Quantitative risk assessment and optimization of process intervention parameters for French raw milk soft cheese

Subhasish Basak, Julien Bect, Laurent Guillier, Fanny Tenenhaus-Aziza, Janushan Christy, Emmanuel Vazquez

To cite this version:
Subhasish Basak, Julien Bect, Laurent Guillier, Fanny Tenenhaus-Aziza, Janushan Christy, et al.. Quantitative risk assessment and optimization of process intervention parameters for French raw milk soft cheese. IDF World Dairy Summit 2022, Sep 2022, New Delhi, India. hal-04017197

HAL Id: hal-04017197
https://centralesupelec.hal.science/hal-04017197
Submitted on 7 Mar 2023

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 Attribution - NonCommercial - ShareAlike 4.0 International License
The aim is to find the optimal values of the process QRA based on model proposed by Subhasish Basak. Quantitative Risk Assessment (QRA) strategies, in order to “economically” reduce the risk of post-harvest module. Farms with cheese module. Consumer module. The primary goal is to establish efficient intervention strategies, in order to “economically” reduce the risk of Haemolytic Uremic Syndrome (HUS) caused by Shiga-Toxin producing Escherichia coli (STEC) present in raw-milk soft cheese. 

Intervention strategies in cheese making: 
• Pre-harvest milk sorting: STEC and E. coli strains follow same fecal route! 
  • A bulk tank of milk is tested with probability \( p_{test} \).
  • Farms with E. coli conc. \( > P^{opt} \) are rejected
  \( C_{r_test} \) : Cost of testing and rejecting bulk tank milk
• Post-harvest cheese sampling: 
  • A batch of cheese is tested with probability \( p_{test}^{cheese} \).
  • From a single batch 100% cheese are tested for presence of STEC
  \( C_{r_test} \) : Cost of testing and rejecting cheese batches

The aim is to find the optimal values of the process intervention parameters \( \{ C_{p_test}, p_{test}, p_{test}^{cheese}, n_{sample} \} \), that minimize the risk of HUS and the costs \( C_{r_test}^{cheese} \).

Quantitative Risk Assessment (QRA) 

QRA based on model proposed by Subhasish Basak. 

• Farm module + Pre-harvest step
  STEC conc. \( Y_{STEC}^{milk} \) in farm milk is computed
  \( Y_{STEC}^{milk} = Y_{STEC}^{milk} \) in farm milk
• Cheese module
  Evolution of STEC is modeled with ODEs
  \( \frac{dY}{dt} = \mu_{Y}(t) - \gamma(t) \cdot \left( Y - \frac{y(t)}{C_{max}} \right) \)
  STEC cells form colonies (clusters) inside cheese
  \( No. \) of colonies (Poisson): \( N_{colony} \)
  \( Size \) of colonies (LogNormal): \( Y_{colony} \)
• Consumer module
  Batch risk is computed using a dose-response model:
  \( \Gamma = \sum_{x} x_{colony} x_{colony} \)
  \( R_{batch} = \sum_{x} \int Y_{STEC}^{milk} \cdot p(Y_{test}) \cdot p(Y_{colony}) \) averaging over consumer age
  \( Post-harvest \) module
  Proportion of rejected batches \( p_{batch}^{rej} \) is computed
  \( p_{batch}^{rej} = \Phi \left( \frac{\Gamma - n_{sample}}{\gamma} \right) \)

Bi-objective optimization 

We consider the bi-objective optimization problem
\[ \min_{x \in X} f(x) \]
where, \( f = (f_1, f_2) \)
• Not necessarily has a unique solution \( x^{opt} \) in \( X \), in presence of conflicting objectives
• The solution set \( P \) consists of Pareto optimal points
\[ P = \{ x \in \mathbb{R}^2 : \forall f_1(x) \prec f_1(y), \forall f_2(x) \prec f_2(y) \} \]
• EVPI (Expected Value of Perfect Information) 
  EVPI = \( E \left( \max_{x \in \mathbb{R}^2} V(x) - \min_{x \in \mathbb{R}^2} V(x) \right) \)

PALS 

Optimization of the QRA simulator 

• It is stochastic and computationally expensive
• Gradient based optimization is not feasible
• Thus we rely on Bayesian approaches
• Pareto Active Learning for Stochastic simulators proposed by Subhasish Basak and extended by Subhasish Basak.
• It uses Gaussian process regression for approximating the simulator function
• Estimates \( P \) by classifying each point in \( X \) as Pareto optimal, Non-Pareto optimal and Unclassified

Experimental results 

• Minimizing \( f \) over the input space \( X \)
  \( n_{sample} \) \( \in \{ 5, 10, 20, 30, 50 \} \)
  \( P_{test}^{cheese} \) \( \in \{ 10, 20, 30, 40, 50 \} \)
  \( P_{test}^{milk} \) \( \in \{ 1/10, 1/20, 1/30, 1/40, 1/50 \} \)
  \( P_{test}^{cheese} \) \( \in \{ 0.1, 0.2, 0.3, 0.4, 0.5 \} \)
• True Pareto front: estimated using 5000 samples for each of 5 \times 5 \times 5 \times 5 = 625 input points

• Pareto optimal (green) and dominated points (red)
• Pareto front estimated using PALS
• Initial design size = 60, evaluation budget = 40
• batch size per iteration = 300

• With PALS using significantly less (\( 100 \times 300 \)) evaluations, the user can provide the following insights
  Most of the dominated (red) points are well classified
  The points corresponding to \( P \) remain unclassified (blue)

References 