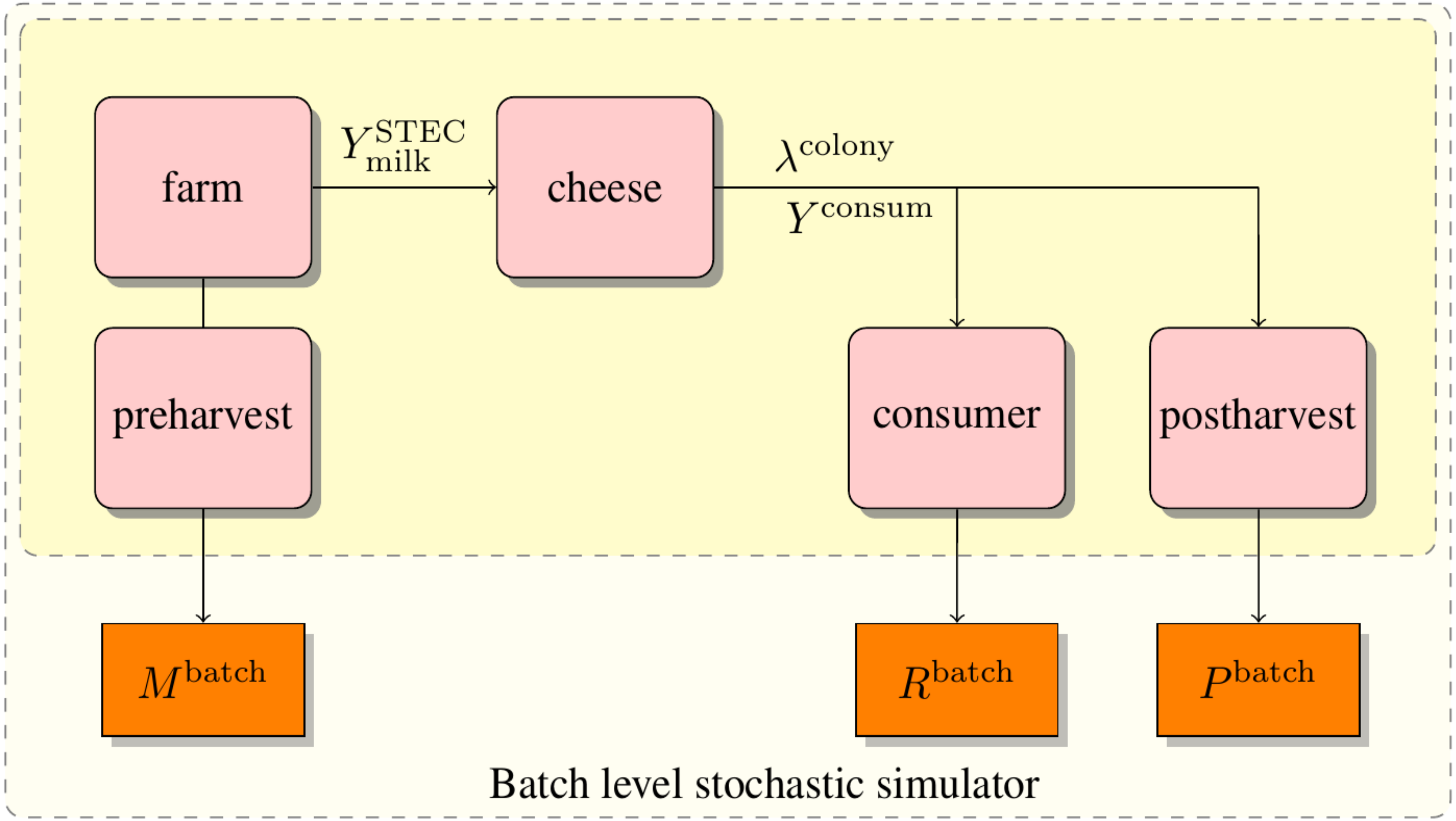
- *E. coli* test frequency:
  f_{sort} days
- Sorting limit for *E. coli*:
  l_{sort} CFU/ml

Rule:

Reject farms with *E. coli* concentration > l_{sort}

Output:

Milk loss per batch (in L)



Quantitative Risk Assessment (QRA)

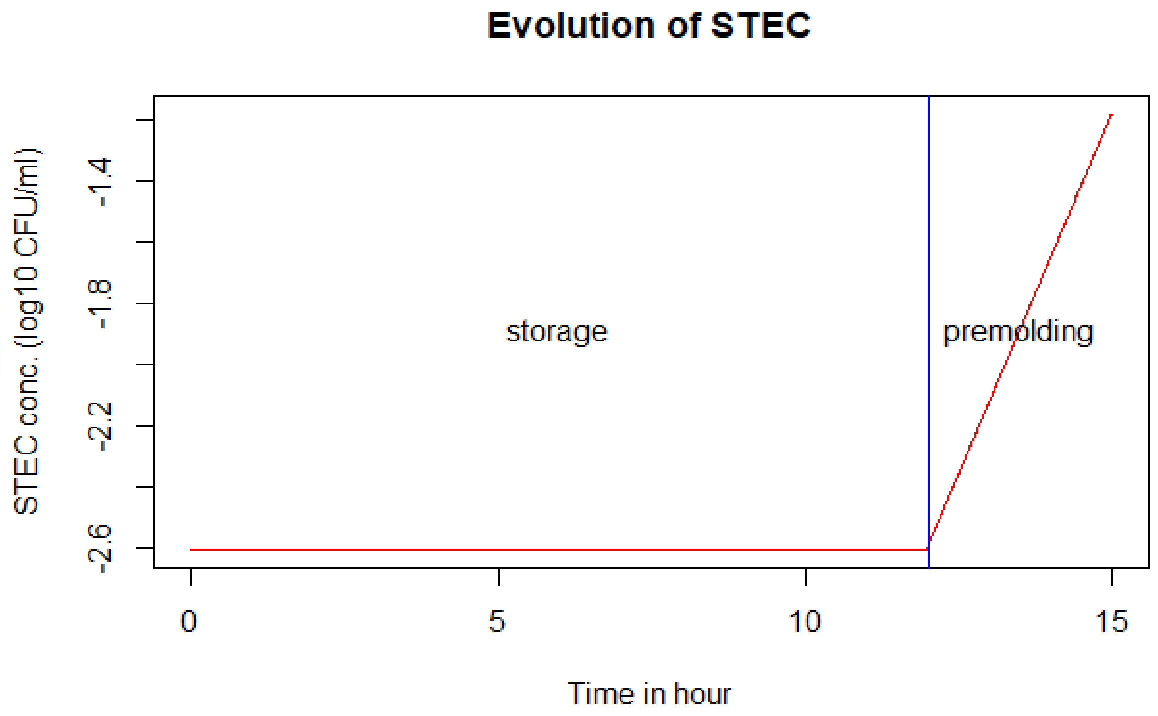
- Cheese module

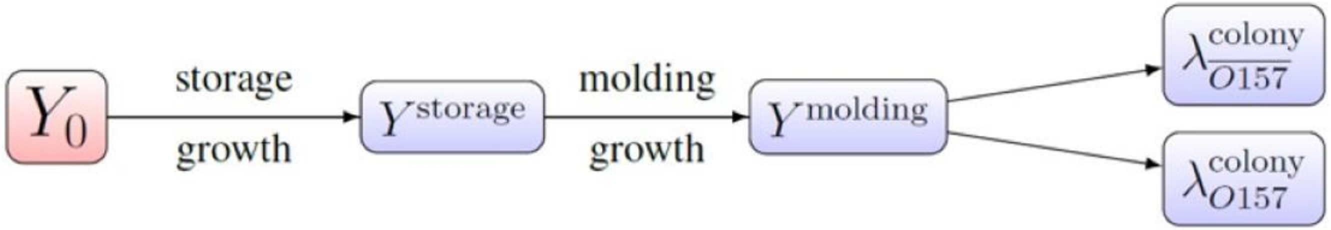
Evolution of STEC:
(liquid phase)

1. Milk storage
2. Molding Moulage

Output:

Average number of colonies of size 1



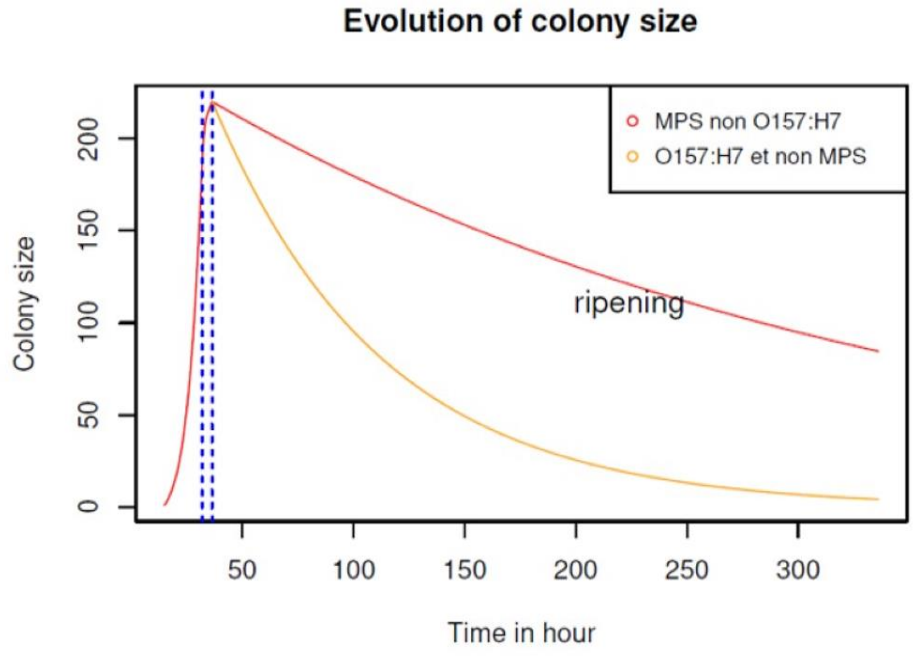


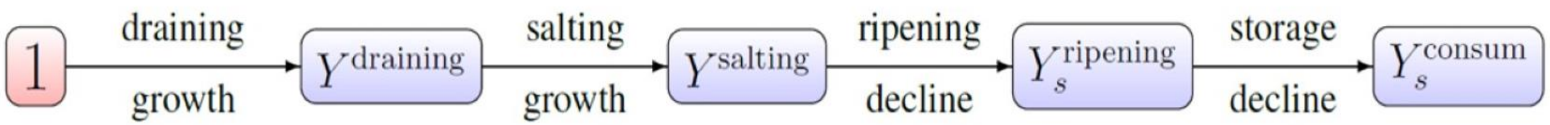
Quantitative Risk Assessment (QRA)

- Cheese module

Evolution of STEC:
(solid phase)

1. Draining/egoutage
2. Salting/salage
3. Ripening/affinage
4. Cheese storage/stockage et distribution





Output:

Size of colonies

Quantitative Risk Assessment (QRA)

- Postharvest step

Parameters:

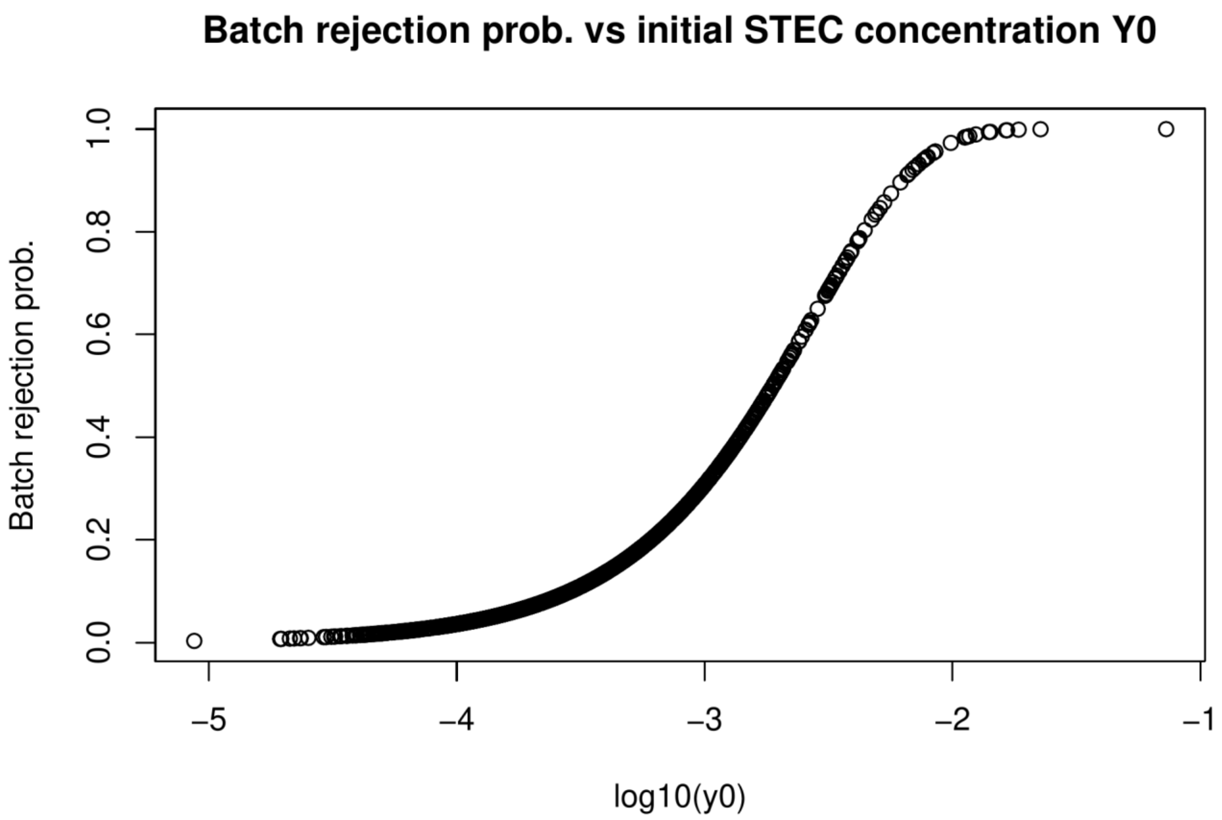
- Frequency of testing a batch : p_{test}
- Number of sample units: n_{sample}

Rule:

Samples of weight 25g are tested and the whole batch is rejected if STEC is detected.

Output:

Batch rejection probability



Probability of rejecting a batch (if tested) with a initial concentration of STEC ***y0*** CFU/ml

Quantitative Risk Assessment (QRA)

- Consumer module

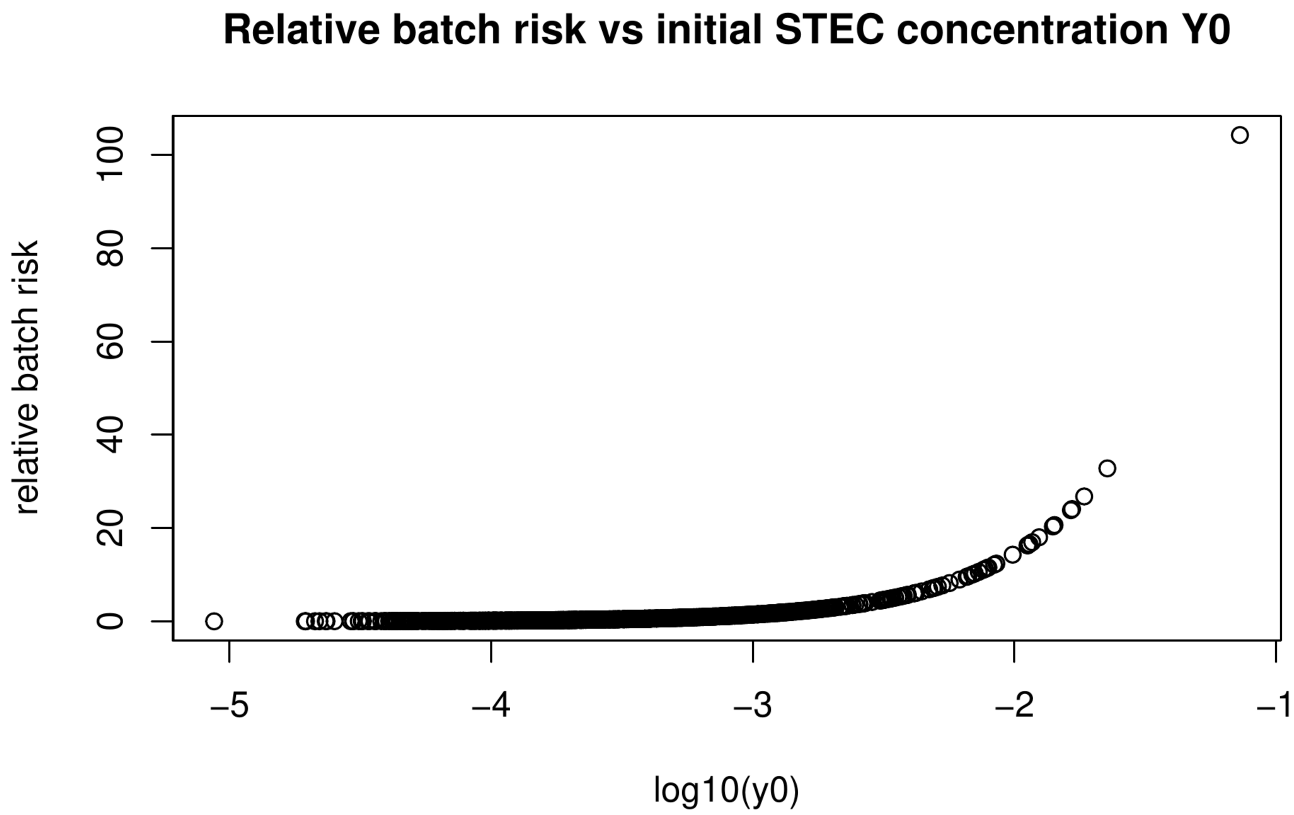
Computation of batch risk:

Averaging over consumtion behaviour of age groups

Output:

Relative risk of HUS =

(Batch risk / baseline risk)



(Baseline risk = risk computed with no intervention)

QRA simulator

The QRA simulator is stochastic

For a fixed set of inputs the simulator produces different outcome

Inputs can be fixed:

Intervention parameters, premolding draining step paramerters etc.

Inputs can be random:

Storage duration, temperature, consumption time etc.

Internal variables are random:

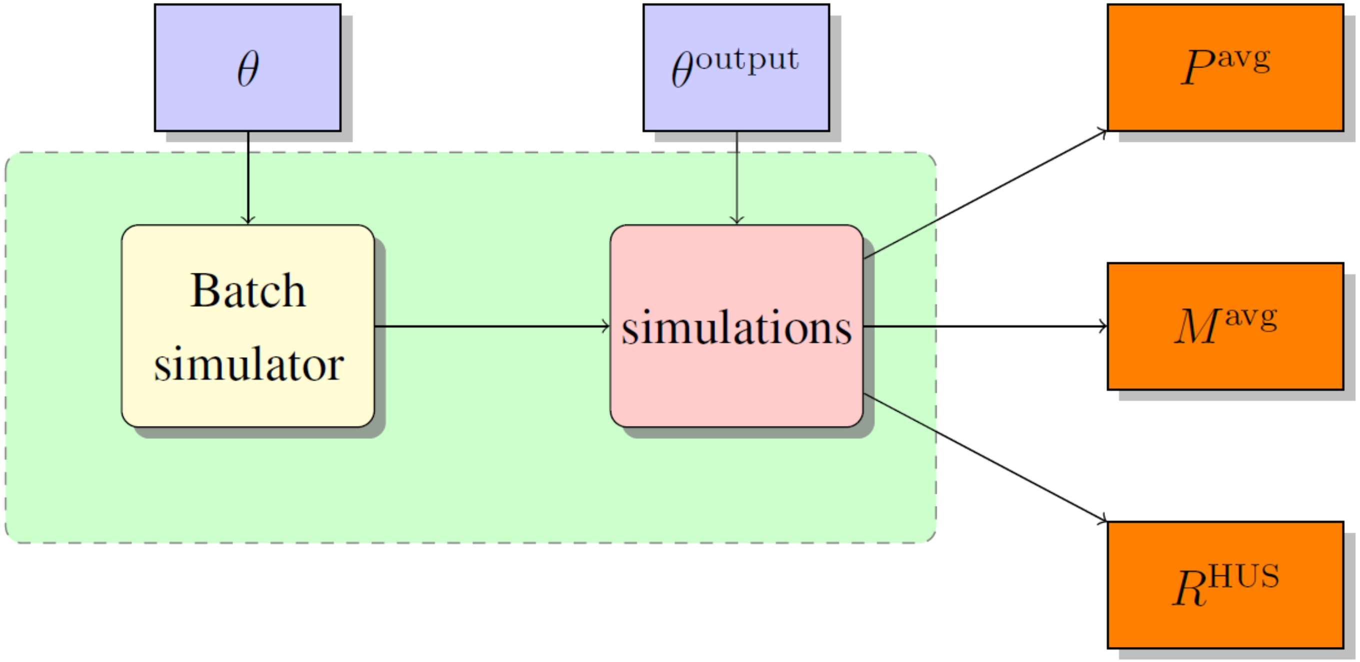
Initial STEC concentration, colony numbers and size etc.

Quantitative Risk Assessment (QRA)

Quantities of interest (QoI)

Simulating several batches

Several batches are simulated and the final QoIs are averages over the batches.



Output:

- Proportion of rejected batches
- Milk loss
- Risk of HUS

QRA model: modifications et improvements

- La variable u_i dénotant la variabilité entre les fromages est Perrin et al. (2014) est supprimé pour assurer la cohérence des résultats.
- Dans le module ferme, les hyper-paramètres de la distribution de la concentration d'E. coli (UFC/mL) dans le tank à lait en vrac sont estimés à l'aide d'une approche Bayésienne basée sur un modèle mixte de Poisson hiérarchique tel que décrit par l'équation (4) dans Perrin et coll. (2014).
- La proportion de vaches infectées par le MPS est estimée à 2,5 %. Cette proportion est utilisée dans le module ferme pour simuler le nombre de vaches infectées par MPS dans chaque ferme.
- Le risque de lot est calculé au moment de la consommation qui inclut l'inactivation (diminution de la concentration) pendant la phase de stockage du fromage.

Intervention strategies (Recap)

Preharvest step

Milk Sorting:

- Test frequency: f_{sort} days
- Sorting limit for E. coli: l_{sort} CFU/ml

Output:

Milk loss per batch (in L)

Postharvest step

Cheese testing:

- Frequency of cheese batch tested: p_{test}
- Number of test samples: n_{sample}

Output:

Batch rejection probability

Intervention strategies: Cost

Preharvest step

Parameters:

- Cost for E. coli testing in milk from one farm: *50€*
- Cost of 1L of milk: *0,3€*

Output:

Total cost of preharvest intervention (in €)

Postharvest step

Parameters :

- Cost of testing one sample: *50€*
- Cost of one Camembert: *3€*
- Cost of one Camembert after STEC is detected: *0,5€*

Output:

Cost of destroying a batch: *10 000 €* (20 000 cheeses)

+ Analytical cost

The Optimization problem

Objectives to minimize:

- Risk of HUS
- Total cost of intervention

Find optimal values of:

- Frequency de test du lait de ferme: f_{sort} days
- Seuil limite de contamination en E coli: l_{sort} CFU/ml
- Frequence de test d’un lot: p_{test}
- Nombres de fromages testés par lot: n_{sample}

plan d'expérience X appliqué pour les simulations est le suivant:

Total 5 x 5 x 5 x 5 = 625 points in X

- f_{sort} {10, 20, 30, 40, 50}
- l_{sort} {10, 20, 30, 40, 50}
- p_{test} {0.1, 0.2, 0.3, 0.4, 0.5}
- n_{sample} {5, 10, 20, 30, 50}

The Optimization problem

Objectives to minimize:

- Risk of HUS
- Total cost of intervention

Find optimal values of:

- Test frequency: f_{sort}
- Sorting limit: l_{sort}
- Test proportion: p_{test}
- Number of samples: n_{sample}

Multiobjective Optimization

- Les objectifs d’optimisation sont « contradictoires »
- Il est possible qu’il n’existe pas qu’une seule solution qui permet de minimiser les objectives

Estimation of Pareto set P

- Le « Pareto set » fournit le sous-ensemble de toutes les solutions optimales
- On appelle P ce sous-ensemble de X

Solution Pareto optimale

Objectives space:

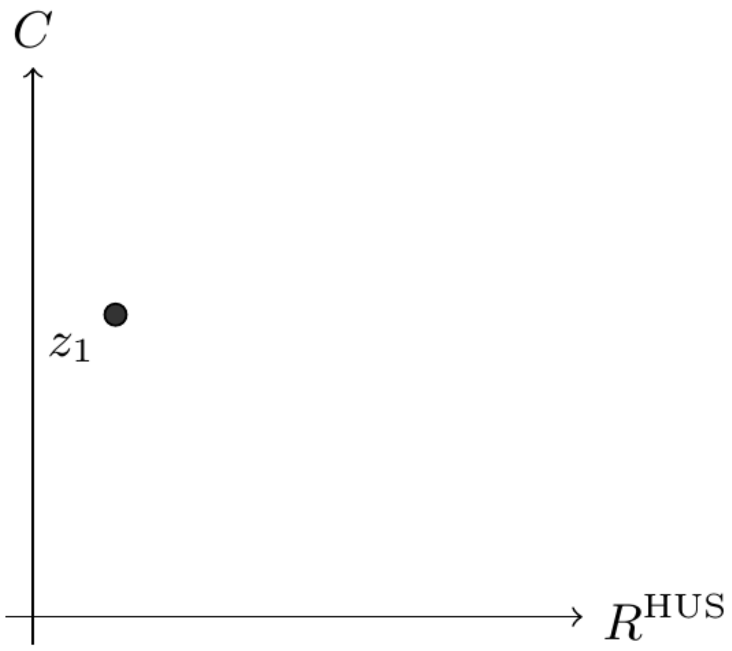
- Axe des abscisses X → Risque de SHU
- Axe des ordonnées Y → Cout des mesures de maitrise

Entrées:

- $f_{sort\_1} = 10$ *days*
- $l_{sort\_1} = 20$ *CFU/ml*
- $p_{test\_1} = 0.3$
- $n_{sample\_1} = 5$

Resultats:

- $z\_1 = (R_{HUS}\_1, C\_1)$



One observation $z_1 = (R_1^{\mathrm{HUS}}, C_1)$

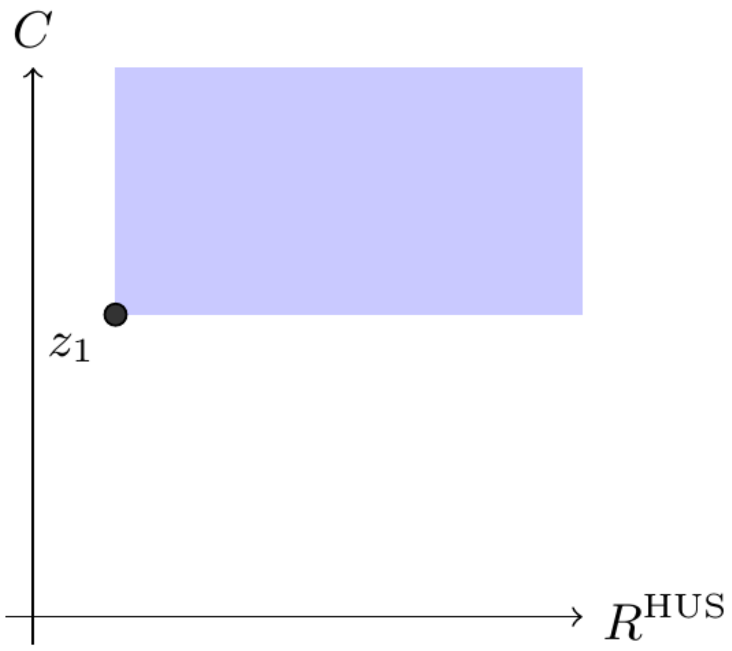
Solution Pareto optimale

Région dominée :

Région colorée dominée par *z_1*

Tout point dans cette région:

- soit est plus risqué
- soit plus coûteux
- Ou les deux



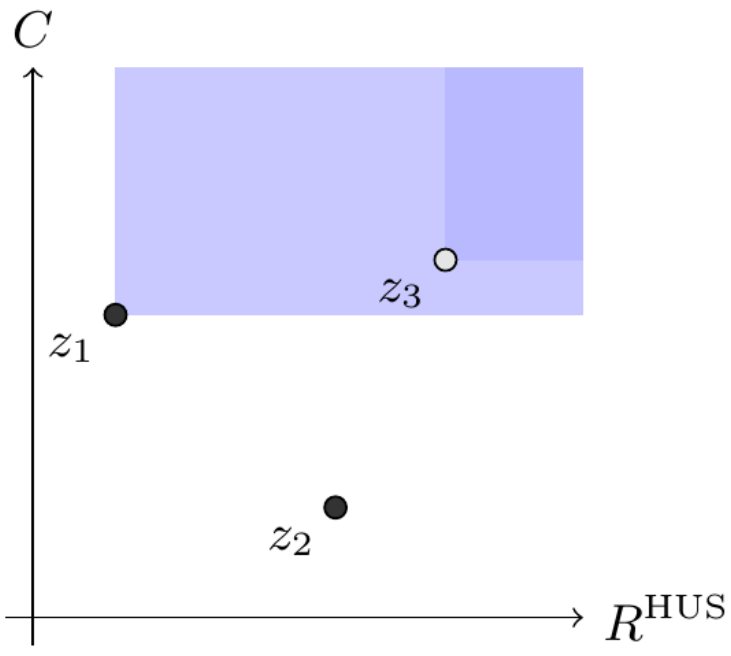
Dominated area by z_1 in objective space

Solution Pareto optimale

Soit deux autres entrées :

- *fsort _2 = 20 jours*
- *lsort _2 = 20 CFU/ml*
- *ptest_2 = 0.2*
- *nsample_2* = 5

- *fsort _3 = 30 jours*
- *lsort _3 = 20 CFU/ml*
- *ptest_3 = 0.5*
- *nsample_3* = 5



z_1 dominates z_3 but not z_2

Résultats correspondants :

- $z\_2 = (R_{HUS\_2}, C\_2)$
- $z\_3 = (R_{HUS\_3}, C\_3)$

z_1 et z_2 sont optimales

- $R_{HUS\_2} > R_{HUS\_1}$
- $C\_2 < C\_1$

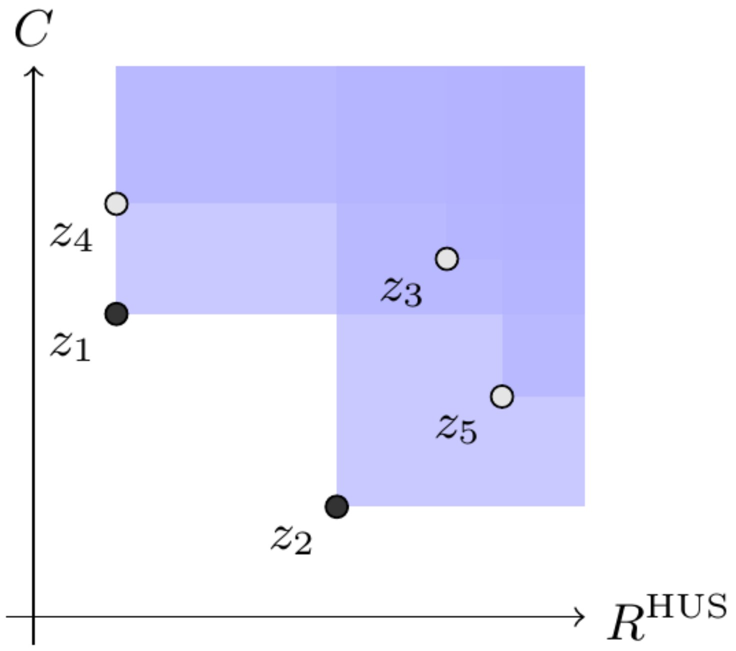
Solution Pareto optimale

Points Pareto optimaux :

- *z_1* , *z_2* est Pareto optimal

Points dominés :

- *z_3* , *z_4* , *z_5* n'est pas optimale



z_3, z_4 & z_5 dominated by z_1 and z_2

Solution Pareto optimale

Espace de départ X:

En tout 5 x 5 x 5 x 5 = 625 points dans X

- f_{sort} {10, 20, 30, 40, 50}
- l_{sort} {10, 20, 30, 40, 50}
- p_{test} {0.1, 0.2, 0.3, 0.4, 0.5}
- n_{sample} {5, 10, 20, 30, 50}

Espace d'arrivée:

- R_{HUS}: Relative risk of HUS
- C: Aggregated cost of intervention

Pareto Set P:

- POINTS OPTIMALE
- Points in X that are Pareto optimal

Pareto front:

- Points in objective space corresponding to P

Monte Carlo Estimate

Estimation du front de Pareto :

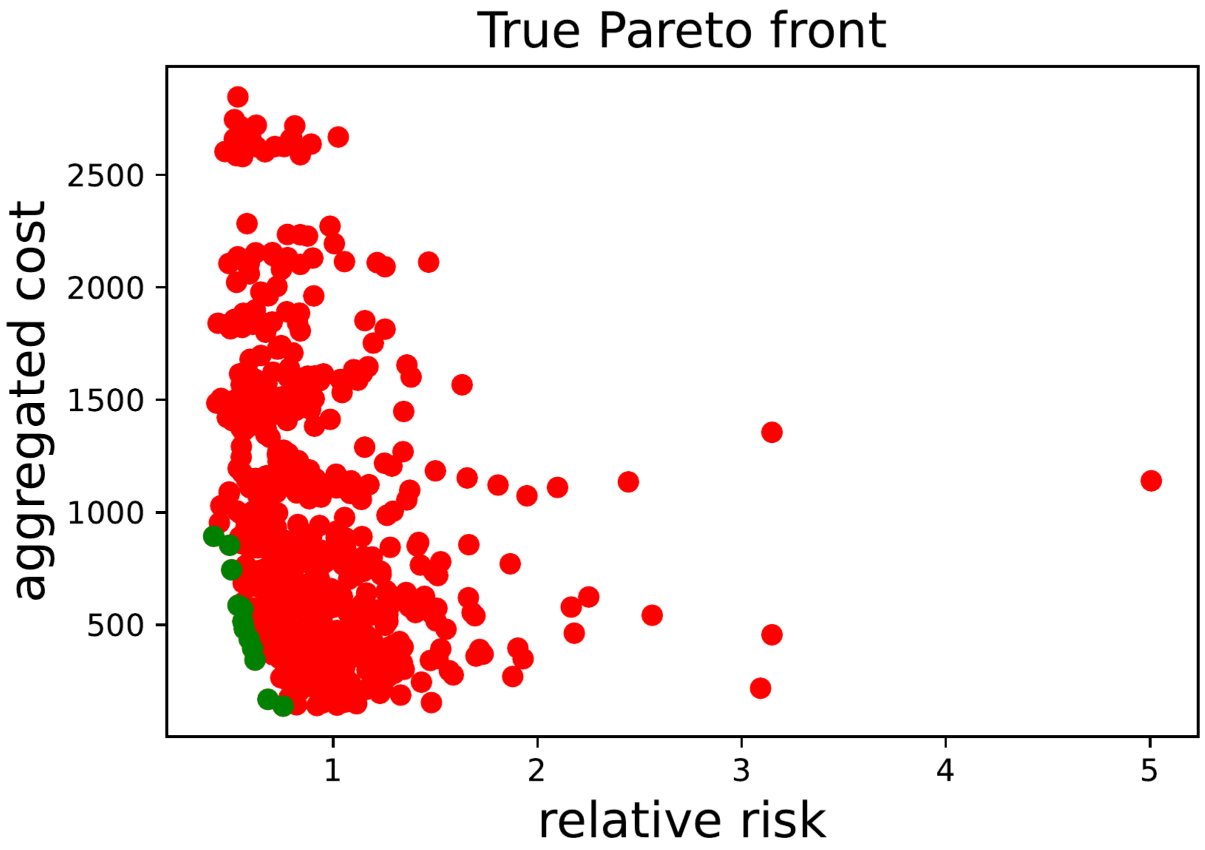
- Évalue le simulateur sur toutes les entrées
- Utilise un grand nombre de simulations
- Coûteux en termes de calcul

Exemple :

- Ça prend 4 jours !!
- Evalué sur 625 points d'entrée

Graphique :

- Points verts: Pareto optimal
- Points rouges : Dominé (non optimal)



PALS Estimate

Stochastic Pareto Active Learning:

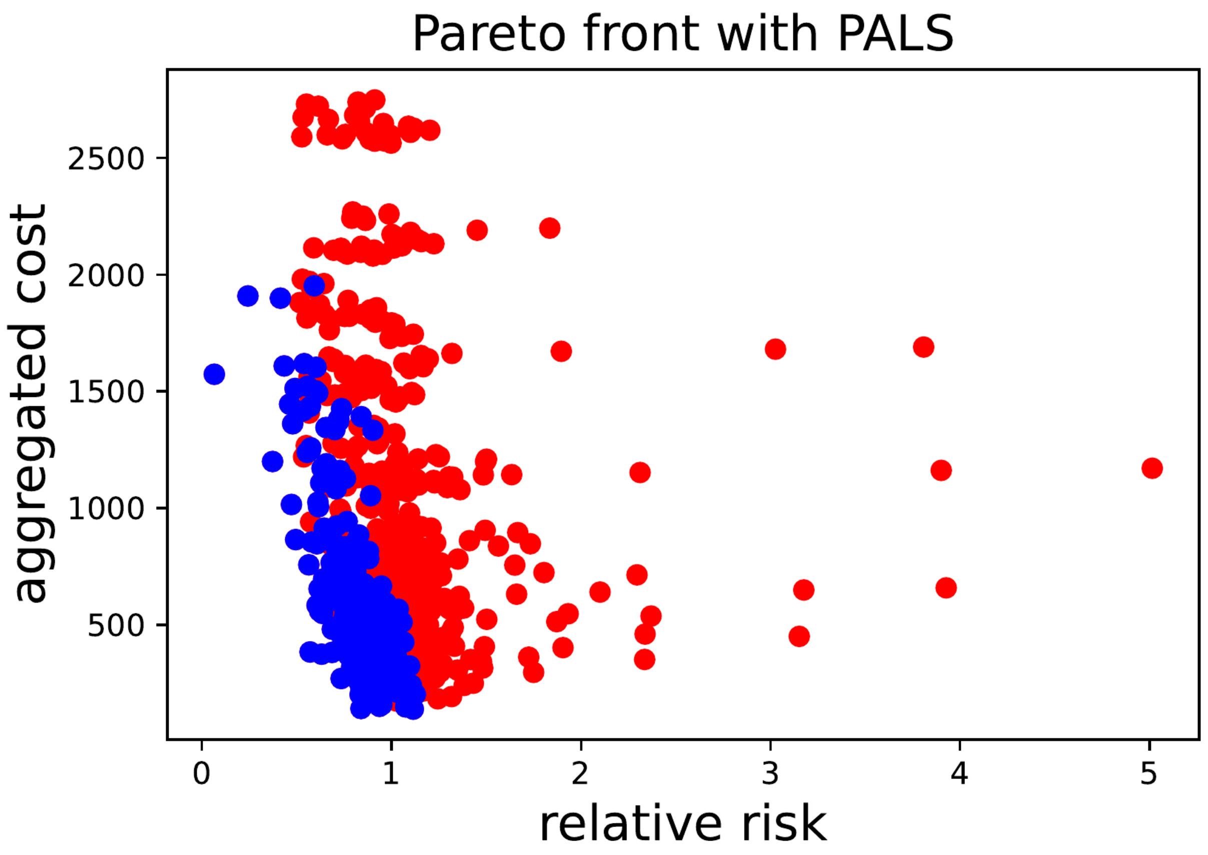
- Proposé par Barracosa et al (2021)
- Échantillonne astucieusement les points à évaluer
- Utilise un petit nombre de simulations
- Moins cher!

Algorithme PALS :

- Prend moins d'une heure (<1h)
- Evalué sur 100 points d'entrée

Graphique:

- Points bleus : Potentiellement Pareto optimal
- Points rouges : Dominé (pas optimal)



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. In SIAM Conference on Computational Science and Engineering (CSE21), Virtual Conference originally scheduled in Fort Worth, Texas, United States, Mar 2021.

WIP: Improving PALS

PALS tel que proposé par Barracosa et al (2021) n'est pas directement applicable à notre problème

- Le risque relatif n'est pas l'espérance du résultat du simulateur.
- Nous proposons d'utiliser les quantiles des quantités d'intérêt pour construire des rectangles de confiance.

Poster : Subhasish Basak, Julien Bect , Laurent Guillier , Fanny Tenenhaus, Aziza, Janushan Christy, Emmanuel Vazquez. Bayesian multiobjective optimization for quantitative risk assessment in microbiology. In PhD students day in the Annual meeting of GdR MASCOT-NUM research Network, June 2022, Clermont Ferrand, France.

PALS n'est pas toujours capable de classer les points lorsque les observations sont trop proches dans l'espace objectif

- Pour résoudre ce problème, nous proposons d'utiliser des simulations conditionnelles sur des rectangles de confiance pour classer les points

Intervention strategies: Cost

Preharvest step

Parameters:

- Cost for E. coli testing in milk from one farm: *50€*
- Cost of 1L of milk: *0,3€*

Output:

Total cost of preharvest intervention (in €)

Postharvest step

Parameters :

- Cost of testing one sample: *50€*
- Cost of one Camembert: *3€*
- Cost of one Camembert after STEC is detected: *0,5€*

Output:

Cost of destroying a batch: *10 000 €* (20 000 cheeses)

+ Analytical cost