

# COPIL ArtiSaneFood France

Subhasish Basak

Maison du Lait, Paris

Octobre 2, 2023

Agence Nationale de Sécurité Sanitaire (ANSES), Maison-Alfort, France  
Université Paris-Saclay, CNRS, CentraleSupélec, L2S, Gif-sur-yvette, 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](#).



## Ma thèse

- **Durée de contrat** : Janvier, 2021 - 2024
- **Encadrants** : E. Vazquez, J. Bect, L. Guillier, F. Tenenhaus-Aziza
- **Proposition de thèse** : Optimisation Bayésienne pour l'appréciation quantitative des risques en microbiologie.
- **Idée générale** :
  - Améliorer les mesures sanitaires dans la production fromagère
  - Trouver les paramètres optimaux pour les mesure de metrise
  - Réduire simultanément :
    - le risque de maladie pour les **consommateurs**
    - le coût des mesures sanitaires pour le **producteur**

## Direction de la recherche

- Construire un modèle pour L'Appréciation Quantitative des Risques
  - Pathogènes :
    1. STEC HP
    2. *Salmonelle*
    3. *Listeria monocytogenes*
  - Calcule le **risque** de maladie et le **coût** de l'intervention (mesures sanitaires)
- Construire un algorithme d'optimisation
  - Pour les fonctions **stochastiques** et **coûteuses en calcul**
  - Optimiser **Plusieurs** fonctions simultanément
  - Des fonctions **contradictaires**

## AQR - Appréciation Quantitative des Risques



## Modèle de base - AQR STEC



Figure 1: Modèle AQR STEC de fromage au lait cru pâte molle

## Risque - STEC

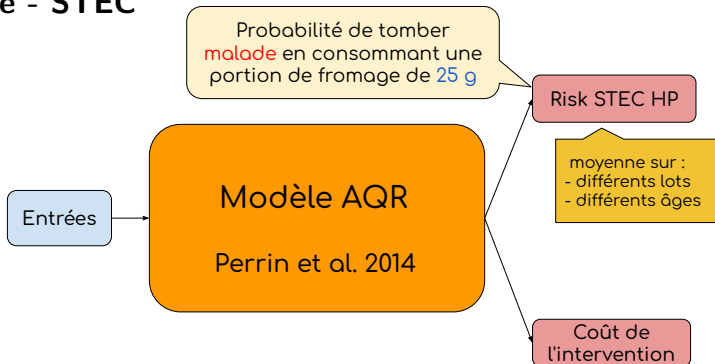


Figure 2: **Risque par lot** : 1 lot  $\sim$  20.000 fromage de 250g

## Côut - STEC

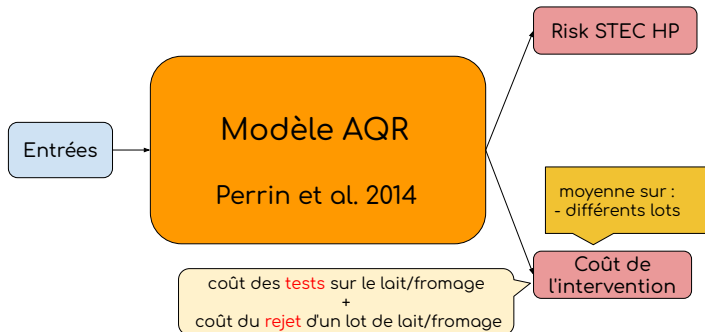


Figure 3: **Côut** correspondant à la fabrication d'un lot de fromage



## Modèle de base - AQR STEC



Figure 4: Modèle AQR STEC de fromage au lait cru pâte molle

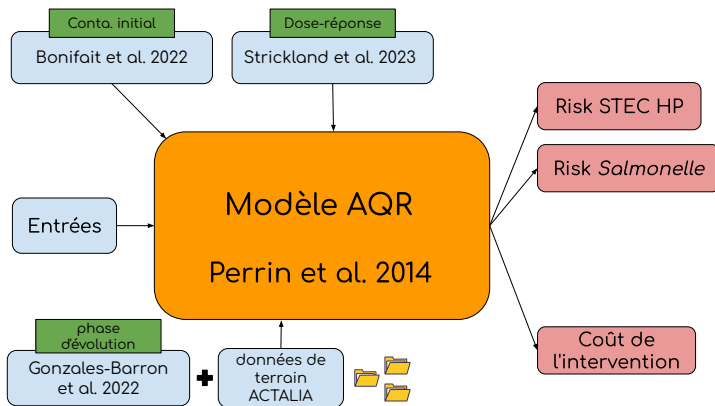


Figure 5: Modèle AQR *Salmonelle* entrées

## Modèle AQR - STEC + Salmonelle

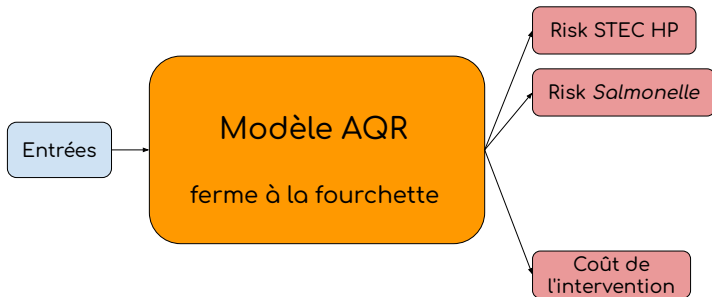


Figure 6: Modèle AQR STEC + *Salmonelle* de fromage au lait cru pâte molle

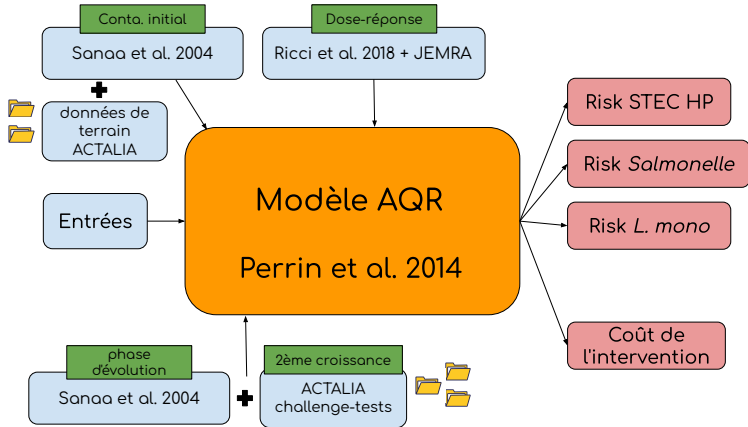


Figure 7: Modèle AQR *L. mono*. entrées

## Modèle multipathogènes : A quoi ça ressemble

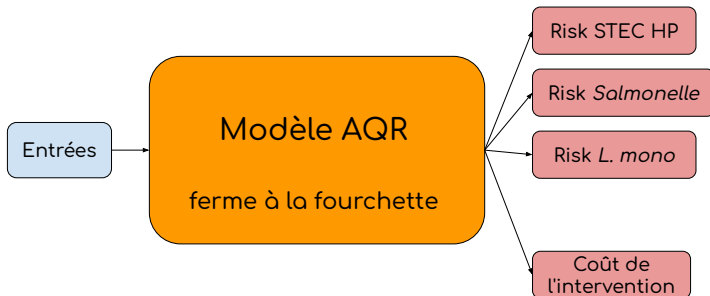


Figure 8: Modèle AQR multipathogènes - Entrées et les sorties

## Mesure de maîtrise

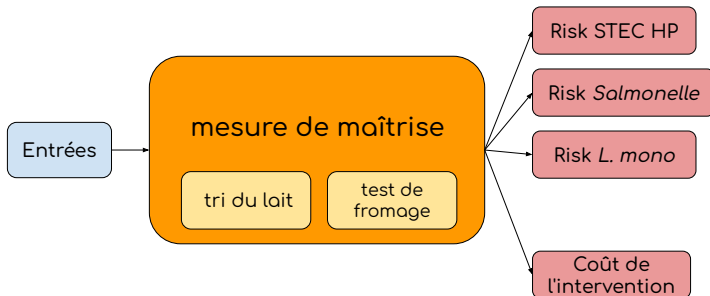


Figure 9: Modèle AQR - les processus d'intervention

## Mesure de maîtrise

### ■ Tri du lait

- lait de ferme sont testés pour la concentration de *E. coli*
- le lait est rejeté si la concentration dépasse un certain limite
  - \* la fréquence de tri
  - \* limite de concentration

### ■ Test de fromage

- les lots de fromages sont testés (entre J3 et J14)
- le lot est rejeté si une contamination est détectée (par l'un des trois pathogènes)
  - \* la proportion de lots testés
  - \* le nombre d'unités d'échantillonnage (25g)

## Mesure de maîtrise

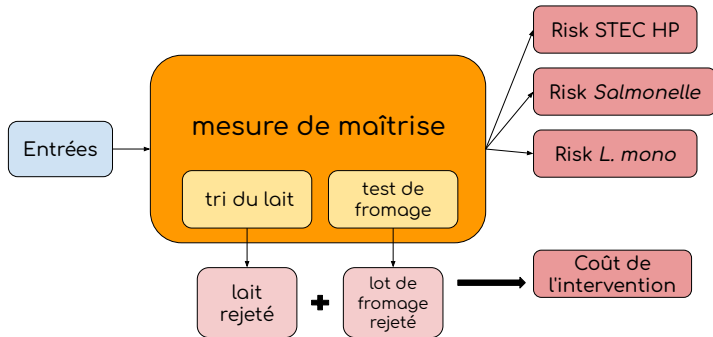


Figure 10: Coût de l'intervention = Coût analytique + Coût de la perte



## Les entrées à optimiser

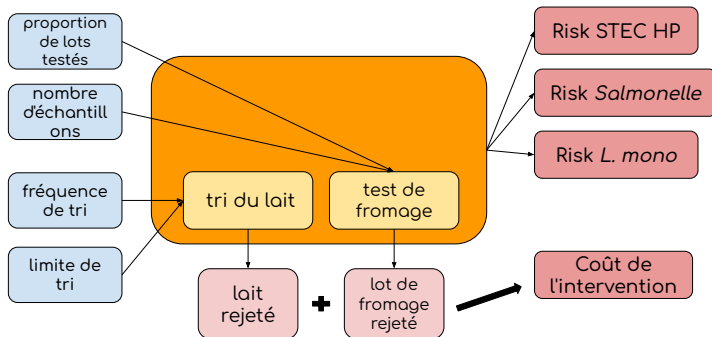
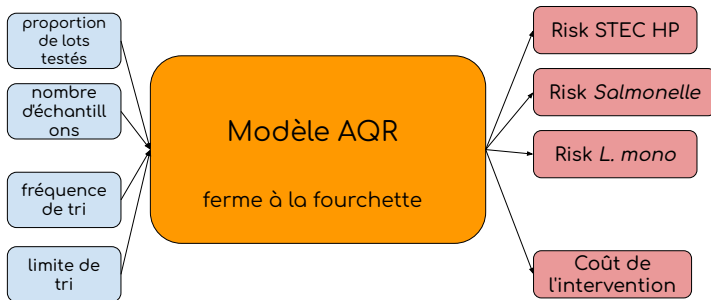


Figure 11: Les paramètres des processus d'intervention

## Résumer - modèle AQR



**Rémark** : On analyse un échantillon de 25 g pour trois pathogènes

## Le risque baseline

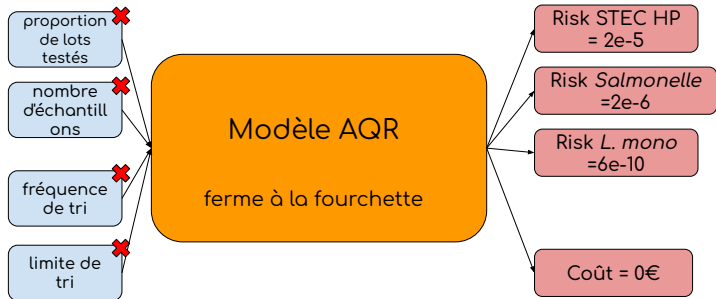


Figure 12: Scénario de base : aucune intervention n'est appliquée

## Scénario 1 : intervention

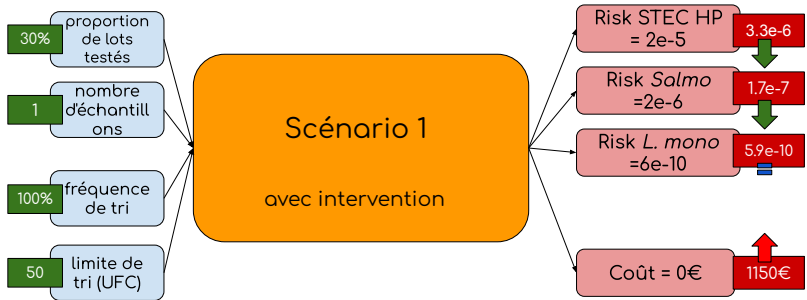


Figure 13: réduit le risque mais augmente le coût

## Une seule métrique pour 3 risques

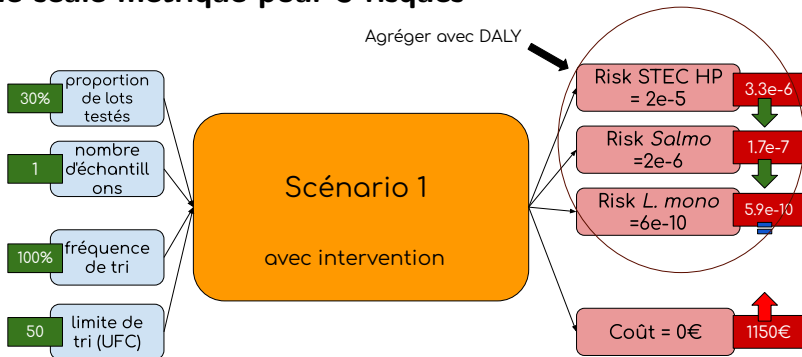


Figure 14: les différents risques peuvent être combinés par DALY

## Qu'est-ce que DALY ?

$$\text{DALY} \text{ Disability Adjusted Life Years} = \text{YLD} \text{ Years Lived with Disability} + \text{YLL} \text{ Years of Life Lost}$$

«Daly» est une mesure de la charge de morbidité globale, exprimée par le nombre cumulé d'années perdues en raison d'une mauvaise santé, d'un handicap ou d'une mort précoce.

Années vécues avec un handicap

Années de vie perdues



- L'idée est de convertir le risque en années de vie

## Calcul du DALY

**DALY** : Espérance de vie corrigée de l'incapacité (exprimées en années)

Pathogènes	YLL (1 cas)	YLD (1 cas)	DALY (1 cas)
STEC HP	0.0411	0.013	0.0541 années
<i>Listeria monocytogenes</i>	3.3	0.4	3.7 années
Non-typhoidal <i>Salmonella</i>	0.015	0.004	0.019 années

- Étape 1 : Pour pathogène (x) calculer,

$$- P_x[\text{observer 1 an de DALY}] = \frac{P[\text{observer 1 cas de maladie}]}{\text{DALY}_x} = \frac{\text{Risque}_x}{\text{DALY}_x}$$

- Étape 2 : Calculer  $P[1 \text{ an de DALY}]$  causé par l'un des 3 pathogènes

## DALY : Scénario de base

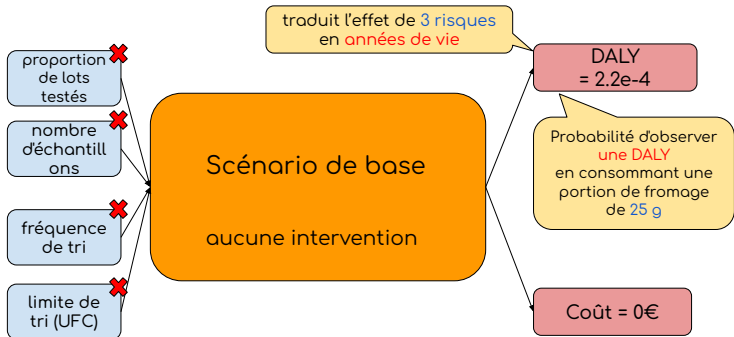


Figure 15: DALY et coût pour scénario baseline



## Scénario 1 : intervention

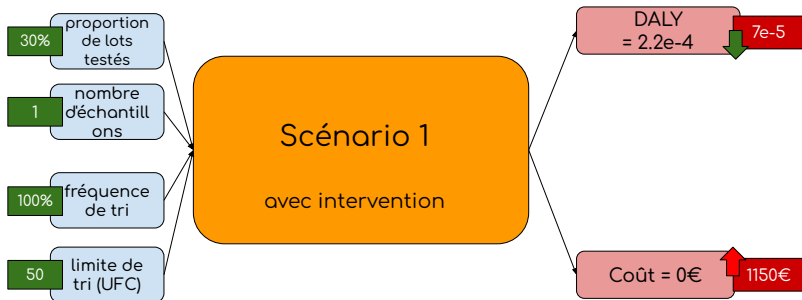


Figure 16: réduit le DALY mais augmente le coût

## Scénario 2 : augmenter le nombre d'échantillons

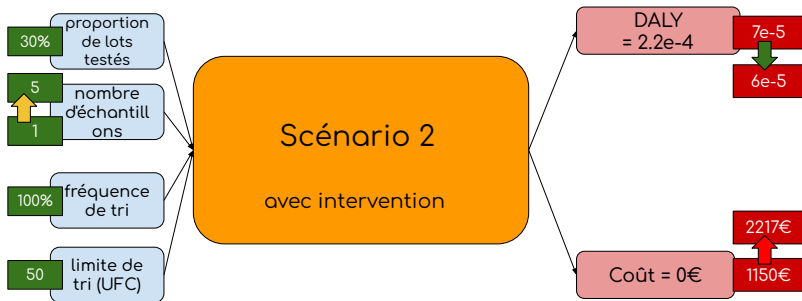


Figure 17: réduit le DALY mais augmente le coût

## Scénario 3 : diminuer la limite de tri

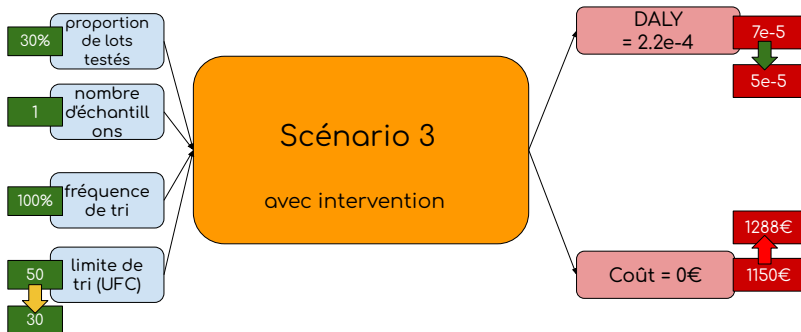


Figure 18: réduit le DALY mais augmente le coût

## Scénario 4 : modifier simultanément deux paramètres

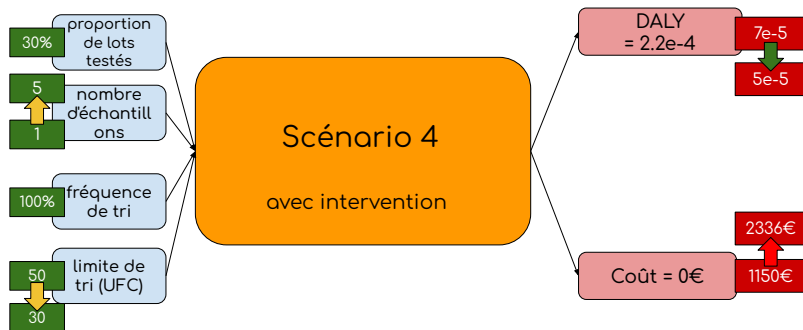
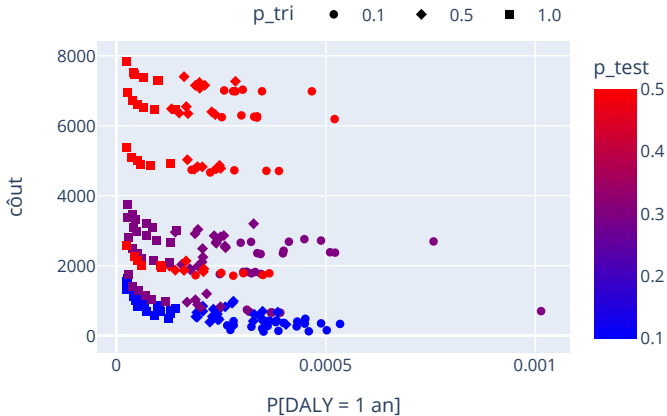


Figure 19: le DALY et le coût sont des objectifs **contradictoires**

## Plusieurs scénarios

- On peut imaginer de nombreux scénarios avec tous les paramètres variables:
  - Proportion de lots testé  $p^{\text{test}}$  : {10%, 30%, 50%}
  - n° d'échantillons  $n^{\text{sample}}$  : {1, 5, 10, 15}
  - Fréquence de tri  $p^{\text{tri}}$  : {10%, 50%, 100%}
  - Limite de tri  $l^{\text{tri}}$  : {10, 20, 30, 50, 100, 200}
- Construire un scénario particulier  $\rightarrow \{p^{\text{test}}, n^{\text{sample}}, p^{\text{tri}}, l^{\text{tri}}\}$
- Toutes les combinaisons possibles avec ça : 216
- Idée : trouver les scénarios qui réduisent à la fois le DALY et le coût

## DALY vs coût pour diff. scénarios (STEC + Salmo + L. mono)



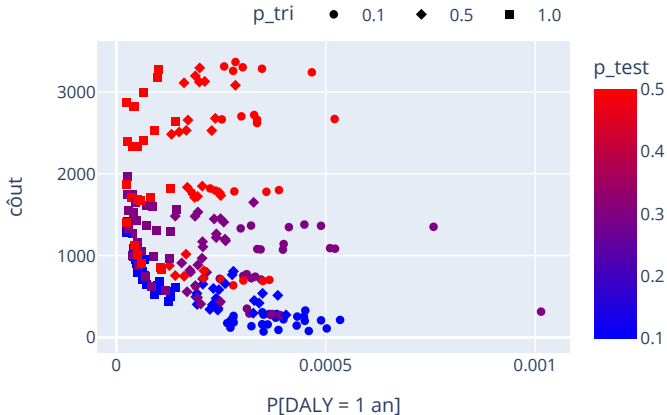
## Redéfinir le plan d'échantillonnage ? (basé sur virulence)

- Le test du fromage n'a pas une efficacité significative pour réduire le DALY
- Une forte prévalence favorise la détection, mais seulement une petite proportion des agents pathogènes hautement virulent provoque la maladie

Prévalence (pourcentage d'une portion de 25 g de fromage avec au moins un colonie au moment de la consommation) dans un scénario de référence

Symbol	Mean	Median	SD	$q_{0.05}$	$q_{0.95}$
MPS STEC	1.97	0.16	8.21	0	7.43
HV <i>Salmonella</i>	0.37	0.02	3.22	0	0.91
LV <i>Salmonella</i>	0.75	0.06	4.67	0	2.27
<i>Listeria</i>	39.47	37.03	21.97	8.37	80.03

## DALY et coût pour différents scénarios (STEC + Salmo)





## Contribution

- Une version modifiée du modèle STEC

*Subhasish Basak et al. Quantitative risk assessment of haemolytic and uremic syndrome (hus) from consumption of raw milk soft cheese. Food and Ecological Systems Modelling Journal, (under review).*

- Implémentation avec R et FSKX
- **Approche bayésienne** : Utilisation des données ACTALIA pour estimer les paramètres d'hygiène des fermes
- Un modèle multipathogen (STEC + *Salmonelle* + *Listeria*)

*Subhasish Basak et al. Minimizing risk of illness and analytical costs using a QMRA model for raw milk cheeses. ICPMF'12, Sapporo, Japon (oral).*

## Optimisation multiobjectif



## Différents scénarios d'intervention

- Considérer le problème de l'optimisation pour STEC
- Espace de paramètres

$$p^{\text{tri}} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$$

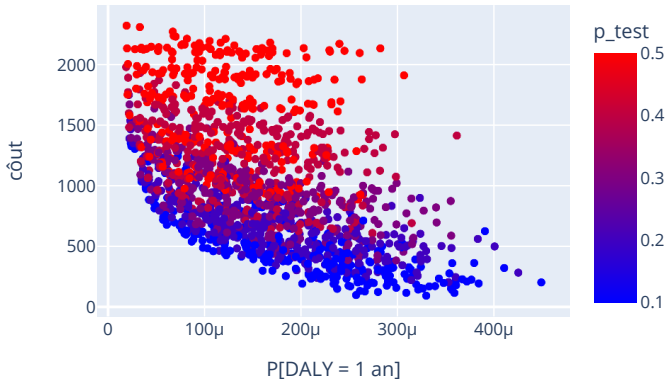
$$l^{\text{tri}} \in \{10, 20, 30, 50, 100, 200\}$$

$$p^{\text{test}} \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$$

$$n^{\text{sample}} \in \{3, 5, 8, 10, 12\}$$

- Toutes les combinaisons possibles donnent 1500 scénarios possible
- Pour chaque scénario le DALY et le coût sont calculés
- Comment trouver les scénarios qui réduisent à la fois le DALY et le coût ?

## DALY vs coût pour diff. scénarios (STEC)



## Comment choisir les scénarios optimaux ?

- Nous utilisons le concept d'optimalité de Pareto
- L'idée est d'obtenir un compromis efficace entre les différents objectifs plutôt qu'une solution unique
- Il cherche des solutions où aucune amélioration n'est possible dans un objectif sans détériorer un autre
- Les solutions Pareto-optimales sont non dominées, offrant un compromis efficace entre les objectifs
- Elles forment le front de Pareto, et les décideurs choisissent la meilleure solution selon leurs préférences

## Solutions Pareto optimale

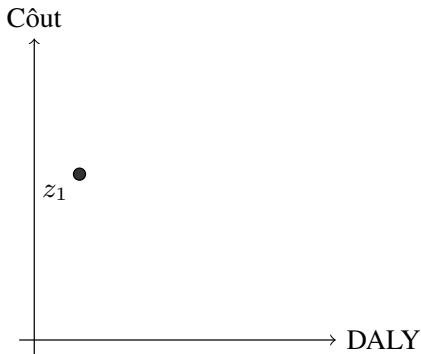


Figure 20: Un scénario particulier :  $z_1$

## Solutions Pareto optimale

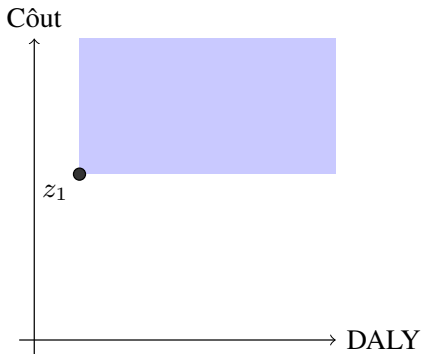


Figure 21:  $D(z_1)$  : Région dominée par  $z_1$

## Solutions Pareto optimale

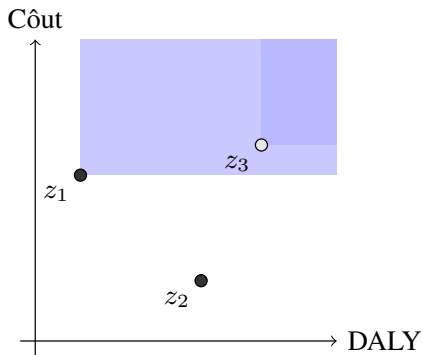


Figure 22: Tout autre point en  $D(z_1)$  aura un coût ou un DALY plus élevé



## Solutions Pareto optimale

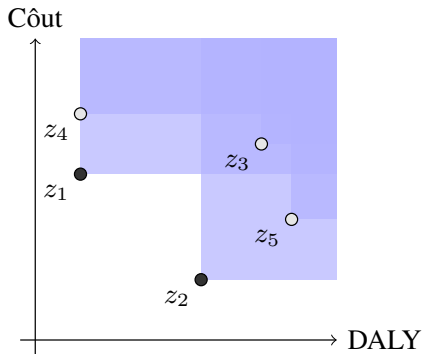


Figure 23:  $z_3$ ,  $z_4$  and  $z_5$  sont dominée par  $z_1$  et  $z_2$

## Le front Pareto est dans la **frontière**

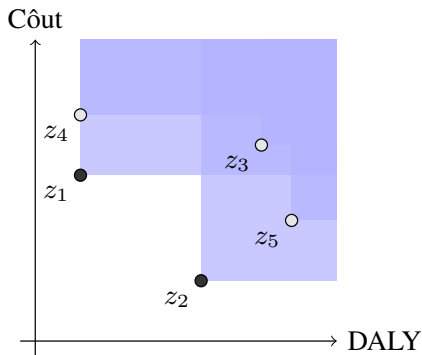
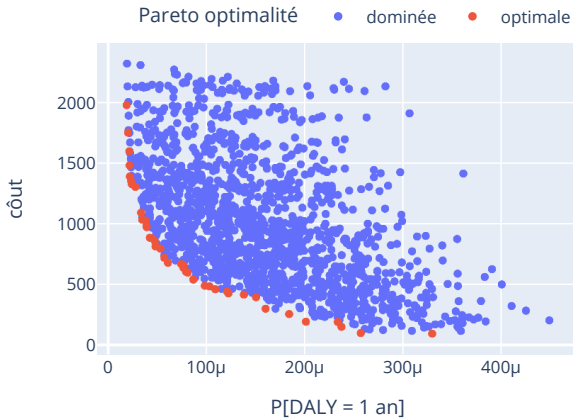


Figure 24:  $z_1$  et  $z_2$  sont Pareto optimale

## Le front de Pareto pour STEC



## Direction de recherche et contribution

- 3 jours de calcul pour évaluer tous les 1500 points
- Proposer un algorithme d'optimisation pour un gagner du temps
- Nos contributions :
  - Nous avons proposé une extension de l'algorithme PALS par Barracosa et al. (2021)  
*Subhasish Basak et al. Bayesian multi-objective optimization for quantitative risk assessment in microbiology. MASCOT-NUM 2022, June 2022, Clermont-Ferrand, France.*
  - Nouvel algorithme : basé sur le principe SUR et w-IMSE  
*Subhasish Basak et al. Bayesian multi-objective optimization for quantitative risk assessment in microbiology. MASCOT-NUM 2023, April 2023, Le Croisic, France.*
- Avec l'algorithme : 4h de calcul avec une précision de 96%

## Calendrier des travaux

- **Fin de thèse** : Decembre 2023
- **Rédaction de manuscript** : En cours ...
  - à soumettre fin octobre
- **Soutenance** : Janvier 2024
- **Projet d'article scientifique**
  - Papier journal sur modèle multipathogene : **Microbiological Risk Assessment Journal - Special Issue - ICPMF'12**

**Merci pour votre attention!**

