



HAL
open science

Modélisation de la contamination par *Listeria monocytogenes* pour l'amélioration de la surveillance dans les industries agro-alimentaires

Natalie Commeau

► **To cite this version:**

Natalie Commeau. Modélisation de la contamination par *Listeria monocytogenes* pour l'amélioration de la surveillance dans les industries agro-alimentaires. Alimentation et Nutrition. AgroParisTech, 2012. Français. NNT : 2012AGPT0043 . pastel-00770790

HAL Id: pastel-00770790

<https://pastel.hal.science/pastel-00770790>

Submitted on 7 Jan 2013

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.



Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

**L'Institut des Sciences et Industries
du Vivant et de l'Environnement**

(AgroParisTech)

Spécialité : Biostatistique

présentée et soutenue publiquement par

Natalie COMMEAU

le 4 juin 2012

**Modélisation de la contamination par *Listeria monocytogenes* pour
l'amélioration de la surveillance dans les industries agro-alimentaires**

Directeur de thèse : **Marie CORNU**
Co-encadrement de la thèse : **Eric PARENT**

Jury

Mme Chantal GUIHENNEUC, Professeur, EA 4064, Université Paris Descartes
Mme Annemie GEERAERD, Associate professor, MeBioS, KU Leuven
M. Jean-Baptiste DENIS, Directeur de recherche, Miaj, INRA
M. Jean-Christophe AUGUSTIN, Maître de conférences, MASQ, ENVA
Mme Marie CORNU, Chargée de recherche, LM2E, IRSN
M. Eric PARENT, ICPEF, MIA 518, AgroParisTech

Rapporteur
Rapporteur
Président
Examineur
Directrice de thèse
Codirecteur de thèse

Remerciements

Avant de commencer à présenter mon travail de thèse, même en résumé, je tiens à remercier toutes les personnes qui m'ont accompagnées, de près ou de loin, durant ces années. Un doctorant seul dans son coin ne peut pas grand-chose, c'est par les échanges, les collaborations et les discussions qu'une thèse se construit.

Pour commencer, je tiens à remercier chaleureusement mes deux directeurs de thèse, Marie Cornu et Eric Parent. Merci à Marie pour son enthousiasme et son énergie, pour m'avoir accompagnée et aidée lorsque je me trouvais en difficulté et d'avoir trouvé du temps pour moi, même après son déménagement dans le Sud de la France et son changement de travail. Ce n'était pas facile de me guider dans le vaste domaine de la modélisation en microbiologie alors que le travail professionnel est tout autre. Marie a su rester disponible malgré un emploi du temps très chargé. Merci également pour le projet ANR Quant'HACCP qui m'a permis de côtoyer des scientifiques d'horizons très variés et de découvrir des domaines que je n'aurais pas abordés sans ce projet. Merci à Eric pour m'avoir fait découvrir le monde des statistiques bayésiennes. Ses nombreuses idées et sa très grande culture dans ce domaine ont enrichi ma thèse et ont permis de débloquent de nombreux problèmes. Merci également pour les discussions m'ayant permis de prendre du recul sur les pratiques statistiques des microbiologistes. Je le remercie enfin pour m'avoir toujours suivie dans mes travaux, même lorsque les données atteignaient une proportion inquiétante de zéros.

Merci également à Laurent Rosso qui m'a orientée vers la modélisation en microbiologie des aliments lorsque je prospectais pour un doctorat et grâce à qui j'ai rencontré mes futurs directeurs de thèse.

Je tiens aussi à remercier Chantal Guihenneuc-Jouyaux et Annemie Geeraerd pour avoir accepté de rapporter ce travail. Merci également à Jean-Christophe Augustin pour avoir bien voulu participer au jury et pour avoir suivi mon travail lors des différents comités de pilotage. Merci à Jean-Baptiste Denis avoir présidé mon jury, pour m'avoir donné de bons conseils durant ses années de thèse et pour les discussions scientifiques par mails.

Je tiens à exprimer toute ma gratitude aux autres membres de mon comité de pilotage : Judith Rousseau, Karine Boquet, Annie Beaufort, Marie-Bénédicte Peyrat, Billy Amzal et Marie-Laure Delignette-Muller pour m'avoir soutenue et accompagnée. Merci également à Billy pour m'avoir fait découvrir le monde de l'expertise au niveau européen en m'ayant fait participer à un groupe de travail à l'EFSA sur le critère microbiologique

des salmonelles dans le poulet. Le travail pratiqué au niveau européen est vraiment différent de ce que j'ai pu faire dans mon travail de recherche en France. Cette découverte m'a beaucoup apportée. Merci à Marie-Laure pour ses conseils notamment pour le deuxième article que j'ai rédigé.

Je souhaite ma profonde reconnaissance à Véronique Zuliani et à Pascal Garry de l'IFIP pour leur savoir sur les pathogènes dans la viande de porc et sur leur connaissance du monde des fabricants industriels de lardons.

Merci à Isabelle Albert pour son aide pour mon premier article en premier auteur.

Merci à Elise Billoir pour la modélisation de la contamination de *L. monocytogenes* dans les lardons qui m'a été bien utile et pour les quelques jours de tourisme à Washington.

Merci à Sophie Ancelet pour avoir en grande partie rédigé le rapport du SP2 pour Quant'HACCP et pour l'organisation des soirées pour jeunes statisticiens. J'espère qu'il y en aura encore beaucoup d'autres à l'avenir.

La vie d'un doctorant ne serait pas la même sans la vie au laboratoire et surtout les discussions avec les différentes personnes les composant. J'ai eu la chance d'être intégrée à la vie de deux laboratoires, ce qui m'a permis de découvrir deux univers scientifiques et deux manières de travailler.

A l'ANSES, merci tout particulièrement au personnel des unités MOB et MAHY pour les discussions sympathiques et animées à la cantine ou autour d'un café (enfin, à petites doses le café). Merci à Laurent Guillier pour nos nombreuses discussions sur la modélisation en microbiologie des aliments mais aussi sur l'art d'avoir un potager. Merci à Annie Beaufort, Hélène Bergis et Sylvie Rudelle pour leurs connaissances des pathogènes, pour m'avoir aidé à trouver des données dans les anciens cahiers de résultats et surtout, pour leur patience lorsque je venais demander pour la n-ème fois de consulter la norme ISO 11290... Merci à Patricia Ng pour m'avoir montré plusieurs fois le travail à la paillasse. Merci à Anne-Laure Lardeux pour les conversations sur la grossesse. Bienvenue dans le monde pas toujours simple des jeunes parents! Merci à Véronique Noël et Elisabeth Morelli pour les informations sur le sous-projet 3 de Quant'HACCP. Merci enfin à Séverine Jaloustre pour avoir partagé les hauts et les bas de la vie d'un doctorant, pour l'ambiance dans le bureau lors des (rares) jours où nous étions présentes toutes les deux et bien sûr pour notre collaboration sur les modèles de température dans les cellule de refroidissement. Je lui souhaite bon courage pour sa nouvelle vie d'après thèse.

A AgroParisTech, merci aux membres des équipes MORSE et Stat & génomes pour leur accueil chaleureux. Merci à mes co-bureaux successifs : Alain Célisse, Marie Morfin puis Sébastien Li-Thiao-Té et enfin Antoine Channarond et Alice Cleynen. Merci à Tristan pour l'appel (muet) à la pause de l'après-midi, à Marie pour son énergie et pour le ressenti de jeune maman, à Aurore pour ses chouquettes, à Michel et Jean-Benoît pour les cours de C++, à Stéphane, Liliane, Gabriel, Julie, Emilie, Marie-Laure et à tous les autres.

Grand merci à Vanessa, Francine et Sophie pour l'aide administrative efficace apportée, pour les ordres de mission et les billets de train à commander avec des délais courts.

Merci à mes parents pour m'avoir accompagnée et soutenue durant mes (longues) années d'études. La vie d'un doctorante est semée de doutes et de programmes qui ne tournent pas et il est précieux de disposer de personnes auprès de qui se plaindre et qui, surtout, acceptent les appels téléphoniques après 22 h.

Merci à Gabriel pour avoir chamboulé ma troisième année de thèse et l'avoir même étirée pour en faire une année double. Ton arrivée n'a pas simplifié la rédaction du manuscrit mais elle m'a fait découvrir un bonheur insoupçonné.

Enfin, merci à toi Kristopher pour le soutien logistique surtout lors des périodes grosses charges de travail, pour ton aide en informatique surtout pour l'élaboration des figures et pour m'avoir soutenue de manière indéfectible durant toutes ces années. Je profite de cette dernière ligne pour te dire à quel point cela m'a été précieux.

Résumé / Abstract

Résumé

Les industriels du secteur agro-alimentaire sont responsables de la qualité des produits mis sur le marché. Un moyen de vérifier cette qualité consiste à déterminer la distribution de la contamination. Dans cette thèse, nous avons utilisé des données portant sur *L. monocytogenes* durant le procédé de fabrication de lardons et du saumon fumé. Nous avons ensuite élaboré des modèles hiérarchiques pour décrire la concentration en prenant ou non en compte diverses variabilités, nous avons estimé les paramètres par inférence bayésienne, puis comparé leur capacité à simuler des données proches des observations. Nous avons également comparé l'estimation de paramètres par inférence fréquentiste sur deux modèles en utilisant les données brutes issues des analyses microbiologiques et ces mêmes données converties en concentration. Par ailleurs, nous avons amélioré un modèle décrivant le devenir de *L. monocytogenes* au cours de la fabrication des lardons.

Le plan d'échantillonnage permettant d'estimer la qualité des produits, nous avons appliqué la théorie de la décision aux couples *L. monocytogenes* / lardons et *L. monocytogenes* / saumon fumé en sortie usine pour déterminer la taille optimale de l'échantillon à prélever par lot de manière à minimiser les coûts moyens supportés par le fabricant. Enfin, nous avons comparé plusieurs plans d'échantillonnage de mesure de la température d'un plat en sauce fabriqué dans une cuisine centrale et placé dans une cellule de refroidissement rapide. L'objectif était de sélectionner le meilleur plan d'échantillonnage en fonction du risque admissible pour le gestionnaire quant à la croissance de *C. perfringens*.

Mots clé : modèles hiérarchiques, inférences bayésienne et fréquentiste, théorie bayésienne de la décision, plans d'échantillonnage, microbiologie des aliments, détection, dénombrement

Abstract

Food business operators are responsible for the quality of the products they sell. A way to assess the safety of food is to determine the contamination distribution. During my PhD thesis, we used data about *L. monocytogenes* during the process of diced bacon and of cold smoked salmon. Then, we constructed several hierarchical models to describe contamination taking or not into account several kinds of variability such as between batches variability. We compared the capacity of each model to simulate data close to the observed ones. We also compared the parameters assessment by frequentist inference using raw data (the results of the microbiological analyses) and concentration-like data. In addition to the models describing the contamination at one step of the process, we improved an existing model describing the fate of *L. monocytogenes* throughout the diced bacon process.

A tool to assess the quality of a product is the sampling plan. We applied the Bayesian theory of decision to the pairs *L. monocytogenes*/diced bacon and *L. monocytogenes*/cold smoked salmon at the end of the process to determine the optimal size of a sample analysed per batch so that the average cost for the manufacturer is as low as possible. We also compared several sampling plans of temperature measurement of a meal cooked in an institutional food service facility and put in a blast-chiller just after cooking. The aim was to select the best sampling plan regarding the risk of *C. perfringens* growth that the manager is ready to take.

Keywords : hierarchical models, Bayesian and frequentist inferences, Bayesian theory of decision, sampling plans, food microbiology, detection, enumeration

Laboratoires d'accueil

INRA/AgroParisTech

UMR 518 MIA

Institut national de recherche agronomique (INRA) / AgroParisTech

16 rue Claude BERNARD

75 231 Paris cedex 05

ANSES

Laboratoire de Sécurité des Aliments

Agence Nationale de Sécurité Sanitaire de l'alimentation, du travail et de l'environnement
(ANSES)

23 avenue du Général de Gaulle

94706 Maisons-Alfort Cedex

Publications scientifiques issues du travail de thèse

Articles publiés

Commeau, N., Parent, E., Delignette-Muller, M.-L., and Cornu, M. (2012b). Fitting a lognormal distribution to enumeration and absence/presence data. *International journal of food microbiology*, 155 :146–152.

Commeau, N., Cornu, M., Albert, I., Denis, J.-B., and Parent, E. (2012a). Hierarchical bayesian models to assess between- and within-batch variability of pathogen contamination in food. *Risk Analysis*, 32(3) :395–415.

Billoir, E., Zuliani, V., Commeau, N., Cornu, M., and Denis, J.-B. (2011). Probabilistic modeling of the fate of *Listeria monocytogenes* in diced bacon during the manufacturing process. *Risk Analysis*, 31(2) :237–254.

Article soumis

Commeau, N. and Jaloustre, S. Impact of temperature sampling strategy on the risk of *Clostridium* growth : application to rapid cooling of food in institutional food service facilities. Soumis à *Food Control*.

Notations

Y, X :	Majuscules désignant des variables aléatoires
y, x :	Minuscules latines désignant les réalisations respectives des variables Y, X
θ :	Lettre grecque désignant par défaut les paramètres inconnus d'un modèle
$[Y]$:	Distribution de probabilité de la variable aléatoire Y
$[Y Z]$:	Distribution conditionnelle de la variable Y , conditionnellement à la variable Z
$[Y = y]$:	Probabilité de l'événement $Y=y$ lorsque Y est une variable aléatoire discrète
$[y]$:	Densité de probabilité en y d'une variable continue Y
$[y z]$:	Densité de probabilité en y d'une variable continue Y conditionnellement à une valeur z de la variable Z

Principales densités de probabilités utilisées

Densités discrètes

Poisson(λ) : Loi de Poisson de paramètre d'intensité $\lambda > 0$
 $[X = k|\lambda] = \mathcal{P}(\lambda)$
 $[X = k|\lambda] = e^{-\lambda} \frac{\lambda^k}{k!}$

pour tout $k \in \mathbb{N}$

Binomiale(n, π) : Loi Binomiale issue de n tirages de Bernoulli avec une probabilité $\pi \in [0, 1]$
 $[X = k|n, \pi] = \mathcal{Bin}(n, \pi)$
 $[X = k|n, \pi] = C_n^k \pi^k (1 - \pi)^{n-k}$
pour tout $k = 0, 1, \dots, n$, avec $C_n^k = \frac{n!}{k!(n-k)!}$

Binomiale Negative (n, π) : Loi Binomiale Négative avec une probabilité $\pi \in [0, 1]$
 $[X = k|n, \pi] = \mathcal{NegBin}(n, \pi)$
 $[X = k|N, \pi] = C_{k+n-1}^k \pi^n (1 - \pi)^k$

pour tout $k = 0, 1, \dots, n$

Multinomiale (n, π_1, \dots, π_k) : Loi Multinomiale de paramètres (n, π_1, \dots, π_k) , avec $\sum_{i=1}^k \pi_i = 1$
 $[X_1 = x_1, \dots, X_k = x_k|n, \pi_1, \dots, \pi_k] = \mathcal{MN}(n, \pi_1, \dots, \pi_k)$
 $[X_1 = x_1, \dots, X_k = x_k|n, \pi_1, \dots, \pi_k] = \frac{n!}{x_1! \dots x_k!} \pi_1^{x_1} \dots \pi_k^{x_k}$,

si $\sum_{i=1}^k x_i = n$ et 0 sinon

Hypergéométrique(N, m, n) : Loi Hypergéométrique de paramètres N, m et n
 $[X = k|N, m, n] = \mathcal{HG}(N, m, n)$
 $[X = k|N, m, n] = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n}$

Cette loi décrit la probabilité de tirer k boules blanches parmi n tirages dans un urne contenant N boules au total dont m blanches et $N - m$ boules d'une autre couleur sans remise entre les tirages.

Densités continues

Uniforme(a,b) : Loi uniforme de paramètres (a, b) avec $a, b \in \mathbb{R}$:

$$[x|a, b] = \mathcal{U}(a, b) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a; b] \\ 0 & \text{sinon} \end{cases}$$

Gamma(a,b) : Loi gamma de paramètres (a, b) avec $a > 0$ et $b > 0$:

$$[x|a, b] = \mathcal{G}(a, b) = \begin{cases} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$$

avec $\Gamma(z) = \int_0^{+\infty} t^{z-1} e^{-t} dt$.

Remarque : Si $X \sim \mathcal{G}(a, b)$ alors $\tilde{X} = 1/X$ est dite Gamma-inverse(a,b).

Normal(μ, σ^2) : Loi normale de paramètres (μ, σ^2) :

$$[x|\mu, \sigma^2] = \mathcal{N}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \quad x \in \mathbb{R}$$

LogNormal(μ, σ^2) : Loi log-normale de paramètres (μ, σ^2) :

$$[x|\mu, \sigma^2] = \mathcal{LN}(\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}x\sigma} e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} \quad x \in \mathbb{R}$$

X suit une loi log-normale de paramètres (μ, σ^2) si $Y = \ln X$ suit une loi normale de paramètres (μ, σ^2) .

Beta(α, β) : Loi bêta de paramètres (α, β) avec $\alpha > 0$ et $\beta > 0$:

$$[x|\alpha, \beta] = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad \text{avec } x \in [0, 1] \text{ et}$$

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Liste des abréviations

ALOP	<i>appropriate level of protection</i>
AQR	appréciation quantitative des risques
BIC	<i>Bayesian informative criterion</i>
BPH	bonnes pratiques d'hygiène
CCP	<i>critical control point</i>
DAG	<i>directed acyclic graph</i>
DIC	<i>deviance informative criterion</i>
FSO	<i>food safety objective</i>
HACCP	<i>Hazard analysis critical control point</i>
ICMSF	<i>International commission on microbiological specifications for foods</i>
LOD	<i>Limit of detection</i>
LOQ	<i>limit of quantification</i>
MC	<i>microbiological criterion</i>
MPN	<i>most probable number</i> (se traduit par NPP en français)
NPP	nombre le plus probable (se traduit par MPN en anglais)
NQA	niveau de qualité acceptable
NQL	niveau de qualité limite
PC	<i>performance criterion</i>
PMS	plan de maîtrise sanitaire
PO	<i>performance objective</i>
TIAC	toxi-infection alimentaire collective

Table des matières

1	Introduction	21
1.1	Principales notions utilisées pour la surveillance des pathogènes	21
1.1.1	Concepts pour la maîtrise des dangers microbiologiques	22
1.1.1.a	ALOP	22
1.1.1.b	FSO et PO	23
1.1.1.c	PC	24
1.1.1.d	BPH et HACCP	26
1.1.2	Critère microbiologique	26
1.1.3	Plan d'échantillonnage	28
1.2	Les plans d'échantillonnage : approche statistique	30
1.2.1	Qu'est-ce qu'un échantillon ?	30
1.2.2	Plans d'échantillonnage à deux et à trois classes	31
1.2.2.a	Contrôle par attribut à deux classes	31
1.2.2.b	Contrôle par attribut à trois classes	31
1.2.2.c	Comment fixer m et M ?	32
1.2.3	Risques fournisseur et client, courbe d'efficacité	33
1.2.3.a	Risque fournisseur et risque client	33
1.2.3.b	Courbe d'efficacité	34
1.2.4	Plan à deux ou à trois classes ?	35
1.2.5	Détermination de n et c	38
1.2.5.a	Par approximation normale d'un test unilatéral	38
1.2.5.b	Proposition de l'ICMSF	39
1.2.5.c	Critique de l'approche classique	41
1.3	Démarche initiale et bases de la théorie statistique de la décision	42
1.3.1	Risques fréquentiste et bayésien	43
1.3.1.a	Risque fréquentiste	44
1.3.1.b	Risque bayésien	45
1.3.2	Quelques fonctions usuelles de coûts	47
1.4	Modèles de contamination existants	49
1.4.1	Prévalence	50
1.4.2	Modèle Log-Normal	51
1.4.3	Modèle Log-Normal et prévalence	51
1.4.4	Modèle de Poisson	51
1.4.5	Modèle Poisson-Gamma ou Binomial Négatif	51
1.4.6	Modèle Poisson-Log-Normale	52

1.4.7	Modèles de type “hurdle”	52
1.4.8	Modèles de type “zero-inflated”	52
1.4.9	Mélange de lois log-normales	52
1.4.10	Estimation des paramètres	54
1.4.11	Modèle hiérarchique	54
1.5	Démarche adoptée	58
1.5.1	Plan général	58
1.5.2	Cas d’application	59
1.5.3	<i>L. monocytogenes</i>	60
1.5.4	<i>C. perfringens</i>	62
1.5.5	Projet Quant’HACCP	63

I DONNÉES ET MODÈLES DE CONTAMINATION 65

2 Données 69

2.1	Types des données utilisées	69
2.1.1	Principe de la détection et du dénombrement	69
2.1.1.a	Détection	69
2.1.1.b	Dénombrement	70
2.1.2	Expression des résultats bruts en concentration et limites de ces méthodes	71
2.1.3	Particularités des données de détection et de dénombrement	73
2.1.3.a	Les données sont imprécises	73
2.1.3.b	Données peu nombreuses	74
2.1.3.c	Autocontrôles difficiles à obtenir	74
2.2	Qu’est-ce qu’un lot ?	74
2.2.1	Lot “industriel” et lot “statistique”	74
2.2.2	Exemple : fabrication de lardons	75
2.2.3	Exemple : fabrication de saumon fumé	77
2.2.4	Fabrication d’un plat en sauce dans une cuisine collective	78
2.2.5	Fabrication de steaks hachés	80
2.2.6	Lot, procédé et données	80
2.3	Acquisition de données dans le cas lardons	81
2.3.1	Choix du plan	81
2.3.1.a	Choix des facteurs	81
2.3.1.b	Plan d’échantillonnage stratifié	82
2.3.1.c	Combien de lots par strate choisir ?	82
2.3.2	Mise en place du plan d’échantillonnage	83
2.3.3	Résultats	85
2.3.3.a	Flore lactique	85
2.3.3.b	<i>L. monocytogenes</i>	86
2.3.4	Conclusion	87
2.4	Acquisition de données dans le cas saumon fumé	88
2.4.1	Prélèvements et protocoles	88

2.4.2	Résultats	90
3	Variabilités de la contamination	91
3.1	Variabilité inter et intra lots	91
3.2	Article : <i>Hierarchical Bayesian models to assess between and within batch variability of pathogen contamination in food</i>	94
3.2.1	Introduction	94
3.2.2	Theoretical framework	96
3.2.2.a	Definition of the contamination structure	96
3.2.2.b	Models	98
3.2.2.c	Prior setting	106
3.2.2.d	Comparison	107
3.2.3	Applications	109
3.2.3.a	Pork breasts	109
3.2.3.b	Cold smoked salmon	114
3.2.4	Discussion	118
3.2.4.a	What is learnt for each case study	118
3.2.4.b	This approach is applicable to other cases	118
3.2.4.c	The study is deeply grounded on priors	119
3.2.4.d	Exploiting the approach for designing effective sampling plans	120
3.3	Retour sur la comparaison	121
3.3.1	Critère modifié de Gelfand and Ghosh (1998)	122
3.3.2	Inspection visuelle	122
3.3.3	En utilisant la p-valeur	123
3.4	Conclusion	124
3.4.1	Courbe d'efficacité et modèle de contamination	124
3.4.2	Utilisation des données de flore lactique	125
4	Données et estimation	129
4.1	Données et estimation des paramètres	129
4.1.1	Positionnement du problème	129
4.1.2	Quel type de données utiliser en fonction du pourcentage d'absences?	131
4.2	Article : <i>Fitting a lognormal distribution to enumeration and absence/presence data</i>	131
4.2.1	Introduction	132
4.2.2	Materials and methods	133
4.2.2.a	How should raw data be converted into concentration-like data?	133
4.2.2.b	Models to estimate the contamination distribution	133
4.2.2.c	Parameters estimation	134
4.2.2.d	Simulating detection and enumeration protocols	135
4.2.2.e	Simulating a sampling plan	136
4.2.2.f	Bias and variance in the estimates of mean and variance	137
4.2.2.g	Estimating the 97.5th percentile	137

4.2.3	Results	137
4.2.3.a	Results of the two inference methods	137
4.2.3.b	Bias and variance of the estimates μ and σ	138
4.2.3.c	Estimating the 97.5th percentile of the total distribution	139
4.2.4	Discussion	141
4.2.5	Appendix A	145
5	Modèle de contamination	147
5.1	Modélisation initiale de la contamination	147
5.2	Modification du modèle initial	148
5.3	Introduction des données pour construire les lois de μ^A et de σ^A	151
5.4	Impact sur la distribution des la contamination après malaxage	154
II	AIDE À LA DÉCISION EN UTILISANT L'ÉCHANTILLONNAGE	159
6	Plans d'échantillonnage et lardons	163
6.1	Description de la surveillance adoptée dans une usine	163
6.2	Optimisation à partir de la prévalence	164
6.2.1	Loi <i>a priori</i> , vraisemblance, décisions et fonction de coût	164
6.2.2	Calcul du risque bayésien	166
6.2.3	Fonction de coût adaptée à l'application	168
6.2.4	Autre détermination de la fonction de coût	172
6.2.5	Une application numérique	174
6.2.6	Variation des valeurs de a , b , ψ_0 et ψ_1	176
6.2.7	Discussion sur les hypothèses du modèle	177
6.3	Le client ne connaît pas exactement la prévalence	178
7	Plans d'échantillonnage et saumon fumé	185
7.1	Principe : deux états et une fonction de coût de type "0-K"	185
7.2	Calcul du risque actualisé sur plusieurs périodes	188
7.3	Intégrons les éléments précédents selon E_t	189
7.3.1	Risque au temps t intégrée selon E_t	190
7.3.2	Risque au temps t actualisé et intégré selon E_t	191
7.4	Applications numériques	191
7.4.1	Cas où $\rho = 0$	191
7.4.2	Cas où $\rho \neq 0$	192
7.5	Perspectives	195
8	Plan d'échantillonnage de températures	197
8.1	Préparation d'un plat en restauration collective	197
8.2	Plan d'échantillonnage thermique	198
8.3	Modélisation de la température	199
8.4	Quel plan choisir ?	200
8.5	Article	201

III	CONCLUSION	219
9	Conclusion	221
9.1	Bilan	221
9.1.1	Travaux accomplis	221
9.1.2	Application de concepts théoriques	223
9.1.3	Critiques de la théorie bayésienne de la décision	228
9.1.4	A côté des statistiques	230
9.2	Perspectives	231
A	Le bayésien à travers les âges	235
B	Fonction d'utilité	245
C	Une application de la théorie de la décision	249
D	Inférence bayésienne et inférence classique	253
D.1	Inférence bayésienne	253
D.2	Inférence fréquentiste	254
D.3	Intervalle de crédibilité et intervalle de confiance	254
D.4	Tests d'hypothèse	255
D.4.1	Approche fréquentiste	255
D.4.2	Approche bayésienne	256
E	Algorithme EM	259
E.1	Principe de l'algorithme EM	259
E.2	Application de l'algorithme EM au modèle \mathcal{M}_{RD}	261
E.2.1	Calcul de $Q(\theta, \theta')$	261
E.2.2	Calcul de Ae_i, Ad_i, Be_j et Bd_j par importance sampling	262
E.2.3	Algorithme obtenu	263
F	Article Billoir et al.	265
F.1	Introduction	265
F.2	The model	267
F.2.1	Overview of the model	267
F.2.2	Step combination	269
F.2.3	Arrival	269
F.2.3.a	Mass m^A	269
F.2.3.b	Contamination c^A_{Lm}, c^A_{LAB}	269
F.2.4	Brining-and-tumbling	270
F.2.4.a	Mass m^{BT}	270
F.2.4.b	Contamination c^{BT}	271
F.2.5	Steaming	272
F.2.5.a	Mass m^S	272
F.2.5.b	Contamination c^S	272
F.2.6	Dicing	273

	F.2.6.a	Mass m^D	274
	F.2.6.b	Contamination c^D	274
	F.2.7	Packaging	275
	F.2.7.a	Mass m^P	275
	F.2.7.b	Contamination c^P	275
F.3		Baseline calibration	275
	F.3.1	Arrival ($\mu_{bm}^A, \sigma_{bm}^A, \mu_{\log \log(\text{conc}^A)_{Lm}}, \sigma_{\log \log(\text{conc}^A)_{Lm}}, tr_{\log(\text{conc}^A)_{Lm}}, \mu_{\log_{10}(\text{conc}^A)_{LAB}}$ and $\sigma_{\log_{10}(\text{conc}^A)_{LAB}}$)	275
	F.3.2	Brining and tumbling ($N^{BT}, m_{bri}, \text{conc}_{bri}, p_{rel}$ and p_{loss})	278
	F.3.3	Steaming ($t^S, T^S, \mu_{D_{ref}}, \sigma_{D_{ref}}, \mu_z$ and σ_z)	278
	F.3.4	Dicing ($\alpha, \beta, gr, ar, sec, \lambda_{dic}$)	280
	F.3.5	Packaging (μ_{pum}^P and σ_{pum}^P)	280
F.4		Design of simulations	280
	F.4.1	Sensitivity analysis towards uncontrolled process parameters	280
	F.4.2	Adaptating the model to various processes	281
	F.4.3	Alternative scenarios	281
F.5		Results	282
	F.5.1	Baseline calibration	282
	F.5.2	Sensitivity analysis towards uncontrolled process parameters	285
	F.5.3	Adaptating the model to various processes	285
	F.5.4	Alternative scenarios	288
F.6		Discussion	288
	F.6.1	What we can learn from the simulations	288
	F.6.2	Assumptions underlying the model	289
	F.6.3	Step combination	290
	F.6.4	Prospects for the use of the model	290
F.7		Acknowledgments	291
G		Modification de l'article de Billoir et al.	295
H		Rapport qualité	297

Chapitre 1

Introduction

Plus particulièrement consacrée à la contamination par des pathogènes d'aliments en cours de fabrication ou en sortie d'usine, ma thèse de statistique appliquée à la microbiologie est constituée de deux parties :

1. recueil de données et modélisation de la contamination par le pathogène *Listeria monocytogenes*^a dans des lardons en cours de fabrication et en fin de procédé et du saumon fumé en fin de procédé ;
2. étude de plans d'échantillonnage microbiologique et thermique sur trois applications. Pour deux d'entre elles, la théorie statistique de la décision est utilisée pour optimiser un plan d'échantillonnage de manière à minimiser les coûts moyens supportés par le fabricant. Dans la troisième application, différents plans d'échantillonnage thermique sont comparés de manière à déterminer le risque de croissance du pathogène *Clostridium perfringens*^b supporté par le fabricant.

Dans cette introduction, je replace les plans d'échantillonnage dans le contexte microbiologique, et pour cela, je décris tout d'abord le contexte général de la surveillance des pathogènes dans les industries agro-alimentaires (partie 1.1). Ensuite, je décris les plans d'échantillonnage sous l'angle statistique (partie 1.2) puis la théorie de la décision (partie 1.3). Je présente les modèles les plus courants de contamination d'un aliment par un pathogène (partie 1.4) avant de terminer par l'annonce du plan de thèse, une brève description des pathogènes sur lesquels j'ai travaillé (*L. monocytogenes* et *C. perfringens*) et le projet de recherche ANR intitulé Quant'HACCP dans lequel s'inscrit une large partie de ma thèse (partie 1.5).

1.1 Principales notions utilisées pour la surveillance des pathogènes dans une usine agro-alimentaire

Les pouvoirs publics et les industriels du secteur agro-alimentaire s'attachent à s'assurer que les microorganismes susceptibles de dégrader la qualité d'un aliment ou d'atteindre la santé du consommateur sont éliminés ou au moins maîtrisés dans les aliments

a. noté *L. monocytogenes* dans la suite

b. noté *C. perfringens* dans la suite

produits. Compte-tenu des enjeux aussi bien sanitaires qu'économiques, un système de gestion de la sécurité des aliments a été élaboré au niveau international, essentiellement au niveau du *Codex alimentarius*^c.

Sans retracer l'histoire de la sécurité sanitaire des aliments et de la gestion du risque, je décris dans cette partie le système de gestion de la sécurité des aliments tel qu'il est défini par le Codex Alimentarius.

1.1.1 Concepts pour la maîtrise des dangers microbiologiques dans les aliments

1.1.1.a ALOP

La sécurité des aliments dans le commerce international est traitée dans l'accord sur l'application des mesures sanitaires et phytosanitaires (WTO (1995)). Cet accord habilite les pays membres de l'organisation mondiale du commerce (OMC) à prendre des mesures d'ordre sanitaire ou phytosanitaire pour s'assurer un niveau de protection approprié, fondé sur une évaluation scientifique des risques. Ce niveau est appelé *ALOP*, pour *appropriate level of protection*. Les aliments importés ne peuvent pas être rejetés s'ils satisfont l'ALOP du pays importateur. Dans le cas d'un refus d'un produit, l'Etat importateur doit prouver, à partir d'une évaluation du risque de manière scientifique, que le produit ne permet pas de respecter l'ALOP. Toujours selon l'accord WTO (1995), l'ALOP doit être déterminé à partir d'une évaluation du risque pour le consommateur (art. 5). Cependant, la manière dont l'ALOP doit être exprimé et le moyen précis de sa détermination ne sont pas détaillés.

Le Codex Alimentarius (2003) précise que l'ALOP peut être exprimé en termes quantitatif ou qualitatif. Par exemple, les Etats-Unis ont fixé pour plusieurs infections d'origine alimentaire des objectifs nationaux de santé (*national health objectives*) pour les années 2010 (CDC (2010)) et 2020 (Anonymous (2010)). Le tableau 1.1 présente quelques objectifs pour des infections dues à trois pathogènes. Par exemple pour les Salmonelles, l'incidence observée durant la période 2006-2008 était de 15,2 (Anonymous (2010)). Manifestement, l'objectif en 2010 était trop ambitieux ; l'horizon 2020 semble plus raisonnable.

En Europe, les Etats n'ont pas fixé d'ALOP. Selon l'EFSA, l'ALOP représente l'état actuel de santé publique. La détermination de sa valeur à partir des réseaux de surveillance de santé publique, en particulier pour évaluer le nombre de malades et de morts attribuables à un pathogène. L'EFSA suggère également de prendre en compte d'autres sources afin d'estimer le nombre de cas non répertoriés par les réseaux institutionnels de surveillance (EFSA (2007)).

L'ICMSF (*International commission on microbiological specifications for foods*^d) a

c. La commission du *Codex Alimentarius* a été créée en 1963 par la FAO et l'OMS. Elle élabore des normes alimentaires, des lignes directrices et des codes d'usages internationaux afin de protéger la santé des consommateurs et d'assurer des pratiques loyales dans le commerce des aliments. Pour en savoir plus, consulter <http://www.codexalimentarius.org/about-codex/fr/>.

d. Créée en 1960, l'ICMSF poursuit les objectifs suivants :

- assembler, corrélérer et évaluer les preuves de la sécurité microbiologique et de la qualité des aliments ;
- considérer si les critères microbiologiques peuvent améliorer et assurer la sécurité microbiolo-

Année	Pathogène	Objectif
2010	<i>L. monocytogenes</i>	0,24
”	Salmonelles	6,28
”	STEC O157 :H7	1,0
2020	<i>L. monocytogenes</i>	0,20
”	Salmonelles	11,4
”	STEC O157 :H7	0,6

TABLE 1.1 – Objectifs nationaux de santé exprimés en nombre de cas pour 100 000 habitants pour les années 2010 et 2020 aux Etats-Unis.

STEC O157 signifie *E. coli* productrice de Shiga toxine de sérotype O157 :H7.

considéré que l’ALOP n’était pas un outil pratique pour les industriels, surtout s’il était exprimé en nombre de malades pour une population donnée.

1.1.1.b FSO et PO

Le *FSO* (*food safety objective*) a alors émergé. Un FSO est défini comme “la fréquence ou la concentration maximale d’un danger^e dans un aliment au moment de la consommation qui assure ou contribue à assurer l’ALOP” (Codex Alimentarius (2004)). Le FSO peut être utile aussi bien pour l’industriel (pour implémenter les mesures adaptées pour contrôler le danger) que pour l’autorité de contrôle (pour définir les procédures permettant d’évaluer les mesures de contrôle adoptées pour l’industriel et pour quantifier l’équivalence des différentes procédures d’inspection dans différents pays). Le FSO est en principe fixé par les pouvoirs publics et l’ICMSF recommande (ICMSF (2002)) d’utiliser des données épidémiologiques pour le déterminer. Les industriels ont alors la charge de s’assurer que les procédés qu’ils utilisent permettent de respecter le FSO. Le FSO peut être vu comme la traduction d’un objectif de santé public (l’ALOP) en une caractéristique du produit fabriqué pouvant être contrôlé par les industriels.

Le FSO est défini au moment de la consommation, alors que l’industriel ne maîtrise pas le devenir du produit sorti de l’usine. Il peut donc décliner le FSO en *PO* (*performance objective*) qui est un FSO à une étape du procédé de fabrication. Un ou plusieurs PO peuvent être définis par procédé. La définition exacte du PO est (Codex Alimentarius (2004)) : “la fréquence ou la concentration maximale d’un danger dans un aliment à une étape définie et qui assure ou contribue à assurer un FSO ou un ALOP selon le cas”.

Voici deux exemples hypothétiques de FSO (ICMSF (2002)) :

-
- giques dans certains aliments ;
 - proposer des critères microbiologiques lorsque c’est approprié ;
 - recommander des méthodes d’échantillonnage et d’analyses.

Les recommandations de l’ICMSF n’ont pas de caractère contraignant. Actuellement, la commission est composée de 18 microbiologistes des aliments. Pour plus d’information, consulter <http://www.icmsf.org/>.

e. Agent biologique, chimique ou physique présent dans un aliment ou état de cet aliment pouvant avoir un effet adverse sur la santé (Codex Alimentarius (1999)). Dans le cadre de cette thèse, seuls les dangers biologiques sont considérés.

- la concentration en aflatoxine dans les cacahuètes ne doit pas excéder 15 $\mu\text{g}/\text{kg}$;
- la concentration en salmonelles dans la poudre de lait doit être inférieure à 1 ufc^f dans 100 kg de poudre.

Un FSO concerne la production d'un aliment particulier, pas seulement d'un lot. L'ICMSF recommande que le FSO soit quantitatif et vérifiable. Il n'est pas toujours nécessaire de définir un FSO pour chaque aliment ou tous les couples aliments/pathogène : il ne présenterait guère d'intérêt par exemple lorsqu'un pathogène représente un risque infime dans un aliment donné. Parfois, les sources de contamination d'un pathogène sont si variées qu'il est impossible de dresser la liste de tous les aliments pour lesquels un FSO pour ce microorganisme serait nécessaire (ICMSF (2002)).

Pour faire en sorte que les FSO soient respectés, l'ICMSF conseille de mettre en place des mesures de contrôle tout au long du procédé. Ces mesures sont de trois sortes :

- contrôle des niveaux de contamination des matières premières, ce qui passe par la sélection des produits entrants dans l'usine avec, éventuellement, des analyses microbiologiques et le rejet des lots trop contaminés ;
- prévention de l'augmentation de la contamination, en empêchant la croissance des pathogènes ou en limitant la contamination croisée par exemple ;
- réduction de la contamination en détruisant les pathogènes ou en les enlevant (par des nettoyages).

Selon les procédés, une ou plusieurs mesures de contrôles sont nécessaires.

Enfin, il faut établir des critères pour s'assurer que les mesures de contrôle permettent bien d'atteindre le FSO.

1.1.1.c PC

Ces critères se nomment *PC* (*performance criteria*) et sont définis comme "l'effet de la fréquence ou de la concentration d'un danger dans un aliment qui doit être atteint par l'application d'une ou plusieurs mesures de contrôle afin d'assurer ou de contribuer à assurer un PO ou un FSO" (Codex Alimentarius (2004)). Un PC est par exemple la réduction de 6 \log^g de la concentration de *L. monocytogenes* dans les aliments prêt à être consommés (Lund et al. (1989)). Pour définir un PC, il faut prendre en compte la contamination initiale et les évolutions de la contamination tout au long de la vie du produit jusqu'à sa consommation (ICMSF (2002)). L'ICMSF (2002) définit également des *process criteria* et des *product criteria* qui sont des traductions opérationnelles de PC, respectivement sur le procédé et sur le produit. Par exemple, un *process criterion* est la pasteurisation du lait à 71,7°C pendant 15 s pour inactiver *Coxiella burnetti* et un *product criterion* est que le pH soit inférieur ou égal à 4,6 dans des boîtes de conserve car à ce pH, la croissance du pathogène *C. botulinum* est inhibée. La figure 1.1 permet de visualiser les endroits de la chaîne où s'appliquent les PO, PC et FSO depuis les matières premières jusqu'à la consommation. Pour appliquer l'ALOP, il faut déterminer le nombre de cas engendrés par l'ingestion d'un micro-organisme donné en ayant recours par exemple à un modèle dose-réponse. Pour que les critères soient respectés, l'industriel applique

f. ufc signifie "unité formant colonie"

g. Dans ce travail, "log" désigne le logarithme décimal et "ln" le logarithme népérien.

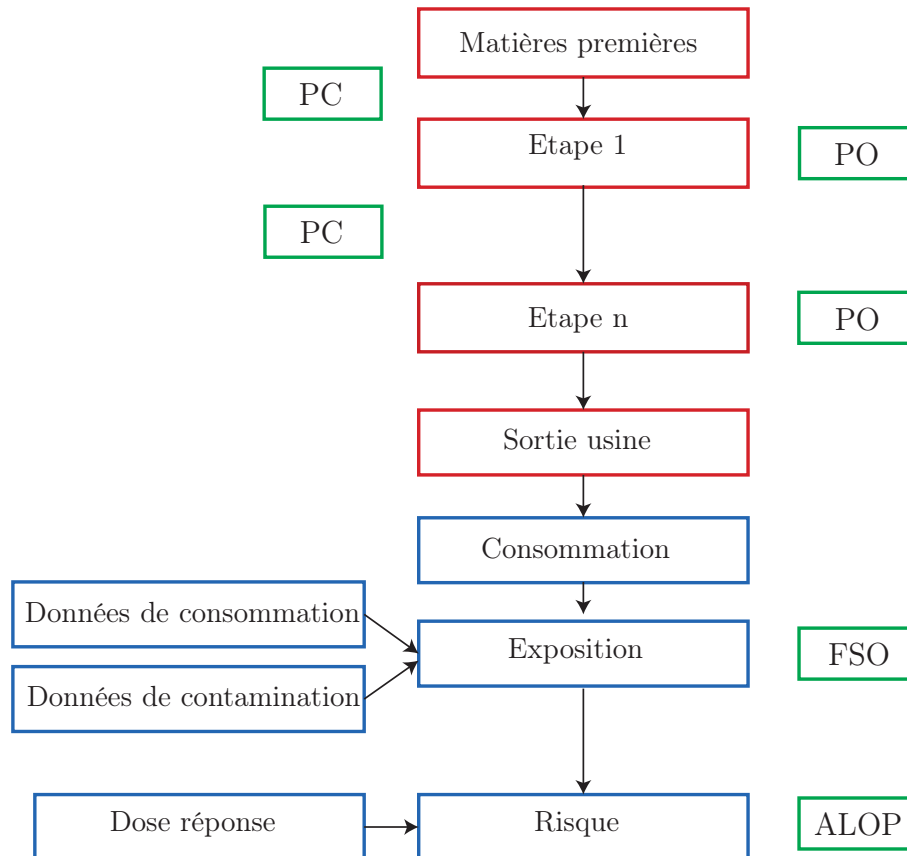


FIGURE 1.1 – Positionnement de l’ALOP, du FSO, des PO et des PC durant la vie d’un produit depuis les matières premières jusqu’à sa consommation et en tenant compte du risque. Les PO et les PC sont liés à un couple produit/pathogène dans un procédé de fabrication, le FSO est lié à un pathogène et un type de produit, l’ALOP est défini pour un pathogène. Les étapes de fabrication sont encadrées en rouge, les étapes de la consommation au risque sont en bleu, les données nécessaires et le modèle dose-réponse sont en orange, l’ALOP, FSO, PO et PC sont en vert.

les bonnes pratiques d'hygiène (BPH) et le système HACCP (pour *Hazard Analysis and Critical Control Point*).

1.1.1.d BPH et HACCP

Les BPH sont les conditions et pratiques sanitaires de base à respecter afin de produire des aliments sains. Le Codex Alimentarius a décrit les composantes principales des BPH (Codex Alimentarius (1997b)). Celles-ci concernent l'hygiène générale des locaux, les manipulations, le stockage et le transport, les équipements, les nettoyages et le personnel. Les BPH constituent un préalable à la mise en place du système HACCP. Initialement créé dans les années 1980 pour assurer la sécurité sanitaire des aliments destinés aux astronautes dans l'espace, ce système est devenu la référence pour la maîtrise de l'hygiène des aliments. Les sept principes constituant ce système sont décrits dans Codex Alimentarius (1997a).

1. procéder à une analyse des dangers ;
2. déterminer les points critiques pour la maîtrise (CCP pour *critical control point*) ;
3. fixer le ou les seuil(s) critiques(s) ;
4. mettre en place des mesures de surveillance permettant de maîtriser les CCP ;
5. établir des mesures correctives à prendre s'il apparaît qu'un CCP n'est pas maîtrisé ;
6. déterminer des procédures de vérification afin de s'assurer que le système HACCP fonctionne ;
7. constituer un dossier regroupant les documentations sur les procédures et tous les enregistrements.

Ce système induit une réflexion globale sur le procédé de production et les éventuelles contaminations. En effet, implémenter les principes HACCP conduit à identifier, évaluer et contrôler systématiquement tous les dangers susceptibles d'être présents dans un produit depuis sa fabrication jusqu'à sa consommation. En particulier, l'HACCP est très utile lorsqu'une des étapes du procédé permet de diminuer fortement voire d'éliminer un pathogène, ce qui est le cas de la cuisson par exemple.

1.1.2 Critère microbiologique

Définition Un *critère microbiologique* (noté MC par la suite pour *Microbiological criterion*) définit "l'acceptation d'un produit ou d'un lot, basé sur la présence ou l'absence d'un microorganisme ou des toxines ou métabolites par unité de masse, volume, aire ou lot" (Codex Alimentarius (1997a)). Un MC est composé^h :

1. du microorganisme (ou toxine ou métabolite) recherché et la raison pour laquelle ce microorganisme est recherché ;
2. du protocole d'analyse utilisé ;
3. du nombre d'unités à prélever pour analyse ainsi que leur taille ;

h. Comme indiqué plus loin, le plan d'échantillonnage est formé par les points 3 à 5 de cette définition.

4. de la limite microbiologique pour l'aliment considéré à une étape spécifiée de la chaîne ;
5. du nombre d'unités qui devraient respecter cette limite.

Un critère microbiologique s'applique pour un aliment donné. Il devrait aussi comporter les actions à entreprendre lorsqu'il n'est pas satisfait (Codex Alimentarius (1997c)).

Utilité L'ICMSF considère qu'un système de management de la sécurité des aliments basé sur les BPH et l'HACCP est plus efficace que l'analyse des aliments en sortie usine pour connaître leur qualité. Néanmoins, l'utilisation des critères microbiologiques est utile dans deux cas (ICMSF (2002)) :

- valider que mesures de contrôle permettent de respecter les critères de performance ;
- déterminer si la contamination d'un aliment est acceptable quand aucun autre moyen de s'assurer que la sécurité est respectée ou en l'absence de connaissance de l'application des BPH et de l'HACCP

D'après ICMSF (2002), un MC doit être fixé de telle manière que le FSO associé au couple aliment/pathogène soit respecté, bien que le FSO n'ait pas été créé pour l'acceptation ou non de lots. FSO et MC diffèrent aussi bien par leur contenu que par leur but : les seconds peuvent être utilisés en l'absence de premiers. Les MC peuvent être définis à tout endroit de la chaîne de production, depuis la production dans les exploitations agricoles jusqu'à la vente ; ils peuvent être utilisés par les autorités de contrôle en l'absence de meilleur outil et lorsque l'on suppose qu'ils peuvent améliorer la protection du consommateur. Si un MC n'est pas respecté, l'autorité de contrôle peut demander à l'industriel par exemple de rappeler, détruire ou trier les produits. Un MC peut être fixé par les autorités.

Exemples Le Règlement (CE) No 2073/2005 (2005) fixe plusieurs MC. Ainsi, pour les "denrées alimentaires prêtes à être consommées destinées aux nourrissons et denrées alimentaires prêtes à être consommées destinées à des fins médicales spéciales" pouvant être contaminées par *L. monocytogenes*, le MC est le suivant : 10 prélèvements, 0 positif dans 25 g autorisé, le protocole d'analyse est la norme ISO11290-1 (1996a). Ce MC est applicable pour les "produits mis sur le marché pendant leur durée de conservation". Le Codex Alimentarius a proposé des MC concernant les produits prêts à consommer pouvant être contaminées par *L. monocytogenes* (Codex Alimentarius (2007)) . Si la croissance de cette bactérie est possible, alors 5 prises d'essai de 25 g doivent être analysées, et toutes doivent être négatives. Si *L. monocytogenes* ne croît pas, le nombre d'analyses reste le même mais la concentration des portions analysées doit être au plus de 100 ufc/g. Dans ce cas, le protocole utilisé est décrit dans la norme ISO11290-2 (1998a). Le FSO concernant ces produits et ce pathogène est : une concentration ne devant pas excéder 100 ufc/g lors de la consommation.

Les MC peuvent également être utilisés par les industriels pour vérifier ou valider l'efficacité des applications des BPH et du système HACCP. Ils peuvent être plus stricts que les MC définis par les autorités de contrôle. Concernant les BPH, un MC peut être défini par exemple pour contrôler des surfaces dans un atelier après nettoyage.

D'après le Codex Alimentarius (1997c), un MC ne devrait être établi uniquement lorsqu'il y a un besoin et que son application est pratique.

Après cette description du contexte, nous arrivons à la définition du plan d'échantillonnage du point de vue microbiologique. Pour l'instant, aucune définition ni aucune propriété statistique n'ont été fournies.

1.1.3 Plan d'échantillonnage

Définition Comme il est écrit plus haut, un plan d'échantillonnage est un sous-ensemble d'un MC. Il est composé du nombre d'unités à prélever ainsi que le nombre maximal d'unités analysées ne respectant pas la limite microbiologique.

Utilisation

Mesure de maîtrise Une *mesure de maîtrise* est une intervention établie en un point du procédé pour prévenir ou éliminer un danger menaçant la sécurité sanitaire de l'aliment. Celle-ci doit être

- validée : il faut établir si la mesure réduit effectivement le danger à condition d'être appliquée correctement ;
- surveillée, généralement en temps réel pour déterminer si la mesure aboutit bien aux résultats escomptés ;
- vérifiée par application de méthodes, procédures ou tests pendant ou après son application pour déterminer si elle a été correctement mise en œuvre.

Le plan d'échantillonnage peut servir à valider les mesures de maîtrise. Il constitue également l'un des outils de surveillance de l'efficacité de ces mesures par l'analyse de quelques produits et de surfaces. Enfin, il peut également servir à la vérification par l'analyse des résultats obtenus pendant la période définie. Les mesures de maîtrise doivent évidemment être adaptées en fonction des résultats de la surveillance. Un schéma des utilisations du plan d'échantillonnage est présenté à la figure 1.2. La validation et son imbrication avec la surveillance et la vérification sont détaillées par Codex Alimentarius (2008).

Un petit détour par le contexte réglementaire français Le contexte réglementaire dans lequel évolue le secteur agro-alimentaire en France. Celui-ci a beaucoup évolué avec le Règlement (CE) No 178/2002 (2002) qui a renforcé l'obligation de résultat. Les objectifs fixés par ce texte et par d'autres sont regroupés sous l'expression "paquet hygiène". Les professionnels sont notamment tenus de garantir l'innocuité de leurs produits et de protéger la santé du consommateur. Ils sont libres de choisir les moyens pour remplir ces objectifs. En France, la stratégie est indiquée dans le *plan de maîtrise sanitaire* (PMS) qui, d'après la définition de l'Arrêté du 8 juin (2006), modifié le 26/11/10, "décrit les mesures prises par l'établissement pour assurer l'hygiène et la sécurité sanitaire de ses productions vis-à-vis des dangers biologiques, physiques et chimiques". Il comprend les documents relatifs aux BPH, aux procédures fondées sur le principe de l'HACCP et la procédure de traçabilité des produits ainsi que la gestion des produits non conformes,

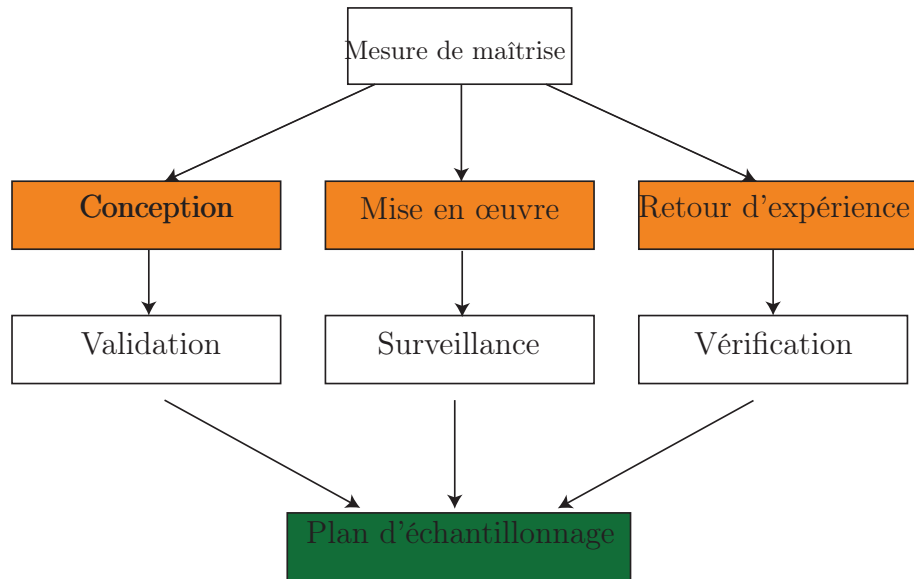


FIGURE 1.2 – Interactions possibles entre une mesure de maîtrise et un plan d'échantillonnage.

comme le retrait et le rappel. Lors de la conception du PMS, à chaque étape du procédé, on recherche :

- les dangers susceptibles d'apparaître,
- les facteurs de risques,
- le risque associé et la mesure de maîtrise à prendre.

De même que pour les mesures de maîtrises, le plan d'échantillonnage peut être utilisé lors de la validation, la surveillance et la vérification du PMS.

ICMSF L'ICMSF considère que la raison pour laquelle on échantillonne doit être définie avant de mener les prélèvements et les analyses et, qu'en général, les raisons d'échantillonner dans le cadre de la gestion de la sécurité des aliments sont au nombre de quatre (ICMSF (2011)) :

- déterminer la sécurité ;
- déterminer si les BPH sont respectées ;
- déterminer l'utilité d'un aliment ou d'un ingrédient pour un but particulier ;
- déterminer si un produit est stable.

Si un procédé ne comporte pas de CCP et s'il n'y a pas d'autres solution pour évaluer la sécurité d'un aliment, l'échantillonnage en fin de chaîne de production reste un moyen à considérer (ICMSF (2011)). Le nombre d'analyses peut être réduit pour la surveillance de routine lorsque tout va bien et être augmentée lorsque la contamination a augmenté ou en cas de TIACⁱ. Différents micro-organismes peuvent être recherchés, selon l'objectif fixé. ICMSF (2011) en distingue trois sortes :

i. toxi-infection alimentaire collective

- les microorganismes d'utilité : ils sont utiles pour évaluer la qualité du produit. Il peut s'agir de décompte de la flore totale ou de microorganismes causant une dégradation particulière d'un aliment comme les lactobacilles dans la mayonnaise ;
- les indicateurs : ce sont des microorganismes non pathogènes dont la présence va de pair avec des bactéries pathogènes. Par exemple, les Enterobacteriaceae sont des indicateurs de salmonelles dans les œufs en poudre. Une information quantitative de ces indicateurs peut être utilisée pour vérifier si un procédé est bien appliqué. Le lien entre indicateur et pathogène dépend du produit fabriqué et du procédé utilisé ;
- les pathogènes.

On voit bien que, pour l'ICMSF, le plan d'échantillonnage est un outil à utiliser faute de mieux, par exemple lors de contrôle aux frontières en l'absence d'information sur tel produit importé, dont il faut décider s'il est autorisé à entrer sur le territoire.

Recherche d'information On peut également avoir recours à l'échantillonnage pour obtenir de l'information (par exemple sur la contamination d'un produit par un pathogène donné). Dans ce cas, aucune limite microbiologique n'est fixée. L'échantillonnage peut aussi être un outil intéressant pour obtenir des informations lorsqu'une épidémie ou une TIAC se produisent. D'autres raisons ont été invoquées par Brown et al. (2000) comme l'acquisition de données microbiologiques afin de déterminer certains CCP lors de la mise en place d'un système HACCP. En effet, ces obtentions sont nécessaires pour savoir à quelle(s) étape(s) les contaminations apparaissent, diminuent ou sont éliminées.

Relation client fournisseur L'échantillonnage peut également être utile dans les relations entre un fournisseur et son client pour les produits périssables. Les analyses peuvent alors porter sur les matières premières utilisées par le fournisseur ou à n'importe quelle étape de fabrication du produit chez le fournisseur.

Maintenant que les plans d'échantillonnage ont été introduits par le contexte de la sécurité des aliments, considérons-les sous l'angle statistique.

1.2 Les plans d'échantillonnage : approche statistique

1.2.1 Qu'est-ce qu'un échantillon ?

Pour un microbiologiste, un *échantillon* est un produit prélevé alors que pour le statisticien, c'est l'ensemble des produits prélevés dans un lot, chaque produit prélevé étant une *unité d'échantillonnage*. La *prise d'essai* est la partie du produit effectivement soumise à l'analyse. Pour éviter toute confusion, dans ce travail, le terme "échantillon" sera considéré au sens statistique et l'unité d'échantillonnage sera confondue avec la prise d'essai. Le terme "lot" est discuté partie 2.2. Pour l'instant, le lot est considéré comme l'ensemble des produits fabriqués durant une période définie et dans des conditions similaires.

Afin d'éviter les biais, le choix des unités doit se faire par tirage au sort. Une fois les produits sélectionnés, comment choisir la prise d'essai ? On peut la prendre en une fois

(ex : une partie de 25 g) ou alors prélever plusieurs petits morceaux du produit (ex : 5 parties de 5 g chacune)

1.2.2 Plans d'échantillonnage à deux et à trois classes

Le principe d'un plan d'échantillonnage est le suivant : un échantillon est prélevé dans un lot. En fonction des analyses, une décision est prise. Dans le contrôle par attribut, les échantillons sont classés en catégories et la décision dépend de la répartition des unités d'échantillonnage dans chacune des classes. Les plans classiquement employés dans l'industrie agro-alimentaire sont les contrôles par attributs à deux ou à trois classes (ICMSF (1986, 2002); ANSES (2008)).

1.2.2.a Contrôle par attribut à deux classes

Comme son nom l'indique, le plan à deux classes permet de répartir les unités d'échantillonnage dans deux groupes : conforme et non conforme. Généralement, l'analyse pratiquée est la détection^j. Cependant, il est possible de faire des analyses de dénombrement par exemple. Dans ce cas, la non conformité est définie par le dépassement d'une certaine concentration définie à l'avance. Pour spécifier un plan, deux nombres sont nécessaires (trois si l'analyse pratiquée évalue la concentration) :

- n : la taille de l'échantillon ;
- c : le nombre maximal d'unités non conformes tel que le lot est jugé conforme ;
- m : la limite de concentration du microorganisme considéré correspondant à une hygiène satisfaisante du procédé examiné. La concentration est habituellement exprimée en ufc/g ou en ufc/cm².

En considérant que X , la variable aléatoire représentant le nombre d'analyses positives, suit une loi binomiale de paramètres n et p (la proportion de produits défectueux dans le lot), alors la probabilité d'acceptation vaut

$$\mathbb{P}(X \leq c|p) = \sum_{i=0}^c C_n^i p^i (1-p)^{n-i}. \quad (1.1)$$

1.2.2.b Contrôle par attribut à trois classes

Pour le plan à trois classes, les unités d'échantillonnage sont réparties en trois catégories. Quatre nombre définissent ce plan :

- n : idem plan à deux classes ;
- m : idem plan à deux classes ;
- M : limite de concentration au-delà de laquelle l'hygiène est jugée insatisfaisante. L'unité est la même que celle de m ;
- c : le nombre maximal d'unités de qualité acceptable, c'est-à-dire d'unités dont la concentration est comprise entre m et M .

j. Les termes *détection* et *dénombrement* sont définis à la partie 2.1.1.

Le lot est jugé conforme si il y a au plus c unité ayant une concentration comprise entre m et M et il est rejeté si au moins une unité d'échantillonnage à une concentration supérieure à M .

Pour le plan d'échantillonnage à trois classes, Legan et al. (2001) définit trois proportions :

- la proportion de produits défectueux (P_d) : c'est la proportion de produits dont la concentration est supérieure à M ;
- la proportion de produits marginalement acceptables (P_m) : c'est la proportion de produits dont la concentration est comprise entre m et M ;
- la proportion de produits acceptables (P_a) : c'est la proportion de produits dont la concentration est inférieure à m .

Bien évidemment, $P_d + P_m + P_a = 1$. Les variables aléatoires X_a , X_m et X_d respectivement le nombre de résultats inférieurs à m , compris entre m et M et supérieurs à M suivent une loi multinomiale de paramètres n , P_a , P_m et P_d . La probabilité d'accepter un lot vaut donc

$$\mathbb{P}(X_a \geq n - c, X_m \leq c, X_d = 0) = \sum_{i=0}^c \frac{n!}{(n-i)!i!0!} P_a^{n-i} P_m^i P_d^0.$$

1.2.2.c Comment fixer m et M ?

D'après l'ICMSF (2002), la valeur de m reflète l'implémentation des BPH au sein d'une usine ou des exigences réglementaires. Par exemple, si on considère un pathogène pour lequel le niveau de tolérance est nul, un plan à deux classes est mis en place avec m fixé de manière à ce que le lot soit accepté seulement si toutes les prises d'essai sont négatives selon l'analyse pratiquée. Dans le Règlement (CE) No 2073/2005 (2005), les prises d'essai font 25 g donc la concentration de chaque prise doit être inférieure à 1 colonie par 25 g. Ainsi, la valeur de m est souvent fixée à 0 ufc/g pour un plan à deux classes (ICMSF (2002)). Toujours d'après la même source, pour un plan à trois classes, m est supérieur à 0.

La limite M sépare les lot d'une qualité acceptable des lots de qualité inacceptable. ICMSF (2002) propose trois approches pour fixer M :

- concentration au-delà de laquelle l'odeur ou la couleur varie (si on s'intéresse aux flores d'altération) ou au-delà de laquelle la durée limite de consommation est jugée trop courte ;
- concentration au-delà de laquelle une bactérie servant d'indicateur d'hygiène est trop élevée : la qualité n'est plus garantie ;
- pour un pathogène, concentration liée à la santé : pour la fixer, considérer des données épidémiologiques, de laboratoire ou autre indiquant la dose dangereuse pour la santé, en tenant compte la taille des portions consommées ainsi que les sensibilités des consommateurs.

La valeur de M est définie par le danger, alors que celle de m est déterminée par les BPH.

Selon ANSES (2008), si le microorganisme d'intérêt est utilisé comme indicateur

d'hygiène des procédés^k la valeur de m dans un plan à deux classes est fixée à partir des autocontrôles acquis dans une entreprise respectant les BPH. La limite m est généralement fixée de manière à ce que 95% au moins des résultats observés d'une entreprise aient une concentration inférieure à m . Ce critère nécessite de bien connaître la distribution de concentration de l'indicateur analysé et donc d'avoir de nombreuses données, ce qui n'est pas toujours facile dans le secteur agro-alimentaire. En revanche, l'intérêt est que l'entreprise sait à peu près combien d'analyses fourniront un résultat supérieur à m .

Toujours d'après l'ANSES (ANSES (2008)), M est habituellement fixée à : $M = m + 1$ exprimé en log ufc/g lorsque les données sont obtenues par dénombrement sur milieu de culture solide. L'intérêt de la définition de m et M dans ANSES (2008) est que, si l'indicateur microbien est adapté^l, un petit échantillon suffit pour discriminer l'état "les BPH sont correctement appliquées" de l'état "les BPH sont mal appliquées" car la proportion d'unités analysées dépassant la valeur de m aura tendance à être supérieure à 5%.

1.2.3 Risques fournisseur et client, courbe d'efficacité

Un plan d'échantillonnage est également déterminé par les valeurs de n et de c . Pour fixer celles-ci, on introduit préalablement les notions de risque fournisseur, risque client, niveau de qualité acceptable et niveau de qualité limite.

1.2.3.a Risque fournisseur et risque client

Client et fournisseur s'accordent pour fixer

- le *niveau de qualité acceptable* (NQA) que le fournisseur peut assurer dans des conditions normales de production ;
- le *niveau de qualité limite* (NQL) tel que le lot est inutilisable par le client en raison de sa mauvaise qualité.

Comme la qualité du lot est jugé sur un échantillonnage des produits le constituant, il est possible de rejeter un lot de bonne qualité ou au contraire d'accepter un lot de mauvaise qualité. La probabilité de rejeter un lot de qualité NQA s'appelle le *risque fournisseur* ou risque de première espèce (noté α). A l'inverse, la probabilité d'accepter un lot de qualité NQL s'appelle le *risque client* ou risque de deuxième espèce (noté β).

k. Un indicateur d'hygiène des procédés est un indicateur de l'acceptabilité du fonctionnement hygiénique du procédé de production ou de distribution. Le non-respect d'un critère microbiologique d'hygiène de procédé entraîne des actions correctives, mais ne permet pas de conclure que l'aliment est impropre à la consommation humaine (règlement (CE) Regulation (CE) No 178/2002 (2002), article 14, section 5 : "Pour déterminer si une denrée alimentaire est impropre à la consommation humaine, il est tenu compte de la question de savoir si cette denrée alimentaire est inacceptable pour la consommation humaine compte tenu de l'utilisation prévue, pour des raisons de contamination, d'origine externe ou autre, ou par putréfaction, détérioration ou décomposition"). Les concentrations bactériennes qui causent de tels défauts sont significativement supérieures aux concentrations m des critères d'hygiène de procédé (ANSES (2008)).

l. par exemple, si sa concentration est inversement corrélée à l'application des BPH

On a donc les deux relations suivantes à partir de l'équation (1.1) :

$$\begin{aligned}\mathbb{P}(X \leq c | p = \text{NQA}) &= 1 - \alpha \\ \mathbb{P}(X \leq c | p = \text{NQL}) &= \beta.\end{aligned}$$

A partir de ces deux équations à six inconnues, on peut déterminer deux grandeurs si les quatre autres sont connues.

1.2.3.b Courbe d'efficacité

La courbe d'efficacité évalue la performance d'un plan d'échantillonnage, notamment sa capacité à discriminer les lots conformes des lots non conformes. Dans un contrôle par attribut, elle représente la probabilité d'accepter le lot en fonction de la proportion p de produits non conformes dans le lots. Idéalement, elle doit valoir 1 tant que p est inférieure ou égale à NQA, et 0 dans les autres cas. Pour obtenir une telle courbe, il faudrait découper le lots en prises d'essai et les analyser toutes. Dans la pratique, la courbe d'efficacité est décroissante sur $[0;1]$ et passe par quatre points :

- $(0; 1)$ car un lot composés de produits conformes est accepté avec une probabilité égale à 1 ;
- $(\text{NQA}; 1 - \alpha)$ car un lot avec $p = \text{NQA}$ est accepté avec une probabilité de $1 - \alpha$;
- $(\text{NQL}; \beta)$ car un lot avec $p = \text{NQL}$ est accepté avec une probabilité de β ;
- $(1; 0)$ car un lot composé de produits non conformes est rejeté avec une probabilité égale à 1 ;

Plus la courbe d'efficacité est plate, moins elle est discriminante. Plus n est élevé ou plus c est faible, plus la courbe a une forme en marche d'escalier, comme l'illustre la figure 1.3.

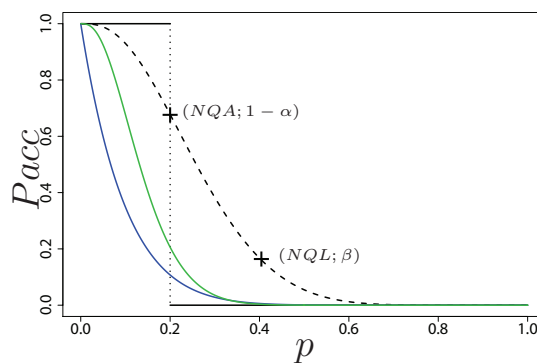


FIGURE 1.3 – Différentes courbes d'efficacité (probabilité d'acceptation en fonction du taux de produits non conformes) : la courbe si toutes les unités du lot sont analysées (trait plein noir), la courbe obtenue avec $n = 10$ et $c = 2$ (trait discontinu) ainsi que deux courbes plus discriminantes que la précédente, l'une avec $n = 10$ et $c = 0$ (trait plein bleu) et l'autre avec $n = 20$ et $c = 2$ (trait plein vert). la courbe en pointillé passe par le point $(\text{NQA}; 1 - \alpha) = (0,2; 0,7)$ et par le point $(\text{NQL}; \beta) = (0,4; 0,18)$.

Bien évidemment, les courbes d'efficacité existent également lorsque l'analyse évalue la concentration en un pathogène et non pas seulement sa présence. Dans ce cas, l'abscisse du graphe est la concentration moyenne (exprimée en log) d'un lot ou le taux de produits défectueux dans le lot, égal à $p = 1 - F(m)$, avec F la fonction de répartition de la loi suivie par la concentration. La probabilité d'accepter un lot est alors égale à la probabilité que la concentration du lot soit inférieure à m : $\mathbb{P}(\text{conc} \leq m)$. Généralement, la loi utilisée pour la concentration est la loi lognormale (ICMSF (2002); Legan et al. (2001)). La figure 1.4 présente différentes courbes d'efficacité pour un lot dont la concentration suit une loi log normale de moyenne μ variant de -3 à 2 log ufc/g et d'écart-type $\sigma = 1$ log ufc/g, avec $n = 30$ et trois valeurs de c . Plus c est élevé, plus le plan d'échantillonnage est sélectif.

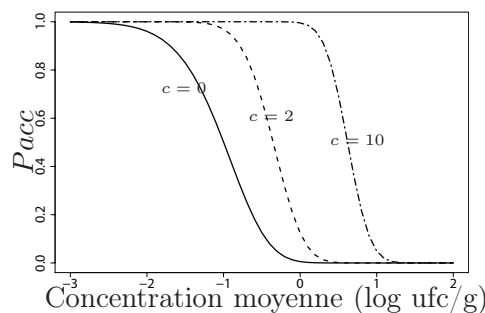


FIGURE 1.4 – Différentes courbes d'efficacité (probabilité d'acceptation en fonction de la log concentration moyenne du lot pour $n = 30$ et trois valeurs de c : 0, 2 et 10).

Les courbes d'efficacité existent également pour les plans d'échantillonnage à trois classes : elle représentent la probabilité d'accepter un lot en fonction du taux de produits acceptables et marginalement acceptables. Comme elles dépendent de deux paramètres, je préfère employer le terme de surface d'efficacité. La figure 1.5 présente un exemple de surface pour $n=30$ et $c=2$. Cette représentation ne permet pas de déterminer graphiquement les probabilités d'acceptation, mais elle a le mérite de souligner que cette probabilité dépend des proportions de produits marginalement acceptable P_m et de produits défectueux P_d . Aux surfaces d'efficacité, l'ICMSF (2002) préfère des *contour maps*, qui sont des lignes de niveau de la probabilité d'acceptation en fonction des proportions de produits défectueux et marginalement acceptables. Les courbes d'efficacité en fonction de la concentration moyenne d'un lot telles que définies pour les plans à deux classes peuvent également être tracées pour les plans à trois classes, ce qui est utile pour les comparaisons de plans.

1.2.4 Plan à deux ou à trois classes ?

Quel type de plan choisir ? Si les analyses fournissent un résultat qualitatif (par exemple présence ou absence d'un pathogène), seul le plan à deux classes est applicable.

Lorsque les résultats des analyses sont quantitatifs, les deux plans sont applicables et peuvent donc être comparés. Dès lors que la concentration est supérieure à m , le plan à deux classes n'opère aucune distinction selon que cette valeur est inférieure ou supérieure

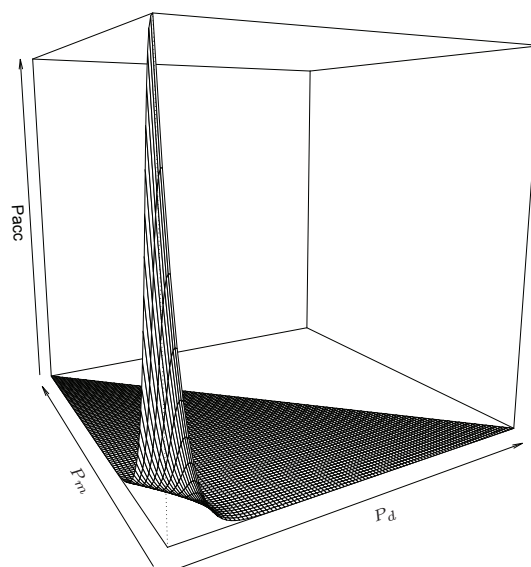


FIGURE 1.5 – Surface d'efficacité pour un plan à trois classes avec $n=30$, $c=2$ en fonction des proportions de produits marginalement acceptable P_m et de produits défectueux P_d .

à M . Pour n fixe et c variable, la surface d'efficacité d'un plan à trois classes varie surtout le long de l'axe des proportions marginalement acceptables. L'ICMSF (2002) interprète ce phénomène en avançant qu'un plan d'échantillonnage peut être analysé comme un mélange de deux plans à deux classes, l'un avec n et c complétés par la limite m , l'autre avec les valeurs n et 0 complétées par la limite M . Lorsque c varie, le second plan du mélange reste inchangé. Il est donc logique de vérifier la quasi-constance de l'acceptation en fonction de la proportion de produits défectueux. Ce phénomène est illustré figure 1.6. Entre les figures 1.6a et 1.6b, la surface varie beaucoup en fonction de P_m mais peu en fonction de P_d . La figure 1.6a est un agrandissement d'une partie de la figure 1.5, car P_{acc} y varie de 0 à 0,15 au lieu de 0 à 1.

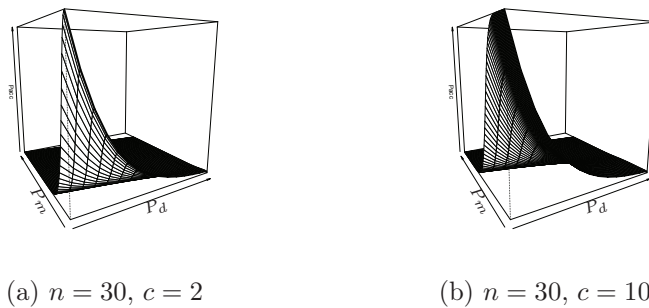


FIGURE 1.6 – Surface d'efficacité pour un plan à trois classes avec $n=30$ et (a) $c = 2$ ou (b) $c = 10$.

Un plan à trois classes est toujours moins discriminant qu'un plan à deux classes si les valeurs de n , c et m sont communes aux deux plans. En effet, le tableau 1.2 récapitule les différentes formules de probabilités d'acceptation d'un lot ainsi que les éléments constituant chaque formule. Il s'ensuit que $1 - p = P_a$ et que $p \geq P_m$. Bien évidemment, cette comparaison ne vaut que si les résultats des analyses sont quantitatifs. La différence entre un plan à deux classes et un plan à trois classes est d'autant plus grande que le logarithme de l'écart-type σ est grand car plus σ est grand et plus $\Phi((M - \mu)/\sigma)$ est éloigné de 1 donc p et P_m d'une part et $1 - p$ et P_a d'autre part prennent des valeurs très différentes donc P_{acc} également.

Plan	P_{acc}	Proportions
2 classes	$\sum_{i=0}^c \frac{n!}{i!(n-i)!} (1-p)^{n-i} p^i$	de défectueux : $p = 1 - F(m)$
3 classes	$\sum_{i=0}^c \frac{n!}{i!(n-i)!} (P_a)^{n-i} P_m^i$	acceptable : $P_a = F(m)$ marginale acceptable : $P_m = F(M) - F(m)$

TABLE 1.2 – Formules des probabilités d'acceptation P_{acc} pour des plans à deux classes et à trois classes et éléments de ces formules en fonction de n , c , m et M . Φ la fonction de répartition de la loi normale centrée réduite.

1.2.5 Détermination de n et c

Supposons que le gestionnaire applique un plan de contrôle à deux classe. Il prend donc une décision dépendant de x , le nombre d'unités analysées non conformes : soit il ne fournit pas le lot qu'il juge de qualité non-acceptable, soit il le livre car le lot est jugé de qualité acceptable. Comment fixer n et c de manière à ce que fournisseur et client soient d'accord sur la qualité des produits livrés ?

1.2.5.a Par approximation normale d'un test unilatéral

Voici une méthode pour trouver la décision optimale :

- fixer les paramètres NQA, NQL, α , et β ;
- définir un modèle probabiliste pour le nombre de non-conformes dans un échantillon de taille n ;
- trouver le plan d'échantillonnage (n^*, c^*) qui permette d'atteindre un risque fournisseur α en $\theta = \text{NQA}$ et un risque client β en $\theta = \text{NQL}$ selon les équations :

$$\mathbb{P}(X \leq c | \theta = \text{NQA}) = 1 - \alpha$$

$$\mathbb{P}(X \leq c | \theta = \text{NQL}) = \beta$$

La variable aléatoire X du nombre d'analyses non conformes parmi n suit une loi hypergéométrique de paramètres N , Np^m et n , avec N le nombre d'unités dans un lot. Si $n \ll N$, on peut approximer la loi de X par une loi binomiale de paramètres n et p .

$$\sum_{k=0}^c C_n^k \text{NQA}^k (1 - \text{NQA})^{n-k} = 1 - \alpha$$

$$\sum_{k=0}^c C_n^k \text{NQL}^k (1 - \text{NQL})^{n-k} = \beta.$$

Si on peut approximer la loi binomiale par une loi normale de paramètres np et $np(1-p)$, il vient :

$$\mathbb{P}(X \leq c | p = \text{NQA}) = \Phi \left(\frac{c - np}{\sqrt{np(1-p)}} \right) = 1 - \alpha$$

$$\mathbb{P}(X \leq c | p = \text{NQL}) = \Phi \left(\frac{c - np}{\sqrt{np(1-p)}} \right) = \beta,$$

Φ étant la fonction de répartition de la loi normale centrée réduite. Par exemple, si les niveaux NQA, NQL et les risques α et β sont fixés, les valeurs de n et c sont déterminées par les équations suivantes :

$$c = z_{1-\alpha} \sqrt{n \text{NQA}(1 - \text{NQA})} + n \text{NQA}$$

$$n = \left(\frac{z_{1-\alpha} \sqrt{\text{NQA}(1 - \text{NQA})} - z_{\beta} \sqrt{\text{NQL}(1 - \text{NQL})}}{\text{NQA} - \text{NQL}} \right)^2,$$

m. Np est le nombre d'unité contaminées dans le lot (cf. les paramètres de la loi hypergéométrique

où z_a désigne le quantile a de la loi normale centrée réduite. Si le client et le fournisseur arrivent à déterminer NQA, NQL ainsi que les risques de première et de deuxième espèce, la taille n de l'échantillon et le seuil c en découlent directement.

Le plan d'échantillonnage découle d'un test d'hypothèse :

$$H_0 : \theta \leq \text{NQA} \quad \text{vs.} \quad H_1 : \theta > \text{NQA}.$$

Le tableau 1.3 présente quelques valeurs de n et de c lorsque NQA, NQL, α et β varient. Plus NQA et NQL sont proches, plus n est élevé.

NQA	NQL	α	β	n	c
0,01	0,45	0,05	0,05	5	0
0,07	0,25	0,05	0,05	40	5
0,06	0,21	0,10	0,05	40	4
0,13	0,30	0,05	0,05	60	12

TABLE 1.3 – Taille n et seuil c du plan d'échantillonnages en fonction du niveau de qualité acceptable NQA, du niveau de qualité limite NQL, du risque fournisseur α et du risque fournisseur β .

1.2.5.b Proposition de l'ICMSF

L'ICMSF (1986), repris notamment par Legan et al. (2001) et ICMSF (2002), a défini 15 plans d'échantillonnages possibles en fonction du degré de danger du microorganisme recherché et en fonction de l'évolution du risque après que le produit est échantillonné. L'échelle des dangers comporte 5 échelons :

- utilité (par exemple contamination générale, altération ou diminution de durée de vie) ;
- indicateur ;
- danger modéré (normalement, pas d'issue fatale pour le consommateur ni de séquelles ; les symptômes sont de courte durée mais peuvent être très incommodes) ;
- danger sérieux : les séquelles sont rares, la durée des symptômes est modérée, ils handicapants mais ne conduisent pas à une issue fatale ;
- danger sévère pour la population générale ou pour une sous-population causant soit des séquelles, une maladie de longue durée ou une issue létale.

Le risque peut évoluer de trois manières différentes entre l'instant où l'analyse a lieu et la consommation :

- réduction du risque ;
- risque inchangé ;
- augmentation du risque.

Le tableau 1.4 présente les différents plans (à deux ou à trois classes) en fonction du type de danger et de l'évolution du risque.

Déterminer le plan d'échantillonnage approprié nécessite de bien connaître le microorganisme analysé ainsi que le devenir normalement prévisible de l'aliment. Certains

Evolution prévisible du risque dans des conditions normales de conservation et de consommation						
Type de danger	m (ufc/g)	M (ufc/g)	Réduction du risque	Risque inchangé	Augmentation possible du risque	
Utilité	10^3	10^4	Cas 1 3 classes, n=5, c=3 $5,13 \times 10^3$ ufc/g	Cas 2 3 classes, n=5, c=2 $3,31 \times 10^3$ ufc/g	Cas 3 3 classes, n=5, c=1 $1,82 \times 10^3$ ufc/g	
Indicateur	10^3	10^4	Cas 4 3 classes, n=5, c=3 $5,13 \times 10^3$ ufc/g	Cas 5 3 classes, n=5, c=2 $3,31 \times 10^3$ ufc/g	Cas 6 3 classes, n=5, c=1 $1,82 \times 10^3$ ufc/g	
Danger modéré	10^3	10^4	Cas 7 3 classes, n=5, c=2 $3,31 \times 10^3$ ufc/g	Cas 8 3 classes, n=5, c=1 $1,82 \times 10^3$ ufc/g	Cas 9 3 classes, n=10, c=1 $5,75 \times 10^2$ ufc/g	
Danger sérieux	0/25g	n.a.	Cas 10 2 classes, n=5, c=0 1 ufc/32g	Cas 11 2 classes, n=10, c=0 1 ufc/83g	Cas 12 2 classes, n=20, c=0 1 ufc/185g	
Danger sévère	0/25g	n.a.	Cas 13 2 classes, n=15, c=0 1 ufc/135g	Cas 14 2 classes, n=10, c=0 1 ufc/278g	Cas 15 2 classes, n=60, c=0 1 ufc/526g	

TABLE 1.4 – Plans d'échantillonnages proposés par ICMSF (1986) en fonction du type du micro-organisme recherché (5 niveaux) et de l'évolution du risque entre l'analyse et la consommation du produit (3 niveaux). La concentration indiquée pour chaque cas est la concentration moyenne pour un lot dont la probabilité d'acceptation est de 0,05 (l'écart-type de la concentration de chaque lot est de $0,8 \log$ ufc/g).

usages sont moins courants que d'autres, mais l'entreprise doit en tenir compte pour déterminer le plan d'échantillonnage adapté, ce qui est d'ailleurs une exigence au sein de l'UE.

1.2.5.c Critique de l'approche classique

D'après ICMSF (2002), le plan d'échantillonnage est indispensable si un Etat reçoit d'un autre pays un lot qu'il doit autoriser ou non le lot à entrer sur le territoire alors qu'il ne dispose guère d'information sur la manière et sur les conditions dans lesquelles ce lot a été fabriqué. Cependant, le plan d'échantillonnage utilisé seul peut donner le sentiment inexact d'un produit sûr quand aucun pathogène n'a été trouvé dans les produits analysés : l'inocuité peut être illusoire si le pathogène d'intérêt présente un risque à faible concentration, surtout lorsque la taille de l'échantillon est faible.

Si le procédé d'un produit a été élaboré en considérant les BPH et le système HACCP et si les matières premières ont été choisies de manière à ce que la contamination soit maîtrisée, on peut se demander si l'échantillonnage est un outil adapté (ICMSF (2002)). En effet, l'ICMSF reproche aux analyses microbiologiques d'être coûteuses en temps, d'avoir une sensibilité et une spécificité faibles et d'avoir une probabilité faible de détecter une contamination lorsque les données acquises précédemment montrent que la proportion de défectueux est faible. La sensibilité d'une méthode d'analyse est la probabilité qu'une prise d'essai comportant le pathogène d'intérêt soit déclarée positive. A l'inverse, la spécificité d'une méthode d'analyse est la probabilité qu'une prise d'essai sans pathogène d'intérêt soit déclarée négative. Les plans d'échantillonnage peu discriminants (par exemple si $n = 5$) ne sont adaptés que si la proportion de produits défectueux est élevée. Augmenter n permet de rendre un plan d'échantillonnage plus discriminant (cf. partie 1.2.3.b). Le coût de l'échantillonnage additionnel doit être comparé avec les gains liés à l'augmentation de la discrimination, ainsi qu'à l'incidence d'une mauvaise décision.

Bien que l'utilisation de tests statistiques conduise en principe à prendre des décisions à risque contrôlé, elle ne prend pas en compte les conséquences socio-économiques de chaque décision (comme par exemple le rejet d'un lot de qualité acceptable ou la mise sur le marché d'un lot de qualité inacceptable), d'autant plus que la pratique est d'opter pour des niveaux forfaitaires de risque fixés à 5% ou 1%. Si un industriel souhaite dimensionner un plan d'échantillonnage permettant de minimiser les conséquences socio-économiques associées à une erreur d'appréciation du risque microbiologique global d'un lot, l'approche statistique présentée ci-dessus n'est pas adaptée.

Dans ce contexte, l'approche statistique bayésienne couplée à la théorie de la décision nous a semblé une voie d'approche privilégiée car elle permet de formaliser les conséquences sous forme de coût et de définir, pour chaque décision envisagée, un nombre unique prenant en compte l'incertitude portant sur la concentration d'un lot et sur les données disponibles. Plus d'informations sur l'origine de l'approche bayésienne sont disponibles à l'annexe A.

La partie suivante présente les bases de la théorie bayésienne de la décision. Deux applications de cette démarche sont présentées aux chapitres 6 et 7.

1.3 Démarche initiale et bases de la théorie statistique de la décision

Dans mon travail de thèse, j'ai appliqué la théorie bayésienne de la décision. Pour ce faire, j'ai déterminé (Robert (2006)) :

1. l'ensemble \mathcal{D} des décisions ;
2. l'ensemble Θ des états de la nature, ainsi que la distribution *a priori* de θ notée $[\theta]$;
3. l'ensemble \mathcal{X} des observations et la famille des distributions, notées $[x|\theta]$;
4. la fonction de coût L associée aux décisions et aux états de la nature.

Avant de présenter l'approche bayésienne, je commence par décrire la théorie de la décision (non nécessairement bayésienne), puis j'introduirai le point de vue bayésien lorsque ce sera nécessaire.

En quoi la théorie statistique de la décision en avenir incertain pourrait-elle nous être utile ?

Prendre une décision en situation certaine est relativement simple : il suffit d'énumérer toutes les options possibles et de choisir celle qui nous convient le mieux (cours de Wilkinson (2008)). Comment faire en situation d'incertitude ? En statistique, la théorie de la décision s'attache à fournir une décision à l'analyste, ou plutôt à son client. Les décisions sont de différents types. Il peut s'agir par exemple de décider si une personne est porteuse d'une certaine maladie, ou, dans un tout autre domaine, de décider ce que je vais faire demain : me promener ou rester chez moi ? Dans ces deux exemples, il faut distinguer deux ensembles :

- celui des décisions
 - Exemple 1 : la personne est mise en arrêt de travail ou considérée comme valide ;
 - Exemple 2 : se promener ou rester chez soi.
- celui des états incertains de la nature (ils existent mais ne sont pas directement ou complètement observés)
 - Exemple 1 : la personne est malade ou non ;
 - Exemple 2 : météo de la journée à venir, que l'on réduit à deux états possibles (présence ou absence de précipitations).

Dans les exemples ci-dessus, les états de la nature ne sont pas connus car l'observation n'a pas encore eu lieu. Plus formellement, l'ensemble des décisions d possibles est noté \mathcal{D} et l'ensemble des états θ de la nature est noté Θ . Ces ensembles peuvent être continus ou discrets. L'ensemble Θ est connu mais on ne sait pas avec quelle valeur θ est instancée.

Le cadre de la théorie de la décision introduit une fonction de coût qui décrit les pertes occasionnées en fonction de l'état de la nature et de la décisionⁿ. Cette fonction sera notée L (comme *loss*) dans la suite de ce travail. Elle est définie de $\mathcal{D} \times \Theta$ dans \mathbb{R}^+ . Comme on le verra par la suite, la fonction de coût peut être difficile à établir, même si \mathcal{D} et Θ sont de petits ensembles. Il est également difficile de déterminer le coût d'un événement extrêmement rare et grave. Par ailleurs, pour un modèle qualitatif, il peut être

n. Il est également possible de traiter les gains au lieu des coûts ; la fonction s'appelle alors une fonction d'utilité, elle est égale à l'opposé de la fonction de coût.

difficile d'ordonner mais surtout de quantifier une décision. Ainsi, même si je préfère rester chez moi alors qu'il fait beau plutôt que d'aller me promener sous la pluie, il n'est pas facile de choisir jusqu'à quel point je préfère la première solution : deux fois plus ? trois fois plus ?

Tout l'enjeu est de déterminer la décision d minimisant la fonction de coût L malgré la connaissance partielle des états de la nature. Pour obtenir des informations sur θ , on utilise des données x , issues de l'observation (Berger (1985)). L'ensemble des données est noté \mathfrak{X} . Les observations et les décisions sont reliées par une règle de décision notée δ et définie de \mathfrak{X} dans \mathcal{D} associant une décision $d = \delta(x)$ pour toute observation x . Dit plus simplement, la règle de décision permet au décideur de prendre la décision d après avoir observé x .

Dans les exemples cités ci-dessus, il y a donc :

- Ensemble des observations :
 - Exemple 1 : résultats de la prise de sang (positif ou négatif) ;
 - Exemple 2 : informations sur la météo concernant le jour où je décide de me promener ou de rester chez moi. Pour simplifier, on peut réduire cet ensemble à deux états : $\mathfrak{X} = \{\text{soleil ; pluie}\}$
- Exemples de règles de décisions
 - Exemple 1 : si le résultat de la prise de sang est positif, un médecin met la personne en arrêt maladie et si le résultat est négatif, le médecin la déclare valide. Une autre règle possible est de faire l'inverse ;
 - Exemple 2 : s'il fait beau, je sors me promener ; s'il pleut, je reste chez moi. Une autre règle de décision est de rester chez soi s'il y a du soleil et de sortir s'il pleut (a priori plus surprenante, cette règle s'explique aisément pour une personne allergique aux pollens).

Il existe également des décisions randomisées : elle sont générées à partir d'une distribution une fois que l'observation est faite.

Si $\Theta = \mathcal{D}$, d est un estimateur de θ . La valeur $L(\delta(x) = d, \theta)$ quantifie donc l'écart entre la vraie valeur θ prise par l'état de la nature et l'estimation d qui en est faite. La décision $\delta(x)$ est alors appelée estimateur de θ . Dans ce cas, prendre une décision revient à proposer une valeur pour l'état inconnu de la nature. La plupart des exemples théoriques traitent ce cas (Robert (2006)), généralement avec une fonction de coût quadratique (cf. partie 1.3.2).

La fonction de coût L doit posséder plusieurs propriétés, détaillées à l'annexe B.

1.3.1 Risques fréquentiste et bayésien

La théorie de la décision est utilisée pour déterminer la décision d^* minimisant la fonction de coût L . Sauf exception, il n'est pas possible de maximiser uniformément en d la fonction L puisque θ est inconnu. Une manière naturelle de procéder est de considérer un coût espéré de prendre la décision d (Berger (1985); Parent and Bernier (2007)) qui minimise L .

1.3.1.a Risque fréquentiste

Les fréquentistes adoptent un coût moyenné selon la variable X . Le décideur cherche à évaluer, pour chaque valeur de θ , combien il s'attend à "perdre" si la décision $\delta(x)$ est prise. Le risque est donc défini de la manière suivante :

$$\begin{aligned} R(\theta, \delta) &= \mathbb{E}_X(L(\theta, \delta(X) = d)) \\ &= \int_{\mathbf{x}} L(\theta, \delta(x) = d)[x|\theta]dx \end{aligned}$$

Dans le cadre de cette thèse, seules les décisions aboutissant à un risque fini sont prises en considération.

Le gestionnaire souhaite prendre la règle de décision δ qui minimise $R(\theta, \delta)$. Comme R dépend de θ qui est inconnu, il faut définir comment ordonner les valeurs de R quand δ varie. Une décision ne minimise pas toujours R uniformément en θ . Ainsi, une fonction de décision δ_1 est dite non admissible s'il existe une fonction δ_2 telle que $R(\theta, \delta_2) \leq R(\theta, \delta_1)$ pour tout θ appartenant à Θ avec une inégalité stricte pour au moins une valeur de θ . Si une telle fonction δ_2 n'existe pas, δ_1 est dite admissible. Il est clair qu'une décision inadmissible ne devrait pas être utilisée puisqu'il existe une fonction de décision aboutissant à un risque moins élevé de façon uniforme en θ , mais déterminer une décision admissible est parfois ardu... Par ailleurs, un problème donné peut admettre plusieurs décisions admissibles, par exemple celles aboutissant à un risque plus petit pour des valeurs différentes de θ . C'est le cas dans l'exemple suivant, repris de Berger (1985).

Soit une variable aléatoire X de loi $\mathcal{N}(\theta, 1)$. On souhaite estimer θ avec la fonction de coût quadratique traditionnelle $L(\theta, \delta) = (\theta - \delta)^2$. Soit la règle de décision (appelée également estimateur dans ce cas) $\delta_c(x) = cx$, alors :

$$\begin{aligned} R(\theta, \delta_c) &= \mathbb{E}_X[L(\theta, \delta_c(x))] \\ &= \mathbb{E}_X[(\theta - cx)^2] \\ &= \mathbb{E}_X[(c(\theta - x) + (1 - c)\theta)^2] \\ &= c^2\mathbb{E}_X[(\theta - x)^2] + 2c(1 - c)\theta\mathbb{E}_X[\theta - x] + (1 - c)^2\theta^2 \\ &= c^2 + (1 - c)^2\theta^2. \end{aligned}$$

Pour $c > 1$, $R(\theta, \delta_1) = 1 < R(\theta, \delta_c)$, donc la fonction de décision δ_c est inadmissible. Si $c \leq 1$, les décisions δ_1 et δ_c ne sont pas comparables et sont toutes les deux admissibles car les fonctions de risque se croisent, ce qui est visible à la figure 1.7 où les estimateurs δ_1 et $\delta_{1/2}$ sont représentés. Il est à noter que l'estimateur δ_0 , pour lequel $\theta = 0$ pour toute valeur de x , est également un estimateur admissible. Qu'une règle de décision estimant le paramètre indépendamment des données soit ici admissible montre un résultat important : l'admissibilité a beau être souhaitable, elle ne garantit en rien une décision raisonnable puisqu'une règle de décision estimant le paramètre indépendamment des données est ici admissible.

L'approche fréquentiste présente plusieurs difficultés, en partie parce que le risque dépend de θ :

- elle suppose que le problème soit rencontré un grand nombre de fois pour donner un sens à l'évaluation en fréquence. Le risque $R(\theta, \delta)$ s'interprète comme le coût

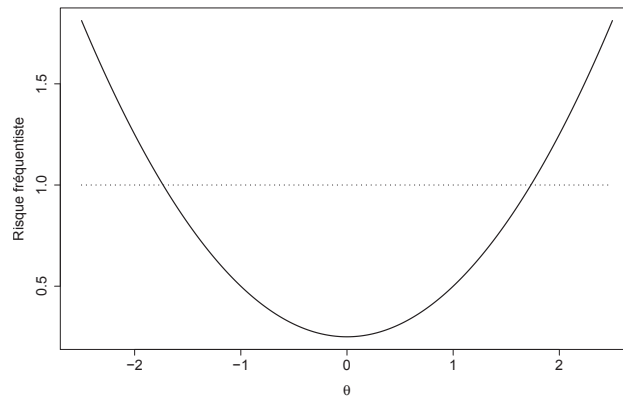


FIGURE 1.7 – Risque fréquentiste de type $R(\theta, \delta_c) = c^2 + (1 - c)^2 \theta^2$ pour $c = 1/2$ (trait plein) et pour $c = 1$ (trait discontinu).

- moyen sur des répétitions i.i.d.^o de la même expérience. Or la répétabilité pratique d'une expérience est fortement controversée (par exemple, Jeffreys (1961a)) ;
- comme écrit ci-dessus, $R(\theta, \delta)$ est une fonction de θ inconnu. Plusieurs règles de décision peuvent cohabiter sans fonction d'ordre total. Une manière de générer toutes des règles admissibles est de pondérer le risque par une mesure positive π_0 que l'on fait varier et de chercher l'optimum du risque ainsi moyenné (l'argument est alors appelé règle de Bayes associée à la mesure π_0). Sous des conditions générale de régularité (théorème de la classe complète de Wald), toute règle admissible est soit une règle de Bayes, soit une limite de règles de Bayes. Se pose alors l'interprétation de la pondération *a priori*. Pour les fréquentistes comme Wald, c'est une technique mathématique de convexification qui permet de générer les règles admissibles. Est bayésien celui qui, comme Savage, considère π_0 comme une distribution de probabilité de θ fidèle aux croyances qu'il a sur θ avant d'examiner les observations.

1.3.1.b Risque bayésien

L'approche bayésienne commence donc par affecter une distribution *a posteriori* au paramètre θ , cette distribution reflétant les croyances qu'a le gestionnaire, par exemple sur les états de la nature.

Ensuite, elle intègre la fonction de coût sur l'espace Θ . Ainsi, le *coût moyen a posteriori* est une fonction définie de $\mathcal{D} \times \mathfrak{X}$ dans \mathbb{R}^+ :

$$\rho(d|x) = \mathbb{E}_\theta(L(\theta, d|x)) = \int_{\Theta} L(\theta, \delta(x) = d)[\theta|x]d\theta,$$

Il s'agit donc de l'espérance de la fonction de coût selon la distribution *a posteriori* de θ . Ce coût moyen est conditionnel à x , qui lui est observé, contrairement à θ .

^o. indépendantes et identiquement distribuées

On définit également le *risque bayésien* comme l'espérance *a priori* du risque fréquentiste :

$$\begin{aligned} r(\delta) &= \mathbb{E}_\theta(R(\theta, \delta)) \\ &= \int_{\Theta} \int_{\mathfrak{X}} L(\theta, \delta(x)) [x|\theta][\theta] dx d\theta \end{aligned}$$

La fonction r peut également s'écrire $\delta \mapsto r(\delta) = \int_{\Theta} \int_{\mathfrak{X}} L(\theta, \delta(x)) [x, \theta] d\theta dx$ ou encore $\delta \mapsto r(\delta) = \int_{\mathfrak{X}} \int_{\Theta} L(\theta, \delta(x)) [\theta|x][x] d\theta dx$ en vertu du théorème de Fubini. L'intérêt de cette définition est d'introduire un ordre total sur l'ensemble des estimateurs : contrairement à la pratique fréquentiste, elle associe à chaque règle de décision un nombre et non une fonction de θ . Les règles peuvent donc être comparées entre elle. Même si $\rho(d|x)$ est défini pour chaque valeur de x alors que $r(\delta)$ est indépendant de x , la décision minimisant ρ minimise également r . En effet, la fonction de décision δ^* qui minimise $r(\delta)$ est obtenue par sélection, pour chaque x de \mathfrak{X} , de la valeur $\delta^*(x)$ qui minimise $\rho(\delta|x)$ car $r(\delta) = \int_{\mathfrak{X}} \rho(d|x)[x] dx$ (avec $[x]$ la loi marginale de x). Ainsi, un *estimateur de Bayes* associé à une distribution *a priori* donnée de θ et à une fonction de coût L est par définition un estimateur minimisant $r(\delta)$. Pour chaque x de \mathfrak{X} , ce dernier est donné par :

$$\delta^*(x) = \arg \min_{\delta} \rho(\delta|x).$$

En outre, ce dernier permet également de travailler lorsque les observations n'ont pas encore été effectuées, c'est-à-dire lorsque l'on se place en prédictif via la distribution des prédictions $[x|\theta]$. Dans ce cas, le décideur doit d'abord choisir le type d'expérimentation e qui lui fournira \mathfrak{X}_e puis x_e . La règle de décision δ_e dépend également de e et associe une décision à prendre après avoir observé x_e . La décision porte sur les observations $x|e$ à venir, il faut donc envisager la règle de décision δ_e qui fait correspondre une action à entreprendre à chaque observation après expérimentation e . Le coût associé est donc $L(\theta, \delta_e(x))$. On cherche alors la règle δ_e^* qui minimise le risque bayésien $r(\delta_e) = \int_{\Theta} \int_{\mathfrak{X}} L(\theta, \delta_e(x)) [x, \theta|e] d\theta dx$. Ce résultat dépend lui aussi de e . On peut ensuite rechercher l'expérimentation e qui minimise $r(\delta_e)$. Un exemple sera examiné dans la partie 6. Choisir e implique que les conséquences du choix du décideur dépendront de la décision prise qui sera soumise d'une part à l'incertitude sur l'état de la nature et d'autre part à l'incertitude sur les observations générées par l'expérimentation e (Parent and Bernier (2007)).

Une fois la décision d prise, comment apprécier sa pertinence ? Elle aurait pu être différente (et peut-être plus adaptée), avec une autre connaissance sur θ . En particulier, si θ était connu, une certaine décision d' serait prise. Comment mesurer l'écart entre d' et la décision prise en état d'incertitude ? Par la fonction de *regret*. Supposons que la fonction L soit suffisamment régulière pour admettre une décision d' qui l'optimise si l'état de la nature θ est connu.

Ainsi, $d' = \arg \min_{\delta} \int \int L(\theta, \delta(\theta)) [x, \theta] dx d\theta^{\mathbb{P}}$. Dans la réalité, θ est inconnu et la décision d est prise. L'écart $\underline{L}(\theta, d) = L(\theta, d) - L(\theta, d'(\theta))$ est appelé *regret*. C'est le

p. Par un abus de notation, d' n'est pas une décision mais une fonction de décision dépendant de θ . Ce n'est pas une règle de décision car d' dépend de θ et non de x .

surcoût par rapport au coût optimal dû à une méconnaissance de θ , c'est-à-dire le surcoût lié au fait d'avoir choisi d au lieu de d' . Bien évidemment, $\underline{L}(\theta, d) \geq 0$. La fonction δ^* minimisant le risque bayésien minimise également le regret moyen (*expected opportunity loss* en anglais), à savoir : $A = \int_{\Theta} \int_{\mathcal{X}} (L(\theta, \delta(x)) - L(\theta, d'(\theta)))[x, \theta] d\theta dx$. En effet, comme δ^* est la fonction qui à tout x associe $\delta^*(x)$ qui minimise $\rho(\delta|x)$ et $r(\delta)$, elle minimise également $\int \rho(\delta|x)[x] dx - \rho(d')$ car $\rho(d')$ est indépendant de x . Cette quantité A mesure la valeur moyenne de l'information, puisqu'elle mesure la perte d'efficacité de la décision du fait de notre connaissance incomplète de l'état de la nature θ .

1.3.2 Quelques fonctions usuelles de coûts

Le coût quadratique C'est le coût communément admis par défaut, il est défini par :

$$L(\theta, d) = (\theta - d)^2$$

La fonction δ est un estimateur de θ et le risque $R(\theta, \delta) = \mathbb{E}_X[(\theta - \delta(X))^2]$ est la variance de l'estimateur. En outre, ce coût est pratique pour les calculs. Ce coût est approprié quand une perte symétrique en $\theta - d$ est acceptable et que la forme exacte de la fonction n'a pas trop d'importance. Il a pourtant été très critiqué (Robert (2006)), notamment parce qu'il pénalise trop fortement les grandes erreurs et qu'il accorde un poids équivalent aux sur et aux sous-estimations. L'estimateur de Bayes associé à cette fonction de coût est la moyenne *a posteriori* $\mathbb{E}(\theta|x)$.

Le coût linéaire La fonction de coût est linéaire par morceaux :

$$L(\theta, d) = \begin{cases} K_0(\theta - d), & \text{si } \theta - d \geq 0 \\ K_1(d - \theta), & \text{si } \theta - d < 0 \end{cases}$$

Ce type de fonction ne pénalise pas de manière exagérée les erreurs élevées mais peu vraisemblables. Les constantes K_0 et K_1 peuvent être égales ; dans ce cas, on obtient une fonction de coût absolu. L'estimateur de Bayes associé à la fonction de coût linéaire est le quantile $\frac{K_0}{K_1+K_0}$ de la loi *a posteriori* de θ , c'est-à-dire : $\mathbb{P}(\theta < d|x) = \frac{K_0}{K_1+K_0}$.

Le coût 0 - 1 Ce type de coût est souvent utilisé dans les tests d'hypothèses (approche classique). Soit le test $H_0 : \theta \in \Theta_0$ contre $H_1 : \theta \notin \Theta_0$. L'ensemble des décisions \mathcal{D} vaut $\{0; 1\}$, 1 est l'acceptation de H_0 et 0 son rejet. La fonction de coût vaut :

$$L(\theta, d) = \begin{cases} 1 - d, & \text{si } \theta \in \Theta_0 \\ d, & \text{sinon} \end{cases}$$

Avec cette fonction de coût, le risque est égal à :

$$\begin{aligned} R(\theta, \delta) &= \mathbb{E}[L(\theta, \delta(X))] \\ &= \begin{cases} \mathbb{P}_\theta(\delta(x) = 0), & \text{si } \theta \in \Theta_0 \\ \mathbb{P}_\theta(\delta(x) = 1), & \text{sinon,} \end{cases} \end{aligned}$$

On retrouve alors les erreurs de première et de deuxième espèce. En pratique, ce coût est peu utilisé, car peu réaliste (Berger (1985)). On peut toutefois élaborer une fonction de perte fonctionnant selon le même principe, mais plus adaptée :

$$L(\theta, d) = \begin{cases} K_0(\theta)(1 - d), & \text{si } \theta \in \Theta_0 \\ K_1(\theta)d, & \text{si } \theta \notin \Theta_0 \end{cases}$$

Dans cette fonction, $K_i(\theta)$ est une fonction de θ , ce qui semble raisonnable, puisque la conséquence d'une mauvaise décision dépendra la sévérité de l'erreur commise. Bien évidemment, K_i peut être indépendant de θ . La décision 1 est la bonne décision à prendre si l'on savait que θ appartient à Θ_0 tandis que 0 est la bonne décision si $\theta \notin \Theta_0$. Le tableau 1.5 présente les valeurs prises par L en fonction de d et de θ (fonction avec K_0 et K_1 indépendants de θ).

		Décisions	
		$d = 1$ (acceptation de H_0)	$d = 0$ (rejet de H_0)
θ	$\theta \in \Theta_0$	0	K_0
	$\theta \notin \Theta_0$	K_1	0

TABLE 1.5 – Valeurs prises par la fonction de coût L .

Revenons à l'exemple 1 sur la prise de sang et la prédiction d'une maladie. On observe un résultat expérimental x , dont la loi $[x|\theta]$ dépend de l'état θ du patient. Le diagnostic du médecin est une règle de décision qui conduit à l'une des deux décisions possibles : $d = 1$, l'analyse est positive et $d = 0$ l'analyse est négative. Par ailleurs, l'état θ (inconnu) de chaque personne est : malade ($\theta \in \Theta_0$) ou non malade ($\theta \notin \Theta_0$), avec Θ_0 l'événement "la personne est malade". Si le résultat de l'analyse est en accord avec l'état réel de la personne (*vrai positif* : l'analyse prédit qu'une personne est malade et elle l'est réellement ou *vrai négatif* : l'analyse prédit qu'une personne en bonne santé n'est pas porteuse de la maladie recherchée), le coût est nul car l'analyse conduit à un résultat juste. En revanche, si l'analyse conduit à un résultat incorrect (l'analyse prédit que la personne est saine alors qu'elle est malade -*faux négatif*, ou l'analyse prédit que la personne est malade alors qu'elle est en bonne santé -*faux positif*), il y a une pénalité. Celle-ci diffère selon les cas, puisque les conséquences pour une personne ne sont pas les mêmes selon qu'elle est diagnostiquée à tort malade ou déclarée à tort en bonne santé. Les valeurs attribuées à K_0 et K_1 permettent de déterminer le pire cas de figure quand l'analyse se trompe. Si $K_0 > K_1$, il est plus grave d'avoir un faux négatif qu'un faux positif.

Un moyen d'évaluer une méthode d'analyse est de déterminer sa *sensibilité* et sa *spécificité*. Si la sensibilité et la spécificité sont égales à 1, le résultat de la prise de sang est toujours juste (l'état de la nature est donc connu avec certitude). Dans ce cas, la meilleure décision δ' est de déclarer que l'état de la personne est celui prédit par l'analyse. Par conséquent, $L(\theta, \delta'(\theta))$ est nulle. La fonction de regret est alors confondue avec la fonction de coût et K_0 et K_1 s'interprètent directement comme les coûts supportés par la méconnaissance de l'état réel de la personne.

Du point de vue bayésien, le coût moyen *a posteriori* vaut :

$$\begin{aligned}\rho(d|x) &= \int L(\theta, d)[\theta|x]d\theta \\ &= \int_{\Theta_0} K_0(1-d)[\theta|x]d\theta + \int_{\Theta \notin \Theta_0} K_1d[\theta|x]d\theta \\ &= K_0(1-d)\mathbb{P}(\theta \in \Theta_0|x) + K_1d\mathbb{P}(\theta \notin \Theta_0|x).\end{aligned}$$

L'estimateur de Bayes associé à la loi *a priori* de θ et à la fonction de coût est :

$$\delta(x) = \begin{cases} 1, & \text{si } K_0\mathbb{P}(\theta \in \Theta_0|x) > K_1\mathbb{P}(\theta \notin \Theta_0|x) \\ 0, & \text{sinon} \end{cases}$$

Par conséquent, $\delta(x)$ vaut 1 si et seulement si $\mathbb{P}(\theta \in \Theta_0|x) > K_1/K_0$: la règle bayésienne se fonde sur le niveau de crédibilité *a posteriori* de l'état Θ_0 .

L'annexe C présente un exemple de fonction de coût linéaire et sa résolution dans un cas d'école (modèle Beta-binomial) où les calculs sont tous explicites.

Dans les cas d'application de cette thèse pour lesquels j'ai eu recours à la théorie bayésienne de la décision, les observations sont des résultats de détection ou de dénombrement de *L. monocytogenes* (Chapitre 2). Les résultats des plans d'échantillonnage (qui correspondent à des simulations de $[x|\theta]$), sont obtenus par simulation. Le paramètre θ représente la contamination et peut être un vecteur. Avant de poursuivre, il faut donc déterminer quel(s) modèle(s) de contamination utiliser, c'est-à-dire quelle(s) vraisemblance(s) utiliser pour relier les observations aux paramètres. La partie suivante présente les modèles de contamination rencontrés dans la littérature.

1.4 Modèles de contamination existants

Pour fixer un MC pertinent et évaluer la qualité d'un plan d'échantillonnage, il est indispensable de connaître la distribution des micro-organismes dans les aliments (ILSI (2004)). Une partie importante du travail de cette thèse a porté sur la modélisation de cette distribution. Je commence donc par les modèles les plus courants dans la littérature.

De nombreux modèles décrivent la présence d'un micro-organisme dans un aliment. D'ailleurs, ces modèles sont utilisés dans des études afin de prédire les conséquences de la consommation d'un aliment donné sur la santé publique (Pouillot et al. (2007) et Pouillot et al. (2009)) ou pour vérifier si des PO ou des FSO sont réalisables (Sosa Mejia et al. (2010)).

Les micro-organismes sont bien évidemment des entités de type discret. On s'attend donc à ce que les modèles de contamination utilisent comme entrée des variables discrètes. Cependant, les données sont généralement exprimées sous forme de concentration (ex. ufc^q/g ou ufc/cm²) et de nombreux modèles ont des variables continues pour entrée. Le tableau 1.6 récapitule les différents modèles présentés dans cette partie.

q. ufc signifie unité formant colonie. Le passage de la donnée brute (décompte de colonies) à l'expression sous forme de concentration est expliquée à la partie 2.1.2.

1.4.1 Prévalence

Avant de décrire les principaux modèles de contamination, j'introduis la *prévalence*, un terme très largement usité en microbiologie. La prévalence est une caractéristique théorique liée à un microorganisme et à un ensemble d'aliments, puisqu'elle est égale à la fraction comportant au numérateur le nombre d'unités contaminées par le microorganisme considéré et au dénominateur le nombre d'unités constituant l'ensemble. Par exemple, on peut s'intéresser à la prévalence de *L. monocytogenes* dans du saumon fumé en sortie usine chaque jour de production. Lorsque l'on parle de prévalence, il faut faire attention à bien définir la population considérée (par exemple : le saumon fumé produit dans une usine particulière pendant un jour ou le saumon fumé produit en France durant une année) et l'entité de base susceptible d'être contaminée (par exemple : un lot de production, une plaque de saumon fumé, une portion de 25 g).

La prévalence peut par exemple être estimée au moyen de données de détection^r. A mon avis, il faut être très attentif à la définition de la prévalence estimée grâce aux données. Par exemple, si on considère n analyses de détection dans des prises d'essai de 25 g d'un même lot menées avec x positives, alors le rapport à x/n estime la prévalence dans les prises de 25 g constituant ce lot, c'est-à-dire le nombre de portions de 25 g contaminées sur le nombre de portions dans un lot. Cette fraction ne fournit qu'une indication de la prévalence du produit (rapport du nombre de produits contaminés dans un lot ou une production nationale par exemple sur le nombre de produits dans un lot ou une production nationale). Il arrive pourtant parfois que la prévalence des prises d'essai soit directement étendu à la prévalence du produit (Dauphin and Ragimbeau (2001) ou Vogel et al. (2001) où la prévalence estimée par des analyses effectuées sur des prises d'essai est considérée comme égale à la prévalence dans le produit fini).

Le fait qu'une prise d'essai ne comporte pas le microorganisme recherché ne signifie pas que la concentration de ce pathogène soit nulle dans la population considérée. Par exemple, si une prise d'essai de 25 g ne contient pas le microorganisme d'intérêt, la concentration de ce microorganisme est considérée inférieure à $1/25 = 0,04$ ufc/g^s. Si on connaît la distribution de la contamination dans la population considérée, alors la prévalence est égale à $1 - F(m)$, avec F la fonction de répartition de la contamination et m la concentration limite telle qu'une prise d'essai dont la concentration est inférieure à m est considérée comme non contaminée. Dans l'exemple présenté ci-dessus, $m = 0,04$ ufc/g. La formule $1 - F(m)$ est déjà apparue dans la partie 1.2.3.b dans l'expression de la proportion de défectueux. Pour les plans d'échantillonnages à deux classes, la proportion de défectueux est une prévalence.

Différentes prévalences peuvent être emboîtées : par exemple, la prévalence des lots contaminés durant un mois de production et la prévalence des unités contaminées dans chaque lot.

Dans le cadre de cette thèse, j'ai systématiquement défini la prévalence considérée.

r. Le principe de la détection est précisé au chapitre 2. Pour l'instant, il suffit de savoir que le résultat d'une analyse de détection est de type binaire : présence ou absence du microorganisme recherché dans une prise d'essai de M g.

s. Cette conversion est expliquée plus en détail partie 2.1.2.

1.4.2 Modèle Log-Normal

Un des modèles les plus simples est le modèle Log-Normal qui fait l'hypothèse que la concentration d'un micro-organisme dans une population donnée (par exemple la production d'une usine pendant six mois) suit la loi log-normale (Busschaert et al. (2010); van Schothorst et al. (2009); Legan et al. (2001); Crépet et al. (2007)). Ce modèle présente divers intérêts (ILSI (2010)), parmi lesquels le fait d'être très largement répandu. Par ailleurs, les microbiologistes substituent volontiers la log-concentration à la concentration observée. Ainsi, les modèles de croissance et de destruction de bactéries représentent la log-concentration en fonction du temps. Cependant, le modèle Log-Normal présente plusieurs inconvénients parmi lesquels celui de ne pas être défini en 0, le modèle faisant l'hypothèse que le produit est forcément contaminé. Cependant, ainsi qu'il est précisé dans l'article présenté au chapitre 4, ce modèle peut tout à fait prendre en compte une prise d'essai ne contenant pas le micro-organisme recherché. Si la concentration d'une bactérie dans un produit est par exemple de 0,01 ufc/g, cela veut dire qu'il y a en moyenne 1 bactérie pour 100 g de produit. Si une prise d'essai pèse 10g, il est tout à fait possible qu'elle ne contienne pas la bactérie d'intérêt comme il est possible qu'elle la contienne.

1.4.3 Modèle Log-Normal et prévalence

Il est possible de concilier produits partiellement contaminés et distribution log-normale des bactéries. Ce modèle est proposé par Pouillot et al. (2007). Il part de l'hypothèse qu'une partie de la production d'un aliment ne comporte pas la bactérie recherchée et que la distribution de la concentration de la bactérie concernant l'autre partie suit une loi Log-Normale. Ce modèle a donc un paramètre de plus que le modèle précédent : la probabilité de non contamination.

1.4.4 Modèle de Poisson

En étudiant le décompte de micro-organismes, un autre modèle simple consiste à considérer que le nombre de bactéries y dans un produit suit une loi de Poisson : $y \sim \mathcal{P}(\lambda K)$. Le paramètre λ du modèle de Poisson est une concentration et K caractérise le protocole (cf. le chapitre 2 et plus particulièrement la partie 2.1.2 pour une décomposition possible de K). Pour l'instant, K sera considéré comme une constante. Cette distribution est appropriée lorsque l'aliment est liquide ou bien mixé (ILSI (2010); Delignette-Muller and Cornu (2008); Jongenburger et al. (2010)). Ce modèle n'est pas très adapté dans les aliments solides, où les bactéries sont la plupart du temps sur-dispersées (c'est-à-dire que l'écart-type est supérieur à la moyenne, ILSI (2010); Gonzales-Barron et al. (2010a)).

Des variantes du modèle Poisson peuvent être utilisées pour obtenir plus de flexibilité.

1.4.5 Modèle Poisson-Gamma ou Binomial Négatif

Pour le modèle Poisson-Gamma, la distribution suivie par λ est une loi Gamma. En intégrant sur la concentration λ la loi du nombre de bactéries y sachant λ , on obtient une loi binomiale négative. Pour cette raison, le modèle Poisson-Gamma est aussi appelé loi

Binomiale Négative. Ce modèle est présenté par exemple par Gonzales-Barron and Butler (2011b); Reinders et al. (2003) et ILSI (2010).

1.4.6 Modèle Poisson-Log-Normale

Une autre distribution possible pour λ est la loi log-normale. Ce modèle présente l'avantage d'autoriser 0 comme valeur et de pouvoir être approximé par la loi log-normale ou par la loi de Poisson selon les valeurs des paramètres (ILSI (2010)).

Il arrive parfois que le nombre de données de décompte nul soit plus élevé que ce qui est permis, même avec un modèle Poisson-Gamma ou Poisson-LogNormal. Gonzales-Barron et al. (2010a) par exemple proposent d'utiliser des mélanges de modèles de deux types : “zero-inflated” et “hurdle”.

1.4.7 Modèles de type “hurdle”

Les modèles de type “hurdle” (Mullahy (1986)) partent de l'hypothèse que les zéros et les données non nulles sont générées de manière différente. Une loi binomiale répartit les valeurs nulles des autres. Les valeurs non nulles suivent une distribution f tronquée en 0. Tous les modèles utilisant des variables discrètes peuvent être utilisés en “hurdle” : Poisson-Hurdle, Poisson-Gamma-Hurdle et Poisson-Log-Normale-Hurdle.

1.4.8 Modèles de type “zero-inflated”

Les modèles de type “zero-inflated” traitent également les zéros différemment des données non nulles mais de manière différente que les modèles de type “hurdle”. Ils reposent sur l'hypothèse qu'il existe deux sortes de zéros : ceux générés par la loi f et ceux en excès de zéros par rapport à la distribution f . Les données non nulles sont modélisées uniquement par la loi f . De même que précédemment, tous les modèles utilisant des variables discrètes peuvent être utilisés en “zero-inflated” : Poisson-Zero-Inflated, Poisson-Gamma-Zero-Inflated et Poisson-Log-Normale-Zero-Inflated. Les modèles de type “hurdle” permettent également de modéliser la sous-dispersion, ce que ne peuvent pas faire les modèles de type “zero-inflated”. Les modèles de type “zero-inflated” sont un cas particuliers des modèles de type “hurdle” avec $\omega_0 = p_0 + (1 - p_0)f(0)$ et ne pouvant pas traiter la sous-dispersion.

1.4.9 Mélange de lois log-normales

Jusqu'à présent, les modèles de mélange présentés étaient composés d'une distribution pour modéliser la contamination (et éventuellement l'absence de micro-organismes) et d'un poids supplémentaire en 0. Une autre possibilité est de considérer deux distributions distinctes, l'une décrivant la contamination faible et l'autre la contamination forte. C'est la solution proposée par Crépet et al. (2007) avec un mélange de deux lois log-normales pour modéliser la concentration de *L. monocytogenes* dans les légumes frais. Ce modèle nécessite de disposer de suffisamment de données, car il a de nombreux paramètres.

Nom	Discret/ Continu	Distributions	Paramètres
Log-Normal	C	$c \sim \mathcal{LN}(\mu, \sigma^2)$	μ, σ
Log-Normal et prévalence	C	$c \sim \mathcal{N}(\mu, \sigma^2)$ pour la proportion p des produits contaminés	p, μ, σ
Poisson	D	$y \sim \mathcal{P}(\lambda \times K)$	λ
Poisson-Gamma	D	$y \sim \mathcal{P}(\lambda \times K)$ $\lambda \sim \mathcal{G}(\alpha, \beta)$	α, β
Binomial Négative	D	$y \sim \mathcal{NB}(n = \alpha, p = \frac{\beta}{\beta+K})$	n, p
Poisson-Log-Normale	D	$y \sim \mathcal{P}(\lambda \times K)$ $\lambda \sim \mathcal{LN}(\mu, \sigma^2)$	μ, σ
hurdle	D	$P(y) = \begin{cases} \omega_0, & \text{si } y = 0 \\ (1 - \omega_0) \frac{f(y)}{1-f(0)}, & \text{sinon} \end{cases}$	ω_0 , paramètres de la loi f
zero-inflated	D	$P(y) = \begin{cases} p_0 + (1 - p_0)f(0), & \text{si } y = 0 \\ (1 - p_0)f(y), & \text{sinon} \end{cases}$	p_0 , paramètres de la loi f
Mélange de Log-Normales	C	$c \sim \beta \mathcal{LN}(\mu_l, \sigma_l^2) + (1 - \beta) \mathcal{LN}(\mu_h, \sigma_h^2)$	$\mu_l, \sigma_l, \mu_h, \sigma_h, \beta$

TABLE 1.6 – Différents modèles décrivant la répartition de micro-organismes dans un aliment.

La notation adoptée dans cette partie pour la loi Gamma est la suivante : m est la moyenne en cfu/g par exemple et $\rho = V/m$ avec V la variance. La fonction de densité de la fonction est $f(\lambda) = \frac{\rho^{m/\rho}}{\Gamma(m/\rho)} x^{m/\rho-1} e^{-\rho x}$.

1.4.10 Estimation des paramètres

Les différents modèles présentés sont utilisés pour évaluer les paramètres à l'aide de données. De nombreux travaux existent sur le sujet, par exemple Pouillot et al. (2007); Crépet et al. (2007); Gonzales-Barron et al. (2010a); Busschaert et al. (2010). Il existe plusieurs techniques pour estimer les paramètres des modèles. Parmi celles-ci, nous citerons l'estimation du maximum de vraisemblance qui consiste à déterminer les valeurs des paramètres pour lesquelles la vraisemblance du modèle est maximale. Cette technique relève des statistiques fréquentistes et a été utilisée par Pouillot et al. (2007); Reinders et al. (2003) par exemple. L'algorithme espérance-maximisation (EM) est un moyen de trouver le maximum de vraisemblance des paramètres lorsque la vraisemblance ne peut pas être écrite de manière explicite (cf. Chapitre 4). Pour fournir un intervalle de confiance la technique du bootstrap a été utilisée par Busschaert et al. (2010).

Une autre possibilité est d'avoir recours à l'approche bayésienne pour estimer les paramètres et leur intervalle de crédibilité. Cette technique a été utilisée par Gonzales-Barron and Butler (2011b) ou Busschaert et al. (2011) par exemple. L'annexe D présente les inférences fréquentistes et bayésiennes.

1.4.11 Modèle hiérarchique

Les différents modèles présentés ci-dessus peuvent être combinés les uns aux autres pour modéliser les regroupements pouvant exister dans les jeux de données. La structure hiérarchique existe dans la vie courante. Par exemple, les enfants sont regroupés dans des classes et dans des écoles, les chercheurs travaillent dans des équipes elle-mêmes regroupées en laboratoire appartenant généralement à un département contenu dans un organisme de recherche par exemple. D'autres hiérarchies dans les données sont possibles, comme les méta-analyses (les données sont regroupées par étude) ou les mesures répétées sur des individus (les mesures d'une certaine quantité sont regroupées par individu). On peut s'attendre à ce que les données ou les individus appartenant au même groupe se ressemblent plus que des individus appartenant à des groupes différents. Ainsi, deux élèves d'une même classe (CP par exemple) se ressemblent plus que deux élèves de deux classes différentes de CP pour plusieurs raisons : les élèves d'une même classe vivent généralement dans le même quartier, de plus la partie du programme étudiée depuis le début de l'année est la même puisque l'enseignant des deux élèves est le même. Dans ces conditions, traiter les données sans faire attention à la structure ou au contraire, n'étudier que les sous-groupes au lieu des individus (les classes au lieu des élèves par exemple) peut ne pas être adapté.

Comment traiter ce type de données ? Grâce aux modèles hiérarchiques. Supposons par exemple un modèle avec deux niveaux d'agrégation, par exemple les élèves i dans une classe j et les classes dans une école ou, pour donner un exemple dans le domaine agroalimentaire : des barquettes de lardons i appartenant à un lot j et les lots fabriqués dans une usine^t. Au niveau des individus (les élèves ou les barquettes de lardons), la variable d'intérêt y_{ij} est modélisée par une certaine distribution :

$$y_{ij}|\phi_j \sim f(\phi_j), \quad i = 1, \dots, N_j \quad j = 1, \dots, J \quad (1.2)$$

t. Le terme *lot* est discuté à la partie 2.2

avec f une distribution de probabilité, ϕ_j un vecteur de variables au niveau du groupe (la classe ou le lot), N_j le nombre d'individus dans le groupe j et J le nombre de groupes. Si on pense que les ϕ_j ne sont pas indépendants entre eux, on peut modéliser cette dépendance par une fonction de distribution g :

$$\phi_j \sim g(\theta) \quad j = 1, \dots, J \quad (1.3)$$

ϕ_j est un paramètre caractérisant la classe j . Si on considère n'importe quelle classe, alors ϕ_j est une variable aléatoire : les variations de ϕ_j caractérisent la variabilité inter-classes de même que les variations de y_{ij} caractérisent la variabilité intra-classe.

La *variabilité* résulte de l'hétérogénéité des composants du système à modéliser et est liée à la stochasticité des phénomènes étudiés et ne peut donc être réduite même si plus d'information sur le système est disponible. L'*incertitude* résulte du manque de connaissance du système : elle peut donc être réduite par l'acquisition de nouvelles connaissances. La variabilité est également appelée incertitude aléatoire ou incertitude par essence et l'incertitude est parfois appelée incertitude épistémique ou incertitude de connaissance.

Ainsi, plusieurs niveaux peuvent être considérés (contre deux seulement dans l'exemple ci-dessus), ϕ_j et θ peuvent être des vecteurs et les distributions f et g peuvent regrouper plusieurs distributions. Les variables ϕ_j sont non observées : elles sont appelées "*variables latentes*". Un modèle hiérarchique est composé *a minima* d'une observable et d'un paramètre. Il repose sur la décomposition d'un phénomène aléatoire complexe en plusieurs phénomènes aléatoires élémentaires. Ces derniers sont agencés par raisonnements conditionnels successifs comme celui présenté aux équations (3.1) et (1.3). Les paramètres, les variables latentes et les observables ont des relations de dépendance orientée dans le même sens et les paramètres sont toujours des grandeurs conditionnantes (Ancelet (2008)).

Les modèles hiérarchiques se prêtent bien à la représentation graphique sous forme de DAG (*directed acyclic graph* ou graphe orienté acyclique, Spiegelhalter et al. (1993)) qui est orienté sans cycle :

- les variables latentes et les paramètres constituent des nœuds représentés par des cercles ou des ellipses,
- les observables constituent des nœuds représentés par des rectangles,
- les dépendances stochastiques sont matérialisées par des flèches.

Les *enfants* d'un nœud sont constitués par tous les nœuds dépendants de ce nœud. A l'inverse, les *parents* d'un nœud sont tous les nœuds dont ce nœud dépend. Les nœuds sans parent sont appelés des paramètres, les nœuds sans enfants sont des feuilles et tous les autres nœuds stochastiques sont des variables latentes (Boreux et al. (2010)). Un DAG à deux niveaux est constitué d'un paramètre, d'une variable latente et d'une observable (cf. figure 1.8)^u. Le modèle le plus simple se réduit à un paramètre et une observable.

Si l'on considère trois variables, le DAG permet de représenter cinq relations possibles (Parent and Bernier (2007)) : l'indépendance des trois variables, la dépendance de deux variables et l'indépendance de la troisième, la chaîne (figure 1.9a), l'héritage (figure 1.9b) et la naissance (figure 1.9c). Le DAG permet de représenter la dépendance conditionnelle. Ainsi, pour l'héritage, une information supplémentaire sur Z ne modifie pas

u. Si la flèche va du paramètre vers l'observable, cela implique que le paramètre est la "cause" et l'observable l'"effet". Dans le cas de l'inférence bayésienne, on cherche à mettre à jour nos connaissances sur le paramètre à partir des observables, ce qui revient à inverser la flèche.

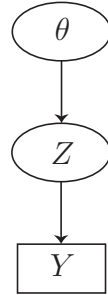


FIGURE 1.8 – DAG d'un modèle à deux niveaux, constitué d'une observable Y , d'une variable latente Z et d'un paramètre θ .

notre connaissance sur Y si on a de l'information sur X . Les structures possibles d'un graphe augmentent avec le nombre de variables les relations "unitaires" définies avec trois variables restent les mêmes. Pour plus d'information sur les modèles graphiques, consulter Boreux et al. (2010).

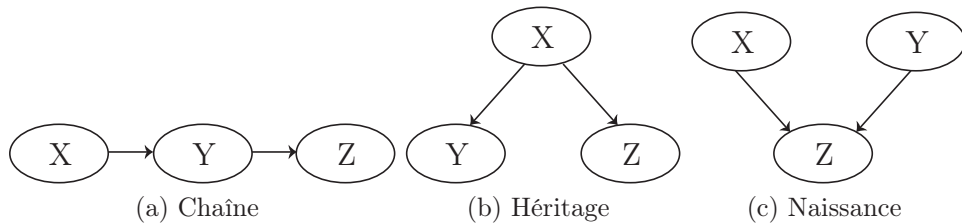


FIGURE 1.9 – Trois relations possibles entre trois nœuds dans un DAG.

Comme écrit plus haut, les DAG permettent de représenter commodément les modèles hiérarchiques. Prenons l'exemple d'un modèle à deux niveaux : on considère une entreprise fabriquant du saumon fumé. Le log de la concentration c_{ij} en *L. monocytogenes* de la portion de 10 g i appartenant au lot j est modélisé de la façon suivante :

1. Niveau de la portion :

$$[c_{ij}] = \mathcal{LN}(\mu_i, \sigma^2)$$

2. Niveau du lot :

$$[\mu_i] = \mathcal{N}(m, s^2).$$

Le DAG de ce modèle est représenté à la figure 1.10. La variance σ^2 représente la variabilité intra-lot et la variance s^2 la variabilité inter-lots. Les observables sont constituées des c_{ij} , avec $i = 1, \dots, N_j$ et $j = 1, \dots, J$, le vecteur θ des paramètres vaut $\theta = (\sigma, m, s)$ et les variables latentes sont les μ_1, \dots, μ_J .

Les hiérarchies successives apparaissent naturellement dans les DAG car les empilements des sous-modèles sont clairement dessinés. Chaque nœud est entièrement défini par ses parents et ses enfants : conditionnellement à ses parents et à ses enfants, un nœud est indépendant de tous les autres nœuds.

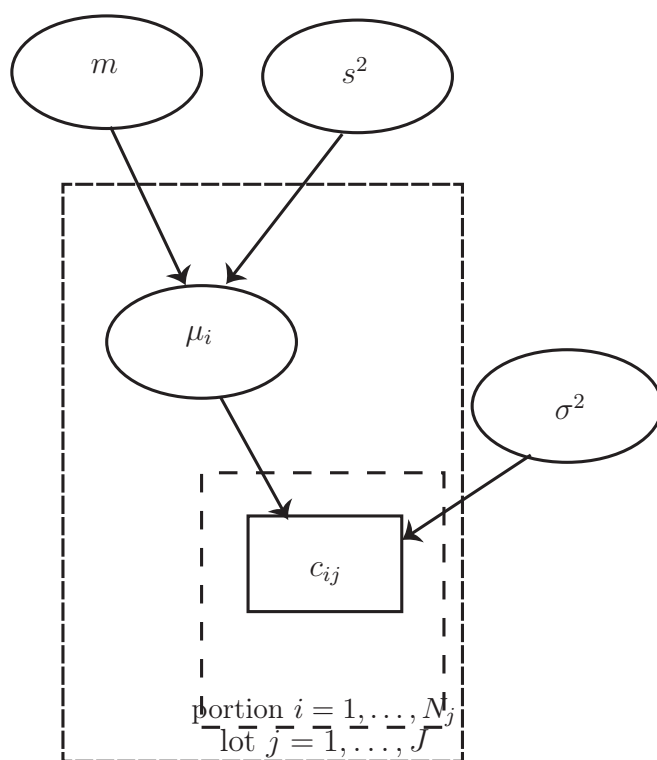


FIGURE 1.10 – DAG d'un modèle à deux niveaux : la portion i et le lot j . Au niveau de la portion, c_{ij} est la concentration. Au niveau du lot, μ_i est la moyenne de la log-concentration.

Dans le domaine de la microbiologie des aliments, les modèles hiérarchiques ont été utilisés en méta-analyse, par exemple chez Crépet et al. (2007) (deux niveaux : (1) mesure appartenant à (2) une étude) ou chez Jaloustre et al. (2011) (modèle à trois niveaux : (1) mesure appartenant à (2) une courbe de croissance appartenant à (3) une étude). Les modèles hiérarchiques ont également été utilisés pour estimer les paramètres de croissance de *L. monocytogenes* et de la flore lactique dans du saumon fumé (Delignette-Muller et al. (2006) avec trois niveaux : (1) une mesure appartenant à (2) une courbe de croissance appartenant à (3) une souche et à un produit).

Se placer dans l'approche bayésienne revient à introduire un niveau supplémentaire, puisque le paramètre θ devient une variable. Cette approche supprime la différence mathématique entre paramètres et variables latentes : seule la place dans le DAG les distingue, puisque les paramètres n'ont pas de parents. Pour le modèle représenté figure 1.10, on peut proposer comme loi *a priori* :

$$\begin{aligned}m &\sim \mathcal{N}(\text{moy}, \text{et}^2) \\ \sigma &\sim \mathcal{U}(a_\sigma, b_\sigma) \\ s &\sim \mathcal{U}(a_s, b_s)\end{aligned}$$

On peut construire des modèles hiérarchiques en utilisant les modèles présentés de la partie 1.4.2 à 1.4.9. Par exemple, on peut combiner le modèle Poisson au niveau des portions et le modèle Log-Normal au niveau du lot. Le résultat est différent du modèle Poisson-Log-Normale présenté partie 1.4.6, intégralement situé à un niveau. Divers modèles hiérarchiques sont présentés au chapitre 3.

1.5 Démarche adoptée

1.5.1 Plan général

Durant la première partie de cette thèse, des données microbiologiques ont été

- acquises grâce au projet Quant'HACCP pour l'application "lardons";
- reprises de la littérature par contact avec les auteurs pour l'application "saumon fumé".

Le mode d'acquisition ainsi que les jeux de données eux-mêmes sont présentés au chapitre 2. Le terme *lot* a été employé dans cette introduction sans véritable définition pour l'instant. Il fait l'objet de la partie 2.2 du chapitre 2 car la notion n'est pas simple à appréhender alors que sa définition précise est indispensable à ma thèse.

Les données des applications "lardons" et "saumon fumé" ont été utilisées pour estimer les lois *a posteriori* de modèles de contamination prenant ou non en compte plusieurs variabilités possibles. Ces modèles ont été comparés pour déterminer le ou les plus adapté(s) aux données (chapitre 3). Les données utilisées peuvent être exprimées sous forme brute (résultat direct des analyses) ou sous forme de concentration. Cette conversion est présentée à la partie 2.1.2. Nous nous sommes intéressés à l'incidence du type de données utilisées sur l'estimation des paramètres de deux modèles de contamination, dont l'un utilise les données brutes observées et l'autre les concentrations. L'estimation a été menée selon l'approche fréquentiste et fait l'objet du chapitre 4. Outre les modèles

de contamination qui viennent d'être introduits, j'ai participé à l'élaboration d'un modèle décrivant la contamination par *L. monocytogenes* dans la fabrication de lardons. J'ai ensuite modifié ce modèle pour prendre en compte une partie des données acquises dans le projet Quant'HACCP, comme il est expliqué chapitre 5.

Dans la seconde partie de ma thèse, j'ai travaillé sur les plans d'échantillonnage. Plus précisément j'ai appliqué la théorie bayésienne de la décision afin de déterminer la taille optimale du plan pour deux applications :

- contamination par *L. monocytogenes* dans des lardons en sortie usine avec modélisation d'une relation entreprise-distributeur (chapitre 6) ;
- contamination par *L. monocytogenes* dans du saumon fumé en sortie usine. La modélisation est dynamique dans la mesure où la décision prise par l'entreprise une semaine donnée dépend des résultats du plan d'échantillonnage conduit la semaine précédente (chapitre 7).

La troisième application concerne le suivi de la température d'un plat en sauce fabriqué dans la cuisine centrale d'un hôpital. Une fois préparé, le plat est placé dans un chariot lui-même refroidi rapidement (la température doit être inférieure à 10°C au bout de 2 h maximum, Arrêté du 21 décembre 2009 (2009)). Les plans d'échantillonnages étudiés sont thermiques (en combien de points mesurer la température dans un chariot ?). Ils sont comparés par rapport au risque que la croissance de *C. perfringens* dépasse 1 log ufc/g entre la cuisson du plat et sa consommation. Cette application est présentée au chapitre 8.

Enfin, ce document s'achève par une conclusion (chapitre 9).

1.5.2 Cas d'application

Trois cas d'application ont été étudiés lors de cette thèse. Les deux premiers concernent la contamination par *L. monocytogenes* de deux produits : le saumon fumé en sortie d'usine et les lardons en cours de procédé et en sortie d'usine. Le saumon fumé est un aliment connu pour être fréquemment contaminé par ce pathogène, ce qui est aggravé par le fait que cet aliment est fréquemment consommé sans cuisson préalable de la part du consommateur. De plus, *L. monocytogenes* croît bien dans ce produit, aussi bien en contaminations artificielle (Szabo and Cahill, 2009; Hudson and Mott, 1993) qu'en contamination naturelle (Jørgensen and Huss, 1998; Dalgaard and Jørgensen, 1998). Plusieurs études ont été conduites pour étudier la prévalence et la concentration dans ce couple pathogène/aliment. Citons Jørgensen and Huss (1998); Dauphin and Ragimbeau (2001); Beaufort et al. (2007) et Uyttendaele et al. (2009).

La prévalence observée est plus faible dans les lardons que dans le saumon fumé. De plus, ce produit est consommé cuit (bien que, d'après un sondage (ANSES, 2009), 14% de la population française consomme occasionnellement des lardons crus). Néanmoins, la croissance de *L. monocytogenes* semble faible d'après Cornu et al. (2011).

La troisième cas d'application concerne la contamination par *C. perfringens* de plats de bœuf en sauce dans le cadre de la restauration collective dans un hôpital car les TIAC^v liées à ce pathogène ont souvent lieu sur des plats à base de viande en restauration

v. TIAC signifie toxi-infection alimentaire collective

collective. Ainsi, selon Crouch and Golden (2005), aux Etats-Unis, dont 46,5% des cas déclarés entre 1990 et 1999 le sont dans des restaurations collectives et 28,8% dans des plats comportant du bœuf.

1.5.3 *L. monocytogenes*

Caractéristiques *L. monocytogenes* est une bactérie non sporulée, mobile, aérobie-anaérobie facultatif, de 0,4 à 0,5 μm de diamètre sur 0,5 à 2,5 μm de longueur. C'est l'une des six espèces du genre *Listeria*. Ce pathogène est ubiquiste, tellurique, largement présent dans l'environnement (lacs, rivières, sol, végétation...). Il peut se retrouver dans presque tous les aliments. La contamination peut être due aux matières premières ou à l'environnement de l'usine : *L. monocytogenes* peut s'établir et persister pendant des mois, voire des années, à un ou plusieurs endroits de l'usine notamment grâce à sa capacité à adhérer à des surfaces et à y former des biofilms, mais également pour sa prédilection pour les endroits humides et frais (ILSI (2005)). Il croît si le pH est compris entre 4,39 et 9,4, une activité de l'eau supérieure à 0,92 (ICMSF (1986)) et une température comprise entre -2°C et 45°C , la température optimale de croissance se situant entre 30°C et 37°C (Jemmi and Stephan (2006)). L'aspect des colonies de *L. monocytogenes* sur un gel dit ALOA est visible (figure 1.11).

Listériose L'ingestion de *L. monocytogenes* peut provoquer une maladie appelée *listériose*. On peut classer cette maladie en deux classes : non-invasive et invasive. Les symp-



FIGURE 1.11 – Photographie d'une cellule de *L. monocytogenes*.

tômes de la première forme évoquent une gastro-entérite. La forme invasive recouvre la septicémie, la méningite et l'encéphalite. Chez la femme enceinte, l'infection se manifeste sous la forme d'un état grippal ; elle peut occasionner un avortement ou une infection néo-natale. La durée moyenne d'incubation est d'un mois (FDA (2003)). La maladie est létale dans 35% des cas pour les formes materno-néonatale et de 20% pour les autres formes. Les personnes à risques sont les personnes âgées, les femmes enceintes et les personnes immunodéprimées (Pouillot et al. (2009)). La contamination par *L. monocytogenes*

peut se produire à tous les stades de la production alimentaire (des matières premières aux manipulations du consommateur). La listériose est une infection reconnue depuis plusieurs dizaines d'années, mais c'est seulement dans les années 1980 qu'un lien entre consommation d'aliments et listériose a été établi. La listériose n'est pas une maladie très répandue dans la population, surtout comparée à d'autres infections d'origine alimentaire. Cependant, comme elle est fortement létale, elle est très surveillée. Le *Center for Disease Control and Prevention* (CDC) considère que *L. monocytogenes* est le troisième pathogène causant le plus d'issues fatales derrière Norovirus et les salmonelles non typhiques (CDC (2011)).

Prévalence de la maladie En France, la listériose fait l'objet d'une déclaration obligatoire depuis 1999. Le nombre de cas déclaré était de 269 en 1999, il a baissé jusqu'en 2001 (188 cas) puis a augmenté avec un pic à 328 cas en 2009. En 2010, 312 cas ont été déclarés. Exprimées en incidence pour un million d'habitants, ces données deviennent 4,5 en 1999, 3,1 en 2001, 5,1 en 2009 et 4,9 en 2010 (InVS (2010)).

Aux Etats-Unis, en 2009, l'incidence était de 0,34 cas pour 100 000 habitants avec une pointe à 2,51 cas pour 100 000 habitants pour les personnes de 80 ans et plus (Food-Net (2009)). Le taux de mortalité était de 13,38% et le taux d'hospitalisation de 94,3%, soit les taux les plus élevés comparés aux autres pathogènes surveillés (*Campylobacter*, salmonelles, *Shigella*, STEC O157, STEC non O157, *Vibrio* et *Yersinia*).

L'impact de la contamination de *L. monocytogenes* sur l'industrie agro-alimentaire est principalement dû aux rappels de lots (Jemmi and Stephan (2006)). A titre d'exemple, au Canada, on a recensé 50 rappels en 2010, dont 35 de classe I (concernant des produits dont la consommation pourrait causer des problèmes de santé graves, voire entraîner la mort); ils ont porté notamment sur du fromage, des sandwiches et de la charcuterie (Food Inspection Agency (2010)). Aux coûts directs du rappel, il faut ajouter les coûts de destruction, des contrôles et des inspections supplémentaires, de retards dans la distribution (Jemmi and Stephan (2006)) et de perte d'image de marque. En plus des coûts directs dus au rappel lui-même, il faut ajouter les coûts de destruction, des contrôles et des inspections supplémentaires, de retards dans la distribution (Jemmi and Stephan (2006)) et de perte d'image de marque.

Contrôles microbiologiques L'ILSI (ILSI (2005)) préconise de mener des contrôles microbiologiques de *L. monocytogenes* dans les usines notamment pour détecter la présence de ce pathogène dans l'environnement de l'usine, comme par exemple le sol ou les tuyaux. Ainsi, l'ILSI recommande d'élaborer des plans d'échantillonnage dans trois situations :

1. normale (surveillance quotidienne ou hebdomadaire) :
2. accrue dans les endroits de l'usine où la présence de *L. monocytogenes* est régulièrement constatée (par exemple dans une trancheuse) ;
3. intense dans les endroits de l'usine où la présence de *L. monocytogenes* est régulièrement constatée, sa source restant inconnue.

Pour identifier les zones à risque, l'ILSI préconise de commencer par effectuer des prélèvements afin de déterminer les lieux de présence de *L. monocytogenes* s'il n'y a aucune

connaissance préalable de la répartition de ce micro-organisme dans l'usine.

L'ILSI ne conseille pas d'échantillonnage systématique en sortie usine, car ce procédé détecte mal la contamination dans des lots dont la prévalence est faible. Il peut être mené pour des produits dont le risque de contamination est élevé, mais plutôt durant la fabrication.

1.5.4 *C. perfringens*

C. perfringens est une bactérie à Gram positif, sporulée^w, à métabolisme anaérobie strict mais aérotolérante (ANSES (2010)). La figure 1.12 présente la forme végétative de *C. perfringens*. Elle est présente dans le sol, mais aussi dans les épices et dans la viande.

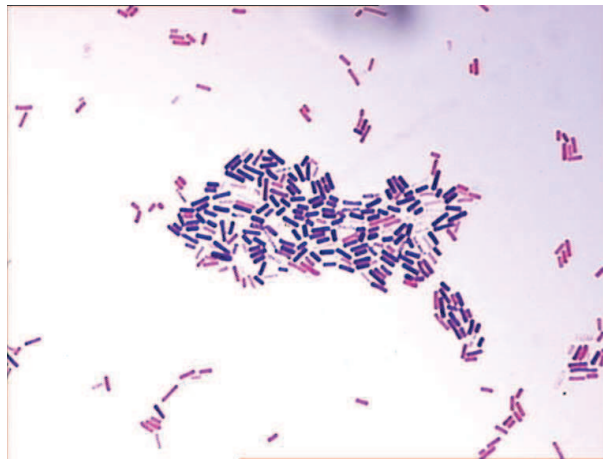


FIGURE 1.12 – Cellules végétatives de *C. perfringens*. Coloration de Gram, visualisation au microscope optique avec grossissement au 1000^{ème}.

Bien que la cuisson détruit la plupart des formes végétatives, elle n'élimine guère les spores. Par ailleurs, l'ébullition permet un dégazage de la préparation culinaire favorisant l'apparition de conditions d'anaérobiose propices à la croissance de *C. perfringens*. Les préparations en restauration collective sont constituées de grands volumes et sont à cet égard particulièrement dangereuses, car la ré-oxygénation par l'air ambiant est plus limitée que pour de petits volumes. *C. perfringens* croît entre 10°C et 53°C. Entre 40°C et 47 °C, avec un pH compris entre 6 et 7 et une $a_w \geq 0.97$, le temps de génération est de 7 minutes (Jaloustre (2011)).

Les symptômes des intoxications provoquées par *C. perfringens* sont généralement une diarrhée avec maux de ventre et nausée, mais sans vomissements ni fièvre, avec une évolution spontanément favorable après 24h. Brève, l'incubation dure 6h à 24h. Ces symptômes sont imputables aux toxines A produites par la forme végétative ; toutes les souches ne produisent pas cette toxine.

w. Ce pathogène existe donc sous forme de cellule (forme végétative) ou sous forme de spore.

En 1999, 19 parmi les 640 TIAC déclarées en France sont dues à *C. perfringens*, ce qui représente 386 malades, 3 hospitalisation et pas de décès (InVS (2004)). Les cas sporadiques dus à *C. perfringens* ne sont pas soumis à déclaration obligatoire, contrairement aux TIAC. En tenant compte de la sous déclaration des TIAC et de la non déclaration des cas sporadiques, l’InVS (2004) estime à 8900 environ le nombre de cas épidémiques et sporadiques dus à *C. perfringens*, correspondant à 110 hospitalisations et 6 décès dans les années 1990 en France.

1.5.5 Projet Quant’HACCP

Une grande partie de cette thèse, notamment les applications lardons/*L. monocytogenes* et plat en sauce/*C. perfringens*, a été réalisée dans le cadre du projet Quant’HACCP financé par l’ANR. Ce projet, piloté par l’AFSSA comporte huit partenaires principaux. Il s’est déroulé sur quatre ans, de 2008 à 2011. Son objectif était de quantifier l’impact et si possible d’optimiser les mesures de maîtrises mises en place par les industriels. J’ai été plus particulièrement chargée de la tâche “optimisation des stratégies d’échantillonnage et de surveillance”.

Pour l’application lardons/*L. monocytogenes*, les tâches principales ont été^x :

- prélèvement et analyse de poitrines de porc et de lardons (cf. chapitre 2) ;
- étude de la croissance de *L. monocytogenes* dans les lardons (Cornu et al. (2011)) ;
- modélisation de la contamination par *L. monocytogenes* dans les lardons (Billoir et al. (2011) et chapitres 3 et 5) ;
- recueil de la température de l’air et dans les poitrines durant l’étape d’étuvage ;
- optimisation du plan d’échantillonnage de surveillance (cf. chapitre 6).

Les différentes étapes du cas plat en sauce/*C. perfringens* ont été :

- collectes de données de température du plat depuis sa fabrication jusqu’à la consommation et analyses microbiologique des matières premières ;
- modélisation de la croissance (Jaloustre et al. (2011)) et de la destruction (Jaloustre et al. (2012)) de *C. perfringens* durant le procédé, établissement d’un modèle dose-réponse ;
- modélisation thermique du procédé ;
- étude de la stratégie de sortie de cellule de refroidissement pour les plats cuisinés en fonction de l’échantillonnage de la température (Poumeyrol et al. (2012) et chapitre 8).

x. Les articles cités entre parenthèses ont été écrit par les membres du projet

Première partie

DONNÉES ET MODÈLES DE
CONTAMINATION

La première partie de cette thèse présente les données qui ont été recueillies dans le cadre du projet Quant'HACCP ou d'un projet antérieur coordonné également par l'ANSES et nommé LmPR. Ces données ont été utilisées pour estimer les lois *a posteriori* de paramètres de modèles de contamination par *L. monocytogenes* de trois produits : des poitrines de porc après malaxage, des lardons et du saumon fumé en fin de procédé de fabrication. Selon les modèles considérés, les variabilités inter-lots, inter-unités et inter-prises d'essai sont prises en compte. Le tableau 1.7 détaille les thématiques abordées dans les chapitres constituant cette partie.

Dans les chapitres 3 et 4 sont reproduits deux articles écrits dans le cadre de cette thèse.

	Chapitre 2	Chapitre 3	Chapitre 4	Chapitre 5
Produit considéré	P, L, S	P, S	-	P
Variabilité et modèles	-	8 modèles V. inter-lots : loi \mathcal{LN} V. inter-unités : loi \mathcal{LN} V. inter-prises d'essai : loi \mathcal{LN} ou loi \mathcal{Bin}	V. inter et intra-lots confondues 2 modèles : Log-Normal ou Poisson-Log-Normal	Modélisation du pro- cédé V. inter-lots : loi \mathcal{LN}
Type de données	Brutes	Brutes/sous forme de concentration	Brutes	Brutes
Inférence	-	Bayésienne	Bayésienne/ Fréquentiste	Bayésienne

TABLE 1.7 – Thématiques de la première partie et chapitres dans lesquels elles sont abordées.

P. signifie “poitrines de porc après malaxage”, L. signifie “lardons en sortie usine”, S. signifie “saumon fumé en sortie usine”, - signifie “sans objet”, V. signifie “variabilité”, \mathcal{LN} signifie “loi log-normale” et \mathcal{Bin} signifie “loi binomiale”.

Chapitre 2

Détection et dénombrement pour caractériser la contamination en un point d'un procédé agro-alimentaire

2.1 Types des données utilisées

Ce chapitre est principalement consacré aux deux types de mesures utilisés dans le cadre de cette thèse : la détection et le dénombrement de *L. monocytogenes*. La détection est une analyse consistant à déterminer si *L. monocytogenes* est présente ou non dans une partie d'un produit, les données se présentent sous forme binaire. Le dénombrement permet de déterminer le nombre d'ufc^a de *L. monocytogenes* dans une quantité définie de produit : les données apparaissent sous forme d'entiers.

2.1.1 Principe de la détection et du dénombrement

2.1.1.a Détection

- Il existe plusieurs protocoles de détection. Brièvement, le principe est le suivant :
- la prise d'essai est placée dans un sachet contenant un milieu d'enrichissement sélectif ;
 - la solution est homogénéisée, par exemple au moyen d'un stomacher (appareil muni de deux pâles venant aplatir alternativement le sachet) ;
 - la solution est incubée pendant 24 h. L'incubation dans un milieu nutritif permet à *L. monocytogenes*, si elle est présente dans la solution, de se développer. Les agents inhibiteurs empêchent les bactéries autres que *L. monocytogenes* de croître. Cette étape est appelée enrichissement primaire ;
 - éventuellement, repiquage d'une partie de la suspension dans un milieu d'enrichissement sélectif pour l'étape d'enrichissement secondaire ;
 - un petit volume est déposé sur une ou plusieurs boîtes de Petri contenant un milieu nutritif sous forme de gélose. La boîte est incubée ;

a. ufc signifie "unité formant colonie"

- les colonies présentant un aspect typique de *L. monocytogenes* sont confirmées au moyen d'essais biochimiques, physiologiques ou morphologiques adaptés.

Mon interprétation de ce protocole est la suivante : si une *L. monocytogenes* vivante est présente initialement dans la prise d'essai, elle se retrouve dans les solutions successives, se multiplie dans les sachets car les milieux d'enrichissement utilisés favorisent la croissance de cette bactérie. Elle est donc présente en très grande quantité au bout de 24 h dans la solution à partir de laquelle un petit volume est déposé sur les boîtes de Petri. Ainsi, la probabilité de déposer des *L. monocytogenes* est quasiment de 1. Les cellules se divisent à nouveau sur les boîtes. Comme les géloses sont solides, les bactéries ne se déplacent pas et des colonies apparaissent.

D'après ce protocole, on en déduit que si des *L. monocytogenes* sont présentes dans la prise d'essai, même en faible quantité, toutes les conditions ont été réunies pour que les concentrations après un ou deux enrichissements soient élevées donc des colonies sont observées. Cependant, si des cellules meurent pendant le protocole, elles n'engendreront pas de colonies. Ainsi, la sensibilité n'est pas égale à 1. À l'inverse, si aucune *L. monocytogenes* n'est présente dans la prise d'essai, elle ne se développera pas au cours des enrichissements et ne sera pas visible sur les boîtes.

Pour simplifier, dans la suite de ces travaux, la spécificité et la sensibilité^b seront considérées égales à 1.

2.1.1.b Dénombrement

Plusieurs protocoles de dénombrement existent, mais le principe général est le suivant :

- la suspension-mère est constituée de la prise d'essai placée dans un diluant et homogénéisée ;
- selon la contamination présumée de la prise d'essai, un volume est déposé directement ou après une ou plusieurs dilutions sur une boîte de Petri contenant un milieu nutritif sous forme de gélose ;
- après incubation, les colonies présumées être des *L. monocytogenes* sont comptées (cf. figure 2.1) puis certaines d'entre elles sont confirmées ;
- le nombre de colonies de *L. monocytogenes* est évalué à partir du nombre total de colonies présumées et de colonies confirmées : si A colonies présumées sont comptées, a sont prélevées pour confirmation et b sont effectivement confirmées, le nombre de colonies de *L. monocytogenes* présentes sur la boîte est estimé à $A \times b/a$.

Les protocoles ne le précisent pas explicitement mais contrairement à la détection, le résultat est fourni au niveau de la fraction étalée sur les boîtes, non au niveau de la prise d'essai elle-même. En effet, il n'y a aucune raison pour que le nombre de bactéries soit le même dans tous les volumes unitaires de la suspension. Si on étalait les 100 mL de la suspension mère sur 100 boîtes de Petri par exemple, le nombre de colonies ne serait pas le même sur toutes les boîtes et par conséquent, les résultats du dénombrement seraient différents bien qu'il s'agisse d'une seule et même prise d'essai.

b. Voir la partie 1.2.5.c pour une définition de ces termes.

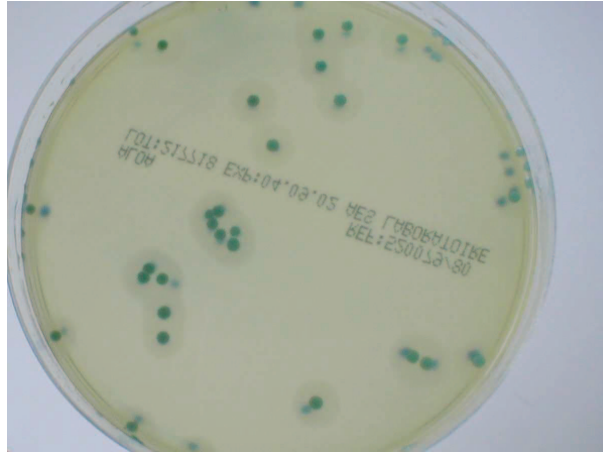


FIGURE 2.1 – Colonies de *L. monocytogenes* sur boîte de Petri sur une gélose dite ALOA (pour “agar Listeria according to Ottaviani and Agosti”).

Comme pour la détection, le nombre de colonies décompté sur les boîtes peut sous estimer le nombre de cellules présentes dans le volume étalé car celui-ci peut contenir des cellules mortes durant le protocole ou des cellules viables non cultivables^c. Par ailleurs, si deux cellules sont déposées quasiment au même endroit sur la boîte de Petri, les deux colonies en résultant seront confondues. Le nombre de colonies décomptés peut également surestimer ou sous-estimer le nombre de cellules dans le volume étalé si l’opérateur se trompe en comptant.

2.1.2 Expression des résultats bruts en concentration et limites de ces méthodes

Les résultats d’une détection (présence ou absence après enrichissement) et d’un dénombrement (nombre de colonies) constituent des *données brutes*. Dans la littérature scientifique et technique, elles ne sont pas souvent communiquées sous cette forme, mais généralement converties en concentration, surtout pour les données de dénombrement. Des données sous forme de dénombrement sont présentées par exemple chez Beaufort et al. (2007), Habib et al. (2008) ou Jørgensen and Huss (1998).

J’explique la conversion de la manière suivante. Supposons que

- la masse de la prise d’essai soit de M' g,
- le volume de la i^{e} dilution soit de V_{di} (le volume de la suspension mère est de V_{d1}),
- le volume prélevé lors de la i^{e} dilution et versé sur une boîte de Petri ou dans un tube pour la $i + 1^{\text{e}}$ dilution soit de V_{ri} .
- il y a I dilutions au total.

c. Les cellules viables non cultivables sont des cellules vivantes qui n’arrivent pas à se développer et à former de colonie sur des media traditionnel de culture. Cependant, elles peuvent reprendre leur croissance sous certaines conditions (Oliver (2010)).

Si y colonies sont décomptées sur les boîtes de Petri, on considère que ces y colonies proviennent des y cellules contenues dans le volume V_{rI} . Le volume V_{dI} contenait $y \times V_{dI}/V_{rI}$ cellules, elles-même provenant toutes du volume V_{rI-1} et ainsi de suite. En poursuivant ce raisonnement à chaque dilution, la suspension mère contenait donc $y \prod_{i=1}^I \frac{V_{di}}{V_{ri}}$ cellules, qui étaient toutes dans les M' g de la prise d'essai. Ainsi, la concentration en bactéries vaut $c = \frac{y}{M' \prod_{i=1}^I \frac{V_{ri}}{V_{di}}}$ ufc/g. Si aucune colonie n'est comptée sur les boîtes, cela veut dire que la concentration est inférieure à celle où une colonie est observée. La concentration c est alors inférieure à $\frac{1}{M' \prod_{i=1}^I \frac{V_{ri}}{V_{di}}}$ ufc/g. Cette valeur constitue la *limite de quantification* ou LOQ. La LOQ est donc la plus petite limite en-dessous de laquelle aucun résultat quantitatif ne peut être rendu. Si aucune colonie n'est observée, c est une donnée censurée à gauche. La formule de conversion revient à considérer qu'il y a exactement le même nombre de cellules dans tous les petits volumes V_{ri} contenus dans le volume V_{di} , même si les microbiologistes savent bien que ce n'est pas le cas.

Il est également possible de convertir un résultat de détection en concentration. Soit M la masse de la prise d'essai. Le résultat est une présence s'il y a au moins une *L. monocytogenes* dans la prise d'essai, donc la concentration minimale détectée est de $1/M$ ufc/g. Cette quantité est appelée *limite de détection* ou LOD. Si le résultat est une absence alors la concentration est inférieure à la LOD (donnée censurée à gauche) et si c'est une présence, la concentration est supérieure à la LOD (donnée censurée à droite).

Prenons un exemple avec les valeurs précisées dans les normes NF EN ISO ISO11290-1 (1996a) et NF EN ISO ISO11290-2 (1998a). En détection, la masse M vaut 25 g, donc la LOD est de 0,04 ufc/g. Pour le dénombrement, la masse M' est de 10 g, le volume de la suspension-mère vaut $V_{d1} = 100$ mL. Si $V_{r1} = 0,1$ mL est prélevé de la suspension-mère et directement étalé sur une boîte de Petri, alors la LOQ est de $\frac{1}{10 \times \frac{0,1}{100}} = 100$ ufc/g. La norme autorise également $V_{r1} = 1$ mL. Dans ce cas, la LOQ vaut 10 ufc/g.

Il est à noter que la norme NF EN ISO ISO11290-2 (1998a) préconise d'utiliser les résultats provenant de deux dilutions consécutives. Dans cette norme, 2 boîtes de Petri sont utilisées par dilution et cette dernière se fait toujours au dixième. Le nombre de colonies par mL de suspension-mère est donc estimé à $\frac{\sum y}{V(n_1+0,1n_2)d}$, avec $\sum y$ la somme des colonies confirmées de *L. monocytogenes* sur les 4 boîtes, n_1 le nombre de boîtes à la dilution la plus élevée ($n_1=2$), n_2 le nombre de boîtes à la dilution la plus faible ($n_2 = 2$), V le volume de l'inoculum appliqué à chaque boîte en mL et d le taux de dilution correspondant à la première dilution retenue.

Je propose la généralisation suivante de la formule : comme le nombre de boîtes de la dilution la plus élevée compte pour V_{rI}/V_{dI} par rapport au nombre de boîtes de la dilution la moins élevée, le nombre moyen de colonies décomptés est $\frac{\sum y}{(n_1+V_{rI}/V_{dI})n_2}$. A cela, il faut ajouter les dilutions précédents la dilution la moins élevée retenue, le nombre de colonies estimé dans la suspension-mère est donc

$$\frac{\sum y}{(n_1 + V_{rI}/V_{dI}n_2) \prod_{i=1}^{I-1} \frac{V_{ri}}{V_{di}}}$$

Pour exprimer le résultat en nombre de colonies par millilitre, il suffit de diviser la formule précédente par V_{d1} , le volume de la suspension-mère.

Les données exprimées sous forme de concentration peuvent être accompagnées d'un intervalle de confiance à 95%. Le calcul suppose préalablement des essais inter et intra laboratoires afin d'estimer les écart-type de répétabilité^d et de reproductibilité^e par type de matrice alimentaire. La norme NF EN ISO ISO19036 (2006) indique la procédure et le calcul à mener. Augustin and Carlier (2006) ont calculé des variances de répétabilité et de reproductibilité pour plusieurs pathogènes dans de la poudre.

Une des possibilités d'obtenir des résultats plus précis est d'abaisser la LOQ, par exemple en augmentant le volume déposé sur les boîtes (et en augmentant le nombre de boîtes) ou en abaissant la dilution. Ces deux techniques ont été utilisées conjointement pour dénombrer des *L. monocytogenes* dans des lardons et des poitrines de porc (cf. partie 2.3.2). Une manière d'augmenter le volume de suspension dénombré est de recourir à la méthode de filtration. Cette technique a été utilisée pour le dénombrement de *L. monocytogenes* dans le saumon fumé (cf. partie 2.4.1).

2.1.3 Particularités des données de détection et de dénombrement

2.1.3.a Les données sont imprécises

La conversion des données aboutit à une perte d'information si le protocole n'est pas entièrement fourni avec les données transformées. En effet, la masse est souvent donnée mais les volumes des différentes suspensions ou le nombre de dilutions ne le sont pas toujours. Il est alors impossible de remonter à la donnée brute. Or, la donnée transformée est imprécise, notamment parce qu'elle extrapole pour le dénombrement un résultat au niveau de la prise d'essai alors que la donnée brute fournit un résultat uniquement au niveau de la fraction. Cette imprécision est plus importante quand le nombre de colonies décomptées est faible. Reprenons les valeurs numériques de la norme ISO11290-2 (1998a) et considérons que 80 colonies ont été comptées après une dilution. La concentration est alors de $\frac{80}{10 \times \frac{1}{100}} = 800$ ufc/g. Si 79 ou 81 colonies avaient été décomptées, la concentration aurait été respectivement évaluée à 790 ou à 810 ufc/g, ce qui revient à une variation de 1,25%. En revanche, si 2 colonies sont comptées, la concentration est estimée à 20 ufc/g. Elle aurait été de 10 ufc/g ou de 30 ufc/g si 1 colonie ou 3 colonies avaient été décomptées, soit une variation de 50%.

De même, la concentration dans une prise d'essai peut être supérieure à la LOQ même si aucune colonie n'est décomptée dans la boîte de Petri. Supposons une prise d'essai contaminée à 30 ufc/g. La suspension-mère contient donc 300 ufc dans 100 mL. Il est communément admis que le nombre de bactéries dans un liquide suit une loi de Poisson (Jongenburger et al. (2010)), chaque petit volume V_{r-1} de 1 mL contient en moyenne 3 ufc et la probabilité qu'il ne contienne aucune bactérie vaut $\exp(-3) = 0,05$. Certes, cette

d. Les conditions de répétabilité sont les conditions où les résultats d'essai indépendants sont obtenus sur la même suspension mère, par la même méthode dans le même laboratoire, par le même opérateur, utilisant le même équipement et pendant un court intervalle de temps.

e. Les conditions de reproductibilité sont les conditions où les résultats d'essai sont obtenus par la même méthode sur des suspensions mères différentes, des laboratoires différents avec différents opérateurs et utilisant des équipements différents.

probabilité n'est pas très élevée, néanmoins elle n'est pas négligeable : dans 5% des cas, la conclusion est que la concentration est inférieure à la LOQ alors qu'elle est en fait de 30 ufc/g.

2.1.3.b Données peu nombreuses

En outre, en raison des coûts (environ 20 € par analyse), les données ne sont généralement pas très nombreuses. Ainsi, Uyttendaele et al. (2009) ont fait une étude pendant trois ans afin de collecter des données de détection de *L. monocytogenes*. Au total, 1974 données sont disponibles, ce qui constitue un jeu assez important mais il faut le découper en trois classes de produits (salades composés en mayonnaise, produits à base de viande cuite et poisson fumé) prélevées dans plusieurs usines de fabrication. Chaque classe est découpée en plusieurs sous-classes de produits. Tout dépend du niveau de précision souhaité mais si l'objectif est d'étudier la prévalence par type de sous-produits et par entreprise, chaque jeu contient une centaine de données au maximum. En général, si les jeux de données sont séparés par produit et par entreprise de production, ils sont rarement composés de plus de 100 données (voir les données de Habib et al. (2008); Beaufort et al. (2007); Jørgensen and Huss (1998) par exemple). Il arrive néanmoins que les données soient plus nombreuses : par exemple Gonzales-Barron et al. (2009) disposent de deux jeux de taille conséquente (590 dénombrements de coliformes totaux et 622 dénombrement d'*E. Coli* sur des carcasses de bœuf).

2.1.3.c Autocontrôles difficiles à obtenir

Il est difficile d'obtenir les données d'autocontrôles des entreprises afin d'évaluer la prévalence ou la distribution de concentration en pathogènes de leurs produits car peu de producteurs sont enclins à communiquer leurs résultats. Il y aurait également un risque de biais si les données transmises étaient préalablement filtrées par l'industriel.

Ces trois particularités des données de détection et de dénombrement ne facilitent pas la tâche du statisticien, surtout que, dans le cas de pathogènes, la concentration n'est pas forcément très élevée, ce qui conduit à des jeux contenant une forte proportion d'absence et encore plus forte de décomptes de 0 colonie. Il est possible d'améliorer la situation, par exemple en augmentant le nombre de prélèvements (si le budget le permet), en modifiant les protocoles d'analyses afin d'abaisser la LOD et la LOQ (cf. les protocoles utilisés aux parties 2.3.2 et 2.4.1) ou en essayant de tirer parti d'informations additionnelles comme des jeux de données antérieurs ou des connaissances d'experts (cf. chapitre 3).

2.2 Qu'est-ce qu'un lot ?

2.2.1 Lot "industriel" et lot "statistique"

Un *lot* est un mot fréquemment rencontré dans le langage courant et dans ce travail de thèse mais ce terme n'est pas si facile à définir. D'après le règlement N° 2073/2005 (CE) (article 2), un lot est " un groupe ou une série de produits identifiables obtenus

par un procédé donné dans des conditions pratiquement identiques et produits dans un endroit donné et au cours d'une période de production déterminée". La définition donnée par l'ICMSF (2002) commence par énoncer que c'est une quantité d'aliments fabriqués et manipulés dans des conditions uniformes, mais elle va plus loin car il est précisé que cette définition implique qu'il y a une homogénéité au sein du lot, par exemple que le logarithme de la concentration suit une loi normale. Cependant, les auteurs remarquent que cette condition d'homogénéité n'est pas toujours vérifiée pour la concentration microbienne car celle-ci peut être très hétérogène. Il est donc suggéré d'adapter la taille du lot en fonction du procédé. Si le procédé est continu, on peut penser qu'il y a une certaine homogénéité, donc le lot peut regrouper une plus grande quantité de produits.

L'homogénéité est indispensable au statisticien pour qu'il puisse concevoir une distribution de contamination permettant, par exemple de décrire la contamination d'un lot (les données sont recueillies par échantillonnage). Or, la définition du statisticien n'est pas forcément compatible avec celle de l'industriel qui détermine le lot selon des critères de traçabilité et d'organisation de la production. L'ICMSF (2002) soulève bien le problème et adopte même un mot pour définir le lot au sens statistique du terme en anglais : il s'agit du mot *lot*, alors que le lot industriel est communément appelé *batch*. Dans cet ouvrage, le mot *lot* représente donc un ensemble de produits dont on peut déterminer s'ils sont acceptables ou non avec un plan à deux classes par exemple, simplement en prélevant un échantillon de cet ensemble puisqu'une distribution de la contamination du lot est disponible.

Par la suite, j'appelle lot "industriel" le lot au sens de l'industriel (terme *batch* pour l'ICMSF (2002)) et lot "statistique" le lot au sens statistique (terme *lot* pour l'ICMSF (2002)).

2.2.2 Exemple : fabrication de lardons

Les notions de lot "industriel" et de lot "statistique" peuvent être difficiles à concilier. Prenons l'exemple de la fabrication industrielle de lardons nature. Les étapes du procédé sont énoncées à la figure 2.2. Brièvement, un procédé possible est le suivant :

- *Malaxage* : les poitrines de porc arrivent désossées et parées en provenance de plusieurs fournisseurs, puis elles sont malaxées. Un malaxeur est un appareil ressemblant à une bétonnière. Lors de cette étape, il est rempli de poitrines de porc et de saumure. Il effectue un certain nombre de rotations durant 20 h. A la fin de l'étape, quasiment toute la saumure a pénétré dans les poitrines. Les parois de l'appareil sont plus ou moins recouvertes de limon.
- *Etuvage* : une fois les poitrines sorties du malaxeur, elles sont accrochées à des chariots mobiles au moyen de crochets puis placées dans un four à 50°C pendant 3 h environ.
- *Refroidissement et raidissage* : à la sortie du four, les poitrines sont refroidies puis placées dans la salle de raidissage où règne une température de -15°C. Cette salle sert également de zone de stockage : les chariots peuvent y rester plusieurs jours en fonction des commandes des clients. Cette étape est nécessaire car le tranchage est plus aisé lorsque la viande est dure.
- *Cubage* : les poitrines sont découpées en lardons au moyen d'une lardonneuse.

- *Conditionnement* : juste après l'étape précédente, les lardons sont placés dans des barquettes de 100 ou de 200 g puis un opercule est scellé. Après être passées par un détecteur de métaux, les barquettes sont mises dans des cartons eux-mêmes déposés dans une zone de stockage avant d'être livrés.

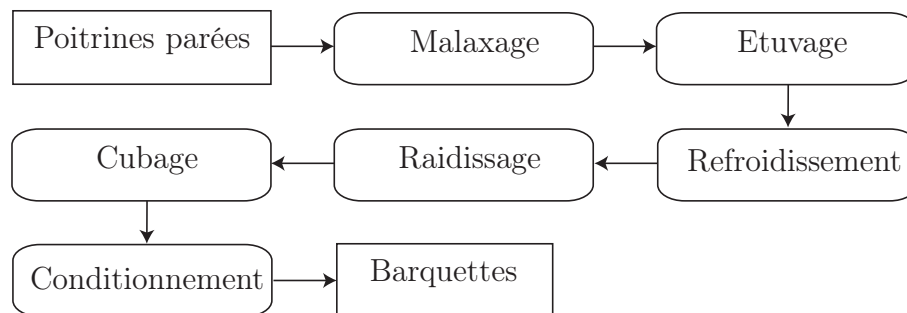


FIGURE 2.2 – Etapes de fabrication de lardons nature industriels. La matière première est constituée de poitrines de porcs parées. Les produits finis sont des barquettes de lardons.

L'entreprise avec laquelle nous avons travaillé a défini deux lots "industriels" : un lot durant le procédé, défini sur les poitrines et un lot pour le produit fini, donc sur les barquettes. Le premier est constitué par les poitrines malaxées par deux malaxeurs en parallèle. Chaque lot reçoit un numéro qui le suit depuis le malaxage jusqu'au raidissage inclus. Le second lot est défini comme le nombre de barquettes fabriquées sur une ligne de production (c'est-à-dire une ligne de cubage et de conditionnement) durant le temps de travail d'une équipe (8 h environ). Un lot de produit fini peut être constitué de plusieurs lots de poitrines. En moyenne, un lot de barquettes est constitué de deux demi-lots de poitrines de porc.

Comment définir le lot "statistique" ? Un malaxeur contient des poitrines provenant de plusieurs fournisseurs, eux-mêmes s'étant fournis auprès de plusieurs élevages. Il n'y a donc aucune raison de penser que le lot "industriel" à l'entrée de l'usine a une quelconque homogénéité. En revanche, il en va tout autrement en sortie du malaxeur. En effet, durant plusieurs heures, les poitrines ont été frottées les unes contre les autres, ce qui leur a permis d'échanger les bactéries se trouvant en surface. Il est donc logique de penser que les concentrations de bactéries en surface de chaque poitrine d'un malaxeur sont proches. C'est d'ailleurs ce qui a été observé en dénombrant la flore lactique des poitrines à la fin du malaxage (cf. partie 2.3). Ainsi, une définition possible d'un lot "statistique" est l'ensemble des poitrines malaxées par un même malaxeur en une fois, c'est-à-dire la moitié d'un lot "industriel" durant le procédé.

Examinons le reste du procédé en suivant le lot "industriel" de poitrines. Celui-ci est étuvé. D'après les expériences menées par le Cemagref dans le cadre du projet

Quant’HACCP, la température en surface des poitrines peut varier en fonction du chariot et de la place d’une poitrine dans un chariot. Une différence de 5°C est possible entre les surfaces de poitrines. En fonction de la température de surface, *L. monocytogenes* peut croître, décroître ou rester stable. On peut donc en conclure que la variabilité intra-lot augmente. Le raidissage a lieu dans une grande salle dans laquelle tous les lots en attente de cubage sont stockés. Comme la température empêche toute croissance de *L. monocytogenes* et de flore lactique, on peut considérer que les différentes variabilités restent inchangées durant cette étape. Lors du cubage, les lots sont redéfinis : le cubage se fait en fonction de la demande des clients et des disponibilités des lignes. Certes, une certaine homogénéisation doit exister au moment du cubage puisque les poitrines sont tranchées avec la même lardonneuse donc des poitrines déposent des bactéries sur les lames et des poitrines tranchées après peuvent récupérer une partie de ces bactéries mais ce phénomène ne se cantonne pas à un lot “industriel” de barquettes et peut s’étendre à tous les lots tranchés entre deux nettoyages, en supposant que le nettoyage arrive à éliminer toutes les bactéries présentes sur la ligne de production, ce qui n’est pas forcément vrai. Dans ces conditions, il est difficile de définir un lot “statistique” au niveau du conditionnement.

Dans ce cas d’étude, le lot ne peut être défini que par une approximation simplificatrice. Ainsi, Billoir et al. (2011) ont modélisé la distribution de *L. monocytogenes* en surface des poitrines de porc, puis des lardons et enfin à l’intérieur des barquettes en suivant le procédé de fabrication. Pour simplifier le problème et pour que la définition du lot ait un sens statistique, le lot considéré est l’ensemble des poitrines malaxées par un malaxeur, en considérant qu’un lot des barquettes est uniquement constitué d’un lot de poitrines.

2.2.3 Exemple : fabrication de saumon fumé

La fabrication du saumon fumé est différente de celle des lardons, en particulier parce qu’il n’y a pas vraiment d’étape “discontinue” comme le malaxage. Les étapes du procédé sont énumérées à la figure 2.3.

Brièvement, les étapes d’un des procédés possibles sont :

- *Filetage/Pelage* : les saumons arrivent entiers dans l’usine. Tout d’abord, les filets sont levés et la peau enlevée. Cette étape se fait en continu sur une ligne de production.
- *Saumurage* : sur cette même ligne, le saumurage se fait par injection : des aiguilles sont plantées dans les filets, la saumure injectée puis les aiguilles retirées.
- *Fumage* : il a lieu dans un four et dure 2 h environ à une température inférieure à 30°C.
- *Raidissage* : il a lieu sur un tapis roulant : c’est donc une étape qui se fait en continu.
- *Tranchage* : il a lieu juste après le raidissage.
- *Conditionnement* : le conditionnement se fait à vide et est en grande partie automatisé.

Ce procédé est donc largement continu, sauf pour le fumage. Le lot “industriel” est le même depuis l’arrivage jusqu’au conditionnement : il est de taille variable (de 200 kg à 3T) et est défini par une marque de vente et une origine de saumon. L’hypothèse faite par

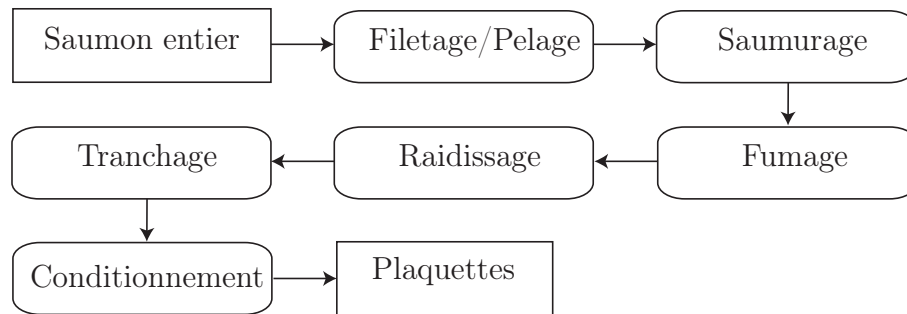


FIGURE 2.3 – Etapes de fabrication de saumon fumé. La matière première est constituée de saumons entiers. Les produits finis sont des plaquettes de saumon fumé.

l’entreprise contactée est que la contamination est apportée par la matière première donc les flux sont séparés au maximum pour éviter les contaminations entre les lignes durant le procédé. Comme ce dernier est continu et que plusieurs lots “industriels” sont fabriqués sur une même ligne de production entre deux nettoyages, il est raisonnable de penser que des contaminations croisées se produisent entre les saumons de ces lots, ce qui contribue à une certaine homogénéisation des contaminations de saumons, bien que faible *a priori*. Ainsi, un lot “statistique” serait l’ensemble des saumons traités sur une même ligne entre deux nettoyages. Si la contamination croisée est faible, on peut même considérer qu’il n’y a pas vraiment de lot car il n’y a pas plus de variabilité entre deux plaquettes de saumon fumé fabriqués entre deux nettoyages que deux plaquettes fabriquées l’une avant et l’autre après un nettoyage.

2.2.4 Fabrication d’un plat en sauce dans une cuisine collective

Le procédé de fabrication d’un plat de bœuf en sauce est celui utilisé par l’hôpital dans lequel des données de température de viande et de sauce ont été acquises dans le projet Quant’HACCP. Les étapes du procédé sont énumérées à la figure 2.4.

Une fois les matières premières sorties de leur emballage et les légumes coupés, les étapes sont les suivantes :

- *Cuisson* : dans une grande sauteuse, le viande est rissolée pendant une demi-heure environ puis les autres ingrédients sont ajoutés pour une cuisson à feu doux pendant à peu près 2 h ;
- *Portionnement* : à la fin de la cuisson, le plat est réparti dans des bacs contenant entre 20 et 25 portions. Les bacs sont operculés puis placés dans un chariot contenant 20 niveaux ;
- *Refroidissement rapide* : le chariot est mis dans une cellule de refroidissement

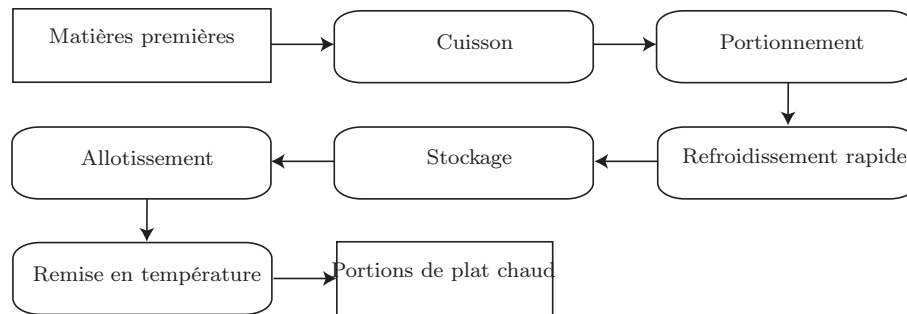


FIGURE 2.4 – Etapes de fabrication d'un plat en sauce en restauration collective.

afin que la température des plats descende à 10 °C en moins de 2h (Arrêté du 21 décembre 2009 (2009)), de manière à ce que la température de la préparation soit le moins longtemps possible comprise entre 10 °C et 63 °C qui est la zone de croissance de *C. perfringens*;

- *Stockage* : lorsque la température de plat est de 10 °C, le chariot est retiré de la cellule de refroidissement et les bacs sont stockés dans une chambre froide au plus pendant 3 jours ;
- *Allotissement* : les bacs sont ouverts et les plateaux-repas des malades constitués. Chaque plateau contient une portion du plat.
- *Remise en température* : les plateaux repas sont placés dans des chariots qui sont déplacés vers les différents services de l'hôpital. Chaque service possède une borne permettant de conserver froids les produits à consommer à température ambiante et de réchauffer le plat à consommer chaud. La remise en température se fait juste avant la distribution des plateaux dans les chambres.

La première étape de ce procédé est discontinuée puisque les aliments sont cuits dans une grande sauteuse. Lorsque le plat est cuit puis refroidi, le personnel de la cuisine centrale se met à la préparation d'un plat différent. En outre, l'étape de cuisson permet aux aliments de se mélanger et d'échanger les éventuelles bactéries se trouvant en surface des matières premières. La cuisson n'élimine pas totalement la présence de pathogène. En particulier, la forme sporulée de *C. perfringens* ne disparaît pas car la température de cuisson n'est pas suffisamment élevée. Après l'étape de cuisson, la préparation culinaire contenue dans la sauteuse est répartie dans des bacs. La température varie d'un bac à l'autre et entre les bacs (au moins dans la cellule de refroidissement et au début du stockage en chambre froide) donc le devenir microbiologique est variable selon les bacs et la localisation des pathogènes dans les bacs. Ensuite, la température devient trop basse pour qu'une croissance soit possible. L'allotissement et la remise en température augmente la variabilité inter-portions puisque la température est variable d'une assiette à l'autre.

Comme la première étape homogénéise la contamination des produits composants le plat, on peut définir le lot “statistique” comme l’ensemble du plat cuit dans la sauteuse. Cependant, si les différentes cellules de refroidissement de l’hôpital n’ont pas la même efficacité, on peut penser qu’une grande variabilité de contamination va apparaître dans les bacs de deux cellules de refroidissement différentes. Nous n’avons pas pu étudier la variabilité inter-cellules car, pour chaque jour de mesure, la température a été mesurée en plusieurs points d’un chariot contenu dans une cellule. La cellule était variable selon le jour de mesure mais cette information n’a pas été conservée. Nous avons donc défini le lot “statistique” comme l’ensemble des plats contenus dans un chariot.

2.2.5 Fabrication de steaks hachés

Ce protocole est discontinu, il fonctionne par “mêlées”, i.e. plusieurs tonnes de viande hachées en même temps avant d’être séparées en steaks. D’après Delignette-Muller and Cornu (2008), la distribution d’*E. coli* O157 :H7 dans une mêlée est Poissonienne, ce qui veut dire qu’il n’y a pas de surdispersion due à la mêlée. Le lot “statistique” peut être identifié comme égal à une mêlée. L’industriel peut lui définir un lot comme étant la somme de plusieurs mêlées.

Quelles conséquences pour l’échantillonnage ? Si un lot vaut une mêlée et qu’une ou deux prises d’essai sont prélevées par lot, on a une indication correcte de la contamination de ce lot. En revanche, si un lot rassemble plusieurs mêlées et que la fréquence d’échantillonnage est la même que précédemment alors les données ne sont représentatives que des mêlées d’où proviennent les prises d’essai donc d’une partie du lot.

2.2.6 Lot, procédé et données

Les quatre exemples ci-dessus illustrent plusieurs types de procédés possibles, depuis le procédé “discontinu” (plat en sauce et steaks hachés) au procédé continu (saumon fumé) en passant par un procédé combinant une étape discontinue et des étapes continues (lardons). La définition du lot “statistique” doit s’adapter à chaque cas : à mon sens, le lot est généralement très dépendant d’une des étapes quand le procédé est au moins discontinu en partie (steaks hachés, lardons et plat en sauce) ou (presque) indépendant du procédé quand ce dernier est continu (saumon fumé). Le lot “industriel”, lui, a une définition beaucoup plus éloignée du procédé (un lot = poitrines malaxées dans deux malaxeurs en parallèle, un lot = un type de saumon fumé et une marque fabriqués durant un jour, etc.).

Une fois le lot “statistique” défini, on peut se demander comment la variabilité inter-lots se comporte. En outre, comme les contaminations observées le sont au niveau d’une prise d’essai ou d’une fraction de suspension-mère, il peut être utile de se pencher sur la structure de contamination d’un lot, i.e. la manière dont on peut modéliser la distribution de contamination d’un lot à partir des observations (prise d’essai et fraction). La définition des lots “statistiques” et la structure de contamination des cas d’applications poitrines de porc et saumon fumé ont été étudiés plus en détail au chapitre 3 en comparant plusieurs modèles possibles de contamination. Connaître la structure de la contamination influence

la manière dont le plan d'échantillonnage est mené, comme discuté au chapitre 3 de cette thèse.

2.3 Acquisition de données dans le cas lardons

Afin de déterminer la distribution de la contamination dans un produit, nous avons acquis des données. Pour l'application *L. monocytogenes*/lardons, celles-ci ont été obtenues conjointement avec l'IFIP^f dans le cadre du projet Quant'HACCP. Il a été décidé de prélever des poitrines après malaxage (pour obtenir des informations sur le lot "statistique") et sur les barquettes de lardons (pour connaître la contamination au niveau du produit fini). Pour des raisons budgétaires, seulement une centaine de poitrines et autant de barquettes ont pu être analysées. La LOD et la LOQ ont été abaissées par rapport aux normes NF EN ISO11290-1 (1996a) et NF EN ISO11290-2 (1998a) afin d'obtenir des informations plus précises sur les contaminations faibles.

2.3.1 Choix du plan

2.3.1.a Choix des facteurs

Pour prélever les poitrines et les barquettes de lardons, le plus simple est de procéder par tirage au sort. Cette technique est certes sans biais mais ne garantit pas que les cas de figure aboutissant à une forte contamination soient pris en compte, surtout parce que le nombre d'unités d'échantillonnage prélevées est très faible devant la production. En outre, il était souhaitable de pouvoir détecter certains effets ayant un impact sur la contamination, si ceux-ci existent. C'est pourquoi nous avons identifié quatre facteurs susceptibles d'expliquer une partie de la contamination :

- fournisseur de poitrines ;
- type de malaxage (semaine ou week-end) ;
- ligne de production (l'usine en possède deux) ;
- tranchage par rapport au moment du nettoyage (un nettoyage des lignes a lieu un jour sur deux).

Dans l'usine étudiée, le malaxage a lieu en semaine (pendant 18 h) ou le week-end (lancement le vendredi en fin d'après-midi et arrêt le lundi matin, durée 60 h). Dans ce cas, les poitrines reposent dans le malaxeur qui commence à tourner uniquement le dimanche dans la journée ; le nombre de rotations est le même que pour un malaxage en semaine. Il a été jugé difficile de choisir des lots en fonction du fournisseur car cela modifiait beaucoup le fonctionnement de l'usine. Il a été également jugé très difficile de pouvoir agir sur l'horaire de passage des poitrines sur une ligne car cette contrainte était trop forte pour l'entreprise. La composition de la cuvée et les horaires de tranchage n'ont donc pas été considérés.

Le principe retenu fut de prélever des poitrines et des barquettes en provenance du même lot de poitrines si possible. Pour cela, les barquettes ont été prélevées lorsqu'un lot de poitrines était en cours de tranchage depuis au moins 45 minutes pour éviter de

f. IFIP signifie "Institut français de l'industrie du porc"

prélever des barquettes contenant des lardons provenant d'un lot de poitrines tranché juste avant. Les poitrines analysées proviennent toutes du même malaxeur (soit un demi-lot de poitrines) afin d'étudier le phénomène d'homogénéisation de la concentration en bactéries à la surface des poitrines en fin de malaxage. Comme après la sortie du malaxeur il est impossible de savoir de quel demi-lot proviennent les poitrines, il est impossible de déterminer de quel demi-lot de poitrines proviennent les barquettes.

2.3.1.b Plan d'échantillonnage stratifié

Nous avons opté pour le plan d'échantillonnage stratifié afin d'avoir la meilleure représentation possible de la contamination des poitrines de porc et des barquettes de lardons et, éventuellement, de mettre en évidence des effets vraiment forts. Dans chaque strate, les unités d'échantillonnage ont été sélectionnées par tirage au sort. Avec deux facteurs possibles, les barquettes se répartissent en quatre strates :

- malaxage semaine, ligne 1
- malaxage semaine, ligne 2
- malaxage week-end, ligne 1
- malaxage week-end, ligne 2

Au niveau des poitrines, un lot correspond uniquement à un type de malaxage. Au niveau des barquettes, il correspond à une type de malaxage \times une ligne.

2.3.1.c Combien de lots par strate choisir ?

Tout d'abord, quelques notations.

La contamination dans chaque strate i , $i = 1, \dots, L$ est représentée par une variable aléatoire X_i . Le nombre de barquettes produites dans la strate i vaut N_i , le nombre total de barquettes vaut N et on note $W_i = N_i/N$. On prélève n_i barquettes dans chaque strate soit un total de n prélèvements.

L'espérance théorique de la strate i vaut : $\bar{X}_i = \frac{\sum_{j=1}^{N_i} x_{ij}}{N_i}$ et la variance théorique $S_i = \frac{\sum_{i=1}^{N_i} (x_{ij} - \bar{X}_i)^2}{N_i}$.

La moyenne empirique de l'échantillon pour la strate i vaut $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$, avec x_{ij} la contamination de la barquette j de la strate i .

La fraction de la strate i échantillonnée est notée $f_i = \frac{n_i}{N_i}$. La contamination moyenne estimée vaut :

$$\bar{x} = \sum_{i=1}^L W_i \bar{x}_i.$$

D'après Cochran (1977), si les \bar{x}_i sont des estimateurs non biaisés des moyennes \bar{X}_i et si les échantillons sont choisis de manière indépendante dans chaque strate, alors \bar{x} est un estimateur non biaisé de la moyenne réelle de la contamination dans la population et sa variance vaut $V(\bar{x}) = \sum_{i=1}^L W_i^2 V(\bar{x}_i)$. Ce résultat est général pour les échantillonnages stratifiés.

Dans le cas d'un échantillonnage stratifié randomisé, la variance de \bar{x} s'écrit (Cochran

(1977)) :

$$V(\bar{x}) = \sum_{i=1}^L \frac{W_i^2 S_i^2}{n_i} (1 - f_i). \quad (2.1)$$

Si $n_i/N_i \leq 0,05$, n_i est négligeable devant N_i et on peut enlever la correction due à la population finie dans le calcul de la variance. Si on fait en outre l'hypothèse que la variance est la même pour chaque strate et vaut σ , l'équation 2.1 devient : $V(\bar{x}) = \sigma^2 \sum_{i=1}^L \frac{W_i^2}{n_i}$. Dans notre cas d'application, n_i est clairement négligeable devant N_i et nous avons supposé que la variance était la même dans chaque strate.

La valeur des n_i est choisie par l'opérateur. Elles peuvent être choisies de façon à minimiser la variance de \bar{x} pour un coût fixé de prélèvement ou de façon à minimiser les coûts pour une variance donnée. Une fonction de coût simple est coût = $c + \sum_i c_i n_i$ avec c_i le coût de prélèvement d'une unité dans la strate i et c un coût systématique. Dans le cas des lardons, le coût de prélèvement d'une barquette est indépendant de la strate à laquelle elle appartient. D'après un théorème présenté par Cochran (1977), la variance de \bar{x} est minimale pour un coût fixé et le coût est minimal pour une variance fixée si $n_i = nW_i$. La variance minimale de \bar{x} vaut alors : $V(\bar{x}) = \sigma^2 \frac{(\sum_i W_i)^2}{n} = \frac{\sigma^2}{n}$. La valeur d'un n_i est choisie proportionnelle à la production de lardons dans la strate i . La répartition de la production de barquettes en mars 2009, seule donnée de ce type dont nous disposons, parmi les quatre strates est présentée au tableau 2.1. En sélectionnant douze lots, la répartition est donnée au tableau 2.2.

		Malaxage	
		sem	w-e
ligne	1	48%	6%
	2	40%	6%

TABLE 2.1 – Répartition de la production de barquettes de lardons en mars 2009

		Malaxage	
		sem	w-e
ligne	1	6	1
	2	4	1

TABLE 2.2 – Répartition de douze lots de barquettes de lardons correspondant à la répartition de production de mars 2009.

2.3.2 Mise en place du plan d'échantillonnage

Comme nous ne pouvons prélever qu'une centaine d'unités de poitrines de porcs et de barquettes, nous avons fixé le nombre de lots à 12 et le nombre d'unités à prélever dans chaque lot à 9. A chaque fois, les 9 poitrines sont issues d'un seul malaxeur (donc un demi-lot de poitrines) et 9 barquettes de lardons issues de cette cuvée devaient être prélevées. Lorsque cela était possible, du limon contenu dans le malaxeur a également été recueilli.

Analyses microbiologiques Pour chaque poitrine, deux prises d'essai ont été prélevées, chacune d'une surface $S' = 100 \text{ cm}^2$ sur la face maigre, auquel a parfois été ajouté un prélèvement sur la face "grasse". Sur une prise d'essai prélevée sur la face maigre, un

dénombrement combiné à une recherche de *L. monocytogenes* a été réalisé. Sur la seconde prise d'essai, ainsi que sur le prélèvement face grasse, la flore lactique a été dénombrée.

Pour chaque barquette prélevée (deux unités de 100 g sécables), une unité de 100 g a été intégralement utilisée pour la recherche et le dénombrement des *L. monocytogenes* et 10 g ont été prélevés dans l'autre afin de dénombrer la flore lactique.

La quantité totale de limon prélevée a été utilisée pour détecter et dénombrer les *L. monocytogenes*.

Protocole analytique La détection et le dénombrement de *L. monocytogenes* a été mené sur la même prise d'essai. le protocole a été inspiré des normes ISO NF EN ISO11290-1 (1996a) et ISO NF EN ISO11290-2 (1998a). Les différents étapes sont les suivantes :

- un prélèvement de surface des poitrines de porc ou une barquette de 100 g de lardons sont mis en suspension dans 100 mL d'un milieu de culture appelé Fraser "base". Ce milieu ne contient pas d'inhibiteur de croissance de bactéries autres que *L. monocytogenes* ;
- homogénéisation de la solution dans un stomacher ;
- partie "dénombrement" du protocole (inspiré du mode opératoire "ALOA one count^g" AES 10/05 - 09/06 validé par Afnor Certification) :
 - 5 mL de la suspension-mère ont été prélevés et chaque mL est étalé en profondeur sur une boîte de Petri (c'est-à-dire que la suspension est versée dans une boîte vide puis de la gélose selon Ottaviani et Agosti sous forme liquide est versé par dessus) ;
 - après une incubation de 24 ou 48h à 37°C, les colonies présentant l'aspect visuel typique des colonies de *L. monocytogenes* sont repiquées sur un milieu nutritif coulé dans une autre boîte de Petri pour ensuite subir des tests de confirmation décrits dans la norme ISO NF EN ISO11290-2 (1998b).
- partie "détection" du protocole :
 - une fois les 5 mL servant au dénombrement retirés, les inhibiteurs sont ajoutés puis la suspension est incubée à 37 °C pendant 24 h ;
 - deux petits volumes (de l'ordre de quelques microlitres) sont prélevés à l'aide d'une anse bouclée ou d'une pipette Pasteur puis déposés sur deux boîtes de Petri contenant chacune un milieu d'isolement sélectif différent sous forme de gélose ;
 - les boîtes sont incubées à 37 °C pendant 24 h et, si nécessaire, c'est-à-dire si les colonies sont petites ou n'apparaissent pas, pendant 48 h ;
 - les colonies présumées être des *L. monocytogenes* sont confirmées au moyen d'essais biochimiques, physiologiques ou morphologiques adaptés.

Enfin, pour les prélèvements de limon, les normes en vigueur ont été appliquées : la détection et le dénombrement de *L. monocytogenes* ont été réalisés en suivant les normes ISO NF EN ISO11290-1 (1996b) et ISO NF EN ISO11290-2 (1998b). La flore lactique a été dénombrée selon la méthode NF EN ISO 15214.

La méthode de dénombrement utilisée a permis d'abaisser la LOQ grâce à une dilution plus faible que celle de la norme ISO NF EN ISO11290-2 (1998b) (dilution au

g. ALOA signifie "agar *Listeria* according to Ottaviani and Agosti".

		Malaxage	
		sem	w-e
ligne	1	<u>1</u>	1
	2	6 + <u>1</u>	1

TABLE 2.3 – Répartition des lots de poitrines pour lesquels des barquettes ont été prélevées. Le lot de poitrines souligné a été partiellement tranché sur les deux lignes.

demi et non dilution au dixième, une seule dilution pratiquée et non pas deux au minimum) et à un volume étalé sur les boîtes plus élevé que la norme précédemment citée (5 mL au lieu de 1 mL). Pour les poitrines de porc, avec les notations utilisées dans la partie 2.1.2, $V_{d1} = 100$ mL, $V_{r1} = 5$ mL et $S = S' = 100$ cm², donc la LOD est de 0,2 ufc/cm². Pour les lardons, $V_{d1} = 400$ mL, $V_{r1} = 4$ mL et $M = M' = 100$ g, la LOD est de 1 ufc/g. La LOQ est de 0,01 ufc/cm² pour les poitrines et de 0,01 ufc/g pour les lardons.

2.3.3 Résultats

L'étude s'est déroulée entre le 13 mai et le 16 juin 2009. Les échantillons ont été réceptionnés par le laboratoire de l'IFIP et analysés le jour ou le lendemain de la réception. Au total, la répartition des lots entre malaxage en semaine et le week-end a été conforme à ce qui était prévu. En revanche, les barquettes ont été prélevées uniquement dans neuf lots de poitrines dont deux malaxés le week-end. Parmi les sept cuvées malaxées en semaine, six ont été tranchées sur la ligne 2 et une a été partiellement tranchée sur les deux lignes. Chacune des deux cuvées malaxées le week-end a été tranchée sur une ligne différente. Le tableau 2.3 synthétise les résultats. En fin de compte, neuf ou huit poitrines ont été prélevées par lot, soit 106 poitrines au total. Des barquettes de lardons n'ont pas été prélevées dans trois lots de poitrines (les seules données disponibles portent sur les poitrines) pour différentes raisons : un lot a été retiré du procédé avant le tranchage, un lot a été tranché en allumettes et un lot a été tranché sur la ligne 2 ; comme quasiment tous les lots sélectionnés ont été tranchés sur cette ligne, il a été jugé inutile de prélever ces barquettes. Pour un lot tranché sur les deux lignes, 14 barquettes ont été prélevées au total (5 sur une ligne et 9 sur l'autre).

2.3.3.a Flore lactique

Poitrines de porc après malaxage Les concentrations en flore lactique exprimées en log ufc/g sont présentées à la figure 2.5a. Pour les neuf poitrines issues d'un même demi-lot, les différences de concentrations mesurées sont inférieures à 1 log ufc/g et la plupart du temps inférieures à 0,5 log ufc/g. Une telle homogénéité des résultats s'explique parce que les poitrines ont échangé des bactéries entre elles au cours du malaxage. Ainsi, à la sortie de cette étape, la flore lactique se répartit de manière presque uniforme à la surface des poitrines. Comme il s'agit d'un phénomène mécanique, on en déduit que cette répartition se fait également pour la flore pathogène. Les concentrations des lots 2 et 5 sont environ dix fois supérieures aux concentrations des autres lots. Ils ont été malaxés le week-end,

ce qui suggère que le malaxage le week-end permet une croissance plus importante des bactéries.

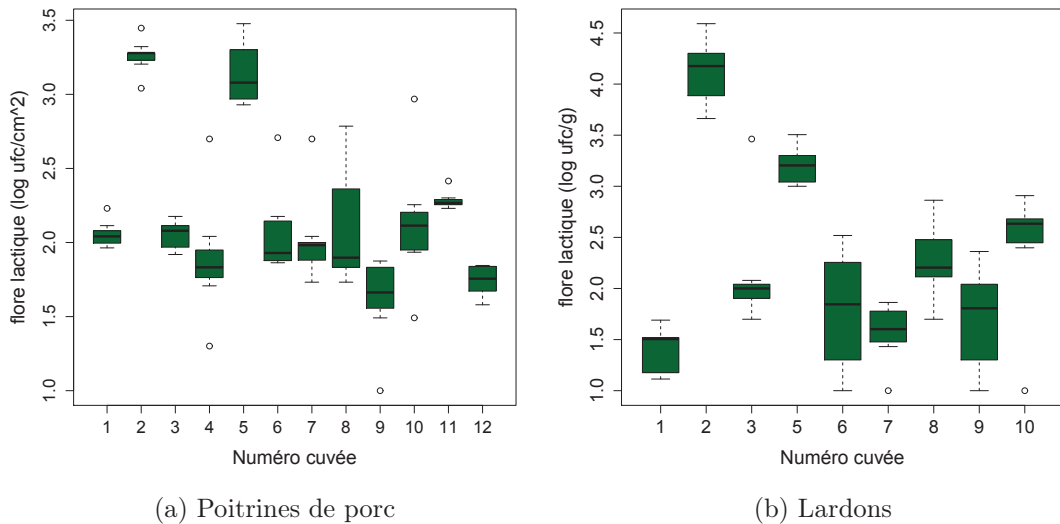


FIGURE 2.5 – Boîtes à moustache des concentrations des prises d’essai (a) de poitrine après malaxage par cuvée ($\log \text{ufc}/\text{cm}^2$) et (b) de lardons après conditionnement ($\log \text{ufc}/\text{g}$).

Lardons après conditionnement Pour un lot de poitrines donné, la différence entre la prise d’essai où la concentration est la plus élevée et celle où la concentration est la plus faible est inférieure à 1 log sauf pour deux un lots où elle est inférieure à 1,5 log. L’homogénéité de la concentration acquise lors du malaxage est en partie conservée. Ce résultat est cohérent puisque les poitrines ne sont plus au contact les unes des autres après le malaxage. La seule étape où des échanges de bactéries peuvent se produire est lors du tranchage mais ce phénomène est plus marginal. En outre, les barquettes proviennent de l’un ou de l’autre des demi-lots composant un lot de poitrines alors que les poitrines ne proviennent que d’un seul demi-lot, il n’est donc pas étonnant que la dispersion soit plus élevée chez les lardons que chez les poitrines. De même que pour les poitrines, les concentrations des lots malaxés le week-end (2 et 5) sont 10 à 100 fois supérieures aux concentrations des autres cuvées.

2.3.3.b *L. monocytogenes*

Poitrines de porc après malaxage La prévalence (ici égale au nombre de détections positives sur le nombre total de détections réalisées) varie de 0 (pour 4 lots) à 1 (pour 1 lot). Au total, il y a 25 présences sur 106. Le dénombrement est toujours inférieur au seuil sauf pour une prise d’essai du lot 3 et toutes les prises d’essai du lot 11, soit 9 résultats différents de 0. Lorsque la concentration d’une prise d’essai est supérieure ou égale à 0,2

ufc/cm², le résultat de la détection est toujours une présence. Les résultats figurent au tableau 2.4.

N° de cuvée	Détection	Dénombrement (ufc)	Nb de poitrines de porc prélevées par lot
1	0-1-1-1-0-1-1-1-1	0-0-0-0-0-0-0-0-0	9
2; 8 à 10	0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	9
3	0-0-0-0-1-0-0-0-1	0-0-0-0-1-0-0-0-0	9
4	0-1-0-0-1-0-0-0-0	0-0-0-0-0-0-0-0-0	9
5	0-0-0-0-1-0-0-0-0	0-0-0-0-0-0-0-0-0	9
6	0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	8
7	0-0-0-1-1-0-0-0-1	0-0-0-0-0-0-0-0-0	9
11	1-1-1-1-1-1-1-1-1	11-9-6-5-12-29-16-3	8
12	0-0-1-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	9

TABLE 2.4 – Données brutes pour la détection et de dénombrement des prises d’essai de poitrines de porc après malaxage. Parmi les lots de poitrines choisis, huit ou neuf poitrines ont été prélevées aléatoirement au sein d’un même demi-lot. La détection et le dénombrement ont eu lieu sur la même prise d’essai de 100 cm². Pour chaque lot, le n^e résultat de détection et le n^e résultat de dénombrement appartiennent à la même prise d’essai.

Pour la détection “0” signifie absence et “1” présence.

Lardons après conditionnement La prévalence (même définition qu’au paragraphe précédent) varie de 0 (pour 5 lots) à 0,66 (pour 1 lot) soit un total de 15 présences. Le dénombrement est toujours inférieur au seuil sauf pour une prise d’essai de la cuvée 10. Comme pour les poitrines, lorsque la concentration d’une prise d’essai est supérieure ou égale à 0,2 ufc/g, le résultat de la recherche est toujours “présence”. Les résultats sont donnés au tableau 2.5.

2.3.4 Conclusion

La prévalence de contamination en *L. monocytogenes* dans les prises d’essai de poitrines de porc après malaxage est de 24% et dans prises d’essai de lardons après conditionnement à 17%. Les résultats diffèrent de manière sensible d’une cuvée à l’autre puisque l’intervalle de confiance à 90% est égal à [0 ; 88%] pour les poitrines après malaxage et de [0 ; 62%] pour les lardons.

Cette étude montre que lorsque les poitrines sont malaxées le week-end, la concentration en flore lactique est dix fois supérieure à celle mesurée après un malaxage en semaine. Il est raisonnable de penser que le même phénomène serait observé pour *L. monocytogenes* mais la trop faible concentration pour ce pathogène (presque systématiquement inférieure à 0,2 ufc/cm²) ne nous a pas permis de le mettre en évidence.

En revanche, nous ne pouvons pas conclure sur l’existence d’un effet “ligne” car trop peu de lardons ont été tranchés sur la ligne 1 lors de la campagne de prélèvements.

D’après les résultats de dénombrement de flore lactique obtenus sur les poitrines après malaxage, la concentration en flore lactique y est homogène. Ce résultat nous semble

N° de cuvée	Détection	Dénombrement (ufc)	Nb de barquettes prélevées par lot
1	0-0-0-0-0-0-0-0-1	0-0-0-0-0-0-0-0-0	9
2;3;8;9	0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	9
5	0-0-0-1-0-1-1-1-1	0-0-0-0-0-0-0-0-0	9
6	0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0-0-0-0-0-0-0	14
7	0-0-0-1-1-0-1-0-0	0-0-0-0-0-0-0-0-0	9
10	0-0-1-1-1-1-0-0-1	0-0-0-0-1-0-0-0-0	9

TABLE 2.5 – Données brutes pour la détection et de dénombrement des prises d’essai de lardons après conditionnement. Parmi les neuf cuvées choisies, neuf barquettes (quatorze pour une cuvée) ont été prélevées dans les mêmes cuvées que les cuvées choisies pour prélever les poitrines après malaxage. La détection et le dénombrement ont eu lieu sur la même prise d’essai de 100 g. Pour chaque cuvée, le n^e résultat de détection et le n^e résultat de dénombrement appartiennent à la même prise d’essai.

Pour la détection “0” signifie absence et “1” présence.

pouvoir être étendu à la flore pathogène : si le malaxeur contient une poitrine contaminée avec *L. monocytogenes*, ces bactéries vont se disperser sur les autres poitrines du malaxeur. Ainsi, la probabilité de sélectionner une poitrine contaminée après malaxage est plus élevée que celle de sélectionner une poitrine contaminée après arrivage (avant malaxage). Par conséquent, effectuer des analyses après malaxage augmente la probabilité de détecter une contamination avec *L. monocytogenes*.

Afin d’évaluer la qualité microbiologique des matières premières livrées par chaque fournisseur, en particulier lors de l’étude de bactérie à faible prévalence, il serait pertinent, pour les cuvées mono-fournisseur, de réaliser les prélèvements après l’étape de malaxage.

2.4 Acquisition de données dans le cas saumon fumé

2.4.1 Prélèvements et protocoles

Les données ont été acquises dans le cadre du projet LmPR, qui s’est déroulé entre 2001 et 2004 et a été coordonné par l’AFSSA. Ce projet avait pour objectif de construire une démarche d’appréciation de l’exposition à *L. monocytogenes* dans les produits réfrigérés dont le saumon fumé. Dans ce cadre, 172 plaques de saumon fumé ont été prélevées par tirage au sort en sortie de procédé dans une usine entre mai 2003 et janvier 2004. Elles étaient réparties en 29 lots “industriels” fabriqués des jours différents, on peut donc considérer qu’elles font également partie de 29 lots “statistiques”.

Pour chaque plaque, une analyse de détection a été faite en suivant le mode opératoire “ALOA one day” N° AES 10/3-09/00 validée par Afnor certification selon la norme NF EN ISO 16140 (2003). Les étapes sont les suivantes :

- de petits morceaux de tranches de saumon d’une même plaque sont prélevés de façon à ce que la masse totale de la prise d’essai soit de 25 g ;
- la prise d’essai est placée dans 225 mL d’un milieu d’enrichissement primaire

- appelé Fraser-demi pour que la suspension-mère soit de 250 mL ;
- la solution obtenue est homogénéisée dans un stomacher ;
- la suspension-mère est incubée à 30°C durant 24 h ;
- un volume de 0,1 mL est prélevé puis étalé sur une boîte de Petri contenant une gélose selon Ottaviani et Agosti ;
- la boîte de Petri est incubée à 37°C pendant 24 h ;
- les colonies caractéristiques de *L. monocytogenes* sont confirmées selon des méthodes non décrites ici.

Pour le dénombrement, Gnanou Besse et al. (2004) ont adapté la technique de filtration sur membrane qui a été reprise dans la norme NFV45-008 (2009) afin d'abaisser la LOQ (par rapport à la LOQ de la norme NF EN ISO11290-1 (1996a)). Brièvement, le protocole est le suivant :

- une prise d'essai est constituée de plusieurs petits morceaux découpés sur une tranche pour une masse totale de 25 g ;
- elle est diluée dans 225 mL de peptone-sel pour obtenir une suspension-mère de 250 mL après stomachage ;
- dans un flacon sont versés 50 mL de la suspension mère ainsi que 5 mL de solution de Tween 80^h et 5 mL de solution de trypsineⁱ ;
- après 20 à 25 min à 37°C passées dans un bain d'eau à agitation,
- trois volumes de suspension (6, 18 et 36 mL) sont successivement filtrés. Une membrane avec des pores de 45 μm de diamètre est utilisée pour chaque volume est prélevé. Les volumes de la suspension mère initiale sont respectivement de 5, 15 et 30 mL ;
- après filtration, chaque membrane est déposée sur une boîte de Petri contenant de la gélose selon Ottaviani et Agosti ;
- les boîtes sont placées dans une étuve à 37°C pendant 24 h (48 h si les colonies sont absentes ou petites après 24 h) ;
- les colonies présentant l'aspect caractéristique des colonies de *L. monocytogenes* (bleu mais sans halo à cause de la présence de trypsine) sont comptées puis 10 colonies (ou toutes les colonies de la boîte s'il y en a moins de 10) sont sélectionnées et ensemencées sur une nouvelle boîte de Petri avec de la gélose selon Ottaviani et Agosti. L'incubation se fait à 37 °C pendant 24 h afin de permettre l'apparition d'un halo typique des colonies de *L. monocytogenes* ;
- Enfin, 5 colonies caractéristiques sont confirmées avec des tests d'identification présentés dans la norme NF EN ISO ISO11290-2 (1998a).

Si tout se passe de manière parfaite et si des *L. monocytogenes* sont présentes sur les membranes, les bactéries se sont déposées sur les boîtes et ont formé des colonies sur les boîtes. Avec ce protocole, la LOQ est très basse : avec les notations de la partie 2.1.2, $V_{d1} = 250$ mL, $V_{r1} = 50$ mL et $M' = 25$ g, donc la LOQ est de 0,2 ufc/g. La valeur de V_{r1} est obtenue en sommant les trois volumes de suspension-mère effective filtrés. La méthode déterminée par Gnanou Besse et al. (2004) permet des résultats plus précis que ceux obtenus avec la norme ISO11290-2 (1998a) lorsque les niveaux de contamination sont

h. La solution Tween 80 permet de faciliter la filtration

i. La trypsine est une enzyme du suc pancréatique qui participe à la digestion des protéines. Elle facilite donc la filtration de la suspension-mère.

faibles puisque l'équivalent de 50 mL (soit 1/5e) de suspension-mère sont étalés sur les boîtes, au lieu de 1 mL et 0,1 mL, soit respectivement 1/100e et 1/1000e dans la norme.

2.4.2 Résultats

Les données ont été publiées dans Beaufort et al. (2007) (usine 8). Les résultats détaillés sont présentés dans le tableau 2.6. La prévalence (nombre de détections positives

N° de lot	Détection	Dénombrement (ufc)	Nb de plaques par lot
1 à 6, 8 à 13;16;20;23;25;27 à 29	0-0-0-0-0	0-0-0-0-0	6
7	1-1-1-1-1	1-0-0-0-0	6
14	1-1-0-0-0	0-0-0-0-0	6
15, 17	1-0-0-0-0	0-0-0-0-0	6
18	1-1-1-0-0	3-2-0-0-0	5
19	1-1-1-0-0	0-0-0-0-0	6
21;24;26	1-1-0-0-0	1-0-0-0-0	6
22	0-0-0-0-0	0-0-0-0-0	5

TABLE 2.6 – Données brutes dans le cas d'application "saumon fumé". Parmi les 29 lots choisis, cinq ou six plaques en sortie usine ont été choisies aléatoirement. Deux prises d'essai de 25 g de chaque plaque ont été analysées, l'une avec un protocole de détection (ISO11290-1 (1996a)), l'autre avec un protocole de dénombrement par filtration (Gnanou Besse et al. (2004)). Pour chaque lot, le n^e résultat de détection et le n^e résultat de dénombrement appartiennent à la même plaque.

Pour la détection "0" signifie absence et "1" présence.

sur nombre total de détections pratiquées) est de 12,8% tous lots confondus et varie de 0 à 1 selon les lots. L'intervalle de confiance à 90% est de [0; 0,56]. Les résultats sont plus homogènes entre les lots que pour les poitrines de porc après malaxage en partie parce que le lot de prévalence 1 pour la détection (lot n° 7) n'est pas celui ayant les résultats de dénombrement les plus élevés.

L'ensemble des données récoltées pour les applications "lardons" et "saumon fumé" nous ont ensuite permis d'étudier la structure de contamination des produits en cours ou en fin de procédé, c'est-à-dire de déterminer la variabilité inter et intra lots "statistiques" par exemple (cf. chapitre 3).

Chapitre 3

Listeria dans les aliments : Modèles bayésiens hiérarchiques pour estimer la variabilité inter et intra-lots

Ce chapitre traite de l'élaboration puis de la comparaison de modèles de contamination qui considèrent ou non plusieurs variabilités (au niveau du lot, de l'unité ou de la prise d'essai). Les données utilisées pour mener l'inférence bayésienne sur les paramètres sont celles décrites au chapitre 2 (poitrines de porc après malaxage et saumon fumé en sortie usine). Ce travail a fait l'objet d'un article paru dans la revue *Risk Analysis* et reproduit à la partie 3.2 du présent chapitre.

3.1 Variabilité inter et intra lots

Lorsque nous avons souhaité modéliser la contamination d'un aliment par un micro-organisme au sein d'une usine, la question du lot s'est posée d'emblée. En effet, la plupart des modèles décrits à la partie 1.4 prennent pour acquis la définition d'un lot au sens d'une population statistique car les données étudiées proviennent de plusieurs lots différents, parfois de plusieurs usines. Il n'est en effet pas toujours possible de connaître l'appartenance exacte des données à un lot "statistique".

Cependant, lorsque l'on travaille au sein d'une même usine, on se demande naturellement si deux produits d'un même lot se ressemblent plus que deux produits de deux lots différents. Pour répondre à cette question, il faut qualifier les variabilités inter et intra-lots. C'est ce que propose ILSI (2010) en comparant quatre cas de figures représentés à la figure 3.1. A la figure 3.1a, les distributions inter lot se chevauchent entre elles. Il est donc difficile de différencier les lots entre eux bien qu'il y ait peu de recouvrements entre les lots les plus et les moins contaminés. La figure 3.1b présente la même variabilité intra lot qu'à la figure précédente mais la variabilité inter-lots a été fortement diminuée. Presque tous les lots, même les plus contaminés et les plus propres se ressemblent. Cette ressemblance persiste si la variabilité intra-lot est divisée par 10 (figure 3.1d). En revanche, si la variabilité inter-lots augmente (figure 3.1c), chaque lot peut être distingué puisque les distributions inter-lots ne se recourent quasiment pas. La variabilité intra-lot est la même

pour les figures 3.1a et 3.1b d'une part et pour les figures 3.1c et 3.1d d'autre part. La variabilité inter-lots est la même pour les figures 3.1a et 3.1c d'une part et pour les figures 3.1b et 3.1d d'autre part.

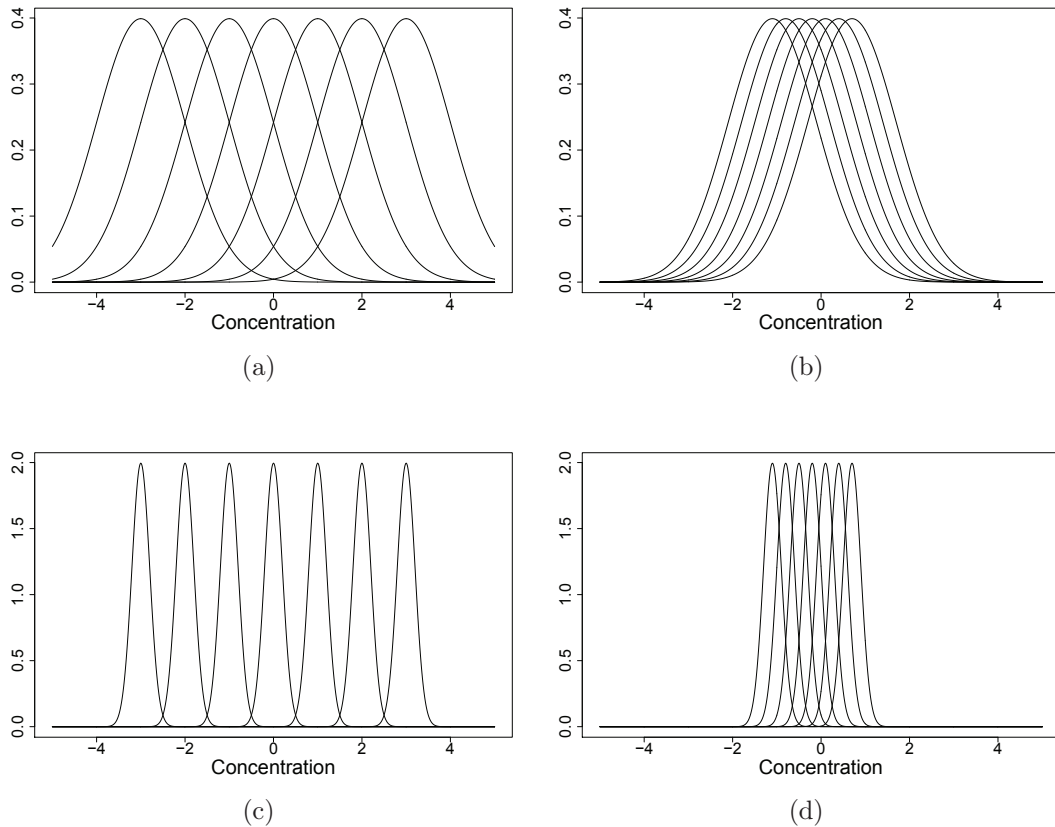


FIGURE 3.1 – Représentation des variabilités inter et intra-lots. Chaque courbe représente la distribution en log ufc/g d'un lot. L'écart-type est de 1 log ufc/g pour les figures 3.1a et 3.1b et de 0,2 log ufc/g pour les figures 3.1c et 3.1d. La variation entre deux moyennes de lots est de 1 log ufc/g pour les figures 3.1a et 3.1c et de 0,3 log ufc/g pour les figures 3.1b et 3.1d.

Si la variabilité inter-lots est supérieure à la variabilité intra-lot (comme représenté à la figure 3.1c par exemple), un petit échantillon d'un lot peut suffire à déterminer la qualité du lot. Dans le cas contraire (figure 3.1b par exemple), un petit échantillon ne donne quasiment aucune information. Si les deux variabilités sont du même ordre de grandeur (comme représenté aux figures 3.1a et 3.1c), un petit échantillon ne permet de distinguer les lots que de manière très imprécise.

Gonzales-Barron and Butler (2011a) ont également étudié la variabilité intra et inter lots et ont proposé plusieurs modèles tous basés sur la même trame. Chaque réplicat k de comptage obtenu sur un produit j (en l'occurrence des carcasses de bœufs à l'abattoir) du lot i suit une loi de Poisson de paramètre $K \times \lambda_{ij}$, avec K une constante liée au protocole utilisé et λ_{ij} une concentration. Cette dernière suit

- soit une loi log-normale de paramètres μ_j et σ (ou σ_j si l'écart-type varie aussi avec le lot) ;
- soit une loi Gamma de paramètres a_j et b_j .

Des lois *a priori* sont proposées pour les paramètres du modèles. Dans les variantes où tous les paramètres varient avec le lot, les auteurs proposent également une version dans laquelle les paramètres (μ_j et σ_j ou a_j et b_j) sont corrélés. Les différents modèles ont été comparés par calcul du DIC^a. Avec leurs jeux de données, lorsque les dénombrements sont élevés (les concentrations varient entre -2 et 7 log ufc/cm²), les modèles de type Poisson-Gamma obtiennent un DIC plus petit que les modèles de type Poisson-Log-Normal. A l'inverse, lorsque les dénombrements sont peu élevés (les données exprimées en concentration varient de -2 à 1 log ufc/cm² et il y a au moins 10% de zéros, cf. Gonzales-Barron et al. (2010a) pour une description des jeux de données) et qu'il y a beaucoup de zéros, les modèles de type Poisson-Log-Normale s'ajustent mieux. De manière générale, la corrélation modifie peu le critère.

Disposant des données de contamination par *L. monocytogenes* dans des poitrines de porc après malaxage et du saumon fumé en sortie usine (tableaux 2.4 et 2.6 respectivement), nous nous sommes penché sur la définition de lot (cf. parties 2.2.2 et 2.2.3). Comment prendre en compte la variabilité inter-lots pour chaque application ? De même qu'au sein d'une usine, on se demande s'il est pertinent d'étudier la variabilité inter et intra lots, nous nous sommes demandé s'il existait une variabilité inter et intra produits dans un même lot.

Ces différentes questions nous ont amenés à définir plusieurs modèles en prenant en compte

- les variabilité inter-lots et inter-produits,
- la variabilité inter-lots,
- la variabilité inter-produits,
- la variabilité inter prises d'essai uniquement.

La présentation et la comparaison des modèles ainsi que l'estimation des paramètres sont décrits dans l'article (partie 3.2).

a. Le *deviance information criterion* (DIC) est un critère courant utilisé pour la sélection de modèles hiérarchiques bayésiens lorsque la distribution *a posteriori* a été obtenue par des *Markov chain Monte-Carlo* (MCMC). Il a été proposé par Spiegelhalter et al. (2002) afin de sélectionner les meilleurs modèles prédictifs. Il est égal à :

$$\begin{aligned} DIC &= D(\bar{\theta}) + 2p_D \\ &= \overline{D(\theta)} + p_D, \end{aligned}$$

où $D(\bar{\theta})$ est la déviance de la moyenne *a posteriori* du paramètre θ , $\overline{D(\theta)}$ est l'espérance *a posteriori* de la déviance et p_D est le nombre effectif de paramètres dans le modèle. $\overline{D(\theta)}$ augmente avec le nombre de paramètres et p_D contrebalance cet effet en favorisant les modèles avec peu de paramètres. Le meilleur modèle est celui dont le DIC est le plus petit.

3.2 Article : *Hierarchical Bayesian models to assess between and within batch variability of pathogen contamination in food*

Abstract Assessing within-batch and between-batch variability is of major interest for risk assessors and risk managers in the context of microbiological contamination of food. For example, the ratio between the within-batch variability and the between-batch variability has a large impact on the results of a sampling plan. Here, we designed hierarchical Bayesian models to represent such variability. Compatible priors were built mathematically to obtain sound model comparisons. A numeric criterion is proposed to assess the contamination structure comparing the ability of the models to replicate grouped data at the batch level using a posterior predictive loss approach. Models were applied to two case studies : contamination by *Listeria monocytogenes* of pork breast used to produce diced bacon and contamination by the same microorganism on cold smoked salmon at the end of the process. In the first case study, a contamination structure clearly exists and is located at the batch level, i.e. between batches variability is relatively strong while in the second a structure also exists but is less marked.

Keywords : Hierarchical models, Bayesian modeling, microbiological contamination, detection & enumeration, predictive approach

Cette partie est la reproduction de l'article écrit en collaboration avec Marie Cornu, Isabelle Albert, Jean-Baptiste Denis et Eric Parent. Il a été publié dans la revue Risk Analysis (32(3) :395-415)

3.2.1 Introduction

Microbial contamination of industrially processed foods is investigated in different contexts, including quantitative microbiological risk assessment (QMRA), and verification by Food business operators (FBOs) of the compliance of their production with microbiological criteria. In the context of QMRA, contamination is studied for a given category of food at a large scale, for instance the national scale (e.g., for cold smoked salmon (Pouillot et al. (2007, 2009)) for two French cheeses (Sanaa et al. (2004)) or the international scale (e.g., for fresh vegetables (Crépet et al. (2007))). At a more focused level, FBOs are concerned with contamination in their plant, to ensure that “foodstuffs comply with the relevant microbiological criteria” (European Regulation (CE) No 2073/2005 (2005), Article 3). This raises the issue of how to define batches within the whole production, and the gap between statistical and practical definitions. In practice batches are defined by FBOs to manage their traceability. Thus, according to European citeCE2005, Article 2, a batch is “a group or set of identifiable products obtained from a given process under practically identical circumstances and produced in a given place within one defined production period”. Ideally, if all types of contamination (physical, biological and chemical hazards) were governed similarly by the circumstances of production, such a definition could match with the statistical point of view. Indeed, for a statistical analyst, a batch

is a portion of production in which the contamination has a steady statistical behavior. As a consequence, any sample is depicted by a stationary probability distribution of independent events. However, in practice, the statistically based definition of batches seldom applies to all types of hazards. Consequently, some misunderstanding may arise between FBOs and statisticians. For a statistical analyst, assessing the effect of heterogeneity is therefore the key issue.

It is very unlikely that different batch parts of the same batch contain exactly the same number of pathogens (ILSI (2010)). Variability is observed among these parts but variability can also be observed at a higher level, for example among batches. The determination of the structure has an impact on the way sampling plans are carried out. For instance, if between-batch variability is relatively strong, all the samples taken from one batch will give information about this batch but not about the whole production. It would help the applied statistician to use models of contamination which take into account different variability structures at the plant level, for instance to correctly assess the population exposure to a pathogen in a given food. The data considered in routine investigation of microbiological contamination are of two kinds : detection (presence absence result) and enumeration (colony count). In many cases, data do not provide much information : data sets are small, most detection data are “absence” and enumeration data are predominantly 0 counts.

When information given by data is poor, it is crucial to use another way to get information. A Bayesian approach offers flexibility in models construction and enables one to incorporate several sources of prior information such as previous data and expert knowledge into the statistical analyses. Another advantage of the Bayesian method is that the separation of variability and uncertainty on the parameters is straightforward. This separation between on the one hand the diversity or heterogeneity that is irreducible by additional measurements and on the other hand the ignorance or lack of perfect knowledge on the part of the analyst was promoted for risk analysis in the 1990s (Hoffman and Hammonds (1994); Cullen (1999)) Hierarchical Bayesian modeling is an efficient framework for taking these two levels into account (Spor et al. (2010)). The first Bayesian implementations of these concepts of uncertainty and variability appeared nearly at the same time in multiple fields, including environmental toxicology (Verdonck et al. (2001)), environmental health (Brand and Small (1995)) ecological modeling (Calder et al. (2003)) climate change (Tebaldi et al. (2005)) and QMRA ((Pouillot et al., 2003; Barker et al., 2005)).

In this article, to check whether it is worth taking the batch structure into account, models describing various levels of aggregation can be compared. Several comparison criteria are available in a Bayesian framework, such as DIC, BIC, the Bayes factor or the predictive posterior approach (Gelfand and Ghosh (1998)). The DIC is often used but it requires to assume that posterior means are good estimates to the stochastic parameters and it uses a global comparison criterion through the deviance criterion (Spiegelhalter et al. (2002)). The Bayes factor compare one model to another by calculating the ratio of observing the data given each model. It is a sort of summary of the evidence provided by the observed data in favor of a statistical model, as opposed to another. It is very sensitive to the choice of priors (Kass and Raftery (1995)). BIC can be seen as a large-sample approximation to Bayes factors (Raftery (1995)). The criterion defined by Gelfand and Ghosh (Gelfand and

Ghosh (1998)) is based on expected losses on replicated data sets under the compared models. We used this last criterion because it allows to choose the loss function that aims to correctly replicate data grouped together at the batch level. Then, this specific predictive posterior approach validates the predictive power of the model as far as grouped structures are concerned. In the Bayesian framework, the comparison criterion is influenced by the priors of each model. In order not to favor a model just because of its priors, so-called compatible priors were developed (Dawid and Lauritzen (2000)). The principle is to choose priors so that simulated data using the priors and the likelihoods of the models have distributions "as close as possible" to each other (Roverato and Consonni (2004)). The aim of this article is to find what structure contamination can be found in the detection and enumeration data of a food/pathogen couple in a plant. The idea is to determine, for example, whether the within-batch structure is worth taking into account or not. This can have important consequences on the efficiency of different sampling schemes for quality control. In this work, particular attention was given to prior construction. In this paper, two cases were considered, both on *L. monocytogenes* contamination, one at the end of the process of cold smoked salmon (CSS) and the other within the diced bacon process. For each application, the objective was to determine the most suitable contamination structure(s). Each datum in the data set is labeled with the batch or sub-batch structure to which it belongs. *L. monocytogenes* is a pathogen which can cause listeriosis and which is regularly isolated from CSS and diced bacon products. The invasive form of this infection particularly affects pregnant women, the elderly and people affected by diseases or conditions such as cancer, diabetes, AIDS, etc. (Ryser and Marth (2007)) Listeriosis is usually associated with consumption of uncooked products, which is clearly the case of CSS. Diced bacon is usually consumed cooked but is occasionally consumed raw by 14% (ANSES (2009)) of people in France.

3.2.2 Theoretical framework

3.2.2.a Definition of the contamination structure

To start with some necessary terminology, in the framework of the contamination structure, let us define what is meant by batch, unit, test portion and fraction. In this section, the word *batch* is used with the statistical definition. In the case studies, the best compromise, according to the authors, between the statistical and practical definitions is defined. A *unit* is a part into which a given food batch is usually divided, corresponding in practice to a retail unit. A *test portion* is a part of a unit on which the microbiological analysis is carried out (usually 10 or 25 g). A *fraction* is the volume from the first dilution pour-plated onto one or several Petri dishes for enumeration. For example, for the CSS, a unit is a packaging unit and a test portion is a part of the unit of 25 g. As a consequence, four nested levels are considered : fraction in test portion in unit in batch. Figure 3.2 depicts these four levels. There can be several units selected from one batch, and there can be several test portions selected from one unit.

Detection and enumeration data for *L. monocytogenes* are obtained through different analyses (ISO11290-1 (1996a); ISO11290-2 (1998a)). To perform detection, a test portion is taken (usually 25 g); it is diluted in a selective broth and stomached before being in-

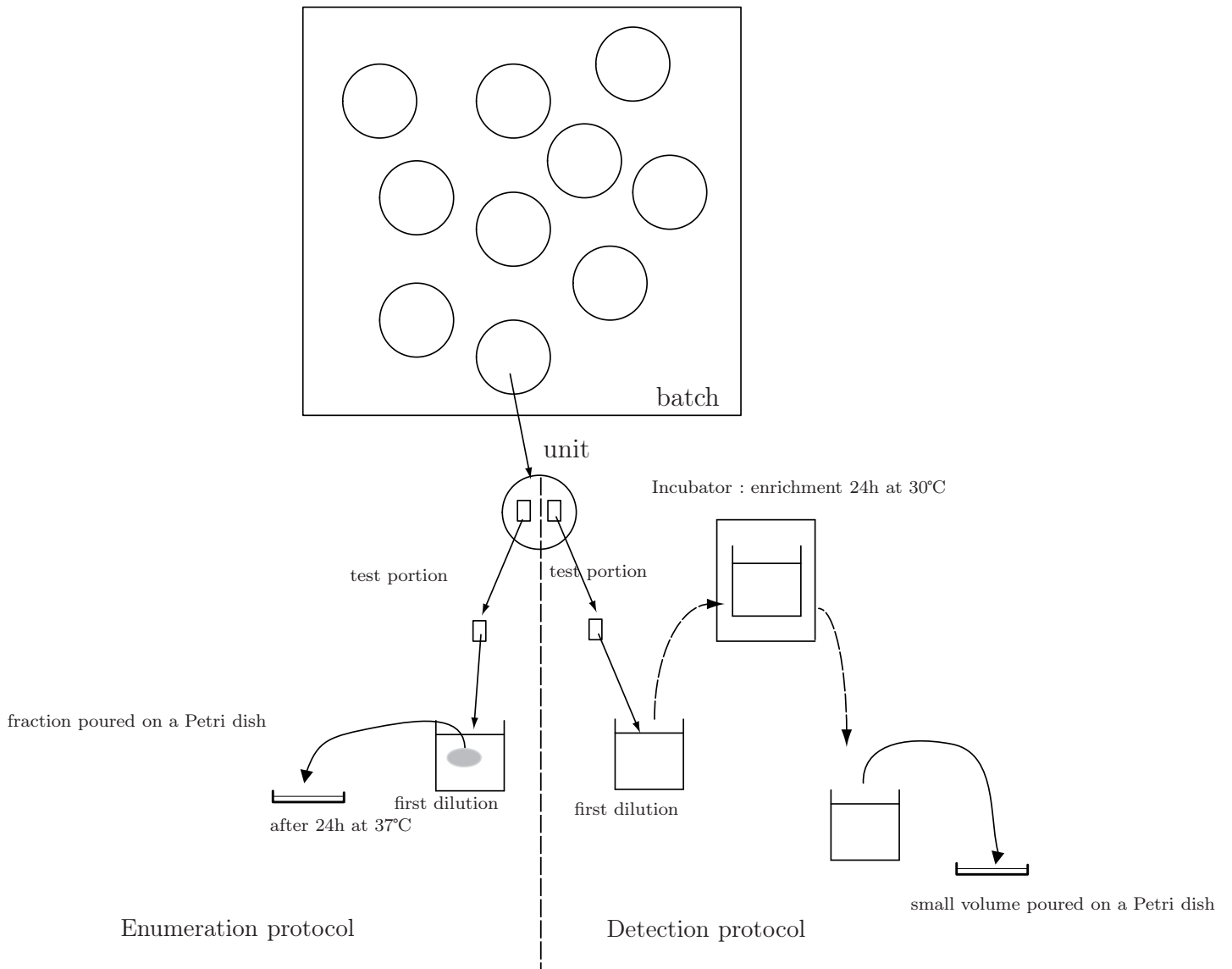


FIGURE 3.2 – Diagram of the different levels of the enumeration and detection protocols : batch, unit, test portion and fraction.

cubated 24h at 30 °C. A small volume is then spread onto a Petri dish and incubated 48h at 37°C. If *L. monocytogenes* is present in the test portion, at least one colony is observed on the Petri dish. In that case, the result is *presence*; if not, the result is *absence*. To perform enumeration, a test portion of a food product is taken and diluted in a solution and stomached for 1 minute. Then, X mL of the solution (the fraction) is removed and pour-plated onto one or several Petri dishes. After 24h at 37°C, colonies are counted on the plates. Enumeration and detection may be carried out or not on the same test portion. Detection (presence or absence) gives a result at the test portion level because of the enrichment step of 24 h : if *L. monocytogenes* was present in the test portion, it was present in the first dilution and multiplied, so the probability of finding cells in the volume spread onto a Petri dish is the same as the probability of finding cells in another volume of the same solution. On the contrary, an enumeration result only gives information about the fraction poured onto a Petri dish, especially if the number of cells in the test portion is low : if there are only 5 cells in a 100 ml solution and if 1 ml is poured onto a dish, the probability of finding at least one cell is $1 - 0.99^5 = 0.049$. In this paper, specificity and sensitivity are supposed to be equal to 1 at the test portion level for detection analysis and at the fraction level for enumeration analysis. In other words, for detection (enumeration) analysis the probability of finding *L. monocytogenes* in the test portion (fraction) when this pathogen is really present and the probability of not finding *L. monocytogenes* in the test portion when it is indeed absent are equal to one.

In food microbiology, there is usually one analysis per test portion and several test portions are taken from the same batch and/or the same unit. This data structure and the way food is processed lead us to propose several models to represent different possible contamination-modeling structures.

For both application studies, data were obtained for different batches (12 for the pork breast study and 29 for the CSS). Several units per batch were taken as well as one test portion per unit (for the pork breast application) and two test portions per unit (for the CSS case). In these data sets, there are huge proportions of absence and of 0 count results.

3.2.2.b Models

General model Models considered in this paper are designed by assembling two independent submodels : a process submodel and an observation submodel. The process submodel refers to the structure of the contamination of a pathogen in food in the plant. The observation submodel refers to how results of the analyses are observed at the test portion and the fraction levels. The distribution of contamination at a certain level given the higher level can be described by a probability density function. In the next equations, p_n , $n = 1, \dots, 5$ are some of the parameters of probability distributions $\mathcal{L}_1, \dots, \mathcal{L}_5$. The subscript of p does not refer to the level it belongs (batch, unit, test portion or fraction) but is arbitrary.

First, let us consider the observation submodel, i.e., the likelihood of the data. As mentioned previously, an enumeration result belongs to a fraction l included in a test portion k included in a unit j included in a batch i . It is denoted y_{ijkl} . A detection result belongs to a test portion k so it is noted x_{ijk} . Both results depend on the concentration c_{ijk} of the test portion. Other parameters can also be taken into account (noted p_1 for y_{ijkl} and p_2

for x_{ijk}). These results follow the distributions respectively :

$$y_{ijkl}|c_{ijk}, p_1 \sim \mathcal{L}_1(c_{ijk}, p_1) \quad (3.1)$$

$$x_{ijk}|c_{ijk}, p_2 \sim \mathcal{L}_2(c_{ijk}, p_2) \quad (3.2)$$

For a single unit j belonging to batch i , one might perform either (i) only detection, (ii) only enumeration, (iii) both enumeration and detection on the same test portion, (iv) detection and enumeration on different test portions taken from the same unit. In case (iv), it means that, for a given test portion k , either y_{ijkl} or x_{ijk} is equal to \emptyset . The concentration of the test portion k depends on θ_{ij} , the vector characterizing the concentration of the unit j and on parameter p_3 :

$$c_{ijk}|\theta_{ij}, p_3 \sim \mathcal{L}_3(\theta_{ij}, p_3) \quad (3.3)$$

A graph of this submodel is represented by the two inside boxes in Figure 3.3. Its inputs are the outputs of the process submodel.

The process submodel considers two levels : the unit and the batch. The unit level, θ_{ij} depends on z_i the vector characterizing the concentration at the batch level and on a parameter p_4 :

$$\theta_{ij}|z_i, p_4 \sim \mathcal{L}_4(z_i, p_4) \quad (3.4)$$

Finally, z_i depends on a parameter p_5 :

$$z_i|p_5 \sim \mathcal{L}_5(p_5) \quad (3.5)$$

The hierarchical structure is presented in Figure 3.3. Of course, it is only worth considering the relevant levels : one or more levels can perhaps collapse from the model depending on circumstances ; the most complex model is not always the best one (van Schothorst et al. (2009)). The next two sections specify different types of process and observation probability distributions currently encountered for the model.

Observation submodel In this section, two different possible parametric families of distributions, denoted A and B, are considered for modeling observed contamination through detection and enumeration. A assumes a Log-Normal distribution while B assumes a Poisson distribution at the test portion level. All the distributions used are shown in Table 3.1.

Observation submodel A The logarithm of the concentration c_s of a test portion k is assumed to follow a normal distribution $[c_s|\phi, p_3 = \rho] = \mathcal{N}(\phi, \rho)$ (distribution functions are noted by squared brackets), where s is the generic subscript and ϕ represents the mean concentration in the level above the test portion (see Table 3.2 for a definition of s and ϕ depending on the process submodel used).

Subscript k refers to the test portion itself while subscript s also mentions the affiliation of the test portion. For instance, if $s = ijk$, that means that the test portion k belongs to unit j which itself belongs to batch i . Concerning detection and enumeration

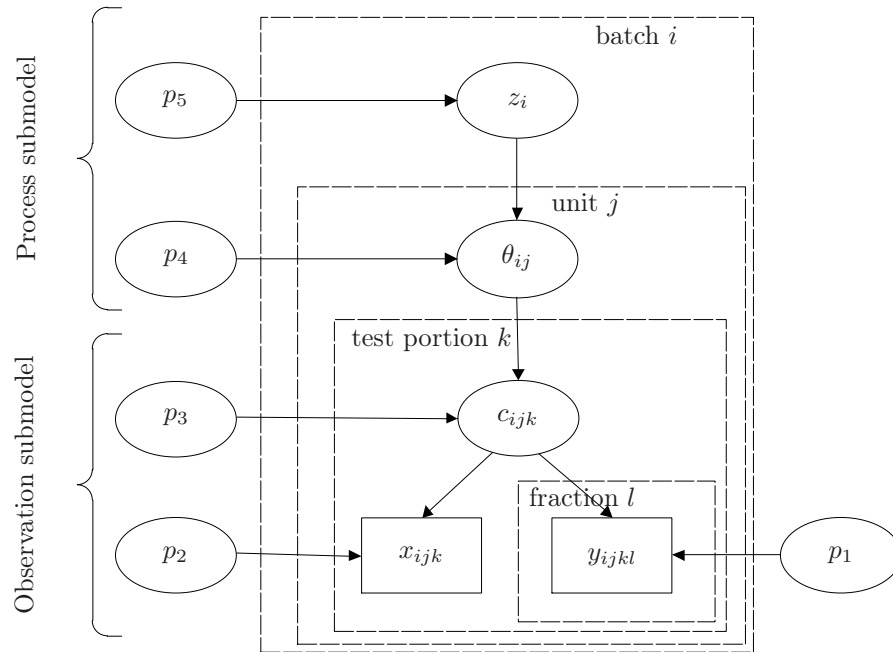


FIGURE 3.3 – Graph of the general model. Four hierarchical levels can be identified : batch, unit, test portion and fraction. The two upper levels are associated with the process submodel and the two lower levels with the observation submodel. The vectors characterizing concentration at the batch and at the unit levels are denoted z_i and θ_{ij} respectively. The concentration at the test portion level is c_{ijk} . An enumeration result (of a fraction) is noted y_{ijkl} and a detection result (at a test portion level) is noted x_{ijk} . The parameters used are noted p_n ($n = 1, \dots, 5$). Arrows indicate a probabilistic dependency from parents to descendants. For example, θ_{ij} has two parents (z_i and p_4) and K children (all the c_{ijk}). Conditional on its parent nodes, a node is independent of all the other nodes, except its own descendants. Nodes without parents are called parameters and nodes with parents are called latent variables ; both are in ellipses. Data are denoted by rectangles. Parameters, latent variables and data of this model are explained in part 3.2.2.b.

TABLE 3.1 – Distributions used in process and observation submodels to take variability into account for each of the eight models. To simplify notation, distributions are given without mentioning the conditioning. Types A and B models differ for distributions $\mathcal{L}1$, $\mathcal{L}2$ and $\mathcal{L}3$. In type A models, c is a concentration in a test portion whereas in type B models, it is a number of cells. Four levels are considered : batch i , unit j , test portion k and fraction l . M_k is the mass of the test portion k and d_l is the ratio between the volume of the fraction l pour-plated on Petri dishes and the volume of the first suspension, NA means “not applicable” and t.p. “test portion”. \mathcal{N} stands for a normal distribution, \mathcal{Ber} for a Bernoulli, \mathcal{P} for a Poisson and \mathcal{Bin} for a binomial distribution.

Models			Variability
General model	A-Reference	A-Batch	
\mathcal{L}_5	$z_i \sim \mathcal{N}(p_5 = (\mu, \sigma))$	$z_i \sim \mathcal{N}(p_5 = (\mu, \sigma))$	between batches
\mathcal{L}_4	$\theta_{ij} \sim \mathcal{N}(z_i, p_4 = \lambda)$	NA, $p_4 = \emptyset$	between units
\mathcal{L}_3	$c_{ijk} \sim \mathcal{N}(\theta_{ij}, p_3 = \rho)$	$c_{ik} \sim \mathcal{N}(z_i, p_3 = \rho)$	between t. p.
\mathcal{L}_2	$x_{ijk} \sim \mathcal{Ber}(1 - \exp(-10^{c_{ijk}} M_k))$	$x_{ik} \sim \mathcal{Ber}(1 - \exp(-10^{c_{ik}} M_k))$	
\mathcal{L}_1	$y_{ijkl} \sim \mathcal{P}(10^{c_{ijk}} M_k d_l)$	$y_{ikl} \sim \mathcal{P}(10^{c_{ik}} M_k d_l)$	
General model	A-Unit	A-Test portion	
\mathcal{L}_5	NA, $p_5 = \emptyset$	NA, $p_5 = \emptyset$	between batches
\mathcal{L}_4	$\theta_j \sim \mathcal{N}(\mu, p_4 = \lambda)$	NA, $p_4 = \emptyset$	between units
\mathcal{L}_3	$c_{jk} \sim \mathcal{N}(\theta_j, p_3 = \rho)$	$c_k \sim \mathcal{N}(\mu, p_3 = \rho)$	between t. p.
\mathcal{L}_2	$x_{jk} \sim \mathcal{Ber}(1 - \exp(-10^{c_{jk}} M_k))$	$x_k \sim \mathcal{Ber}(1 - \exp(-10^{c_k} M_k))$	
\mathcal{L}_1	$y_{jkl} \sim \mathcal{P}(10^{c_{jk}} M_k d_l)$	$y_{kl} \sim \mathcal{P}(10^{c_k} M_k d_l)$	
General model	B-Reference	B-Batch	
\mathcal{L}_5	$z_i \sim \mathcal{N}(p_5 = (\mu, \sigma))$	$z_i \sim \mathcal{N}(p_5 = (\mu, \sigma))$	between batches
\mathcal{L}_4	$\theta_{ij} \sim \mathcal{N}(z_i, p_4 = \lambda)$	NA ($p_4 = \emptyset$)	between units
\mathcal{L}_3	$c_{ijk} \sim \mathcal{P}(10^{\theta_{ij}} M_k)$	$c_{ik} \sim \mathcal{P}(10^{z_i} M_k)$	between t. p.
\mathcal{L}_2	$x_{ijk} \sim \mathcal{Ber}(1 - \exp(-10^{\theta_{ij}} M_k))$	$x_{ik} \sim \mathcal{Ber}(1 - \exp(-10^{z_i} M_k))$	
\mathcal{L}_1	$y_{ijkl} \sim \mathcal{Bin}(c_{ijk}, d_l)$	$y_{ikl} \sim \mathcal{Bin}(c_{ik}, d_l)$	
General model	B-Unit	B-Test portion	
\mathcal{L}_5	NA	NA	between batches
\mathcal{L}_4	$\theta_j \sim \mathcal{N}(\mu, \lambda^2)$	NA	between units
\mathcal{L}_3	$c_{jk} \sim \mathcal{P}(10^{\theta_j} M_k)$	$c_k \sim \mathcal{P}(10^\mu M_k)$	between t. p.
\mathcal{L}_2	$x_{jk} \sim \mathcal{Ber}(1 - \exp(-10^{\theta_j} M_k))$	$x_k \sim \mathcal{Ber}(1 - \exp(-10^\mu M_k))$	
\mathcal{L}_1	$y_{jkl} \sim \mathcal{Bin}(c_{jk}, d_l)$	$y_{kl} \sim \mathcal{Bin}(c_k, d_l)$	

TABLE 3.2 – Subscript for the number c of cells in a test portion and definition of ϕ used in the observation submodels depending on the process submodel used. The subscript s mentions the affiliation of the test portion. For instance, if only the batch structure is interesting, it is necessary to know theta test portion k belongs to batch i , so $s = ik$.

Process submodel	Subscript s for c	ϕ
Reference	ijk	θ_{ij}
Batch	ik	z_i
Unit	jk	θ_j
Test portion	k	μ

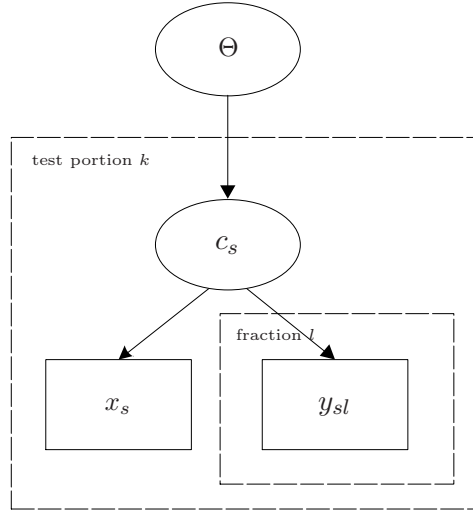


FIGURE 3.4 – Graph of the observation submodels. For the A observation submodel, $\Theta = (\phi, \rho)$ where ρ represents the between-test portion variability, and for the B observation submodel, $\Theta = \phi$. The definition of ϕ depends on the process submodel used (see Table 3.2). An enumeration result of a fraction l is noted y_{sl} , a detection result is noted x_s , c_s is the logarithm of the concentration of pathogens (for the A observation submodel) or a number of cells (for the B observation submodel) in a test portion. See Table 3.2 for a definition of s .

analysis, a Bernoulli $[x_s|c_s] = \mathcal{B}er(1 - \exp(-10^{c_s} M_k))$ and a Poisson distribution : $[y_{sl}|c_s] = \mathcal{P}(10^{c_s} M_k d_l)$ respectively are used. The mass M_k of the test portion k is independent from the batch or the unit to which it belongs, so the batch or the unit affiliation in the subscript are not mentioned, so subscript k is written and not subscript s . The ratio between the fraction l of the volume pour-plated onto a Petri dish and the first suspension in the enumeration analysis is denoted d_l . This ratio depends only on the fraction and not on the batch, unit or test portion. Here, $p_1 = p_2 = \emptyset$.

Observation submodel B Observation submodel B differs from submodel A at the test portion level. Assuming that the distribution of the cells in a unit follows a Poisson distribution, the probability of having at least one CFU in a sample of M_k g is equal to $1 - \exp(-10^\phi M_k)$. The detection result x_s follows a Bernoulli distribution : $[x_s|\phi] = \mathcal{B}er(1 - \exp(-10^\phi M_k))$ and $p_2 = \emptyset$ (see Table 3.2 for definitions of s and ϕ). Let c_s be the number of cells in a test portion k . Assuming that the number of cells follows a Poisson distribution, c_s is modelled as a Poisson distribution $[c_s|\phi] = \mathcal{P}(10^\phi M_k)$, ($p_3 = \emptyset$). The number y_{sl} of cells in a fraction l is a fraction of the cells contained in the test portion, so the result is a Binomial distribution : $[y_{sl}|c_s, d_l = p_1] = \mathcal{B}in(c_s, d_l)$. It is easy to get $[y_{sl}|\phi] = \mathcal{P}(10^\phi M_k d_l)$ by summing on c_s (see Appendix A). Graphs of observation submodels can be seen in Figure 3.4.

In fact, after writing the distributions of x_s and y_{sl} in both observation submodels,

the conclusion is that the observation submodel B is nested in observation submodel A, as observation submodel B is equal to observation submodel A when $\rho = 0$.

Process submodels Four variants of the process submodels can be described. Table 3.1 presents all the distributions used in these submodels.

Reference submodel In the “reference” variant of the process submodel presented in this paper, between-unit and between-batch variability is modeled. Here, z_i is the logarithm of the mean concentration of product in the batch i and θ_{ij} is the logarithm of the mean concentration of product in the unit j . The corresponding graph of this submodel is represented in Figure 3.5a. The parameters μ , σ and λ are supposed to be uncertain and are described by a probability distribution.

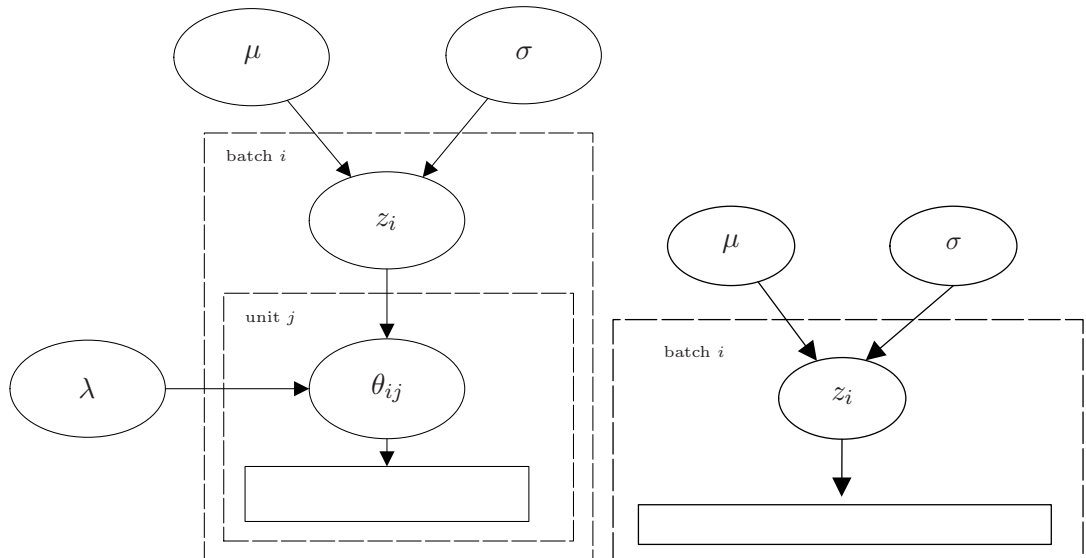
The next three variants presented below are nested in the reference variant.

Batch submodel For the “batch” variant, there is no “unit effect” : two test portions from the same unit are no more similar than two test portions from different units, so only between batch variability is considered. The graph of this model is presented in Figure 3.5b.

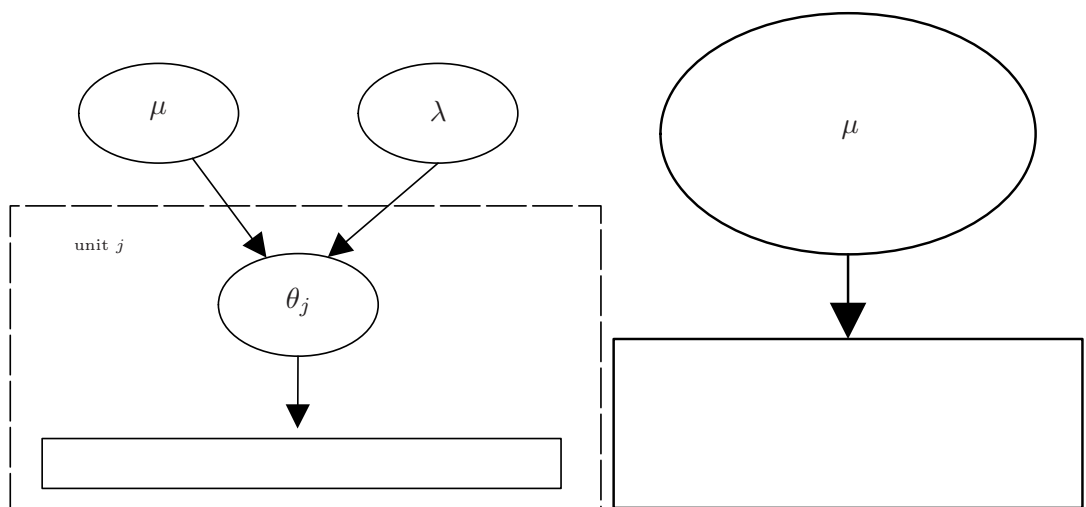
Unit submodel Here, it is assumed that there is no “batch effect”, that is to say that two units from different batches are just as similar as two units from the same batch. However, there is a “unit effect”. Figure 3.5c shows the graph of the model. This submodel only takes into account between units variability. This submodel can be used when the batch level is not relevant from a microbiological point of view, for example, when accidental contamination can occur on one unit without contaminating the others. It has as many latent variables as units, whereas the batch process submodel has as many latent variables as batches.

Test portion submodel Finally, the simplest model is the one that states that all the differences observed in data are only due to the Poisson dispersion of the cells within the considered food : there is variability only between test portions (and neither batch nor unit variability). The only parameter is μ , the logarithm of the mean concentration in food. The graph of this model is shown in Figure 3.5d.

Assembling a process and an observation submodel A complete model is composed by hierarchically assembling a process submodel and an observation submodel. Its name is formed from the names of the two submodels from which it is made. For example, the model A-unit is composed of the unit process submodel and the A observation submodel. The graph of this model is shown in Figure 3.6. As there are four process submodels and two observation submodels, eight models are possible. All the models are nested in the A-reference submodel. The parameters used in the considered models are summarized in Table 3.3. Although they are different because they are model specific, the same names are kept whenever possible to help understand what they mean.



(a) Reference : both between and within batch variabilities (b) Batch : between batch variability only



(c) Unit : between units variability only

(d) Test portion : neither batch nor unit variability

FIGURE 3.5 – Graphs of the four process submodels : reference (a), batch (b), unit (c) and test portion (d). Each node is represented by an ellipse. The empty rectangle represents the observation submodel of the model described in Figure 3.4. For the unit, θ_{ij} is the mean log concentration, z_i is the mean log concentration of the batch, μ is the mean of the mean log concentrations between batches, σ represents the between batches variability, λ the between units variability.

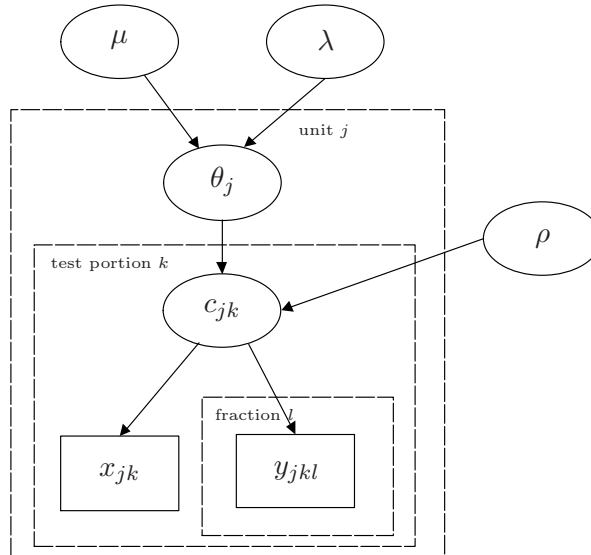


FIGURE 3.6 – Graph of the A-unit model. It is composed of the process submodel unit and the observation submodel A. The mean log concentration at the batch is denoted θ_j . The log concentration at the test portion level is c_{jk} , λ the between units variability and ρ the between test portions variability. An enumeration result (of a fraction) is noted y_{jkl} and a detection level (at a test portion level) is noted x_{jk} .

TABLE 3.3 – List of the parameters for the models presented. The common unit for all of them is $\log(cfu/g)$. Parameter μ is the concentration mean, σ the between-batch concentration standard deviation, λ is the within-batch concentration standard deviation and ρ is the within-unit concentration standard deviation. “X” means that the parameter is present in the model and “-” that it is absent.

Models	Parameters			
	μ	σ	λ	ρ
A-reference (Batch and unit effects)	X	X	X	X
A-batch (Batch effect only ($\lambda = 0$))	X	X	-	X
A-unit (Unit effect only ($\sigma = 0$))	X	-	X	X
A-test portions (No batch and no unit effect ($\lambda = \sigma = 0$))	X	-	-	X
B-reference (Batch and unit effects)	X	X	X	-
B-batch (Batch effect only ($\lambda = 0$))	X	X	-	-
B-unit (Unit effect only ($\sigma = 0$))	X	-	X	-
B-test portion (No batch and no unit effect ($\lambda = \sigma = 0$))	X	-	-	-

Superscripts are not mentioned to simplify notation. In fact, the concentration mean in the reference process submodel is the batch mean, and the concentration mean in the unit submodel is the mean between units, so parameter μ could have been written as μ^B and μ^U respectively. Table 3.1 shows how the different submodels described above deal with different variabilities.

3.2.2.c Prior setting

Setting priors is an important topic in Bayesian analysis. Priors reflect prior beliefs about the parameters before studying the data. It can happen that no previous information is available, so that a vague prior is chosen which means that the prior contains no or very little information about a parameter (Crépet et al. (2007)). Priors can also be chosen for convenience to facilitate the calculation of the posterior distribution (Brand and Small (1995)). In that case, a conjugate prior can be used, which means that the prior and the posterior distributions are in the same class (Berger (1985)). Priors can also be constructed to respect some model considerations, for instance permutation invariance constraints (Forster (2010)). When previous data, different from the ones used in a study, are available they can be used to build a prior distribution. The Bayesian approach can also include information obtained from expert knowledge but one should be very careful about the way questions are asked of the expert (Kunhert et al. (2010)). Methods to elicit multivariate prior distributions is also an active field of present research (Moala and O'Hagan (2010)). Fifteen other techniques to determine priors can be found in Berger (Berger (1985)).

In the two application studies previous available information including expertise was used to encode the prior probabilistic bets for the parameters of the general model only (i.e. the model in which all the other models are nested, see parts 3.2.3.a and 3.2.3.b for application). However, for every nested model, we rely on the so-called compatible priors to balance the a priori chances among them to replicate the observed data correctly from predictive prior distribution.

Compatible priors For each case study, all the models are compared with a numeric criterion (to assess the contamination structure) based on predictive posterior approach (see section 3.2.2.d). The predictive approach is highly prior sensitive, as is the case for Bayes factors (Kass and Raftery (1995)). Indeed, if priors of two models are not well calibrated, a model can be favored only because of the choice of its prior. When comparing two models, the result of the criterion should be least influenced by the two priors because of different constructions and should mainly represent the influence of the data for each model. To avoid this unfair effect, priors are chosen with great care so as to be so-called *compatible*. Several methods have been developed to construct compatible priors, such as projection (Dawid and Lauritzen (2000); Celeux et al. (2006)) and conditioning (Dawid and Lauritzen (2000); Roverato and Consonni (2004)). In these theoretical articles, the priors of the nested model can be calculated analytically. To bring the models into balance before adding data, one of the ideas is to give such priors to nested models so that the predictive prior distributions of the data simulated with the nested models are as close as possible to the distribution of the data simulated with the general model. Sometimes, it

is possible to find a formula defining the priors of the nested models. For example, Celeux et al. (2006) find a way to calculate priors of the nested regression models using the priors of the general linear model by minimizing the Kullback-Leibler distance between the predictive distributions of both models. Unfortunately, with the models presented here, no compatible prior could be derived in closed form. The strategy chosen here is to minimize a “compatibility score”, so that predictive prior distributions for all models be as similar as possible. This score is composed of two components, one for each type of data : detection or enumeration. The prior of a nested model is adapted so that the score is as low as possible.

A detection result can take one of two values : 0 (absence) or 1 (presence). Detection data are first simulated from the general model and from a second model (one of the nested models). The first component used is the difference between the proportion of “absence” in both models :

$$D = |P(x_G = 0) - P(x_N = 0)|,$$

where $P(x_G = 0)$ is the proportion of result “absence” in the data set replicated from the general model and $P(x_N = 0)$ the proportion of result “absence” in the data set replicated from the nested model. This is equivalent to calculating the difference between the proportions of “presence” in both models.

For enumeration, data were first simulated from the general model and from the model whose prior was to be tuned. The 5th, 10th, ..., 95th quantiles were then calculated for each replicated data set. Let $q_{G5}, q_{G10}, \dots, q_{G95}$ be the 5th, 10th, ..., 95th quantiles of simulated enumeration data from the general model and $q_{N5}, q_{N10}, \dots, q_{N95}$ be the 5th, 10th, ..., 95th quantiles of enumeration simulated data from the nested model.

To quantify the proximity between cumulative functions, the following distance component was designed :

$$E = \frac{|\sum_i (\ln(q_{Gi} + 1) - \ln(q_{Ni} + 1))|}{\sum_i \ln(q_{Gi} + 1)}.$$

The greater the similarity between the predictive distributions of both models, the lower the values of E and D . If distributions are equal, $E = 0$ and $D = 0$.

The enumeration and detection distributions from a nested model should be as close as possible to the distributions from the general model. The score defined is a compromise between the two components :

$$TD = \gamma E + (1 - \gamma)D, \tag{3.6}$$

where γ is a number between 0 and 1 and represents the preference for one predictive criterion. Here, $\gamma = 0.5$ was merely chosen.

The minimum of TD was calculated by brute force search by varying the prior parameters using the “optimize” function of the software R (R Development Core Team (2011)).

3.2.2.d Comparison

In this section, it is proposed to base the choice of one model on the data. For each datum item it is known from which batch, unit and test portion it comes. Thanks to the Bayesian approach, these data were used to update prior knowledge resulting from expert

assessment to get posterior distributions of the parameters.

To compare models, a criterion must be defined. As sampling plans are often carried out at the batch level, the ability of the different models to replicate data at this level is estimated at this level. Below, one criterion of interest and a “visual inspection” of the posterior distributions are described.

Predictive posterior loss approach Once posterior distributions are obtained for all the parameters, data can be replicated using these distributions and the likelihoods of the models. A criterion, based on simulated data at the batch level, is defined so that if observed and simulated data are close to each other, the criterion is close to zero. For each model, the criterion chosen assesses how close replicated data y_{rep} are from observed data y_{obs} . To do so, following the reasoning of Gelfand and Ghosh (1998) (see Appendix B), the criterion is equal to :

$$C(M) = \sum_{l=1}^L C_l(M) \quad (3.7)$$

$$C_l(M) = \sigma_l^{2(M)} + \frac{k}{k+1} \left(\sum_{n=1}^{N_l} y_{l,n,obs} - \mu_l^{(M)} \right)^2, \quad (3.8)$$

where L is the number of batches, $y_{l,n,obs}$ (resp. $y_{l,n,rep}$) is the n^{th} observed (resp. replicated) datum of batch l , k indicates the relative regret for departure from $\sum_n y_{l,n,obs}$ compared with departure from $\sum_n y_{l,n,rep}$, N_l is the number of observations for the batch l , $\sigma_l^{2(M)} = Var(\sum_n y_{l,n,rep} | \sum_n y_{l,n,obs}, M)$ and $\mu_l^{(M)} = E(\sum_n y_{l,n,rep} | \sum_n y_{l,n,obs}, M)$. Here, subscripts used for the batch are different from the ones used in the presentation of the models to underline that here a datum belongs to a batch, even if this level is not considered in a model. Each criterion $C_l(M)$ is composed of two parts, the first assesses the variance of the sum of replicated data for the batch l and the second assesses its squared bias. The criterion $C(M)$ is similar to the mean square error, a criterion used in the frequentist approach, but here data are summed at the batch level. Here, the chosen value of k is high so that an equal weight is given to the variance and the squared bias.

Visual inspection Another way to pool data at the batch level is to work by direct inspection of posterior predictive replications. For each model, data using parameters sampled from their posterior distributions are simulated. For each value of these parameters :

- $n_i \times n_j \times n_k$ of detection data and $n_i \times n_j \times n_k$ of enumeration data are generated, where n_i is the number of batches, n_j the number of units per batch and n_k the number of test portions per unit in the observed data ;
- simulated data are grouped together per batch ;
- for detection, the number of batches having only positive data, only negative data and the remaining batches (the ones which have positive and negative data) were retained ;
- for enumeration, the number of batches having only data equal to 0, only data different from 0 and the remaining batches were retained.

TABLE 3.4 – Raw data in the pork breast case study. Among the 12 batches chosen, 8 ou 9 pork breasts were randomly taken and one test portion of 100 cm² from each breast was analysed for both detection and enumeration. For each batch, the nth result of detection and the nth result of enumeration belong to the same test portion. For detection, “0” means absence and “1” means presence.

Batch number	Detection	Enumeration (cfu)	Nb of pork breasts per batch
1	0-1-1-1-0-1-1-1-1	0-0-0-0-0-0-0-0-0	9
2; 8 to 10	0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	9
3	0-0-0-0-1-0-0-0-1	0-0-0-0-1-0-0-0-0	9
4	0-1-0-0-1-0-0-0-0	0-0-0-0-0-0-0-0-0	9
5	0-0-0-0-1-0-0-0-0	0-0-0-0-0-0-0-0-0	9
6	0-0-0-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	8
7	0-0-0-1-1-0-0-0-1	0-0-0-0-0-0-0-0-0	9
11	1-1-1-1-1-1-1-1-1	11-9-6-5-12-29-16-3	8
12	0-0-1-0-0-0-0-0-0	0-0-0-0-0-0-0-0-0	9

Finally, the mean number of batches in each group is calculated, as well as the 50% and the 95% credibility intervals. For comparison with the observed data, these results were plotted, following the seminal work of Gelman (2007), which recommends posterior predictive checks. This visual inspection helps to assess the capacity of the model to simulate the batch structure correctly.

3.2.3 Applications

3.2.3.a Pork breasts

Data In this case study, a batch is defined as the set of pork breasts tumbled in the same tumbler at the same time, a unit as a pork breast and a test portion as 100 cm² of meat excised from the pork breast. After arrival, pork breasts are put into a tumbler, a device in which hundreds of pork breasts and brine are mixed for several hours. The samples were taken from the same French diced bacon plant between May and June 2009. From among 12 batches, 106 units were selected (8 or 9 units per batch). For each of them, detection and enumeration of the pathogen *L. monocytogenes* were performed on the same test portion of 100 cm². Here the weight M_k is replaced by the surface $S_k = 100$ cm². To comply with the protocol, $d_i = 0.05$. There are 25 positive results in detection and 9 positive results in enumeration out of 106 (Table 3.4).

In this case study, there is only one test portion per unit and one test portion for two analyses (detection and enumeration) so it is impossible to estimate within-unit variability. Therefore, the observation submodel A cannot be applied and the B-reference model is the most general one (the three remaining models are nested in this one).

Priors Prior distributions of parameters for the B-reference model (the general one) were determined using simulation data obtained from the model of the process according

TABLE 3.5 – Informative prior distributions for the B-reference model (obtained from the model of the process according to Billoir et al.⁽³⁴⁾) and compatible prior distributions for the three nested models for the pork breast study case. Descriptive statistics of posterior distributions are also given. “dist.” stands for “distribution”, “perc.” stands for “percentile” and “S.D.” for “standard deviation”. \mathcal{N} stands for a normal distribution and \mathcal{G} for a gamma distribution.

Model	Parameter	Prior dist.	Descriptive statistics of the posterior dist.				
			Mean	S.D.	2.5th perc.	50th perc.	97.5th perc.
B-reference	μ	$\mathcal{N}(-3.03, 0.5^2)$	-3.09	0.53	-4.25	-3.05	-2.15
	σ	$\sigma^{-2} \sim \mathcal{G}(0.5, 0.5)$	1.55	0.49	0.89	1.45	2.77
	λ	$\lambda^{-2} \sim \mathcal{G}(3.81, 0.56)$	0.38	0.08	0.25	0.36	0.57
B-batch	μ	$\mathcal{N}(-3.44, 1.13^2)$	-3.12	0.51	-4.21	-3.09	-2.18
	σ	$\sigma^{-2} \sim \mathcal{G}(1.05, 3.90)$	1.72	0.47	1.06	1.63	2.86
B-unit	μ	$\mathcal{N}(-3.43, 0.17^2)$	-3.51	0.15	-3.81	-3.51	-3.21
	λ	$\lambda^{-2} \sim \mathcal{G}(1.32, 7.03)$	1.99	0.24	1.59	1.97	2.51
B-test portion	μ	$\mathcal{N}(-3.85, 3.42^2)$	-0.94	0.005	-0.95	-0.94	-0.93

to Billoir et al. (2011). The data used in the study of Billoir et al. (2011) are different from the ones used in this article. This process-based model produces the empirical distributions for contamination of *L. monocytogenes* after every step of the process, which is described in the article (Billoir et al. (2011), part 2.1). Using this model, simulated data were kept (after the tumbling step) for 1000 batches of 500 pork breasts so as to calibrate informative priors for the general model. Associated compatible priors were then calculated by minimizing the TD criterion (Equation (3.6)) for models B-batch, B-unit and B-test portion. Obtained distributions of priors are shown in Table 3.5.

Bayesian inference Bayesian inference on parameters was performed using the software OpenBugs (Thomas et al (2006)). For each model, after a burn-in of 100,000 iterations, samples of the posterior distributions were obtained from iterations so as to get an efficient sample size of 10,000. Because the autocorrelation in the parameter values was different for each model, the space interval was adapted for every model. For example, the sample of 10,000 posterior values for parameter μ in model B-test portion was built by taking one value out of 10 from the posterior chain given by the software, so in that case, 100,000 iterations were performed. Convergence of the Markov chain Monte Carlo algorithm was assessed using three parallel chains starting at different initial values for each parameter and applying the Gelman and Rubin diagnosis, as modified by Brooks and Gelman (1998). The OpenBugs code is available from the corresponding author upon request.

Results Quantiles of the posterior distributions are shown in Table 3.5.

The posterior distribution of μ belonging to the B-test portion model is unrealistically very narrow, because this model has only one parameter. Given all the data, and assuming such a parsimonious model, there is little posterior uncertainty for just one parameter. Except for the B-test portion model, the 95% credibility interval of μ remains between

-4.2 and -2.2. However, μ has a different meaning for each model so the different values of μ are not directly comparable.

To get an idea of the posterior concentration, posterior densities of θ_1 and θ_{91} (defined in equation (3.4) for the B-reference and B-unit models can be seen in Figure 3.7. Pork breast 1 belongs to a batch for which the data are similar to most other batches (7 presences and no colony counted in enumeration). Pork breast 91 belongs to the most contaminated sampled batch : all the detection results are “presence” and the enumeration results range from 5 to 29 CFUs. In pork breast 1, the results for enumeration and detection were 0 and “absence” respectively. In pork breast 91, the results for enumeration and detection were 9 CFUs and “presence” respectively.

For pork breast 1, posterior distribution of θ_1 in the B-reference model is narrower than the distribution in the B-unit model. This is due to the fact that in the B-reference model, θ_1 depends on the mean concentration z_1 of batch 1 whereas θ_1 in the B-unit model does not. As there are 7 presences in batch 1, θ_1 in the B-reference model is influenced by all these presence results in batch 1, so its distribution is more precise than the distribution of θ_1 in the B-unit model. For pork breast 91, posterior distributions of θ_{91} are almost the same in both models. Concentration θ_{91} in the B-reference model is influenced by the mean concentration of the batch but only slightly because the enumeration result for batch 91 is equal to the mean of enumeration results of that batch. Concentration θ_{91} in the B-unit model is not influenced by the results for other pork breasts from the same batch, so distributions of θ_{91} in both models are similar.

Figure 3.8 shows the results of visual inspection. The B-batch model is the model which best replicates the data : all the observed data are within the 50% credibility interval of the replicated data, except for the batches for which detection results are “absence” or “presence” (column “Other”). The B-reference model is less successful at replicating detection data but it is still not far from the B-batch model (data not shown). The models without a batch structure (B-unit and B-test portion) cannot replicate data correctly, especially the B-test portion model. These results clearly indicate that a batch structure exists in data for the pork breasts case study. This result is consistent with the fact that pork breasts from the same batch are tumbled together, thus swapping bacteria during the tumbling step.

The results of the expected posterior predictive loss approach with grouped data (Equation (9.2)) are shown in Table 3.6. Both the visual inspection and the numeric criterion leads to the same conclusion for the B-test portion model. The only difference between the conclusions for the two criteria is that the B-unit model is close to B-batch and B-reference models for the posterior predictive loss approach. The B-unit model replicates the number of colonies observed in the test portions well (in that case criterion C is low) but not the number of batches in which not all the test portions have *L. monocytogenes* (observed data are always outside the 95% credibility interval of the replicated data).

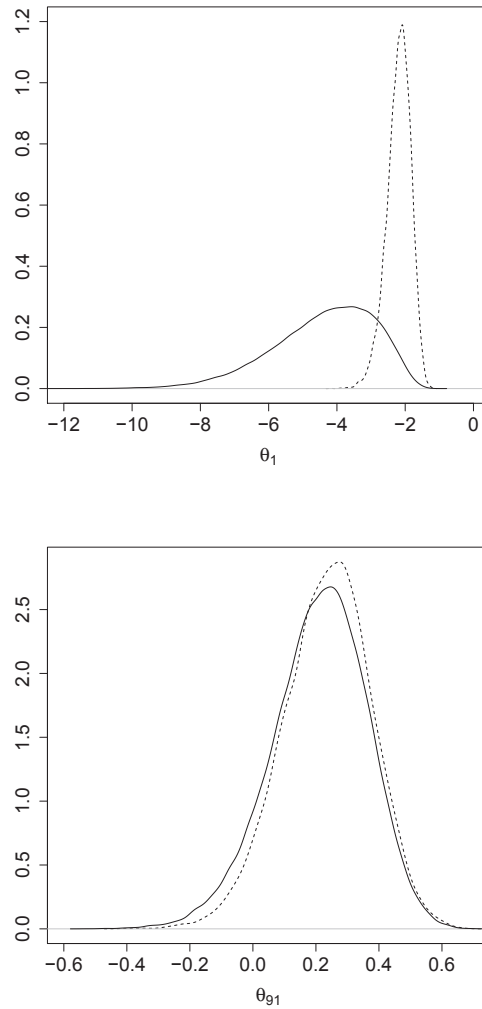


FIGURE 3.7 – Examples of posterior distributions of the mean concentrations θ_1 (left) and θ_{91} (right) in pork breasts 1 and 91 respectively, using two models described in parts 3.2.2.b and 3.2.2.b : the B-unit (plain line) and the B-reference (dotted line) models. The unit of θ is $\log(\text{CFU}/\text{cm}^2)$.

TABLE 3.6 – Results of the posterior predictive loss approach (Equation (9.2)) for the pork breast case study.

Model M	C(M)
B-reference	216
B-batch	204
B-unit	209
B-test portion	8443

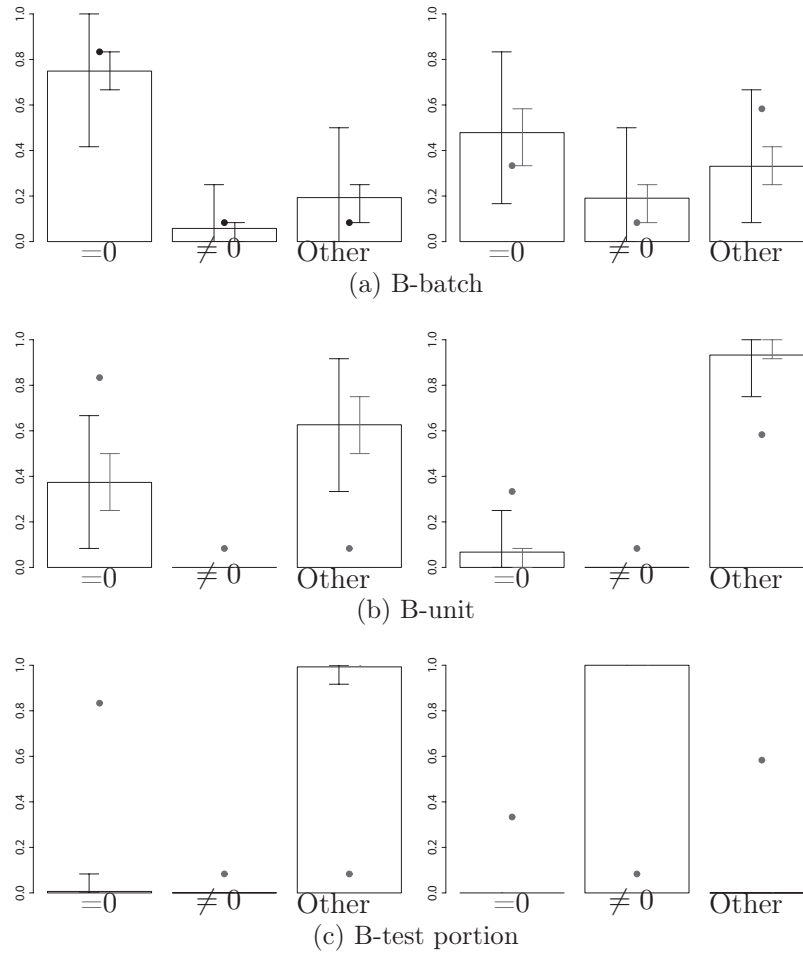


FIGURE 3.8 – Pork breast case study : simulated and observed data for three type B models : B-batch (a), B-unit (b) and B-test portion (c). Criteria for enumeration are illustrated on the left whereas detection data are on the right.

The bars represent the mean of each group (“= 0” : proportion of batches with only negative data, “≠ 0” : proportion of batches with only positive data and “Other” : all the other batches). The segments on the right give the 50% credibility intervals and the segments on the left give the 95% credibility intervals. The mean and the credibility intervals were calculated for replicated data. The points are the observed data.

TABLE 3.7 – Raw data in the CSS case study. Among the 29 batches chosen, 5 or 6 packaging units were randomly taken and two test portions of 25 g from each slice were analysed, one for detection and one for enumeration. For each batch, the n^{th} result of detection and the n^{th} result of enumeration belong to the same packaging unit. For detection, “0” means absence and “1” means presence.

Batch number	Detection	Enumeration (cfu)	Nb of packa- ging units per batch
1 to 6, 8 to 13;16;20;23;25;27 to 29	0-0-0-0-0-0	0-0-0-0-0-0	6
7	1-1-1-1-1-1	1-0-0-0-0-0	6
14	1-1-0-0-0-0	0-0-0-0-0-0	6
15, 17	1-0-0-0-0-0	0-0-0-0-0-0	6
18	1-1-1-0-0	3-2-0-0-0	5
19	1-1-1-0-0-0	0-0-0-0-0-0	6
21;24;26	1-1-0-0-0-0	1-0-0-0-0-0	6
22	0-0-0-0-0	0-0-0-0-0	5

3.2.3.b Cold smoked salmon

Data The data set is described in Table 3.7 and was taken from Beaufort et al. (2007). It consists of 344 results of enumeration and detection of the pathogen *Listeria monocytogenes* in cold smoked salmon. Here, a unit was considered to be a packaging unit and a batch corresponded to the commercial batch : all the products were packaged the same day and with the same raw salmon origin. The samples were taken from the same plant between May 2003 and January 2004 (Beaufort et al. (2007), plant 8). In all, 172 units were selected from 29 batches (5 or 6 units per batch). For each of them, detection was performed on 25g and enumeration was also carried out on 25g, so $M_k = 25$ g. The enumeration protocol described by Gnanou Besse et al. (2004) was applied, so $d_l = 0.2$. There are 22 positive results in detection out of 172 and 6 positive results in enumeration out of 172.

Priors Unfortunately, little information was available to specify prior distributions for the A-reference model : only monthly prevalence results were available from the same plant in 2003 and 2004 carried out on 25 g test portions. Assuming a Poisson distribution of *L. monocytogenes* in CSS, the relationship between prevalence and concentration is :

$$p = 1 - \exp(-25c),$$

where p is the prevalence and c the concentration in a one-month production period. Concentration was calculated from prevalence and converted into a logarithm before adjusting for normal distribution. The latter was used for the prior of μ : $\mu \sim \mathcal{N}(-2.4, 0.85^2)$. All the priors concerning precision parameters (i.e., 1/variance) were vague : a Gamma distribution was set with shape and rate equal to 0.5. The “TD-compatible” priors (equation (3.6)) for the other models are shown in Table 3.8.

TABLE 3.8 – Prior distributions for the A-reference model and associated compatible prior distributions for the seven nested models in the A-reference model for the CSS study case. Descriptive statistics of posterior distributions of parameters for all the models. “dist.” stands for “distribution”, “perc.” stands for “percentile” and “S.D.” for “standard deviation”. \mathcal{N} stands for a normal distribution and \mathcal{G} for a gamma distribution.

Model	Parameter	Prior dist.	Descriptive statistics of posterior dist.				
			Mean	S.D.	2.5th perc.	50th perc.	97.5th perc.
A-reference	μ	$\mathcal{N}(-2.4, 0.85^2)$	-3.61	0.42	-4.52	-3.59	-2.87
	σ	$\sigma^{-2} \sim \mathcal{G}(0.5, 0.5)$	1.51	0.44	0.85	1.44	2.56
	λ	$\lambda^{-2} \sim \mathcal{G}(0.5, 0.5)$	0.81	0.22	0.44	1.80	1.30
	ρ	$\rho^{-2} \sim \mathcal{G}(0.5, 0.5)$	0.56	0.16	0.33	0.54	0.93
A-batch	μ	$\mathcal{N}(-3.6, 5.5^2)$	-4.75	0.95	-7.00	-4.57	-3.40
	σ	$\sigma^{-2} \sim \mathcal{G}(0.6, 0.1)$	1.92	0.74	0.92	1.79	3.72
	ρ	$\rho^{-2} \sim \mathcal{G}(1.1, 7.6)$	1.34	0.30	0.90	1.28	2.05
A-unit	μ	$\mathcal{N}(-4.1, 6.1^2)$	-4.31	0.75	-6.14	-4.19	-3.25
	λ	$\lambda^{-2} \sim \mathcal{G}(0.6, 0.1)$	1.39	0.56	0.45	1.34	2.66
	ρ	$\rho^{-2} \sim \mathcal{G}(1.1, 7.4)$	1.35	0.33	0.89	1.29	2.18
A-test portion	μ	$\mathcal{N}(-4.3, 4.5^2)$	-4.01	0.68	-5.71	-3.88	-3.08
	ρ	$\rho^{-2} \sim \mathcal{G}(0.7, 8.2)$	1.78	0.48	1.12	1.68	2.95
B-reference	μ	$\mathcal{N}(-4.05, 5.5^2)$	-3.49	0.41	-4.37	-3.46	-2.77
	σ	$\sigma^{-2} \sim \mathcal{G}(0.4, 0.1)$	1.44	0.41	0.81	1.37	2.41
	λ	$\lambda^{-2} \sim \mathcal{G}(1.4, 6)$	0.89	0.20	0.55	0.87	1.32
B-batch	μ	$\mathcal{N}(-3.5, 5.5^2)$	-3.14	0.47	-4.27	-3.06	-2.47
	σ	$\sigma^{-2} \sim \mathcal{G}(0.4, 0.2)$	1.22	0.42	0.66	1.13	2.27
B-unit	μ	$\mathcal{N}(-2.12, 1.0^2)$	-3.08	0.29	-3.74	-3.05	-2.59
	λ	$\lambda^{-2} \sim \mathcal{G}(0.2, 0.1)$	1.09	0.22	0.71	1.07	1.58
B-test portion	μ	$\mathcal{N}(-4.1, 7.4^2)$	-2.20	0.08	-2.37	-2.20	-2.06

TABLE 3.9 – Results of the posterior predictive loss approach (Equation (9.2)) for the CSS case study.

Model M	C(M)
A-reference	351
A-batch	33
A-unit	38
A-test portion	39
B-reference	94
B-batch	46
B-unit	51
B-test portion	115

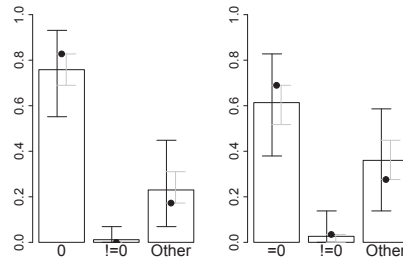
Bayesian inference Inference was performed as for the pork breast case, with the same software and with the same number of simulations of posterior distributions for each model.

Results Some quantiles of posterior distributions are shown in Table 3.8. As in the pork breast case, posterior distribution of μ in the B-test portion model is also unsurprisingly very narrow, for the same reason.

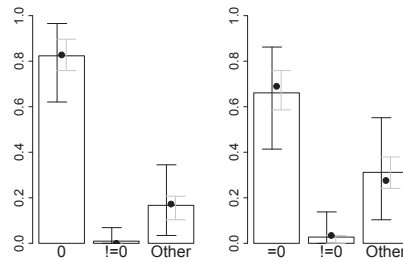
Figure 3.9 gives the posterior predictive visualisation that is useful for model validation. The results for models sharing the same process submodel are closely grouped. For instance, the results for the A-reference and the B-reference models are similar, while models with the B observation submodel give slightly less good results. In addition, the A and B-test portion models and the A and B-unit models have similar results. The closest fit between simulated results and observed data is obtained with the batch process submodel, i.e., with an aggregation at the batch level whichever observation submodel is used. However, and contrary to the previous case study, the observed data are within the 95% credibility intervals of the replicated data for all the models, including the simplest one.

Results with the posterior predictive loss approach (see Table 3.9) confirm that the aggregation at the batch level leads to the closest fit between simulated results and observed data, but that the aggregation at the unit level is not very different. Comparing A and B observation sub-models, this numerical criterion gives another perspective : it appears more accurate to choose the A observation submodel, i.e. modeling the logarithm of the concentration of a test portion as a normal distribution (Section 3.2.2.b). The best is A-batch, but the A-unit and A-test portion models are slightly less good. The B-batch and B-unit models are slightly worse than these models. The A-reference model is surprisingly bad compared to the visual inspection.

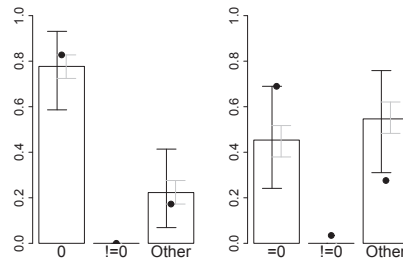
The results suggest that there is a batch effect for smoked salmon but this effect is less clear than for the pork breast case study. Since models with the test portion process submodel replicate the data reasonably well, these models can be chosen to replicate contamination data for CSS since they are simpler than those using the batch process submodel.



(a) A-Reference



(b) A-Batch



(c) A-test portion

FIGURE 3.9 – CSS case study : simulated and observed data for three type A models : A-reference (a), A-batch (b), and A-test portion (c). Enumeration data are illustrated on the left whereas detection data are on the right. The bars represent the mean of each group (“= 0” : proportion of batches with only negative data, “≠ 0” : proportion of batches with only positive data and “Other” : all the other batches). The segments on the right of each bar give the 50% credibility intervals and the segments on the left give the 95% credibility intervals. The mean and the credibility intervals were calculated for replicated data. The points are the observed data.

3.2.4 Discussion

In this study, the models were built according to the possible levels of aggregation (batch, unit, test portion) and taking into account the kind of data usually available to assess *L. monocytogenes* contamination (detection and enumeration). The aim was to determine whether the complexity of a model is justified according to the needs of an FBO, a risk assessor or manager and according to the observed data. There were not enough data in both case studies to assess the parameters without being prior-insensitive. As a consequence, attention was given to the prior construction (informative priors of the general models and compatible priors for the nested models).

3.2.4.a What is learnt for each case study

The ranking of models is different for the two cases, which supports the conclusion that the contamination structure is different for each process, product and plant. In the pork breast case study, a contamination structure clearly exists and is located at the batch level, while in the CSS application a structure also exists but is less marked. The approach proposed here offers a way to quantify what can be qualitatively observed from the data set (Table 3.4 and 3.7) : in the first application, all the batches but one have zero counts and one batch has no zero count and all the detecting results are presence, clearly suggesting at least a batch structure. In the second application, several batches have zero counts and absence results and the others have partly presence results and one or more 0 counts. No clear batch effect appears.

3.2.4.b This approach is applicable to other cases

The proposed modeling approach could be applied to other case studies provided that raw data and sufficient details about the sampling and analytical protocols (traceability about batches and units, mass or surface of the test portion, plated volumes) are available. The choice between models depends on each case study and the amount of data. Results should be checked again as soon as a major change occurs in the plant such as process modifications, new equipment, etc.

A hierarchical Bayesian model can be easily modified (1) by adding a node between the parameters or the latent variables and the data, (2) by adding or replacing a node at any level, (3) by adding an upper level.

What if there is an additional step to get the data? Each model can be modified by adding an additional step. Here, contamination at a step of a process is modeled without considering its source. Another way is to model the origin of the contamination, such as raw material contamination or cross-contamination (Billoir et al. (2011); Aziza et al. (2006)) The proposed approach directly uses colony counts and integrates details about the sampling and analysis protocols in the modeling. This protocol-based description is then easily transposable to any specific way of collecting data. For example a final analytical step of confirming the colonies by biochemical tests could be included (Clough et al. (2005)). This was not necessary in either of the case studies about *L. monocytogenes* since specific plating media was used on which all presumptive *L. monocytogenes*

colonies were indeed confirmed to be of this species. For case studies about *Escherichia coli* O157:H7 or *Staphylococcus aureus*, it might be useful (Clough et al. (2005)).

What if another type of data is added? Enumeration data are not available for all pathogens. For example, Salmonella concentration is estimated using the MPN method (MPN=most probable number, method based on serial dilutions). The method is roughly the following : several dilutions are made from the same first suspension. Usually, for each suspension, three tubes are prepared. The raw result is the number of tubes containing pathogens for each dilution. In that case, the observation submodel presented in this article can be changed to take MPN data into account. A raw result from a test portion k , with N_k dilutions would be the vector $y_{ijk} = (y_{ijk1}, \dots, y_{ijkN_k})$, where each y_{ijkl} is the number of positive tubes for dilution l among n_l tubes. Then, to link y_{ijk} to the pathogen concentration of the test portion, a binomial distribution could be used. As in the model presented here, the process submodels can all be used with MPN data.

What if one is interested in another scale? QMRA usually concerns the national or international scale. In such a context, it would be possible to use the models proposed in this article, by adding one or two upper layers. In that case, parameter μ becomes a latent variable depending on the plant : the distribution of μ represents between-plant variability. Consequently, this modeling approach can be used to assess distributions of bacterial concentrations in contexts of many different QMRAs. The methods of QMRA can also be applied at the plant scale (Aarnisalo et al. (2008)).

3.2.4.c The study is deeply grounded on priors

Working out informative priors In food microbiology, the number of data is usually too low to use models with many parameters. One of the advantages of the Bayesian approach is that it can partially counterbalance the lack of data by introducing information into the prior distributions of the parameters. These distributions should be carefully monitored with experts because they influence the posterior distributions. Prior distributions can be set using expert elicitation : it is quite easy to elicit a mean but much more complicated to elicit a variance (O'Hagan et al. (2006)). In the model presented, several variances needed to be elicited. In addition, these variances characterize dispersion between batches, between units and between test portions, which makes them more difficult to elicit. In the pork breast study, it was possible to build informative priors for each parameter because a computer model was available (Billoir et al. (2011)) to simulate *L. monocytogenes* concentration after tumbling. For the CSS case study, it was not possible to find informative priors for the variance parameters. Moreover, once priors are built for the general model, another difficulty arises because the calculation to get compatible prior distributions for all nested models using the prior distributions of the general model is not straightforward. Nonetheless a tractable way was found to get reasonable compatible priors, so some information was transferred in all our models through prior distributions of the general model.

The unavoidable sensitivity to priors As there were not many data in each application (212 in the pork breast case study and 344 in the CSS case study) and several parameters and latent variables in the models, it is crucial to have informative priors. If priors are changed, posteriors change as well and this change is more important for models with more parameters and latent variables. For instance, in the CSS application for the A-batch models with priors $\mu \sim \mathcal{N}(-2.4, 0.85)$, $\sigma^{-2} \sim \mathcal{G}(0.5; 0.5)$ and the same distribution for ρ^{-2} , the 95% credibility intervals of the posteriors are narrower than the ones presented in Table 3.8 : $I_{Cr}(\mu) = [-4.15, -2.69]$, $I_{Cr}(\sigma) = [0.69, 2.18]$ and $I_{Cr}(\rho) = [0.35, 0.94]$. As expected, the model B-test portion is not sensitive at all to the prior distribution of its single parameter. The A-reference model has many parameters and latent variables and few data are available to assess posterior distributions without influence of the priors. As a consequence, this model is highly sensitive to priors, which should be elicited with great care and considered as a very important part of the modeling process.

3.2.4.d Exploiting the approach for designing effective sampling plans

Food business operators (FBOs) are required to take measures (HACCP and good hygiene practices) to ensure that their foodstuffs comply with the relevant European microbiological criteria. In order to verify these procedures, they must perform microbiological measurements and decide the appropriate sampling frequencies (European Regulation (CE) No 2073/2005 (2005), Article 4).

Taking into account between and within-batch variabilities to design sampling plans To determine this frequency and the way sampling plans are made, it is necessary to take into account the ratio between the within-batch variability and the between-batch variability which has a large impact on the effectiveness of sampling (ILSI (2010)). If between-batch variability is by far higher than the within-batch variability, it is better to sample many batches but to take few or even one sample per batch. Conversely, if the within-batch batch variability is higher, it is difficult to distinguish between two batches. Samples from one batch also give information about pathogen concentration of other batches.

Qualitative assessment of between and within-batch variabilities by studying the process It is possible to get some information about within or between-batch variability by considering the heterogeneity of raw materials, the process and the environment. For instance, in the pork breast study (Billoir et al. (2011)), pork breasts from several farms arrive at the plant and are grouped together into one batch so one may think that the within-batch variability is high. The first step of the process is tumbling : pork breasts exchange the micro-organisms on their surfaces during several hours so after tumbling the within-batch variability is small. If a pork breast is highly contaminated, the pathogens will spread onto the other pork breasts and all the batches will thus be contaminated. On the contrary, if no pork breast is contaminated or contaminated at a low level, the contamination of the tumbled pork breasts will be low. As a consequence the between-batch variability can be high. For the CSS application, it is different because there is no step in which salmon or salmon slices can exchange their micro-organisms. *L. monocytogenes* is

present on the skin of the fish so contamination can occur if the upper part of the skins touches the flesh of the fish during scalding (first step of the process). Of course, cross-contamination can occur, for instance during slicing but it is not linked to a batch : a blade can contaminate salmon from many batches. In that case, between-batch variability seems to be small.

Using contamination models to simulate sampling plans Our proposal is to use the models proposed in this article to determine the best structure and to optimize further the size of the sampling plan. Once one of the models proposed in this article is chosen for a given case study on the basis of historical data, it could be used to simulate detection or enumeration data. These simulations could help to compare different sampling plans by varying the sample size and the decision rules. Thus, FBOs could have an idea about the efficacy of a sampling plan. Instead of trying all the possible sample sizes, an algorithm could be used to find the best sampling plan. Amzal et al. (2006) propose such an algorithm but it requires that a utility function be defined, which can be rather challenging.

Information given by a sampling plan depending on the contamination structure Indeed a routine sampling plan gives different information according to the structure of contamination. For instance, if B-batch is by far the most relevant model, a test portion gives a good idea about the batch to which it belongs but less accurate information about the other batches. For example, the posterior means of μ and σ in the pork breast application is -3.12 and 1.72 respectively. The mean log concentration z of a batch ranges from $-3.17 - 3 \times 1.72 = -8.3$ to $-3.17 + 3 \times 1.72 = -2.0$. Depending on the value of z , enumeration results are all equal to 0, partly equal to zero and partly equal to non-zero and there is no link between the contamination of one batch and the contamination of the next one. If the batches are not all sampled at least once, a high contamination can be missed. Conversely, if the simplest model (B-test portion) is good enough, microbiological measurements are directly interpretable at the plant level, to verify the food safety measures, rather than strictly at the batch level. This illustrates the predictive interest of studying the structure of the contamination.

Acknowledgments

We thank Marie-Laure Delignette-Muller for useful comments. This work was supported by a grant from the Agence National de la Recherche (ANR) (France), as a part of the Quant'HACCP project. We are grateful to the FBOs which answered our enquiry.

3.3 Retour sur la comparaison

Dans l'article, les modèles élaborés ont été sélectionnés à partir de deux critères, tous deux basés sur la prédiction *a posteriori*. Dans l'approche bayésienne, il est naturel d'utiliser les distributions prédictives pour critiquer le modèle à partir des données observées (Box (1980)). La comparaison des distributions prédictives *a posteriori* des différents

modèles est possible alors que celle entre les distributions *a posteriori* des paramètres ne l'est pas toujours (Gelfand and Ghosh (1998)).

3.3.1 Critère modifié de Gelfand and Ghosh (1998)

Le critère modifié de Gelfand and Ghosh (1998) est composé de deux termes : l'un est constitué par la variance de la somme des données simulées dans un lot et l'autre est le biais entre la somme des données simulées et la somme des données observées dans un lot.

3.3.2 Inspection visuelle

Cette inspection est intéressante pour visualiser rapidement l'adéquation des données simulées *a posteriori* aux données observées. Elle est construite de la manière suivante. Pour chaque modèle, des valeurs de paramètres sont simulées dans leur loi *a posteriori*. Pour chaque valeur, un jeu complet de données est simulé. Par exemple, pour le modèle B-reference dans le cas d'application "saumon fumé",

- pour chaque valeur du triplet (μ, σ, λ) , 29 valeurs de z sont simulées,
- pour chaque valeur de z , 5 ou 6 valeurs de θ sont simulées,
- pour chaque valeur de θ , une valeur de détection et une valeur de dénombrement sont simulées.

Pour chaque jeu de données simulées, je compte :

- pour les analyses de détection
 - le nombre de lot dont tous les résultats sont des absences ;
 - le nombre de lots dont tous les résultats sont des présences ;
 - le nombre de lots dont les résultats sont soit des présences, soit des absences ;
- pour les analyses de dénombrement
 - le nombre de lot pour lesquels aucune colonie n'est observée ;
 - le nombre de lots pour lesquels au moins une colonie est observée ;
 - le nombre de lots au moins un résultat est "0 colonie" et un résultat est "présence de colonie".

Comme ce processus est répété autant de fois que de simulation de paramètres, je peux calculer la médiane et les intervalles de crédibilité à 50% et à 95% des nombres de lots dans chaque catégorie. Ainsi, aux figures 7 et 8 de l'article reproduit ci-dessus, les histogrammes indiquent la médiane du nombre de lot dans chacun des six cas pour différents modèles. Les segments gris représentent les intervalles de crédibilité à 50%, les segments noirs les intervalles de crédibilité à 95% et les points représentent les données observées.

Cette inspection permet de visualiser rapidement la capacité des modèles à répliquer des lots avec des résultats positif, négatif ou un mélange des deux. Elle peut être utile pour les industriels mettant en place un plan d'échantillonnage à deux classes par exemple, pour les aider à évaluer le nombre de lots avec des résultats entièrement négatifs (qui passent l'inspection), le nombre de lots avec des résultats entièrement positifs (qui sont rejetés) et les autres lots (qui peuvent pour certains être acceptés si c est positif).

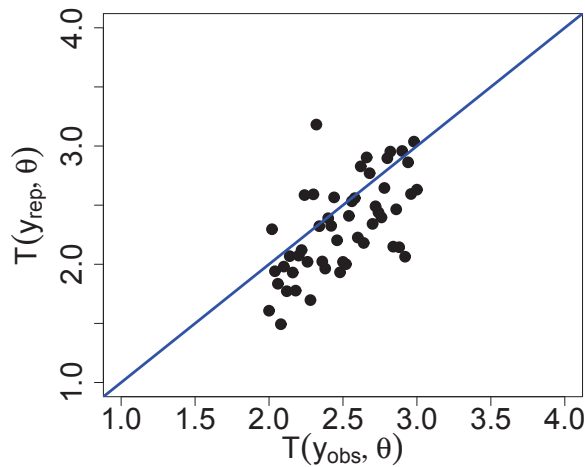


FIGURE 3.10 – Valeurs de $T(\theta, y_{rep})$ en fonction de $T(\theta, y_{obs})$. La valeur de la p-valeur est estimée en comptant le nombre de points au-dessus de la première bissectrice.

3.3.3 En utilisant la p-valeur

Il est également possible de recourir à un critère de vérification par l'approche présentée par Gelman et al. (1996). L'idée est de se rapprocher de la p-valeur telle qu'elle définie dans l'approche fréquentiste (voir annexe D.4.1). Le principe est d'élaborer une statistique T dépendant des données y et des paramètres θ puis de la calculer pour plusieurs valeurs de θ tirées dans les lois *a posteriori* des paramètres, pour les données observées et pour les données répliquées à partir des lois *a posteriori*. Les différentes valeurs prises par la statistique sont placées sur un graphe à deux dimensions comme celui présenté à la figure 3.10. La p-value est estimée par la proportion de valeurs de $T(\theta, y_{rep})$ dépassant $T(\theta, y_{obs})$. Pour d'autres définitions de la p-valeur, consulter par exemple Bayarri et al. (2007). Dans l'approche fréquentiste, la distribution de la p valeur sous l'hypothèse nulle est une loi uniforme sur $[0; 1]$. Dans l'approche bayésienne, la distribution uniforme n'est pas assurée. Ainsi, en reprenant l'exemple de la figure 3.10, la p-valeur est de 0,24. Cependant, la distribution de p est nécessaire pour savoir si la valeur trouvée est petite ou non. Cette approche présente donc deux difficultés : le choix de la statistique T et la détermination de la distribution de la p-valeur. Gelman et al. (1996) proposent par exemple la distance du χ^2 : $X^2(y; \theta) = \sum_{i=1}^n \frac{(y_i - E(y_i|\theta))^2}{Var(y_i|\theta)}$ b. Avec les modèles présentés dans l'article reproduit à la partie précédente, le calcul de $E(y_i|\theta)$ et de $Var(y_i|\theta)$ pour chaque jeu de données (simulées ou observées) ne s'écrit pas de manière analytique. Il faudrait donc choisir une statistique plus adaptée en prenant soin qu'elle prenne en compte les données au niveau du lot. Nous n'avons pas eu recours à cette approche car la construction de la statistique et la détermination de la distribution de la p-valeur sont complexes et que nous avons préféré nous consacrer à d'autres sujets.

b. Les notations sont celles de Gelman et al. (1996).

3.4 Conclusion

Quel est l'impact du modèle de contamination utilisé sur la courbe d'efficacité? Ce point est traité à la section 3.4.1. La comparaison des modèles est-elle plus aisée si les données comportent moins de zéros? Je réponds à cette question à la section 3.4.2 en utilisant des données de flore lactique.

3.4.1 Courbe d'efficacité et modèle de contamination

L'approche décrite dans l'article est applicable à d'autres cas, y compris pour des données de plusieurs entreprises à l'échelle d'un pays. Pour traiter cet aspect, il suffit d'étudier la variabilité inter-entreprises.

L'intérêt de l'étude des différentes variabilités réside en partie dans l'utilisation des plans d'échantillonnage. En effet, l'efficacité d'un plan peut varier énormément en fonction du modèle étudié. Prenons l'exemple du cas des poitrines de porc. Si on décide d'utiliser les modèles B-test portion ou B-unit, il est impossible de tracer une courbe d'efficacité en fonction de la concentration moyenne du lot comme cela a été fait à la partie 1.2.3.b car les modèles B-test portion et B-unit n'intègrent pas le lot dans leur construction. En revanche, les courbes d'efficacité en fonction de la log-concentration moyenne d'un lot existent pour les modèles B-batch et B-reference. Pour le premier modèle, une seule courbe est définie (puisque le nombre de colonies suit une loi de Poisson au niveau du lot) alors que pour le second, on peut tracer un faisceau de courbes en tenant compte de l'incertitude sur λ , l'écart-type inter-unités. La figure 3.11 présente la courbe d'efficacité pour le modèle B-batch ainsi que la médiane et l'intervalle de crédibilité à 95% de la courbe d'efficacité du modèle B-reference. La médiane de la courbe d'efficacité du modèle B-reference et la courbe d'efficacité du modèle B-batch sont quasiment confondues.

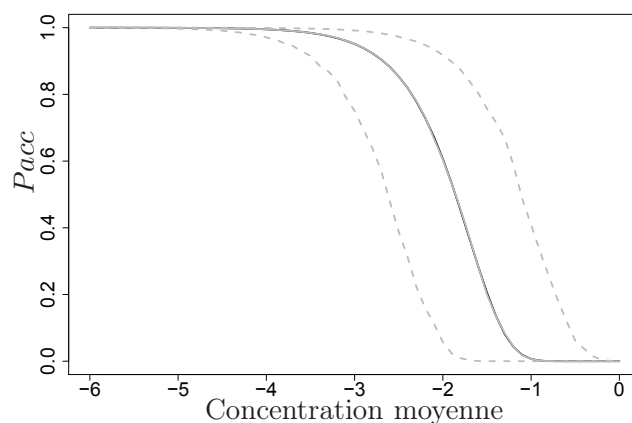


FIGURE 3.11 – Courbes d'efficacité pour un plan à deux classes avec $n=10$, $c=0$ pour les modèles B-batch (noir) et B-reference (gris) en fonction de la log concentration moyenne par lot (en log ufc/cm²). Les lignes en pointillés sont les 2,5e et 97,5e percentiles des courbes d'efficacité pour le modèle B-reference.

Modèle	Paramètre	Dist. <i>a priori</i>	Statistiques desc. de la dist. <i>a posteriori</i>				
			Moy	S.D.	2.5e perc.	50e perc.	97.5e perc.
B-batch	μ	$\mathcal{N}(0, 10^3)$	2.19	0.16	1.86	2.19	2.52
	σ	$\sigma^{-2} \sim \mathcal{G}(10^{-3}, 10^{-3})$	0.55	0.13	0.36	0.53	0.87
B-unit	μ	$\mathcal{N}(0, 10^3)$	3.18	0.05	3.08	3.19	3.29
	λ	$\lambda^{-2} \sim \mathcal{G}(10^{-3}, 10^{-3})$	0.53	0.04	0.48	0.53	0.62

TABLE 3.10 – Distributions *a priori* pour les modèles B-batch et B-unit pour le cas “poitrines de porc” et les dénombrements de flore lactique. Des statistiques descriptives des distributions *a posteriori* sont présentées pour tous les paramètres des modèles. “dist.” signifie “distribution”, “perc.” - “percentile” et “S.D.” for “écart-type”. \mathcal{N} signifie for distribution normale et \mathcal{G} - distribution gamma.

3.4.2 Utilisation des données de flore lactique

Une des difficultés rencontrées dans l’article est le fait que les jeux de données comportent de nombreux 0. C’est pourquoi un soin particulier a été apporté à l’élaboration des lois *a priori* pour le modèle le plus général et des lois *a priori* compatibles ont été introduites pour les modèles emboîtés.

Pour le cas “poitrines de porc”, nous disposions également de dénombrement de flore lactique qui est présente en grande quantité à la surface des poitrines (cf. résultats partie 2.3.3.a). Pour chaque prise d’essai, il y a au moins deux décomptes de colonies avec une dilution d’écart.

J’ai utilisé ces données pour estimer les paramètres de contamination de la flore lactique sur les modèles B-batch afin de déterminer si la structure de contamination apparaissait de manière plus nette qu’en utilisant les données de *L. monocytogenes*, autrement dit si le critère et l’inspection visuelle séparaient mieux les deux modèles. Je suis partie de l’hypothèse que les structures de contamination sont les mêmes pour la flore lactique et pour *L. monocytogenes* car elles sont dues à des actions mécaniques lors du malaxage. Des lois *a priori* peu informatives ont été choisies car comme le jeu de données des flores lactiques ne comporte quasiment aucun zéro, l’information apportée par la loi *a priori* est moins cruciale que celle pour le jeu de données avec *L. monocytogenes*. D’ailleurs, les lois *a posteriori* sont les mêmes lorsque les lois *a priori* sont modifiées, à condition bien sûr que les supports de ces dernières ne soient pas trop éloignés des données. Le tableau 3.10 présente quelques statistiques sur les lois *a posteriori* des paramètres. Les écarts-type sont bien plus petits que pour les paramètres de contamination avec *L. monocytogenes* : l’incertitude est moins élevée. Ceci est dû au jeu de donnée qui est bien plus informatif puisqu’il ne comporte quasiment aucun zéro.

Le tableau 3.11 et la figure 3.12 présentent les résultats du critère numérique et de l’inspection visuelle respectivement. On constate que la différence entre les modèles est flagrante, ce qui n’était pas aussi net avec les données de *L. monocytogenes*.

La méthodologie présentée dans cet article peut également être utilisée pour déterminer l’existence ou non d’une variabilité inter ou intra-lot de différentes caractéristiques physico-chimique comme le pH, l’activité de l’eau (a_w) ou le taux de chlorure de sodium par exemple. Vermeulen et al. (2011) montrent l’existence d’une variabilité inter-lots pour ces caractéristiques ainsi que pour le taux de croissance de *L. monocytogenes*. De même

Modèle M	C(M)
B-batch	$3,0 \cdot 10^4$
B-unit	$3,9 \cdot 10^8$

TABLE 3.11 – Résultats du critère numérique défini à la partie 2.4.1 de l'article (partie 3.2 de la thèse) pour deux modèles en utilisant les dénombrements de flore lactique sur des poitrines de porc après malaxage.

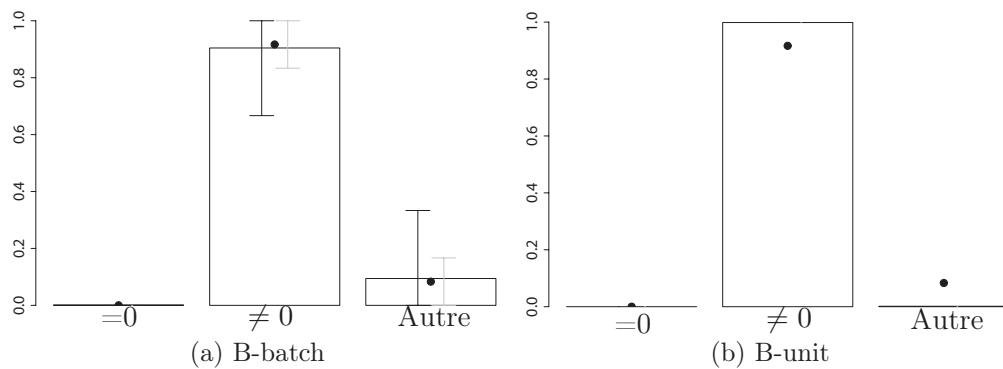


FIGURE 3.12 – Données de flore lactique sur des poitrines de porc en sortie malaxage simulées et observées pour les modèles B-batch (a) et B-unit (b).

Les histogrammes représentent la moyenne de chaque groupe (“=0” : proportion de lots ne comportant que des zéros, “≠ 0” : proportion de lots ne comportant que des dénombrements non nuls, “Autre” : tous les autres lots. Les segments de droite représentent les intervalles de crédibilité à 50% et les segments de gauche les intervalles de crédibilité à 95%. Les points sont les données observées.)

que pour l'étude de la flore lactique, les jeux de données de physico-chimie ne comportent pas de zéros, ce qui facilite grandement l'étude de ces différentes variabilités.

Chapitre 4

Choisir entre les données brutes et les données transformées

Ce chapitre étudie la qualité des estimations des paramètres par inférence fréquentiste en ayant recours soit aux données brutes, soit aux données exprimées sous forme de concentration. Ce travail a été publié dans la revue *International Journal of Food Microbiology*. L'article est reproduit à la partie 4.2.

4.1 Données et estimation des paramètres

4.1.1 Positionnement du problème

Comme il a déjà été précisé dans le chapitre 2, les données de détection et de dénombrement de micro-organismes sont utilisées telles quelles (données brutes) ou sont exprimées sous forme de concentration (données transformées). Si la concentration est faible, des données brutes sont nulles (aucune colonie n'est observée sur les boîtes de Petri). L'expression en concentration est alors une donnée censurée à gauche (elle est inférieure à la LOQ). De même une données de détection est forcément censurée si elle est exprimée en concentration (cf. partie 2.1.2 du chapitre 2). La censure nous paraissait une perte d'information : les données exprimées sous forme de concentration ne permettent pas toujours de remonter aux données brutes car le protocole n'est pas toujours communiqué complètement.

D'autre part, les données sous forme de concentration font penser que les données sont continues, d'autant plus qu'un modèle de contamination classiquement employé est le modèle Log-Normal. Si ce modèle est utilisé pour simuler des données une fois que ses paramètres ont été estimés, cela peut conduire à la génération de données impossibles.

A mon sens, les données exprimées sous forme de concentration peuvent donner l'illusion d'être précises justement parce que la concentration est une variable continue. Reprenons les valeurs de la norme NF EN ISO ISO11290-2 (1998a) : 10 g de produit sont dilués dans 90 mL pour obtenir une solution de 100 mL dont 0,1 mL est prélevé puis étalé sur boîtes de Petri. Une colonie décomptée correspond à une concentration de 100 ufc/g, 2 colonies à 200 ufc/g et ainsi de suite. Si on ne fait pas attention, on considère qu'une donnée égale à 100 ufc/g est une concentration précise au sens où 120 ufc/g par

exemple est une concentration différente alors qu'en fait le protocole utilisé ne permet pas de distinguer ces deux concentrations (la donnée brute correspondant à 100 ufc/g sera presque toujours la même que la donnée brute correspondant à 120 ufc/g), à savoir une colonie décomptée.

Les différents modèles de contamination présentés à la partie 1.4 utilisent l'un ou l'autre type de données. Quel type de données (et donc de modèles) faut-il utiliser pour estimer au mieux les paramètres ?

Partant donc du principe que des données brutes sont plus informatives que des données exprimées sous forme de concentration, nous avons cherché à déterminer l'impact de cette perte d'information sur la qualité d'estimation des paramètres de modèles de contamination. Nous avons étudié deux modèles : le modèle \mathcal{M}_{RD} , basé sur la loi log-normale pour les données exprimées sous forme de concentration et le modèle \mathcal{M}_{CLD} basé sur la loi Poisson-Log-Normale pour les données brutes. Ces modèles présentent l'avantage d'avoir les mêmes paramètres (log concentration moyenne μ et log écart-type σ). Nous avons tenu compte à la fois des données de détection et de dénombrement et nous avons considéré les données censurées en tant que telles (c'est-à-dire sans les remplacer par une valeur fixe par exemple). Afin de mieux étudier l'estimation des paramètres, nous avons travaillé sur des données simulées de manière à contrôler exactement les vraies valeurs des paramètres. Pour chaque valeur du couple (μ, σ) , 1000 jeux de données ont été simulés avec 100 données de détection et autant de données de dénombrement que de détections positives afin de simuler un type d'acquisition de données (trouvé par exemple chez Beaufort et al. (2007) ou Uyttendaele et al. (2009)) puis les paramètres ont été estimés en utilisant la technique du maximum de vraisemblance. Cette technique a pu être utilisée directement pour le modèle \mathcal{M}_{RD} . La vraisemblance de l'autre modèle ne pouvant être calculée analytiquement, l'algorithme EM a été utilisé. Les détails de cet algorithme se trouvent à l'annexe E.

Nous souhaitons donner un caractère opérationnel à nos résultats, c'est pourquoi fournir les résultats uniquement en fonction des valeurs réelles de μ et de σ ne nous semblait pas adapté puisque ces valeurs sont justement inconnues dans un jeu de données non simulé. C'est pourquoi nous avons également présenté les résultats en fonction du pourcentage de résultats "absence" dans les données de dénombrement car cette caractéristique influe sur la qualité des estimations. En effet, plus il y a d'absence et

- moins il y a de données de dénombrement (d'après le type de simulations choisi),
- plus les données de dénombrement comportent de 0 en données brutes et de données censurées lors de leur conversion en concentration.

Par ailleurs, dans le domaine de l'analyse des risques, les valeurs de μ et de σ ne sont pas les seules valeurs intéressantes à considérer. En effet, les percentiles élevés de concentration sont utiles à considérer car ce sont eux qui ont le plus d'effets sur l'exposition (Pouillot et al. (2009)). Le 97,5^{ème} percentile a donc été estimé pour chaque jeu de données et chaque modèle et comparé au 97,5^{ème} percentile réel en fonction du pourcentage d'absences.

4.1.2 Quel type de données utiliser en fonction du pourcentage d'absences ?

De manière générale, le modèle \mathcal{M}_{RD} estime mieux les paramètres ainsi que le 97,5^{me} percentile que le modèle \mathcal{M}_{CLD} . La meilleure estimation dépend du pourcentage d'absence, ainsi :

- s'il y a moins de 15% d'absences, les deux modèles fournissent des résultats quasiment identiques en pratique. Il vaut mieux utiliser le modèle \mathcal{M}_{CLD} car les estimateurs de μ et σ sont plus faciles à programmer : il suffit de recourir "directement" au maximum de vraisemblance, cette dernière s'écrivant facilement ;
- s'il y a entre 15 et 75% d'absence, il vaut mieux utiliser le modèle \mathcal{M}_{RD} car les estimations des paramètres et du 97,5^{me} percentile de la concentration sont meilleures ;
- au-delà de 75% d'absence, les deux modèles n'arrivent pas à bien estimer les paramètres et le 97,5^{me} percentile de concentration car il y a vraiment peu de données informatives. Néanmoins, si on souhaite avoir une estimation ou plutôt une idée d'estimation, il vaut mieux utiliser le modèle \mathcal{M}_{RD} qui est le moins mauvais des deux.

Bien évidemment, ces conclusions sont liées au protocole de simulation de données utilisé.

4.2 Article : *Fitting a lognormal distribution to enumeration and absence/presence data*

Abstract To fit a lognormal distribution to a complex set of microbial data, including detection data (e.g. presence or absence in 25 g) and enumeration data (e.g. 30 cfu/g), we compared two models : a model called \mathcal{M}_{CLD} based on data expressed as concentrations (in cfu/g) or censored concentrations (e.g. < 10 cfu/g, or > 1 cfu / 25 g) versus a model called \mathcal{M}_{RD} that directly uses raw data (presence/absence in test portions, and plate colony counts). We used these two models to simulated data sets, under standard conditions (limit of detection (LOD) =1 cfu / 25 g; limit of quantification (LOQ) =10 cfu/g) and used a maximum likelihood estimation method (directly for the model \mathcal{M}_{CLD} and via the Expectation-Maximisation (EM) algorithm for the model \mathcal{M}_{RD}). The comparison suggests that in most cases estimates provided by the proposed model \mathcal{M}_{RD} are similar to those obtained by model \mathcal{M}_{CLD} accounting for censorship. Nevertheless, in some cases, the proposed model \mathcal{M}_{RD} leads to less biased and more precise estimates than model \mathcal{M}_{CLD} .

Keywords : Microbial contamination assessment, limit of detection, limit of quantification, maximum likelihood estimation, EM algorithm

Cette partie est la reproduction de l'article écrit en collaboration avec Eric Parent, Marie-Laure Delignette-Muller et Marie Cornu. Il a été publié dans la revue International Journal of Food Microbiology (155(3) :146-152)

4.2.1 Introduction

Detection and enumeration of contaminating microorganisms are regularly collected nationally and internationally for quantitative risk assessments of food products (Beaufort et al., 2007; Uyttendaele et al., 2009; Habib et al., 2008; Burfoot et al., 2010). However, in many cases, particularly when the target microorganism of interest is a pathogenic species, available data are limited because analyses are labour-intensive, and they are poorly informative due to a high proportion of non-detects (i.e. when a detection has been performed and the result is absence) and zero counts (i.e. an enumeration was performed and zero colony was enumerated).

How can this type of complex data set be analysed to properly assess microbial contamination? Assuming that concentrations (e.g. in cfu/g) in the contaminated products follow a lognormal distribution, it was frequent in the 1990s and early 2000s to calculate a log mean concentration based only on enumeration results (i.e. ignoring non-detects; “log” refers to the logarithm to base 10), see Abadias et al. (2006). Zero counts were either ignored, substituted with the limit of quantification (LOQ) or a fraction of the LOQ. Since then, this practice has been criticised for the bias it introduces (Helsel, 2006; Lorimer and Kiermeier, 2007) and the maximum likelihood estimation (MLE) approach combined with statistical censoring has been introduced to fit a single lognormal distribution to the full data set which can then be applied to all food products from which the samples have been taken (Lorimer and Kiermeier, 2007; Busschaert et al., 2010; Pouillot and Delignette-Muller, 2010). In statistics, the term “censoring” is used when an observation has not been quantified, but is known (or presumed) to exceed or to be less than a threshold value. In left-censoring, a data point is presumed to be below the threshold value, whereas in right-censoring the data point is above the threshold value. Thus, for MLE fits as proposed by Lorimer and Kiermeier (2007); Busschaert et al. (2010); Pouillot and Delignette-Muller (2010), positive enumeration data are expressed as concentrations (in cfu/g), and all other data are expressed as censored concentrations : a zero count is considered as below the LOQ (left censoring), presence is considered as above the limit of detection (LOD)(right censoring) and absence is considered as a concentration below the LOD (left censoring). The LOD is the minimal theoretical concentration required for a test portion so that the result of detection is “presence”.

These concentration-like data, i.e. the results expressed as concentrations (in cfu/g) for positive enumeration results, and as censored concentrations in all other cases, may actually convey less information in statistical analysis than directly using raw data generated in the laboratory, i.e. either colony counts per plate or binary presence/absence data. A few methods are currently available for assessing the distribution of bacterial concentrations from raw data Commeau et al. (2012a), but they are rarely used in practice.

The objective of this work was to determine in which conditions the use of concentration-like data leads to less accurate estimation of the parameters of the lognormal distribution than the use of raw data. Models and estimation methods are illustrated in this paper with simulated data sets to study the statistical properties of the estimates. Realistic parameters have been chosen with regard to microbiological contamination in the food industry. After presenting the models and estimation methods, we suggest some criteria for comparing the methods’ performances.

4.2.2 Materials and methods

4.2.2.a How should raw data be converted into concentration-like data ?

This section describes the notations used below to denote the current ways in which microbiology laboratories convert raw data into concentration-like data. To illustrate our notations, numerical values have been taken from the methods for *Listeria monocytogenes* (ISO11290-1, 1996a; ISO11290-2, 1998a), but can be applied to other target microorganisms. For detection analysis, if result x (coded as 0=absence and 1=presence) is presence, it means that the target has grown during the enrichment step then it is likely that, before the enrichment step, there was at least one cell in the test portion of mass $M = 25$ g, and so the concentration is presumed to be greater than the LOD= $1/M = 0.04$ cfu/g. If the result is absence, the concentration is presumed to be below the LOD.

Regarding enumeration, (ISO11290-2, 1998a), once a test portion of $M' = 10$ g has been taken, it is diluted in a solution so that the final volume is $V_d = 100$ ml from which a volume $V_r = 1$ ml is removed and pour-plated onto a Petri dish containing the appropriate medium. Finally, after 24 h or 48 h of incubation, y colonies of the target [microorganism] are counted. If $y \neq 0$, then these colonies are assumed to come from the multiplication of y cells initially present in V_r . The number of cells in the volume V_d is considered to be proportional to y and is equal to $y/V_r = y/0.01$ cfu. Because all these cells come from the test portion of mass M' g, the concentration of the target in the test portion is $y/(M'V_r/V_d) = y/0.1$ cfu/g. If $y = 0$, then it is considered that the number of cells in the test portion is less than 1 colony in a volume V_r , thus the concentration is below the LOQ= $1/(M'V_r/V_d) = 10$ cfu/g. Note that in this case study, colonies are not confirmed.

For enumeration, if n dilutions are made, the conversion formula becomes $y/(M' \prod_{i=1}^n V_{ri}/V_{di})$, where V_{di} is the volume of the i th dilution V_{ri} is the volume removed at the i th dilution. This case is not considered in the article to simplify notations.

4.2.2.b Models to estimate the contamination distribution

Fitting a lognormal distribution to a data set, obtained as described in the previous section, relies on assumptions, which can be conceptualized under a model.

The most simplistic models, ignoring non detects and zero counts, or substituting them with an arbitrary value, are not considered in this study because they are not appropriate models to estimate the parameters of contamination properly (Helsel (2006); Shorten et al. (2008); Busschaert et al. (2010)).

The recommended current practice It is possible to take all the available data into account, censored and non-censored. The approach is described in Lorimer and Kiermeier (2007); Busschaert et al. (2010); Pouillot and Delignette-Muller (2010). A concentration datum c follows a lognormal distribution :

$$c \sim \mathcal{LN}(\mu, \sigma).$$

In this case, a detection datum can be expressed as a censored concentration (or concentration-like datum) : less than the LOD (absence) or greater than the LOD (presence). An enu-

meration concentration-like value is either the value of c , with $c \geq LOQ$ or the information $c < LOQ$ when censored. This model is denoted \mathcal{M}_{CLD} , as data in this model are concentration-like ones.

A physically based approach As mentioned above, data expressed as concentrations, concentration-like data, are obtained from raw data, i.e. colony counts for enumeration (y) or presence/absence for detection (x). We therefore designed a hierarchical model based on raw data and that mimics the “real” assay as it is actually carried out in laboratories. Enumeration data follow a Poisson distribution (P), detection data follow a Bernoulli distribution (Ber) and the unknown concentration data λ follow a lognormal distribution (LN) :

$$\begin{aligned} y &\sim \mathcal{P}(\lambda M' \frac{V_r}{V_d}) \\ x &\sim \mathcal{Ber}(1 - \exp(-\lambda M)) \\ \lambda &\sim \mathcal{LN}(\mu, \sigma). \end{aligned}$$

The term $1 - \exp(-\lambda M)$ is the probability of obtaining at least one colony when cells follow a Poisson distribution with parameter λM . Conditionally to the concentration λ , detection result x and enumeration result y are independent. We call this hierarchical model \mathcal{M}_{RD} , as it links raw data to the parameters μ and σ . It corresponds to the model A-test portion in Commeau et al. (2012a). If ignoring detection results, this model is known as the Poisson LogNormal model (ILSI, 2010; Reinders et al., 2003).

To use this model, it is necessary to have the protocol, or at least the LOD and the LOQ. If only the concentration-like data are given, it is not always possible to get the LOQ so there can be a loss of information when converting raw data into concentration-like data.

4.2.2.c Parameters estimation

For the most simplistic models as well as for model \mathcal{M}_{CLD} , maximum likelihood can be used analytically to estimate parameters μ and σ , for example with the package `fitdistrplus` described in Pouillot and Delignette-Muller (2010). The likelihood of this model is equal to :

$$\begin{aligned} L(C|\mu, \sigma) &= \prod_{i=1}^{i=N_{abs}} F(C_i \leq LOD) \prod_{j=1}^{j=N_{pres}} F(C_j > LOD) \prod_{k=1}^{k=N_0} F(C_k < LOQ) \prod_{l=1}^{l=N_{\neq 0}} F(C_l = c_l) \\ &= \prod_{i=1}^{i=N_{abs}} \Phi\left(\frac{\log LOD - \mu}{\sigma}\right) \prod_{j=1}^{j=N_{pres}} (1 - \Phi\left(\frac{\log LOD - \mu}{\sigma}\right)) \\ &\quad \prod_{k=1}^{k=N_0} \Phi\left(\frac{\log LOQ - \mu}{\sigma}\right) \prod_{l=1}^{l=N_{\neq 0}} \frac{1}{c_l \sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\log c_l - \mu)^2}{2\sigma^2}\right), \end{aligned}$$

where N_{abs} is the number of absence results, N_{pres} the number of presence results, N_0 the number of enumeration results with no colony and $N_{\neq 0}$ the number of enumeration result

with at least one colony present in a data set. The cumulative distribution function (cdf) of the log-normal distribution is denoted F and the standard normal cdf is denoted Φ . The random variables are denoted C_i, C_j, C_k and C_l and the values taken by C_l are noted c_l .

The likelihood cannot be analytically calculated for model \mathcal{M}_{CLD} . The likelihood of stochastic variables X and Y conditionally to λ is equal to :

$$\begin{aligned} L(X, Y | \lambda) &= \prod_{i=1}^{i=N_{abs}} \mathbb{P}(X_i = 0) \prod_{j=1}^{j=N_{pres}} \mathbb{P}(X_j = 1) \prod_{k=1}^{k=N_0} \mathbb{P}(Y_k = 0) \prod_{l=1}^{l=N_{\neq 0}} \mathbb{P}(Y_l = y_l) \\ &= \exp(-N_{abs}\lambda M)(1 - \exp(\lambda M))^{N_{pres}} \exp(-N_0\lambda M' \frac{V_r}{V_d}) \\ &\quad \prod_{l=1}^{l=N_{\neq 0}} \exp(-\lambda M' \frac{V_r}{V_d}) (\lambda M' \frac{V_r}{V_d})^y \frac{1}{y!}, \end{aligned}$$

where X_i, X_j, Y_k and Y_l are stochastic variables and y_l are the values taken by Y_l . The likelihood of X and Y conditionally to μ and σ is equal to $L(X, Y | \lambda)$ integrated on λ :

$$\begin{aligned} L(X, Y | \mu, \sigma) &= \int L(X, Y | \lambda) [\lambda] d\lambda \\ &= \int L(X, Y | \lambda) \frac{1}{\lambda \sqrt{2\pi\sigma^2}} \exp\left(\frac{-(\log \lambda - \mu)^2}{2\sigma^2}\right) d\lambda. \end{aligned}$$

Bayesian inference has been proposed by Commeau et al. (2012a). When writing the likelihood, the latent variable λ appears. To find the maximum of the likelihood, the Expectation-Maximisation (EM) algorithm is used, which is an iterative method for finding maximum likelihood estimates of parameters, where the model depends on unobserved latent variables. The EM iteration consists of two processes, the E-step, which computes the expectation of the log-likelihood evaluated using the current estimate for the parameters, and an M-step, which consists in finding the values of parameters maximizing the expected log-likelihood found on the E-step. These parameter-estimates are then used to determine the distribution of the latent variables in the next E-step. The EM algorithm was first described in Dempster et al. (1977). In this work, the E-step was performed using importance sampling (Robert and Casella, 2004) and the M-step was performed using a simple gradient search in a bidimensional space. This algorithm is written in R (R Development Core Team, 2011) and is available upon request from the first author. The algorithm is easy to code because the derivative of $Q(\theta, \theta') = \int \ln[y, \lambda | \theta] [\lambda | \theta', y] d\lambda$ is easy to calculate.

4.2.2.d Simulating detection and enumeration protocols

When carrying out detection and enumeration protocols, several measurement errors can occur : on mass M and volumes V_d and V_r in the enumeration protocol and in the detection protocol for mass M' . All these errors were described as follows, assuming plausible coefficients of variation (0.1% for test portion mass and 2% for dilution volume and masses) commonly found in standard laboratory working conditions.

- Enumeration errors
 - Test portion mass : $m' \sim \mathcal{N}(M', 0.001M')$
 - Dilution volume : $v_d \sim \mathcal{N}(V_d, 0.02V_d)$
 - Removed volume : $v_r \sim \mathcal{N}(V_r, 0.02V_r)$
- Detection errors
 - Test portion mass : $m \sim \mathcal{N}(M, 0.02M)$

Here, upper-case letters for volume and mass refer to the intended masses and volumes (as written in the protocol) and the lower-case letters refer to the mass and the volume actually simulated to model what is obtained in a laboratory. The standard deviations of the normal distributions are expressed in terms of the relative accuracy of the device used. We considered that there is no error in the detection and enumeration results. For example, if six colonies are counted on a Petri dish that means that there were six cells in the pour-plated volume V_r .

To model enumeration result y , a test portion mass m' , a dilution volume v_d and a removed volume v_r were simulated from their distributions and then an enumeration result y was simulated from the Poisson distribution with parameter $\lambda m' \frac{v_r}{v_d}$, where λ is the concentration of the test portion and $\lambda \sim \mathcal{N}(\mu, \sigma)$. Finally, the raw datum y was converted into a concentration $c_y : c_y = \frac{y}{M'V_r/V_d} = y \times LOQ$, if $y \neq 0$ and $c_y < \frac{1}{M'V_r/V_d} = LOQ$, if $y = 0$. Detection result x was simulated from a Bernoulli distribution with parameter $1 - \exp(-\lambda m)$, with $\lambda \sim \mathcal{N}(\mu, \sigma)$ and m simulated from its distribution as described above. The result was converted into concentration using the following formula : if $x = 1$, then $c_x \geq \frac{1}{M} = LOD$, else, $c_x < LOD$.

The intended values of masses and volumes were used in these conversion formulae for both enumeration and detection to mimic real laboratory [assay] conditions. For simulation purposes, values of the mean log concentration μ range from -3 to 4 log cfu/g with an increment of 0.5 log cfu/g between two values. For the standard deviation of the log concentration σ , values 0.7, 1 and 1.5 log cfu/g were chosen.

4.2.2.e Simulating a sampling plan

For each parameter value, 1000 data sets were simulated. In the simulations, we mimicked the strategy applied in the experimental plan of Beaufort et al. (2007) or Uytendaele et al. (2009). In these articles, detection analyses were first performed. Each time the result for one test portion was presence, another test portion (somehow linked to the positive test portion) was enumerated. In these simulations, first detection data were simulated and then as many enumeration data as positive detection results were simulated. Therefore, each data set had 100 detection results and from 0 to 100 enumeration data. The putative link between the positive test portion and the consecutively enumerated test portion is not explicitly taken into account, neither in the simulations, nor in the estimations. This link is not considered because the food is solid and microorganisms do not move inside it. It is therefore possible that a test portion taken from a food sample has a colony on it (the detection result is then 'presence', so the concentration-like datum is right-censored and above the LOD) and that another test portion removed from the same sample has no colony at all so the enumeration result is 'no colony' and is left-censored (below the LOQ). As a consequence, we do not consider that the concentration

is between the LOD and the LOQ. We consider instead that we have two different data, one right-censored with a concentration above the LOD while the second is left-censored with a concentration below the LOQ. This approach is different from the one chosen in Busschaert et al. (2010) in which a presence detection result and a 0-colony enumeration result on two different test portions from the same sample are considered as one interval-censored datum (between the LOQ and the LOD).

The intended values of volumes and masses were : $V_d = 100$ ml, $V_r = 1$ ml, $M = 25$ g and $M' = 10$ g. According to these numerical values, the LOD was equal to 0.04 cfu/g and the LOQ, 10 cfu/g. These values are commonly used, e.g. for *L. monocytogenes* (ISO11290-1, 1996a; ISO11290-2, 1998a).

4.2.2.f Bias and variance in the estimates of mean and variance

To assess the quality of each estimate, its bias and variance were calculated. Bias is the difference between the mean of the estimate and the true value of the parameter. For instance, for μ , the formula is $E(\hat{\mu} - \mu)$, where $\hat{\mu}$ is the result of an estimation of μ for a given data set. Biases for μ and for σ were calculated for each data set and for each model.

The sum of the bias and the variance is equal to the mean square error (MSE). For instance, for μ , it is equal to $MSE(\hat{\mu}) = (E(\hat{\mu}) - \mu)^2 + V(\hat{\mu}) = E((\hat{\mu} - \mu)^2)$. The MSE is also equal to the expected value of the squared error loss. The smaller the MSE are, the better the inferences are.

For the second parameter, the MSE was used on $\ln(\hat{\sigma})$: $MSE(\ln(\hat{\sigma})) = E((\ln \hat{\sigma} - \ln \sigma)^2) = (E(\ln \hat{\sigma}) - \ln \sigma)^2 + V(\ln \hat{\sigma})$, where “ln” is the natural logarithm. The logarithm was taken because σ is always positive. Moreover, it is the order of magnitude that is important. Therefore, taking the logarithm of σ instead of σ is useful for penalising estimations of σ that are too low.

4.2.2.g Estimating the 97.5th percentile

In addition to assessing μ and σ , the 97.5th percentile on the total population was studied for the two models.

4.2.3 Results

4.2.3.a Results of the two inference methods

Figure 4.1 shows the estimates of μ and of σ when $\sigma = 1.5$ log cfu/g. The simulated data set inferred a higher concentration and was less dispersed than the real data set. The pattern was quite similar for $\sigma = 0.7$ and $\sigma = 1$ (results not shown), although $\hat{\mu}$ and $\hat{\sigma}$ were less dispersed. For low values of μ , there was a correlation between $\hat{\mu}$ and $\hat{\sigma}$. Information on estimate bias can be seen in Figure 4.1. This bias decreased when μ increased and was virtually nil when μ was sufficiently high (the exact value depended on the model and on σ ; for instance, for $\sigma = 1.5$ log cfu/g and for model \mathcal{M}_{CLD} , the bias was nil for $\mu > 2$ log cfu/g). The higher the value of σ was, the higher the bias for the same value of μ was (results not shown). The model with the smallest bias was model \mathcal{M}_{RD} . Overall,

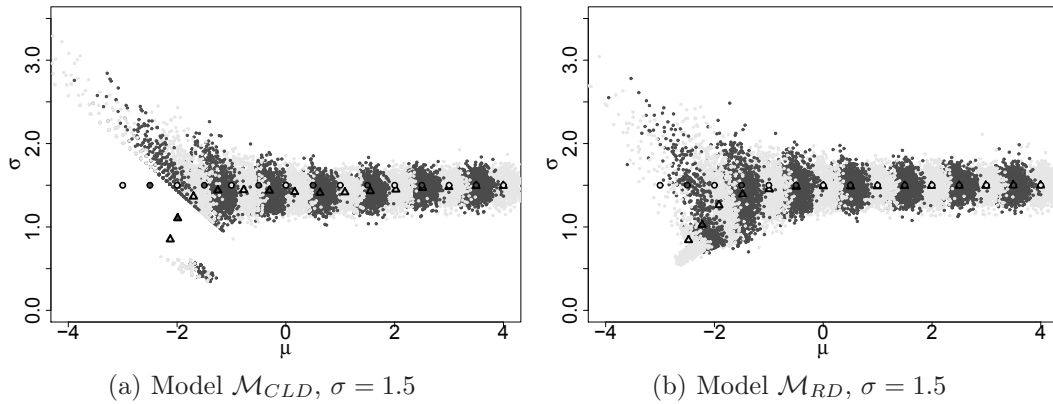


FIGURE 4.1 – Joint distribution of inference on μ (the mean of the log concentration in log cfu/g) and σ (the standard deviation of the log concentration in log cfu/g). Each shade of grey (light or dark grey) corresponds to a value of μ , from -3 to 4 with an increment step of 0.5. Each point represents an inference of μ and σ . Filled circles are the true values of μ and σ . Filled triangles are the mean of the estimates for each value of μ . For each value of the parameter pair (μ, σ) , 100 data sets were simulated. Each data set contains 100 detection results and as many enumeration results as presence detection data. Information on estimate bias can be seen as the distance between the squares and the triangles of the same color.

models \mathcal{M}_{CLD} and \mathcal{M}_{RD} led to overestimation of low μ values. Overestimation was worse for model \mathcal{M}_{CLD} than for model \mathcal{M}_{RD} . Standard deviation σ was underestimated with both models and on average, the estimates of σ with model \mathcal{M}_{CLD} were higher than the estimate using model \mathcal{M}_{RD} , except for $\sigma = 0.7$ log cfu/g.

For a given value of σ , prevalence in detection and percentage of non-zero counts in enumeration results decreased when μ decreased. This is expected since, when μ decreases, the concentration decreases as well; therefore, there are more and more zero-count enumeration results and absence-of-detection results.

Sometimes, for $\mu = -3$ log cfu/g, all 100 simulated detection data points were “absence”. Under the assumed sampling plan, this means that no enumeration was simulated. In that case, estimations were not made for either of the models.

4.2.3.b Bias and variance of the estimates μ and σ

Figure 4.2 shows the results of MSE for μ . The $\text{MSE}(\hat{\mu})$ for $\sigma = 1.5$ log cfu/g are not shown because they are very close to the $\text{MSE}(\hat{\mu})$ when $\sigma = 1$ log cfu/g. When $\mu \geq 1$ cfu/g, models \mathcal{M}_{CLD} and \mathcal{M}_{RD} gave $\text{MSE}(\hat{\mu})$ almost equal to 0. When μ was greater than 0.5 log cfu/g, the MSE for $\ln \hat{\sigma}$, whatever the value of σ , were almost equal in both models and very close to zero (not shown).

When using real data sets, the true values of μ and σ are of course unknown. Available information is the percentage of presence in detection results (i.e. observed prevalence).

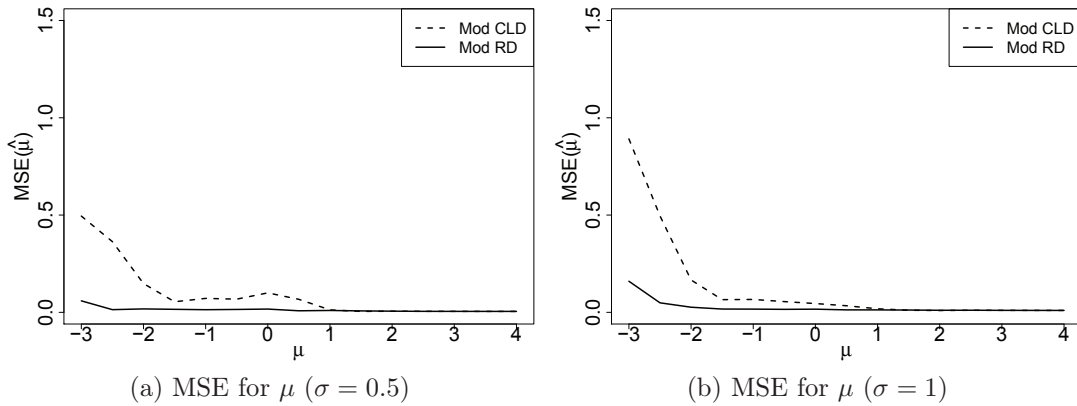


FIGURE 4.2 – $MSE(\hat{\mu})$ (in $(\log \text{ cfu/g})^2$) for the two models and for $\sigma = 0.5 \log \text{ cfu/g}$ (a) and $\sigma = 1 \log \text{ cfu/g}$ (b). Formula is as follows : $MSE(\hat{\mu}) = E((\hat{\mu} - \mu)^2)$, where $\hat{\mu}$ is the estimate of μ . The parameter μ ranges from -3 to 4 $\log \text{ cfu/g}$ with an increment of 0.5 $\log \text{ cfu/g}$ between two values.

For this reason, we also represented the MSE as a function of this percentage (Figure 4.3).

When the prevalence was equal to 100%, models \mathcal{M}_{CLD} and \mathcal{M}_{RD} were similar (see Appendix A). Note that in such a case the simplistic models (ignoring non detects and zero counts, or replacing them with arbitrary values) also provide similar estimates. When prevalence decreased from 100% to zero, the MSE of μ for models \mathcal{M}_{CLD} and \mathcal{M}_{RD} remained low, whatever the value of σ . Nevertheless, the MSE for these two models increased when prevalence was below 40%, especially for model \mathcal{M}_{CLD} . The MSE of μ for model \mathcal{M}_{CLD} reached $0.25 (\log \text{ cfu/g})^2$ only when the prevalence was near 40%. The same level was attained for a prevalence of 25% with model \mathcal{M}_{RD} . When there was more than 85% of absence, the MSE for these two models was higher than $0.4 (\log \text{ cfu/g})^2$, at least for $\sigma = 1.5 \log \text{ cfu/g}$.

4.2.3.c Estimating the 97.5th percentile of the total distribution

When studying microbial risk, it is interesting to examine the high percentiles of pathogen concentration, especially if data are collected at the moment of consumption or at the end of shelf life. In fact, high concentrations are the ones that have the greatest impact on the exposure (Pouillot et al., 2009). For each couple (μ, σ) , 1000 values of the 97.5th percentile were estimated. Figure 4.4 represents the 950th value of the 97.5th percentile estimated with both models. This value is higher for model \mathcal{M}_{CLD} than for model \mathcal{M}_{RD} . The real 97.5th percentile was always lower than the ones estimated with both models (the estimation is fail-safe). The estimation using model \mathcal{M}_{RD} was closer to the real value of the percentile as shown in Figure 4.4.

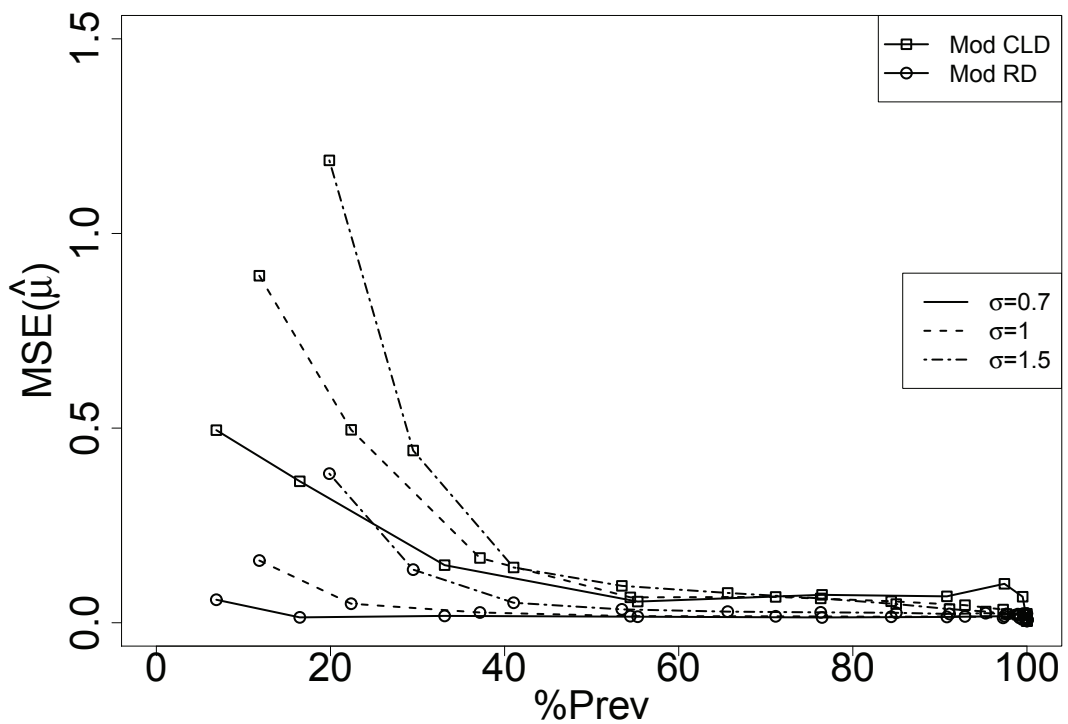


FIGURE 4.3 – MSEs for the four models and for μ with respect to prevalence. Each point couple (line, point) represents a model. Data sets were simulated with a log mean μ which varies from -3 to 4 log cfu/g and a log standard deviation σ which is equal to 0.7 (plain line), 1 (—) and 1.5 (-.-.-) log cfu/g.

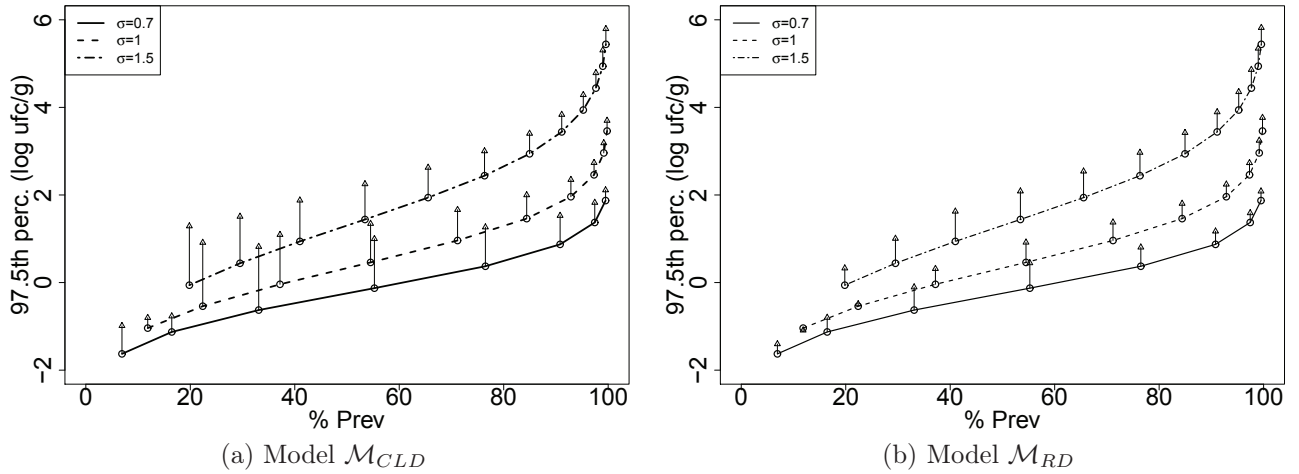


FIGURE 4.4 – 97.5th percentile for the total population (in log cfu/g), depending on the percentage of absence results. Open circles are the real values of the percentiles and triangles are the 950th value of the assessed percentiles for each simulated data set. Vertical lines connect the real percentile to its corresponding mean estimated percentile. Data sets were simulated with a log mean μ which varies from -3 to 4 log cfu/g and a log standard deviation σ which is equal to 0.7 (plain line), 1 (—) and 1.5 (-.-.) log cfu/g.

4.2.4 Discussion

Under the assumption that “true” contamination follows a lognormal distribution, pathogen contamination can be properly assessed from a combination of enumeration and detection results, using methods accounting for zero counts and non-detects. These methods include the model \mathcal{M}_{CLD} with an MLE approach that considers censored data and model \mathcal{M}_{RD} that applies an EM algorithm to raw data. The main objective of this article was to demonstrate the added value of accounting for raw data, i.e using \mathcal{M}_{RD} , instead of \mathcal{M}_{CLD} , which is based on concentration-like data.

As illustrated by each figure, model \mathcal{M}_{RD} provides slightly more precise and accurate estimations of parameters μ and σ of the lognormal contamination (and, consequently, of a high percentile) than those of model \mathcal{M}_{CLD} . The main advantage of the model using the MLE technique directly is that it is very easy to use : for example, using R software with the `fitdistrplus` package. For model \mathcal{M}_{RD} , to our knowledge, no package exists in R to easily apply the EM algorithm specially devoted to this model structure. A function was written in R for this study and is available upon request from the first author. Regarding computational burden, the MLE algorithm runs three times faster than the EM algorithm in model \mathcal{M}_{RD} .

In addition to the likelihood-based frequentist methods discussed in this study, a Markov-Chain Monte Carlo (MCMC) algorithm can also be applied to the so-called models \mathcal{M}_{CLD} and \mathcal{M}_{RD} . This type of Bayesian analysis has recently been published (Busschaert et al., 2011; Gonzales-Barron and Butler, 2011b; Commeau et al., 2012a). Codes are very easy to write for software such as OpenBugs (Thomas et al, 2006), but this program takes longer

to run than the EM algorithm. These MCMC analyses were not included in the present comparison, due to this computational burden. Nevertheless, they are promising for regression analyses, i.e. to consider additional information, e.g. between- and within-batch variability (Gonzales-Barron and Butler, 2011a; Commeau et al., 2012a), the company from which food samples are taken, or the food preparation method (Busschaert et al., 2011). Finally, in Bayesian analysis, additional information can be obtained through expertise and used as priors.

We excluded simplistic models from our comparison. Ignoring censored values or replacing enumeration censored data by the LOD is a source of bias as previously stated (Helsel, 2006; Lorimer and Kiermeier, 2007; Shorten et al., 2008) for both μ and σ estimators : the estimated mean is greater than the real mean, whereas the estimated standard deviation is less than the true one. This result is obvious given that the fitting procedure only applies to a subpopulation of the results, the “positive” subpopulation.

In some cases, the data sets were so uninformative that even the “best” method could not extract any information. In the sampling strategy simulated here, enumeration was performed only as many times as there were presence data. When observed prevalence is low, this leads to few enumeration results, most (or all) of them being zero counts. For some simulated data sets of this study, there were no presence results at all and all estimates were then either impossible or aberrant. This is an extreme case but can occur in real situations, for instance for samples naturally low in contamination, when the LOD is high. For instance, Cornu et al. (2005) show that the 99.9th confidence interval for prevalence (percentage of presence in 25 g) of *L. monocytogenes* in French pork rillettes at the end of production is between 0.10% and 0.50%. In this case, presence is rare and positive enumeration results even rarer. Turning to more informative protocols (by lowering both LOQ and LOD so as to get less zero count or absence results, see for example Gnanou Besse et al. (2004)) would be a promising way to estimate the concentration parameters better, but this would increase the cost of microbiological analyses. To assess the distribution of a microbial population in a food product at a given level of accuracy, it may be necessary to analyse samples with low LOD and LOQ. A cost-benefit analysis should be performed to determine whether it is worth lowering the LOD and LOQ. Data obtained with various protocols (and therefore various LOD and/or LOQ values) can be combined in complex data sets. Methods based on models \mathcal{M}_{CLD} and \mathcal{M}_{RD} can be easily used to analyse a complex data set with various LODs and LOQs, as illustrated by Busschaert et al. (2010) for model \mathcal{M}_{CLD} .

Assessing accurately and precisely a statistical distribution characterising microbial counts is essential in numerous applications of food microbiology, e.g. for risk assessments, for characterising contamination of the hazard at the initial step of a modelled pathway in an exposure assessment, or in risk management, for use in application risk-based metrics, such as Performance Objectives (PO) and Food Safety Objectives (FSO). Nevertheless, this estimation is blurred by diverse sources of uncertainty, resulting from measurement errors in the raw data, possible loss of information from raw data to the data set used in the fitting, sampling uncertainty and the intrinsic properties of the fitting method. Given that routine data sets in food microbiology are usually obtained from low numbers of samples and analytical methods have poor precision, targeting perfectly accurate and precise estimations seems unattainable. For this study, realistic targets were chosen, relative

to the current practise used routinely in laboratories. Thus, it was considered here that the bias in the estimation of a 97.5th percentile should be less than twice the between-laboratory standard deviation for the enumeration of *L. monocytogenes*, as assessed by Augustin and Carlier (2006) at 0.28 log cfu/g. Then, in practice, the 97.5th percentile is considered in this study to be “correct” as long as the difference between the real value of the 97.5th percentile and 95% of the estimates of the 97.5th percentile is below 0.5 log cfu/g.

According to this decision rule applied to the estimation of the 97.5th percentile, and taking all other results and considerations discussed above into account, we make the following recommendations to fit a lognormal distribution properly to a data set obtained under the assumptions of the simulations :

1. if prevalence is greater than 85% (i.e. less than 15% of absence), all estimates of the 97.5th percentile are “correct”. MSE values for the estimation of the log mean using \mathcal{M}_{CLD} and \mathcal{M}_{RD} are very low. model \mathcal{M}_{CLD} However, the model \mathcal{M}_{CLD} is preferred because it is easy to run ;
2. if prevalence is between 25% and 85%, MSE values on the estimation of the log mean are greater for model \mathcal{M}_{CLD} than for model \mathcal{M}_{RD} , and estimates of the 97.5th percentile are closer to the real value for the model \mathcal{M}_{CLD} . Our recommendation is to use only model \mathcal{M}_{RD} .
3. below prevalence of 25%, models \mathcal{M}_{CLD} and \mathcal{M}_{RD} both give poor results, with high MSE values on the estimation of the log mean, and many “incorrect” estimates of the 97.5th percentile. This is easily explained by the poor informativeness of the sample, under the conditions tested here (data set of 100 detection, then less than 25 “presence”, and less than 25 enumerations performed, most or all of them being zero counts). Our best recommendation would be to analyse additional samples, if possible with lower LOQ and LOD. Nevertheless, in this situation, if a statistical analysis has to be done, model \mathcal{M}_{RD} gives slightly better estimates than the other models.

All conclusions developed here are contingent on the way data sets were simulated. For instance, errors in the test portion mass, in the first suspension volume and the volume pour-plated on Petri dishes were considered. However, the model used to simulate data does not consider any errors in colony counting. In reality, this type of error can occur, especially when there are a high number of colonies on plates. If this type of error affects only high numbers of colonies, the inference on μ and σ is practically unchanged. If there are errors in counts of small numbers of colonies, the parameter estimates are more biased (result not shown).

Another key feature is the simulated sampling procedure. In this article, it mimics a practise described in Beaufort et al. (2007) : first, detection is performed on a test portion ; then, if the result is positive, another test portion is taken for enumeration analysis. As a consequence, enumeration is carried out as many times as there are positive detection results. This does not provide much information when μ is low as there is only a small number of enumeration results. If there were as many enumeration analyses as detection analyses, the inference would be slightly better for low values of μ . For instance, if each data set had 100 detection data points and 100 enumeration data points, then for $\sigma = 1.5$

log cfu/g, both the bias and variance of the estimates are lower when $\mu \leq -1.5 \log \text{cfu/g}$ for models \mathcal{M}_{CLD} and \mathcal{M}_{RD} (results not shown).

Here it was assumed that the second test portion (for enumeration) is independent of the first one (for detection), even if enumeration is performed only when detection is positive. An other assumption was made by Busschaert et al. (2010), who assumed that the results are completely correlated : if an enumeration result leads to 0 colony, detection and enumeration results are considered to be one interval-censored concentration datum between the LOD and the LOQ. In this paper, it is considered that there are two data, one above the LOD and the other below the LOQ. Thus, it is possible that a test portion has a lot of *L. monocytogenes* on it (for e.g. a colony) so that the detection result is “presence” whereas a test portion next to the first has no *L. monocytogenes* so the enumeration result is 0. The real concentration of both test portions is then high above the LOQ. This may happen because the food is solid. In reality, it is likely that a correlation exists between the results of two test portions from the same unit of food but it is neither 0 (our assumption), neither 1 (the assumption of Busschaert et al. (2010)). The ideal case is when both detection and enumeration can be performed on the same test portion, the only reliable way to derive interval-censoring. In Commeau et al. (2012a), we present additional Bayesian methods to account for variability between test portions within a food unit (and other levels of variability as well).

An assumption for both models \mathcal{M}_{CLD} and \mathcal{M}_{RD} is that the microbial concentration in a test portion is at least below the LOD but both models can handle pathogen free test portions. As the “concentration \times test portion mass = number of *L. monocytogenes*”, it is possible that a test portion has no *L. monocytogenes* while the concentration of the food sample is 0.01 cfu/g. In that case, if test portions have a mass of 10 g, there will be on average one test portion in ten contaminated with one cell. Another way to model 0 colony test portion is to consider a three-parameter model. The third parameter is the proportion of uncontaminated food in the product considered. This kind of model is presented in Gonzales-Barron et al. (2010a) with various distributions such as the Poisson and the negative binomial distributions.

The LOD and LOQ are fixed through the protocol. However, it is possible that a contaminated test portion with a concentration above the LOD (resp. above the LOQ) gets an absence result (resp. gets a 0 colony result). For instance, if a test portion of 25 g is contaminated with concentration equal to 0.1 cfu/g, the probability that a detection analysis leads to the result ‘absence’ is $\exp(-0.1 \times 25) = 0.08$. Model \mathcal{M}_{CLD} cannot take into account this situation whereas model \mathcal{M}_{RD} can : if a detection result $y = 1$ cfu, the corresponding unobserved log concentration λ can be below the LOQ.

Taking into account both detection and enumeration data is very important because often few data are available to assess a food-borne pathogen contamination. Models presented in this work can handle with both kinds of data. Model \mathcal{M}_{RD} together with the EM algorithm is a way to estimate the mean and the standard deviation of a contamination without being time consuming while being more accurate than the model \mathcal{M}_{CLD} .

4.2.5 Appendix A

When there is no zero count, both models \mathcal{M}_{CLD} and \mathcal{M}_{RD} have the same expectation and similar variances. The expectation and the variance of model \mathcal{M}_{CLD} are as follows :

$$E(c) = e^{\mu' + (\sigma')^2/2},$$

where $\mu' = \mu \ln(10)$ and $\sigma' = \sigma \ln(10)$, and

$$V(c) = e^{2\mu' + (\sigma')^2} (e^{(\sigma')^2} - 1).$$

For model \mathcal{M}_{RD} , the expectation of $K = \frac{y}{LOQ}$ is as follows :

$$\begin{aligned} E(K) &= \frac{E(y)}{LOQ} \\ &= \frac{1}{LOQ} E(E(y|\lambda)) \\ &= E(\lambda) = e^{\mu' + (\sigma')^2/2}, \end{aligned}$$

thus $E(c) = E(K)$. The variance of K is equal to the following :

$$\begin{aligned} V(K) &= \frac{V(y)}{(LOQ)^2} \\ &= \frac{1}{(LOQ)^2} (E(V(y|\lambda)) + V(E(y|\lambda))) \\ &= \frac{1}{(LOQ)^2} (E(\lambda LOQ) + V(\lambda LOQ)) \\ &= \frac{1}{(LOQ)^2} (LOQE(\lambda) + (LOQ)^2 V(\lambda)) \\ &= \frac{e^{\mu' + (\sigma')^2/2}}{LOQ} + e^{2\mu' + (\sigma')^2} (e^{(\sigma')^2} - 1) \\ &= \frac{e^{\mu' + (\sigma')^2/2}}{LOQ} + V(c). \end{aligned}$$

For a given value of σ , when μ is sufficiently high so that the probability of having zero counts is negligible, the term $\frac{e^{\mu' + (\sigma')^2/2}}{LOQ}$ is much smaller than $V(c)$. In fact, the limit of $V(K)/V(c)$, as μ approaches infinity is unity. In this case, models \mathcal{M}_{CLD} and \mathcal{M}_{RD} are similar because their first two moments are almost equal.

Chapitre 5

Modélisation de la contamination par *L. monocytogenes* lors de la fabrication industrielle de lardons

Ce chapitre présente les modifications d'un modèle décrivant la contamination par *L. monocytogenes* durant le procédé de fabrication de lardons (depuis l'arrivage jusqu'à la sortie usine). Ce modèle présente deux insuffisances qui n'avaient pas pu être surmontées lors de son élaboration faute de données :

- la distribution extrêmement piquée de la contamination à l'arrivée des poitrines de porc qui ne nous semblait pas refléter la réalité ;
 - la non prise en compte de la variabilité inter-lots alors que les résultats d'analyses sur la flore lactique (partie 2.3.3.a) indiquent que cette variabilité est importante.
- Le modèle a été remanié sur ces deux points.

5.1 Modélisation initiale de la contamination

Dans le projet Quant'HACCP la modélisation de la contamination par *L. monocytogenes* et par la flore lactique depuis les poitrines à leur arrivée à l'usine jusqu'au conditionnement a été réalisée parallèlement à l'acquisition de données de contamination de *L. monocytogenes* dans les poitrines et les lardons (cf. partie 2.3). Les objectifs étaient d'une part d'estimer la contamination actuelle dans l'usine et d'autre part de quantifier les modifications de distribution après avoir fait varier un ou plusieurs paramètres, comme par exemple une augmentation substantielle de la contamination de la matière première.

Ce modèle est publié dans Billoir et al. (2011) et reproduit à l'annexe F. Brièvement, le principe de la modélisation est le suivant : tout d'abord, nous faisons l'hypothèse que l'ensemble des poitrines contenues dans un malaxeur constituent un lot "statistique" et que ce lot est conservé jusqu'au conditionnement. Le modèle est composé de sous-modèles, chaque sous-modèle représentant une étape du procédé de fabrication (arrivage, malaxage, étuvage, cubage et conditionnement). La contamination initiale suit une certaine distribution qui évolue dans chaque sous-modèle car les sorties du sous-modèle m constituent les entrées du sous-modèle $m + 1$. En parallèle, la masse de l'entité (poitrine, lardons puis

barquette) est également caractérisée par une distribution à chaque étape.

Le DAG (cf. partie 1.4.11) du modèle est représenté à la figure 5.1.

5.2 Modification du modèle initial

Une fois le modèle opérationnel, des modifications ont été introduites pour plusieurs raisons (cf. tableau 5.1).

	Modèle de Billoir et al. (2011)	Modifications du modèle
Prise en compte de la variabilité inter-lots	Non	Oui
Distribution de la contamination initiale	Loi log-log-normale	Loi log-normale inter-lots et Poisson intra-lot
Données utilisées	Autocontrôles de 4 entreprises	Acquises lors du projet Quant'HACCP
Etape d'étuvage	Etape de réduction de la contamination	Etape de réduction ou de croissance de la contamination selon la température
Après la sortie d'usine	Pas de modélisation	Modélisation jusqu'à la consommation

TABLE 5.1 – Différences entre le modèle de Billoir et al. (2011) et le modèle modifié.

J'ai traité les trois premiers points et les modifications qui en découlent sont présentées dans ce chapitre. Dans le modèle initial, comme les données acquises dans le cadre du projet Quant'HACCP n'étaient pas encore disponibles, les données utilisées pour construire la loi *a priori* de la concentration^a $conc_i^A$ en *L. monocytogenes* dans la poitrine *i* à l'entrée de l'usine sont les données issues des contrôles de routine (détection dans 25 g de poitrine) effectués pas quatre usines. Il se trouve que 95 % des données sont des absences. Avec ces données peu informatives, il se trouve que la loi paramétrique trouvée pour $conc_i^A$ est une loi log-log-normale, c'est-à-dire que

$$\log(\log(conc_i^A) + tr_{\log(conc^A)})|_{i=1,\dots,N^A} \sim \mathcal{N}(m = \mu_{\log \log(conc^A)}, sd = \sigma_{\log \log(conc^A)}),$$

pour une poitrine *i* parmi les N^A poitrines composant le lot, avec $tr_{\log(conc^A)}$ un terme de translation de manière à ce que $\log(conc_i^A) + tr_{\log(conc^A)}$ soit positif.

Paradoxalement, la distribution de $conc_{Lm}^A$ est très piquée (cf. figure 3 de l'article Billoir et al. (2011)) alors que les données utilisées pour les estimées sont peu informatives. Cela entraîne une très grande homogénéité de la contamination initiale entre les poitrines. Or, cette contamination est encore plus homogénéisée après le malaxage, de par la nature de cette étape. Les distributions de la contamination à chaque étape sont donc également piquées. Comme la variabilité nous semblait sous-estimée, nous avons décidé d'introduire les données acquises lors du projet Quant'HACCP pour construire une loi plus réaliste

a. Les notations utilisées sont celles de l'article de Billoir et al. (2011).

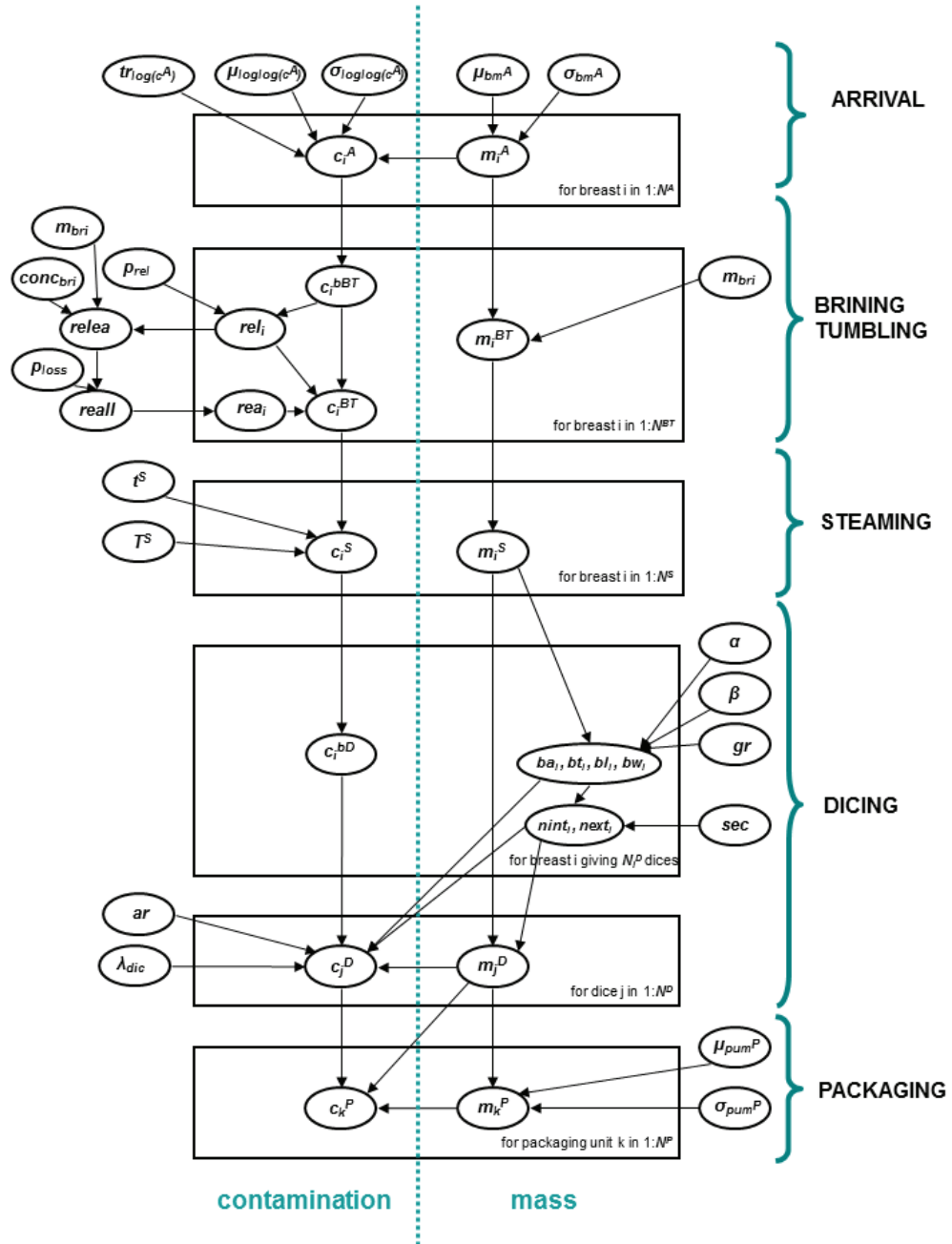


FIGURE 5.1 – Graphe du modèle décrivant le procédé de fabrication des lardons depuis les matières premières (poitrines de porc) jusqu’au produit fini (barquettes de lardons). La partie gauche de la figure décrit la contamination par *L. monocytogenes* et la partie droite la masse du produit (poitrine, lardon ou barquette) selon l’étape du procédé. C’est la figure 1 de l’article de Billoir et al. (2011).

pour $conc_i^A$: ces données sont plus informatives que les autocontrôles car les LOD et LOQ sont plus basses (LOD=0,01 ufc/cm² et LOQ=1 ufc/cm² pour les données de détection et de dénombrement de *L. monocytogenes* dans les poitrines après malaxage acquises lors de cette thèse ; LOD=0,04 ufc/g et LOQ=100 ufc/g pour les données de détection et de dénombrement de *L. monocytogenes* dans les poitrines lors de l'arrivage, données utilisées dans Billoir et al. (2011)). En outre, nous en avons profité pour modéliser la distribution de concentration de manière plus classique, c'est-à-dire par une loi log-normale. Par ailleurs, comme nous disposons alors de données sur les poitrines de porc réparties dans plusieurs lots de malaxage et de plusieurs données par lot, nous avons décidé d'introduire la notion de variabilité intra et inter-lots de malaxage. Ainsi, le modèle de Billoir et al. (2011) a été modifié comme suit à l'étape d'arrivée des poitrines de porc :

- nombre c_{ij}^A de *L. monocytogenes* en surface d'une poitrine de porc j de masse m_{ij}^A appartenant au lot de malaxage j :

$$c_{ij}^A \sim \mathcal{P}(conc_i^A \times m_{ij}^A), \quad (5.1)$$

- concentration moyenne $conc_i^A$ en *L. monocytogenes* dans le lot i :

$$\log(conc_i^A) \sim \mathcal{N}(\mu^A, \sigma^A). \quad (5.2)$$

L'étape "malaxage" a été modifiée uniquement par l'introduction du lot i . Les distributions sont donc les mêmes que dans Billoir et al. (2011) à ceci près que l'indice i dans l'article a été remplacé par l'indice ij :

- perte de rel_{ij}^{bBT} cellules de *L. monocytogenes* de la poitrine j lors du malaxage du lot i :

$$rel_{ij} \sim \mathcal{Bin}(c_{ij}^{bBT}, p_{rel}),$$

avec $c_{ij}^{bBT} = c_{ij}^A$ le nombre de *L. monocytogenes* présentes à la surface de la poitrine j juste avant le malaxage et p_{rel} la probabilité qu'une cellule en surface d'une poitrine soit relâchée dans le malaxeur ;

- nombre total $relea_i$ de cellules relarguées dans le malaxeur par les poitrines :

$$relea_i = \sum_{j=1}^{N_i^{BT}} rel_{ij} + c_{bri},$$

avec N_i^{BT} le nombre de poitrines dans le lot i lors du malaxage et c_{bri} le nombre de cellules apportées par la saumure ;

- nombre total $reall_i$ de cellules réallouées sur les poitrines (les autres partent dans le limon déposé sur les parois du malaxeur en fin de malaxage) :

$$reall_i \sim \mathcal{Bin}(relea_i, 1 - p_{loss}),$$

avec p_{loss} la probabilité qu'une bactérie soit prise dans le limon,

- nombre rea_{ij} de cellules récupérées par la poitrine j du lot i au cours du malaxage :

$$rea_{i1}, \dots, rea_{iN_i^{BT}} \sim \mathcal{MN}(reall_i, rep(\frac{1}{N_i^{BT}}, N_i^{BT})),$$

avec $rep(\frac{1}{N_i^{BT}}, N_i^{BT})$ le vecteur composé de N_i^{BT} fois la valeur $1/N_i^{BT}$;

- le nombre c_{ij}^{BT} de cellules à la surface de la poitrine j du lot i en fin de malaxage est égal au nombre initial de cellule moins le nombre de cellules perdues auxquels on ajoute le nombre de cellules déposées sur la poitrine :

$$c_{ij}^{BT} = c_{ij}^{bBT} - rel_{ij} + rea_{ij}.$$

Le graphe du modèle modifié est présenté à la figure 5.2. Seules les modifications aux étapes “arrivage” et “malaxage” sont considérées dans cette thèse. Les autres modifications sont présentées à l’annexe G.

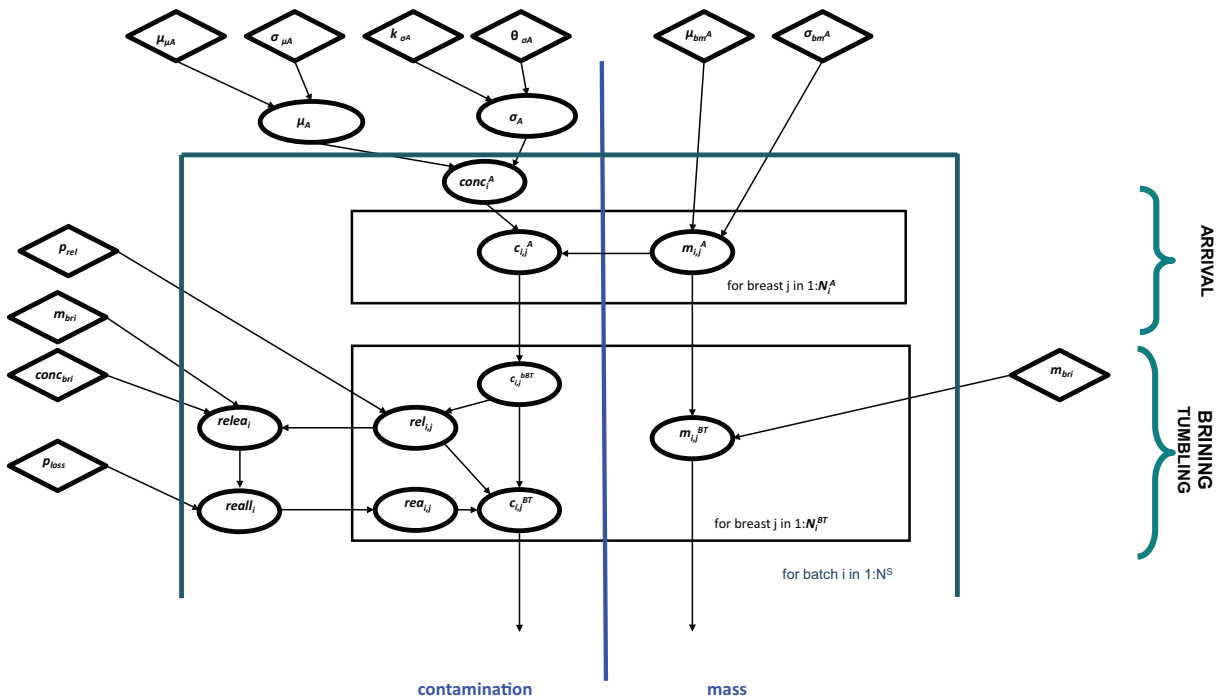


FIGURE 5.2 – Graphe du modèle modifié décrivant le procédé de fabrication des lardons depuis les matières premières (poitrines de porc) jusqu’à la fin du malaxage. A l’intérieur des chaque cadre noir est représenté ce qui se passe sur chaque poitrine d’un lot. Le cadre bleu en trait plus épais représente ce qui se passe dans un lot. Les losanges représentent les constantes des modèles.

5.3 Introduction des données pour construire les lois de μ^A et de σ^A

Une fois le modèle modifié, il reste à déterminer les lois de μ^A et de σ^A dans l’équation (5.2). Le principe retenu a été de considérer le modèle composé uniquement des deux premières étapes et d’injecter les données de détection et de dénombrement des poitrines après malaxage afin d’obtenir des lois *a posteriori* pour μ^A et σ^A . Ces lois seront ensuite utilisées pour simuler la concentration de chaque lot.

Les données ont été introduites au niveau de la variable c_{ij}^{BT} représentant le nombre de bactéries présentes en surface de la poitrine j du lot i en fin de malaxage. La distribution suivie par un résultat de détection x_{ij} (codé avec 0 pour “absence” et 1 pour “présence”) est une Bernoulli de paramètre $1 - \exp(-100/2382 \times c_{ij}^{BT})$. Ce paramètre est égal à la probabilité d’obtenir un résultat non nul (donc au moins une bactérie) en faisant l’hypothèse que la distribution à la surface des poitrines est une loi de Poisson de paramètre $100/2382 \times c_{ij}^{BT}$. Cette formule est une conversion du nombre de cellules c_{ij}^{BT} présente sur la surface de la poitrine divisé par la surface moyenne d’une poitrine, estimée à 2 382 cm² par Billoir et al. (2011). Comme la prise d’essai analysée est de 100 cm², la concentration $c_{ij}^{BT}/2382$ est multipliée par 100. Ainsi,

$$x_{ij} \sim \text{Bin}(1, 1 - \exp(-\frac{100}{2382} \times c_{ij}^{BT}))$$

La distribution suivie par un résultat brut de dénombrement y_{ij} est une loi de Poisson :

$$y_{ij} \sim \mathcal{P}(\frac{5}{2382} \times c_{ij}^{BT}).$$

La prise d’essai est de 100 cm² mais comme 5% du volume de la suspension mère a été prélevé (cf. partie 2.3.2), cela revient à considérer que le dénombrement a été fait sur 5 cm².

Des lois *a priori* peu informatives ont été choisies pour μ^A et pour σ^A (cf. tableau 5.2). L’inférence bayésienne a été réalisée au moyen du logiciel JAGS (version 1.0.3, Plummer (2008)). Au total, 10 000 000 itérations ont été effectuées et une valeur toutes les 1000 itérations a été prise en compte en raison d’une très forte autocorrélation. En effet, chaque lot de malaxage comporte 1000 poitrines, or nous ne disposons de données que pour 8 ou 9 d’entre elles selon les lots (cf. partie 2.3.3.b). Le modèle a été extrêmement long à tourner (plus d’un mois). Pour cette raison, une seule chaîne a été considérée au lieu de trois. La loi multinomiale a été remplacée par une collection de binomiales indépendantes car il n’y a pas d’échantillonneur dans JAGS^b pour cette loi (au moins dans la version utilisée à l’époque). Ainsi, $rea_{ij} \sim \text{Bin}(rea_i, 1/N_i^{BT})$. En outre, la binomiale a été utilisée uniquement pour les 8 ou 9 poitrines pour lesquelles nous disposons de données (donc pour les variables rea_{i1} à rea_{i8} ou rea_{i9}) et la loi binomiale m’a semblé une bonne approximation de la loi multinomiale puisque le nombre de données par lot est petit devant la taille du lot (j’ai utilisé les lois marginales des $rea_{i1}, \dots, rea_{iN_i^{BT}}$ au lieu de la loi jointe).

Quelques statistiques descriptives des lois *a posteriori* de μ^A et de σ^A sont présentées dans le tableau 5.2. Malgré le peu de données, l’incertitude sur ces deux variables a été réduite (cf. figure 5.3). La moyenne *a posteriori* de μ^A est de -2,0 log ufc/cm², ce qui signifie que chaque lot a une log concentration moyenne de -2,0 log ufc/cm² (c’est-à-dire une cellule par 100 cm² de surface). Cependant, la variabilité inter-lots (représentée par σ^A) est importante puisque dans 95% des lots, la log concentration moyenne varie entre $\mu^A - 1,6 \log \text{ ufc/cm}^2$ et $\mu^A + 1,6 \log \text{ ufc/cm}^2$.

b. Les logiciels WinBugs et OpenBugs ne conviennent pas non plus car le paramètre n de la multinomiale(n, p) doit être une constante alors que c’est la variable aléatoire $real_i$ dans le modèle.

Paramètre	dist. <i>a posteriori</i>	Statistiques descriptives de la dist. <i>a posteriori</i>				
		Moyenne	E.t.	2.5e perc.	50e perc.	97.5e perc.
μ^A	$\mathcal{N}(0, \sqrt{10})$	-2,0	0,24	-2,5	-2,00	-1,5
σ^A	$\sigma^{-2} \sim \mathcal{G}(0, 5; 0, 5)$	0,8	0,2	0,5	0,8	1,3

TABLE 5.2 – Distributions *a priori* et statistiques descriptives des lois *a posteriori* de la moyenne μ^A et l'écart-type σ^A exprimés en log ufc/g de la log concentration moyenne d'un lot.

“dist.” signifie “distribution”, “perc.” signifie “percentile” et “E.t.” “écart-type”.

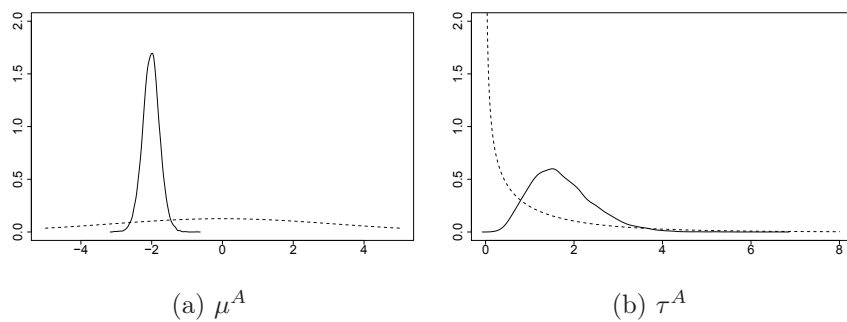


FIGURE 5.3 – Lois *a priori* (trait pointillé) et *a posteriori* (trait plein) de la moyenne μ^A et de la précision $\tau^A = 1/\sigma^{A2}$ de la log concentration moyenne d'un lot de malaxage.

5.4 Impact sur la distribution des la contamination après malaxage

Pour évaluer les conséquences des modifications apportées au modèle de Billoir et al. (2011) sur la contamination lors de l’arrivage et du malaxage, la log concentration en *L. monocytogenes* a été simulée pour 1000 lots de poitrines en tenant uniquement compte de la variabilité inter-lots et pas de l’incertitude (μ^A et σ^A sont fixés à leur moyenne *a posteriori*). Les fonctions de répartition de chaque lot sont représentées pour l’arrivée (figure 5.4a) et après malaxage (figure 5.4b). Ces figures sont à comparer aux figures obtenues avec le modèles de Billoir et al. (2011) (figures 5.4c et 5.4d).

La modification du modèle introduit une augmentation de la variabilité aussi bien intra-lot (la fonction de répartition pour chaque lot est plus “large”) qu’inter-lots (les fonctions de répartitions couvrent un intervalle plus grand que dans l’article de Billoir et al. (2011)). Ce résultat était attendu puisque la distribution de conc^A a une variance bien plus élevée dans le modèle de Billoir et al. (2011) modifié que dans le modèle initial (cf. figure 5.5b).

On s’attend à ce que l’étape de malaxage réduise la variabilité intra-lot puisque les poitrines de porcs sont frottées entre elles et s’échangent les bactéries présentes à leur surface. Or les fonctions de répartition des figures 5.4a et 5.4b sont quasiment identiques, ce qui, au passage, était déjà le cas dans le modèle initial (figures 5.4c et 5.4d). A mon avis, cela s’explique par le fait que la variabilité intra-lot lors de l’arrivage est modélisée par une loi de Poisson (équation (5.1)). La variabilité intra-lot est donc faible et la manière dont le malaxage est modélisé ne diminue pas cette variabilité. Le modèle pourrait être modifié en proposant une modélisation de la variabilité intra-lot plus large (par exemple par une loi Poisson-Log Normal au lieu d’une distribution de Poisson uniquement) mais cela nécessiterait d’estimer encore plus de paramètres, ce qui est difficile compte-tenu du petit nombre de données.

Les modifications du modèle de Billoir et al. (2011) nous ont permis :

- de modéliser la variabilité inter-lots ;
- de tenir compte des données de détection et de dénombrement de *L. monocytogenes* dans les poitrines de porc après malaxage acquises durant le projet Quant’HACCP pour inférer les paramètres de moyenne et d’écart-type de la log concentration moyenne d’un lot.

Avec le modèle modifié, il est donc possible d’étudier l’impact d’un scénario (par exemple, la durée de l’étuvage est réduite d’une heure) sur des lots différemment contaminés. Par ailleurs, les distributions des paramètres *a posteriori* reflètent mieux la contamination de l’usine que les valeurs de paramètres déduites dans le modèle de Billoir et al. (2011) car les données utilisées lors des modifications du modèle ont des LOD et des LOQ plus petites que les données utilisées pour l’inférence du modèle initial.

Le modèle tel que décrit dans Billoir et al. (2011) a été publié avant les modifications apportées car il formait un ensemble satisfaisant même sans les apports détaillés dans ce chapitre. De plus, ne pas avoir incorporé dans le modèle publié les données de poitrines de porc acquises dans le cadre du projet Quant’HACCP m’a permis d’utiliser ce modèle pour construire les lois *a priori* des modèles de contamination présentée au chapitre 3. Comme

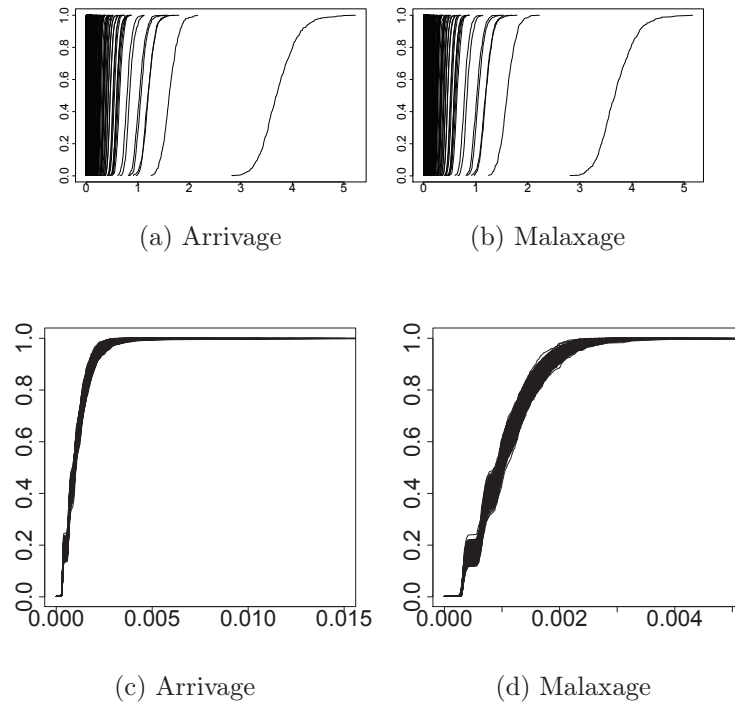


FIGURE 5.4 – Fonctions de répartition empiriques de la concentration en ufc/g de *L. monocytogenes* pour 1000 lots de poitrines lors de l’arrivage (figures de gauche) et après malaxage (figures de droite). Les fonctions de répartition sont issues du modèle modifié (figures du haut) et du modèle initial (figure du bas).

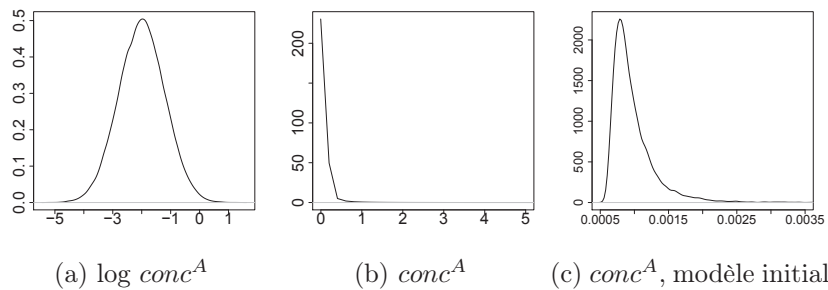


FIGURE 5.5 – Distribution de la log concentration en log ufc/g (a) et de la concentration en ufc/g (b) $conc^A$ moyenne dans un lot de malaxage du modèle modifié et distribution de la concentration en ufc/g pour le modèle initial (c).

j'ai ensuite utilisé ces mêmes données pour mener l'inférence, ces données auraient été utilisées deux fois si j'avais construit les lois *a priori* avec le modèle modifié.

Dans cette première partie de thèse, nous avons présenté des données microbiologiques acquises par nous ou récupérées d'un projet de recherche datant de quelques années et nous avons élaboré plusieurs types de modèles de contamination. Le premier type décrit la contamination à une étape donnée du procédé ou de la vie d'un produit (modèles présentés aux chapitres 3 et 4). Le second type décrit le devenir de la contamination durant tout le procédé (chapitre 5) d'un produit.

Le premier type de modèle peut être utile lors de l'élaboration d'un plan d'échantillonnage car il permet d'adapter ce dernier à la structure de contamination. En effet, s'il y a une structure de lot, échantillonner un lot sur deux par exemple donnera peu ou pas d'information sur les lots non échantillonnés. En revanche, si aucune structure n'est apparente, alors l'échantillonnage d'un lot sur deux renseignera aussi sur la contamination des lots non retenus pour l'analyse. Ces modèles peuvent également être utilisés lors de l'étape d'appréciation de l'exposition d'une appréciation quantitative des risques (AQR), qui est composée d'une étude de la contamination d'un aliment par un pathogène et d'une étude de la consommation de cet aliment.^c En effet, la distribution de la contamination des produits est différente selon le modèle considéré (avec structure d'unité, de lot, etc.). Ne pas prendre en compte la structure adéquate peut conduire à une mauvaise estimation de la distribution de contamination.

Le second type de modèle peut être utile pour identifier les étapes du procédé susceptibles d'avoir un impact déterminant sur la contamination. Il permet également d'envisager plusieurs scénarios possibles de variation de procédé et d'étudier les impacts sur la contamination, comme cela a été fait dans l'article de Billoir et al. (2011).

Ces types de modèles peuvent être utiles pour relier les états de la nature aux observations lors de l'étude de plans d'échantillonnage, ce qui constitue l'une des étapes pour appliquer la théorie de la décision. Plus simplement, ces modèles permettent également de simuler des résultats obtenus avec différents plans d'échantillonnage.

c. L'AQR est une méthode permettant d'estimer le risque du nombre de malades ou de morts liés à la consommation d'un aliment et à sa contamination par un pathogène. Elle est divisée en quatre étapes : identification des dangers, appréciation de l'exposition, l'appréciation des effets et l'estimation du risque. Pour plus de détails, consulter Codex Alimentarius (1999).

Deuxième partie

AIDE À LA DÉCISION EN UTILISANT L'ÉCHANTILLONNAGE

Cette seconde partie présente plusieurs exemples d'aide à la décision en utilisant des plans d'échantillonnage. Dans les deux premières applications, concernant la fabrication de lardons et de saumon fumé, le plan d'échantillonnage est utilisé pour la surveillance microbiologique (en l'occurrence, de *L. monocytogenes*). Elles permettent de confronter la théorie bayésienne de la décision, présentée à la partie 1.3, à la pratique. Dans le dernier exemple, fabrication de bœuf en sauce dans la cuisine collective d'un hôpital, la surveillance porte sur la température de la viande au cours du temps entre la fin de l'élaboration du plat et la sortie de la cellule de refroidissement. Ce cas illustre les conséquences de plusieurs stratégies de surveillance sur la croissance ultérieure du pathogène d'intérêt (ici *C. perfringens*) et n'a pas recours à la théorie bayésienne de la décision.

Tout comme dans la première partie, la notion de lot est présente puisqu'un échantillonnage est effectué dans chaque lot. Le tableau 5.3 détaille les thématiques abordées dans les chapitres constituant cette partie.

TABLE 5.3 – Thématiques de la seconde partie et chapitres dans lesquels elles sont abordées.

	Chapitre 6	Chapitre 7	Chapitre 8
Enjeu de la surveillance	Vérification du PMS	Contrôle libératoire	Pilotage du procédé
Population d'intérêt (=lot)	Production sur une période de temps	Production d'une journée	Production contenue dans un chariot
Objet surveillé	Prévalence de <i>L. monocytogenes</i> dans 25 g de lardons en sortie usine	Prévalence de <i>L. monocytogenes</i> dans 25 g de saumon fumé en sortie usine	Température de la viande dans un chariot durant le refroidissement rapide
Méthode d'aide à la décision	Théorie bayésienne de la décision		Comparaison de stratégies vis-à-vis d'un critère adapté

Chapitre 6

Optimisation du plan d'échantillonnage appliquée à la fabrication de lardons

Ce chapitre présente l'application de la théorie bayésienne de la décision : quelle taille d'échantillon adopter de manière à ce que le coût moyen supportés par en entreprise fabriquant des lardons soit le plus faible possible ?

6.1 Description de la surveillance adoptée dans une usine

A l'annexe C, j'ai présenté un contrôle de type libératoire pour chaque lot c'est-à-dire que l'usine attend d'avoir les résultats concernant le lot avant de décider de le vendre ou de le détruire. Dans l'usine fabriquant des lardons que nous avons contactée, le contrôle est d'un autre ordre car tous lots "industriels" (cf. partie 2.2.2 pour sa définition) ne sont pas analysés. Par ailleurs, les lots analysés ont déjà quitté l'usine lorsque les résultats sont disponibles. Dans ces conditions, il est inapproprié d'appliquer la même fonction de coût que dans l'exemple de l'annexe C.

Dans cette application, la première tâche a été de définir l'entité sur laquelle porte l'échantillonnage. En considérant la pratique de l'entreprise d'une part et les différents usages des plans d'échantillonnage d'autre part, nous avons décidé de modéliser le plan d'échantillonnage comme outils de vérification du PMS^a durant une période donnée, par exemple un ou six mois. Le principe retenu est le suivant : n échantillons sont prélevés durant la période définie. La décision porte sur des résultats acquis durant cette période, nous avons donc redéfini le lot pour l'échantillonnage : il s'agit de la production durant laquelle n échantillons sont analysés. La valeur numérique de n est déterminée par la théorie bayésienne de la décision. Le moyen de la déterminer est le sujet de ce chapitre.

A la fin d'une période, les résultats sont analysés. En fonction de ces derniers, l'entreprise peut prendre plusieurs décisions, lesquelles, pour les besoins de la modélisation, ont été réduites à trois :

- ne rien faire car le PMS est vérifié ;
- entreprendre une action corrective mineure car le PMS n'est pas vérifié mais le niveau de contamination n'est pas très élevé ;

a. PMS signifie plan de maîtrise sanitaire et est défini à la partie 1.1.3.

- entreprendre une action corrective majeure car le PMS n'est pas vérifié et le niveau de contamination est élevé.

Limiter le nombre de décisions à trois nous a semblé pertinent par réalisme pour tenir compte de plusieurs décisions sans pour autant en avoir trop (parcimonie). Nous avons également opté pour des actions correctives fixes (quand l'entreprise décide d'adopter une action corrective mineure par exemple, c'est toujours la même) pour que le coût de la décision soit toujours le même.

6.2 Optimisation à partir de la prévalence

6.2.1 Loi *a priori*, vraisemblance, ensemble des décisions et fonction de coût

Les modèles de contamination décrits aux chapitres 3 et 4 portent sur la concentration d'un pathogène dans une prise d'essai, une unité ou un lot. Il est également possible de modéliser la contamination par sa prévalence. Comme nous partons du principe que la prévalence est estimée grâce à des analyses de détection sur 25g de produit, la prévalence théorique d'un lot est donc le nombre de morceaux de 25 g contaminés divisé par le nombre total de morceaux de 25 g composants un lot. Ne connaître que la prévalence d'un lot et non sa distribution de concentration conduit à une perte d'information. En effet, si une personne consomme un produit contaminé, elle peut tomber malade et la probabilité de ce phénomène dépend de la quantité de bactéries pathogènes ingérées, information apportée par la concentration et non la prévalence. De plus, on peut très bien imaginer un lot où toutes les parties sont contaminées avec une concentration faible (prévalence égale à 1) ou, au contraire, un lot où presque toutes les parties sont non contaminées et une toute petite fraction avec une concentration élevée en pathogènes (prévalence proche de 0). Evidemment, ces cas sont extrêmes. Travailler avec la prévalence est constituée une perte d'information sur la contamination, cependant, cela est couramment utilisé car les contrôles effectués par les entreprises prennent souvent la forme d'analyse de détection plutôt que de dénombrement, notamment parce que la LOD est plus basse que LOQ d'après les protocoles décrits dans les normes NF EN ISO11290-1 (1996a) et NF EN ISO11290-2 (1998a) (la LOD vaut 0,04 ufc/g et la LOQ vaut 10 ou 100 ufc/g selon le protocole utilisé). Enfin, il est pratique d'utiliser la prévalence au lieu de la concentration : si on choisit une loi beta comme loi *a priori* et une loi binomiale pour modéliser les résultats de détection, alors la loi *a posteriori* de la prévalence est conjuguée : c'est également une loi beta de paramètres connus.

Dans ce chapitre, nous travaillerons donc à partir de la prévalence. Elle est notée ψ et varie d'une période à l'autre par nature, il s'agit donc de déterminer une loi de probabilité caractérisant les prévalences des différentes périodes. La loi Beta est choisie comme loi *a priori* de ψ car sa forme est très souple selon les valeurs des paramètres. Pour l'instant, on note simplement :

$$[\psi] = \mathcal{Beta}(a, b).$$

Puisque seule la prévalence est considérée, il est naturel d'effectuer des analyses de détection sur les échantillons. On note x le nombre de résultats positifs parmi les n

analysés. La loi de x choisie est une binomiale :

$$[x|\psi] = \text{Bin}(n, \psi).$$

En toute rigueur, il aurait fallu choisir la loi hypergéométrique puisque les prises d'essai sont prélevées dans le lot et n'y sont pas remises. Cependant, comme n (le nombre de portions prélevées) est négligeable devant le nombre de portions de 25 g constituant le lot. Trois décisions sont possibles :

$$\begin{cases} d_0 : \text{PMS vérifié} \\ d_1 : \text{PMS non vérifié, une action corrective mineure est apportée} \\ d_2 : \text{PMS non vérifié, une action corrective majeure est apportée} \end{cases}$$

Dans cette application, nous avons choisi de modéliser la fonction de coûts par les coûts monétaires des décisions et des conséquences possibles. Dans ce cadre, il nous faut donc définir les coût des actions correctives apportées.

L'entreprise vend sa production à un client qui est, dans cette application, un distributeur. On suppose que ce dernier connaît la prévalence du lot. En fonction de ψ et de la décision prise par l'entreprise, le client peut imposer une pénalité s'il considère que le cahier des charges n'est pas respecté. La modélisation proposée peut également s'étendre au cas où le client est également un industriel du secteur agro-alimentaire qui utiliserait les lardons comme matière première, par exemple pour fabriquer des quiches. Nous faisons ici le choix de modéliser ces pénalités de manière non continue par rapport à la prévalence : définissons deux seuils ψ_0 et ψ_1 tels que la prévalence est considérée comme

- faible si $\psi \in \Psi_0 = [0; \psi_0]$;
- moyenne si $\psi \in \Psi_1 =]\psi_0; \psi_1]$;
- élevée si $\psi \in \Psi_2 =]\psi_1; 1]$.

Les valeurs de ψ_0 et de ψ_1 peuvent être fixées en lien avec des exigences réglementaires. Des exemples sont données partie 6.2.5. De manière générale, la fonction de coût est définie dans neuf cas différents, c'est-à-dire pour toutes les combinaisons possibles des valeurs prises par la décision de l'entreprise et l'intervalle dans lequel se situe la prévalence (cf. Tableau 6.1).

Décisions		d_0	d_1	d_2
ψ				
$\psi \in \Psi_0$		$C(0, 0)$	$C(0, 1)$	$C(0, 2)$
$\psi \in \Psi_1$		$C(1, 0)$	$C(1, 1)$	$C(1, 2)$
$\psi \in \Psi_2$		$C(2, 0)$	$C(2, 1)$	$C(2, 2)$

TABLE 6.1 – Valeurs prises par la fonction de coût L .

Si la prévalence était connue avec certitude de la part de l'entreprise, alors la meilleure règle de décision δ' à prendre serait :

- $\delta'(\psi) = d_0$ si $\psi \in \Psi_0$;
- $\delta'(\psi) = d_1$ si $\psi \in \Psi_1$;
- $\delta'(\psi) = d_2$ si $\psi \in \Psi_2$.

Autrement dit, les coûts les plus faibles ligne par ligne doivent se trouver sur la diagonale du tableau 6.1. La *fonction de regret* par rapport au coût optimal dû à la méconnaissance de ψ (cf. partie 1.3.1.b) s'écrit $\underline{L}(\psi, \delta(x)) = L(\psi, \delta(x)) - L(\psi, \delta'(\psi))$. Si on adopte cette fonction, on cherche la décision δ définie de \mathfrak{X} dans \mathcal{D} qui minimise \underline{L} . La décision qui minimise L minimise également \underline{L} . Le tableau 6.2 reprend le tableau 6.1 mais avec les coûts exprimés en terme de regret.

		Décisions		
		d_0	d_1	d_2
ψ	$\psi \in \Psi_0$	0	$K(0, 1)$	$K(0, 2)$
	$\psi \in \Psi_1$	$K(1, 0)$	0	$K(1, 2)$
	$\psi \in \Psi_2$	$K(2, 0)$	$K(2, 1)$	0

TABLE 6.2 – Coûts exprimés en terme de regret pris par la fonction de regret \underline{L} . Pour i et j variant de 0 à 2, le regret $K(i, j)$ vaut $C(i, j) - C(i, i)$, avec $C(i, j)$ le coût défini au tableau 6.1.

6.2.2 Calcul du risque bayésien

La fonction de regret $\underline{L}(\psi, \delta_n(x))$ s'écrit

$$\begin{aligned} \underline{L}(\psi, \delta(x)) &= Cn + \mathbf{1}_{\delta(x)=d_0}(K(1, 0)\mathbf{1}_{\psi \in \Psi_1} + K(2, 0)\mathbf{1}_{\psi \in \Psi_2}) \\ &\quad + \mathbf{1}_{\delta(x)=d_1}(K(0, 1)\mathbf{1}_{\psi \in \Psi_0} + K(2, 1)\mathbf{1}_{\psi \in \Psi_2}) \\ &\quad + \mathbf{1}_{\delta(x)=d_2}(K(0, 2)\mathbf{1}_{\psi \in \Psi_0} + K(1, 2)\mathbf{1}_{\psi \in \Psi_1}). \end{aligned}$$

Cette fonction tient compte du coût C d'une analyse de détection. Ici, les données n'ont pas encore été observées et leur observation dépend de la taille n de l'échantillon puisque celle-ci est égale au nombre de données (c'est en fait l'expérimentation e définie partie 1.3.1.b) puisque $\mathfrak{X} = [0; n]$. La règle de décision dépend de n et sera notée dorénavant δ_n . La fonction $\underline{L}(\psi, \delta_n(x))$ intégrée selon ψ se calcule aisément :

$$\begin{aligned} \int \underline{L}(\psi, \delta_n(x))[\psi|x, n]d\psi &= Cn + \mathbf{1}_{\delta_n(x)=d_0}(K(1, 0)P_1 + K(2, 0)P_2) \\ &\quad + \mathbf{1}_{\delta_n(x)=d_1}(K(0, 1)P_0 + K(2, 1)P_2) \\ &\quad + \mathbf{1}_{\delta_n(x)=d_2}(K(0, 2)P_0 + K(1, 2)P_1), \end{aligned}$$

avec $P_i = \mathbb{P}(\psi \in \Psi_i|x, n)$ ($i = 0, \dots, 2$) la probabilité *a posteriori* que ψ soit comprise dans l'intervalle Ψ_i . La probabilité P_0 vaut :

$$\begin{aligned} P_0 &= \int_0^{\psi_0} [\psi|x, n]d\psi \\ &= I_{\psi_0}(a + x, b + n - x), \end{aligned}$$

où $I_\alpha(k, l)$ est la fonction de répartition d'une loi beta de paramètres k et l calculée en α . De même,

$$\begin{aligned} P_1 &= I_{\psi_1}(a+x, b+n-x) - I_{\psi_0}(a+x, b+n-x) \\ P_2 &= 1 - I_{\psi_1}(a+x, b+n-x). \end{aligned}$$

La fonction de regret intégrée sur ψ pour les trois décisions d_0, d_1 et d_2 s'écrit :

$$\begin{aligned} \int \underline{L}(\psi, \delta_n(x) = d_0)[\psi|x, n]d\psi &= Cn + K(1, 0)P_1 + K(2, 0)P_2 \\ \int \underline{L}(\psi, \delta_n(x) = d_1)[\psi|x, n]d\psi &= Cn + K(0, 1)P_0 + K(2, 1)P_2 \\ \int \underline{L}(\psi, \delta_n(x) = d_2)[\psi|x, n]d\psi &= Cn + K(0, 2)P_0 + K(1, 2)P_1. \end{aligned}$$

Si on minimise l'espérance du regret, alors la décision d_0 est choisie lorsque la fonction de regret intégrée selon ψ et prise en d_0 est plus petite que celles prises en d_1 et en d_2 . Les valeurs de $\underline{L}(\psi, d_0)$, $\underline{L}(\psi, d_1)$ et de $\underline{L}(\psi, d_2)$ dépendent de n , des seuils ψ_0 et ψ_1 .

On souhaite montrer que choisir une décision se fait en fonction de la valeur de x :

$$\begin{cases} \delta_n(x) = d_0 & \Leftrightarrow x \leq c_1 \\ \delta_n(x) = d_1 & \Leftrightarrow c_1 < x \leq c_2 \\ \delta_n(x) = d_2 & \Leftrightarrow x > c_2, \end{cases}$$

avec c_1 et c_2 deux seuils tels que $c_1 < c_2$. Un ou plusieurs seuils peuvent ne pas être définis. Par exemple, si la décision d_2 n'est jamais prise lorsque x varie de 0 à n , le seuil c_2 n'est pas défini. Les seuils c_1 et c_2 , s'ils existent, sont déterminés pour chaque valeur de n , il faudrait donc les noter $c_1(n)$ et $c_2(n)$ ce qui n'est pas fait ici pour ne pas alourdir les notations. Si c_1 existe, c'est la valeur telle que $\underline{L}(\psi, \delta_n(c_1) = d_0) \leq \underline{L}(\psi, \delta_n(c_1) = d_1)$ et $\underline{L}(\psi, \delta_n(c_1 + 1) = d_0) > \underline{L}(\psi, \delta_n(c_1 + 1) = d_1)$ avec ces mêmes probabilités calculées avec $x = c_1 + 1$. Le calcul de c_2 se fait de la même manière avec $\underline{L}(\psi, d_1)$ et $\underline{L}(\psi, d_2)$.

Faute de démontrer la forme de la décision dans le cas général, on se contentera de deux cas : (1) n très supérieur à a et b et (2) a et b très supérieurs à n .

(1) Si n est très supérieur à a et b , lorsque x est petit, la distribution beta *a posteriori* de ψ est concentrée vers 0 donc P_0 s'approche de 1. Comme $\underline{L}(\psi, d_0)$ est le seul regret sans P_0 , il est le plus petit donc la décision d_0 est prise. Lorsque x augmente, le poids de la distribution *a posteriori* de ψ se concentre entre les valeurs de ψ_0 et ψ_1 donc P_1 tend vers 1 donc la décision d_1 est choisie car $\underline{L}(\psi, d_1)$ ne dépend pas de P_1 . Lorsque x est proche de n , la distribution *a posteriori* de ψ est concentrée en 1 donc P_2 en vient à valoir 1 et la décision d_2 est prise. Dans ce cas, le basculement d'une décision à l'autre dépend de la position de x par rapport à c_1 et c_2 .

(2) Si a et b sont très supérieurs à n , x n'a que peu d'influence sur la loi *a posteriori* de ψ donc les valeurs de P_0, P_1 et P_2 sont constantes ou presque quelle que soit la valeur de x . Dans ce cas, le même regret est toujours plus petit que les deux autres donc une seule décision est prise; les seuils c_1 et c_2 ne sont pas définis.

Une fois la fonction de regret intégré en ψ , il reste à l'intégrer en x pour calculer le risque bayésien. Pour cela, il faut déterminer la loi marginale de x . Dans le cas présent, cela se fait sans difficulté en utilisant le théorème de Bayes :

$$\begin{aligned} [x|n] &= \frac{[\psi][x|\psi, n]}{[\psi|x, n]} \\ &= \frac{\Gamma(a+b)\Gamma(a+x)\Gamma(b+n-x)\Gamma(n+1)}{\Gamma(a)\Gamma(b)\Gamma(a+b+n)\Gamma(n-x+1)\Gamma(x+1)}. \end{aligned}$$

On reconnaît la loi dite de Polya, que l'on notera $\mathcal{Po}(a, b, n)$ ^b. Les seuils c_1 et c_2 prennent chacun une valeur par valeur de n donc l'intégrale de la fonction de regret selon ψ et x (i.e. le risque bayésien) s'écrit :

$$\begin{aligned} \underline{r}(\delta_n) &= Cn + \int \left(\int \underline{L}(\psi, \delta_n(x))[\psi|x, n]d\psi \right) [x|n]dx \\ &= Cn + \int (1_{\delta_n(x)=d_0}(K(1,0)P_1 + K(2,0)P_2) \\ &\quad + 1_{\delta_n(x)=d_1}(K(0,1)P_0 + K(2,1)P_2) + 1_{\delta_n(x)=d_2}(K(0,2)P_0 + K(1,2)P_1)) [x|n]dx \\ &= Cn + \sum_{x=0}^{c_1} (K(1,0)P_1 + K(2,0)P_2)[x|n] \\ &\quad + \sum_{x=c_1+1}^{c_2} (K(0,1)P_0 + K(2,1)P_2)[x|n] + \sum_{x=c_2+1}^n (K(0,2)P_0 + K(1,2)P_1)[x|n] \end{aligned}$$

Une fois intégré selon ψ et x , le risque ne dépend plus que de n . Il est donc possible de déterminer la taille n optimale de l'échantillon pour laquelle le risque est le plus petit. Pour mener le calcul à son terme, il faut déterminer les coûts tels que présentés au tableau 6.1.

6.2.3 Fonction de coût adaptée à l'application

Les pénalités du client imposées à l'entreprise dépendent de l'intervalle dans lequel ψ se trouve et de la décision prise par l'entreprise. Nous avons défini la fonction L dépendant de $\delta_n(x)$ et de ψ de la manière suivante :

- $\psi \in \Psi_0$ (situation sous contrôle)
 1. la décision d_0 est prise, ce qui ne coûte rien à l'entreprise ;
 2. la décision d_1 est prise : l'usine mène une action corrective mineure, ce qui lui coûte r ;
 3. la décision d_2 est prise : l'usine mène une action corrective majeure, ce qui lui coûte R ;
- $\psi \in \Psi_1$ (contamination moyenne)

b. La loi dite de Polya est la loi indiquant la probabilité qu'il y ait $a+x$ boules rouges au n^e tirage dans une urne qui comptait initialement a boules rouges et b boules noires. A chaque tirage, une boule est extraite de l'urne puis remise dedans avec une boule supplémentaire de la même couleur.

ψ	Décisions		
	d_0	d_1	d_2
$\psi \in \Psi_0$	$0 + 0$	$r + 0$	$R + 0$
$\psi \in \Psi_1$	$0 + pb$	$r + pb_1$	$R + pb_2$
$\psi \in \Psi_2$	$0 + PB$	$r + PB_1$	$R + PB_2$

TABLE 6.3 – Valeurs de la fonction de coût L en fonction de la décision d_0 , d_1 ou d_2 prise et de l'intervalle dans lequel se trouve la prévalence ψ de la période de production considérée. Dans chaque case, le coût de gauche est le coût lié à la décision de l'entreprise et le coût de droite est la pénalité imposée par le client. Le coût d'échantillonnage n'a pas été indiqué, il est bien sûr le même dans chaque case et vaut Cn .

1. la décision d_0 est prise, ce qui coûte cher à l'entreprise car elle vend des produits contaminés et ne s'en est pas rendue compte. Elle doit donc faire face à des coûts supplémentaires que l'on note pb pour l'instant et que l'on peut imaginer dépendants ou non de ψ , par exemple une pénalité imposée par le client qui a détecté des produits positifs ;
 2. la décision d_1 est prise, l'usine mène donc une action corrective mineure (coût r) mais elle subit également des coûts supplémentaires car elle a sous-estimé la contamination, ce qui coûte $pb_1 < pb$ car le client impose moins de pénalité que dans le cas où l'entreprise n'a rien détecté ;
 3. la décision d_2 est prise, l'usine mène alors une action corrective majeure (coût R) et elle subit également des coûts supplémentaires car elle a sous-estimé la contamination, ce qui coûte $pb_2 < pb_1$, car l'entreprise a entrepris une action corrective majeure. On peut donc imaginer qu'il lui sera plus facile de négocier avec son client pour que les pénalités ne soient pas trop élevées.
- $\psi \in \Psi_2$ (haut niveau de contamination)
1. la décision d_0 est prise, ce qui coûte très cher à l'entreprise car elle a vendu des produits très contaminés et ne s'en est pas rendue compte. Elle doit donc faire face à des coûts supplémentaires que l'on note PB (également dépendants ou non de la valeur de ψ , par exemple une pénalité imposée par le client qui a détecté des produits positifs) ;
 2. la décision d_1 est prise : l'usine ne mène qu'une action corrective mineure (coût r). Elle subit également des coûts supplémentaires car elle a sous-estimé la contamination, ce qui coûte néanmoins $PB_1 < PB$ car le client impose moins de pénalité que dans le cas où l'entreprise n'a rien détecté ;
 3. la décision d_2 est prise, alors l'usine doit mener une action corrective majeure (coût R) et subit également un coût supplémentaire. Mais elle n'a pas sous-estimé la contamination, néanmoins le client impose des pénalités d'un montant PB_2 inférieur à PB_1 , car le client tient compte de l'action corrective majeure mise en œuvre par l'entreprise.

Les valeurs prises par L sont résumées dans le tableau 6.3.

Elicitation de la fonction de coût Eliciter la fonction de coût signifie déterminer, le plus souvent à partir d'un entretien avec un expert, les valeurs des coûts pour chaque condition (décision prise et segment d'appartenance de ψ). Les questions peuvent être difficiles (comme par exemple évaluer la proportion d'une population consommant un aliment donné, Ross and Sumner (2002)), de plus on demande souvent à l'expert de donner également des bornes maximale et minimale ainsi que des valeurs de percentiles (O'Hagan et al. (2006)). L'élicitation de coût peut se faire grâce à des estimations d'experts (Gorton et al. (1997)), par des sondages pratiqués dans la population (Teisl and Roe (2010)) ou même par consultation de revues spécialisées ou de rapports officiels (van der Gaag et al. (2004)).

Nous avons eu recours à un expert de l'Institut français de l'industrie du porc (IFIP). L'exercice était double : il s'agissait de définir les termes (action corrective mineure, action corrective majeure, pénalités) et de déterminer leurs coûts. L'action pouvait recouvrir un ensemble de mesures à condition que cet ensemble soit toujours le même chaque fois que la décision correspondante était prise. De même, la pénalité encourue dans un cas de figure pouvait être un ensemble de pénalités pourvu que cet ensemble soit le même à chaque fois que le cas de figure se produit. L'expert devait ensuite fixer les coût de chaque action et de chaque pénalité du client. Nous n'avons pas réussi à interroger directement un responsable qualité de l'usine avec laquelle nous avons été en contact, qui nous a fourni des informations pour la modélisation du procédé (cf. chapitre 5) et chez laquelle nous avons acquis des données de contamination (cf. partie 2.3.2). L'expert qui nous a répondu ne connaissait pas précisément les coûts de chaque action, cependant il avait une bonne vision de la production des lardons en France : les coûts qu'ils nous a énoncé ne sont donc pas propres à une usine mais reflètent une situation plausible dans ce secteur.

Action corrective mineure (coût r) Cette action a été décomposée par l'expert en deux mesures. D'une part un nettoyage renforcé est effectué pendant une semaine. Deux produits sont utilisés (un détergent et un désinfectant) par deux personnes pendant une heure. On estime le coût supplémentaire à 50 € par heure, soit 250 € au total pour la semaine.

D'autre part, 20 lots de chaque fournisseur sont contrôlés afin de déréférencer les fournisseurs dont les poitrines de porc sont les plus contaminées. En prenant l'hypothèse que l'entreprise a 10 fournisseurs, le coût d'analyses supplémentaires s'élève à 20 analyses \times 20 € \times 10 fournisseurs = 4000 €.

Ainsi, $r = 4000 + 250 = 4250$ €.

Action corrective majeure (coût R) L'usine est fermée pendant 24 h pour démontage des appareils et nettoyage. Elle perd donc une journée de production. L'usine produit 500 tonnes de lardons par an et travaille 250 jours, elle fournit en moyenne 2 tonnes de lardons par jour. Si le coût de vente des lardons est de 7€/kg et le coût de la matière première de 2€/kg, la perte s'élève à 2000 kg \times 5€/kg = 10000€. Par ailleurs, le nettoyage très renforcé dure 8h et occupe 10 personnes. Le coût est donc de 8h/pers \times 10 personnes \times 50€/h=4000 €. Soit un total de $R = 10000 + 4000 = 14000$ €.

Le client détecte que ψ appartient à Ψ_1 Si le client détecte que la prévalence est moyenne et que l'usine n'a pris aucune action corrective, alors il impose une pénalité fixe à l'usine pour livraison d'une marchandise contaminée d'un montant de 4000 €. En outre, le client procède à 10 analyses supplémentaires à la charge de l'entreprise soit 20€/analyse \times 10 analyses = 200 €. Ainsi, $pb = 4000 + 200 = 4200$ €.

Si l'usine a mené une action corrective mineure, le client maintient la pénalité mais ne fait que 5 analyses supplémentaire, donc $pb_1 = 4000 + 100 = 4100$ €.

Enfin, si une action corrective majeure a été menée, le client ne procède pas à des analyses supplémentaires, donc $pb_2 = 4000$ €.

Le client détecte que ψ appartient à Ψ_2 Si le client détecte que la prévalence est élevée et que l'usine n'a pris aucune action corrective, alors il impose la même pénalité fixe à l'usine que lorsque $\psi \in \Psi_1$ pour livraison d'une marchandise contaminée de 4000 €. En outre, le lot dont est issu l'échantillon incriminé est généralement composé de plusieurs lots "industriels" de malaxage (cf. le procédé de fabrication partie 2.2.2). Il faut donc rappeler tous les lots "industriels" de barquettes contenant les lots de malaxage entrant dans la composition du lot contaminé. D'après les données dont nous disposons un lot de barquettes est composé en moyenne de fractions de deux lots de poitrines de porc et deux lots de poitrines entrent en moyenne dans la composition de quatre lots de barquettes, donc ces quatre lots sont rappelés. Le coût du rappel peut se décomposer comme suit : perte de la production (5€/kg), logistique (transport et destruction du lot : 5€/kg) et mise à disposition d'une personne pendant quelques heures par jour pendant 3 jours pour répondre au numéro vert mis en place (350€/par jour). Le montant total du rappel est $4 \times 2000 \text{ kg} \times (5\text{€/kg} + 5\text{€/kg}) + 350\text{€/j} \times 3\text{j} = 81\,050$ €. Le client impose également un contrôle libératoire à l'entreprise pendant un mois pour tous les lots envoyés au client avec $n = 5$ (détection et dénombrement). Si l'usine livre un lot par jour, le montant s'élève à : $1 \text{ lot/j} \times 20\text{j ouvrés} \times 40 \text{ €} \times 5 = 4\,000$ €. Enfin, le client mène un audit dans l'entreprise à la charge de cette dernière pour 1 000 €. Au total, $PB = 4000 + 81050 + 4000 + 1000 = 90\,050$ €.

Si l'usine a mené une action corrective mineure, le client maintient toutes ses décisions sauf l'audit donc $PB_1 = 89\,050$ €.

Enfin, si une action corrective majeure a été menée, le client n'impose ni l'audit ni le contrôle libératoire. Ainsi, $PB_2 = 85\,050$ €.

Le tableau 6.4 reprend le tableau 6.3 mais avec des valeurs numériques. En passant

ψ	Décisions		
	d_0	d_1	d_2
$\psi \in \Psi_0$	0	4250	14000
$\psi \in \Psi_1$	4200	8350	18000
$\psi \in \Psi_2$	90050	93300	99050

TABLE 6.4 – Valeurs numériques prises par la fonction de coût L (sans tenir compte du coût d'analyses).

au regret, le tableau 6.4 devient le tableau 6.5.

ψ	Décisions		
	d_0	d_1	d_2
$\psi \in \Psi_0$	0	4250	14000
$\psi \in \Psi_1$	-4150	0	9650
$\psi \in \Psi_2$	-9000	-5750	0

TABLE 6.5 – Valeurs numériques prises par la fonction de regret \underline{L} (sans tenir compte du coût d’analyses).

Un regret négatif ne devrait pas se produire : cela signifie qu’en choisissant une décision sub-optimale, le coût est moindre qu’avec la décision présentée par l’expert comme optimale ! Cela signifie surtout que la règle de décision optimale a été mal choisie. Avec les valeurs numériques telles que définies, l’optimisation est facile : la décision d_0 serait la moins coûteuse pour l’entreprise, quelle que soit la valeur de ψ . Dans ce contexte, il n’est pas nécessaire d’analyser des produits puisque la valeur de l’état sanitaire ψ n’a aucune influence sur les coûts.

Cependant, la fonction δ' telle que définie plus haut devrait être la règle de décision optimale : il est naturel de considérer que si l’entreprise constate une dérive de la prévalence, elle souhaite y remédier en apportant une action corrective adaptée. Le fait que la décision d_0 entraîne les coûts les moins élevés pour l’entreprise signifie simplement, qu’avec les valeurs numériques et donc le scénario définis plus haut, l’entreprise n’a aucune raison d’effectuer des contrôles. Or, il se trouve que l’entreprise procède à des analyses sur ses produits finis. Cela nous amène à nous interroger sur la construction de la fonction de coût.

Dans le cas présent, la diminution des pénalités de la part du client quand l’entreprise a entrepris une action corrective ($pb - pb_1$ par exemple) ne contrebalance pas l’augmentation du coût pour l’entreprise d’une décision autre que d_0 . Ainsi, le coût de passage de la décision d_0 à d_1 coûte $r = 4250 \text{ €}$ à l’entreprise alors que la réduction de la pénalité n’est que de $pb - pb_1 = 100 \text{ €}$ ou de $PB - PB_1 = 1000 \text{ €}$ selon le cas.

Ce résultat est aussi du au fait qu’il est très difficile pour un producteur de conceptualiser pourquoi il échantillonne et surtout de le conceptualiser sous cette forme-là.

Cependant, si nous faisons l’hypothèse que la fonction de coût a été correctement élicitée, il n’est pas nécessaire que l’entreprise contrôle sa production puisque la décision d_0 est systématiquement la meilleure, quelle que soit la prévalence. Comme nous souhaitons obtenir un cas de figure où la taille de l’échantillon n’est pas déterminée de manière triviale, nous présentons une élaboration de fonction de coût légèrement différente dans la partie suivante.

6.2.4 Autre détermination de la fonction de coût

Dans cette partie, l’angle d’approche pour déterminer la fonction de coût est modifié afin que les regrets soient positifs. Seules la première ligne (correspondant à $\psi \in \Psi_0$) et la première colonne (correspondant à $d = d_0$) du tableau 6.4 sont conservées, la modélisation des coûts étant modifiée pour les autres cases. Le principe de cette nouvelle approche est

le suivant : chaque décision engendre un coût intrinsèque, qui est le coût lorsque l'état sanitaire est le meilleur ($\psi \in \Psi_0$). Chaque état sanitaire engendre un coût, celui lorsque la décision d_0 est prise. Les décisions d_1 et d_2 diminuent l'impact du coût lié à l'état sanitaire et d_2 le diminue plus que d_1 . Le tableau de coûts est présenté au tableau 6.6.

θ \backslash Décisions	d_0	d_1	d_2
$\psi \in \Psi_0$	0	4 250	14 000
$\psi \in \Psi_1$	4 200	$4\,250 + \alpha \cdot 4\,200$	$14\,000 + \beta \cdot 4\,200$
$\psi \in \Psi_2$	90 050	$4\,250 + \alpha \cdot 90\,050$	$14\,000 + \beta \cdot 90\,050$

TABLE 6.6 – Valeurs des coûts en fonction de la décision et de l'état sanitaire. Chaque décision engendre un coût (déterminé par le coût lorsque $\psi \in \Psi_0$). Un état sanitaire occasionne une dépense pour l'entreprise mais ce coût est modulé par la décision prise (réduction de 0 pour d_0 , de $1 - \alpha$ pour d_1 et de $1 - \beta$ pour d_2 , avec $0 < \beta < \alpha < 1$).

Le tableau 6.6 ne permet pas de satisfaire les hypothèses nécessaires pour que la fonction de regret soit toujours positive ou nulle car 4200 est toujours inférieur à $4250 + \alpha \cdot 4200$ ($\alpha \geq 0$). Pour y remédier, le coût de l'état sanitaire $\psi \in \Psi_1$ lorsque d_0 est prise est augmenté de 2000 € , ce qui revient à considérer que le client augmente la pénalité de ce même montant. Il faut par ailleurs que

$$\begin{aligned}
 4\,250 + \alpha \cdot 6\,200 < 6\,200 &\Leftrightarrow \alpha < 0,31 \\
 4\,250 + \alpha \cdot 6\,200 < 14\,000 + \beta \cdot 6\,200 &\Leftrightarrow \beta > -1,57 + \alpha \\
 14\,000 + \beta \cdot 90\,050 < 90\,050 &\Leftrightarrow \beta < 0,84 \\
 14\,000 + \beta \cdot 90\,050 < 4\,250 + \alpha \cdot 90\,050 &\Leftrightarrow \beta < \alpha - 0,11
 \end{aligned}$$

En prenant par exemple $\alpha = 0,3$ et $\beta = 0,15$, les inégalités de construction sont respectées. Le tableau 6.7 reprend le tableau 6.6 exprimé cette fois en coût de regret et en utilisant ces valeurs numériques.

θ \backslash Décisions	d_0	d_1	d_2
$\psi \in \Psi_0$	0	4 250	14 000
$\psi \in \Psi_1$	90	0	8 820
$\psi \in \Psi_2$	62 542	3 757	0

TABLE 6.7 – Valeurs des coûts de regret en fonction de la décision et de l'état sanitaire. Chaque décision engendre un coût (déterminé lorsque $\psi \in \Psi_0$). Un état sanitaire occasionne une dépense pour l'entreprise mais ce coût est modulé par la décision prise (réduction de 0 pour d_0 , de $1 - \alpha = 0,70$ pour d_1 et de $1 - \beta = 0,85$ pour d_2).

6.2.5 Une application numérique

La distribution de ψ représente la distribution de prévalences d'une période de production à l'autre. Pour l'estimer, il faudrait disposer des valeurs empiriques de prévalence de plusieurs périodes. Nous ne disposons pas de telles données. Les données de lardons en sortie usine présentées à la partie 2.3.3.b ne conviennent pas car ce sont des données acquises durant un mois et demi. A la rigueur, elles peuvent servir à estimer la prévalence sur une période si celle-ci dure plus d'un mois et à estimer la prévalence sur deux périodes si la période est définie comme un mois de production. En gardant la première possibilité, la prévalence durant la période où nous avons prélevé des lardons est estimée à 17 %. Nous pourrions également construire une loi *a priori* pour ψ en utilisant les autocontrôles des industriels mais nous en avons également très peu. Pour l'instant et pour mener le calcul à son terme, nous prenons pour Ψ une distribution Beta de paramètres $a = 2$ et $b = 3$ (cf. figure 6.1). La moyenne *a priori* de ψ est donc égale à 0,4.

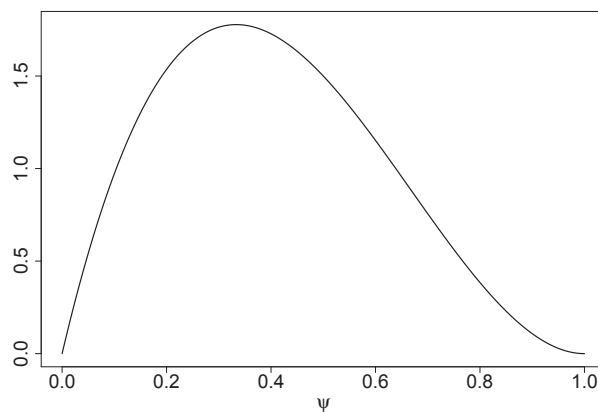


FIGURE 6.1 – Distribution *a priori* de la prévalence ψ .

Nous choisissons de fixer la limite ψ_0 d'un point de vue microbiologique en la choisissant telle que quand $\psi < \psi_0$ on respecte le critère microbiologique défini par le règlement européen Regulation (CE) No 2073/2005 (2005), annexe 1, c'est-à-dire qu'il y ait absence de pathogène dans cinq prises d'essai de 25 g. La limite serait donc que lorsque $\psi = \psi_0$, il y a exactement une prise d'essai positive parmi cinq de 25 g chacune. De manière simplifiée, on peut alors considérer que $\psi_0 = 0,2$. Pour l'autre limite, on choisit $\psi_1 = 0,6$. Si on décide que la période de production considérée dure 2 mois, alors la quasi totalité des prévalences estimées par les autocontrôles de l'industriel chez qui nous avons prélevé les lardons et les poitrines (cf. chapitre 2) sont inférieures à ψ_0 et aucune n'est supérieure à ψ_1 .

Le coût C d'une analyse est fixé à 20 €. Pour commencer, étudions les seuils c_1 et c_2 . La figure 6.2 représente les décisions prise en fonction de n et de x/n . Lorsque $n = 1$, seules deux décisions sont possibles au maximum puisque x ne peut prendre que deux valeurs. Avec les valeurs numériques choisies, la fonction de regret en d_1 est toujours plus petite que la fonction en d_0 et en d_2 , c_1 n'est pas défini et $c_2 = n$: la décision d_1 est donc toujours prise. Pour les autres valeurs de n , les trois décisions sont prises, selon la valeur

de x . Les seuils exprimés en fonction de n tendent vers : $c_1 = 0,2n$ et $c_2 = 0,6n$ lorsque n augmente. On retrouve les valeurs de ψ_0 et de ψ_1 , ce qui est logique puisque c_1/n et c_2/n sont des estimateurs de la prévalence, l'un quand $\psi = \psi_0$ et l'autre pour $\psi = \psi_1$. Lorsque n est petit, les valeurs de c_1/n et de c_2/n varient beaucoup que lorsque n est élevé puisque le pas de variation de x/n est grand (il vaut par exemple 0,2 lorsque $n = 5$). Lorsque n est suffisamment grand (ici, pour $n > 20$), le pas de variation est peu élevé et les valeurs de c_1/n et de c_2/n varient peu lorsque n augmente de 1. A titre d'exemple, le tableau 6.8

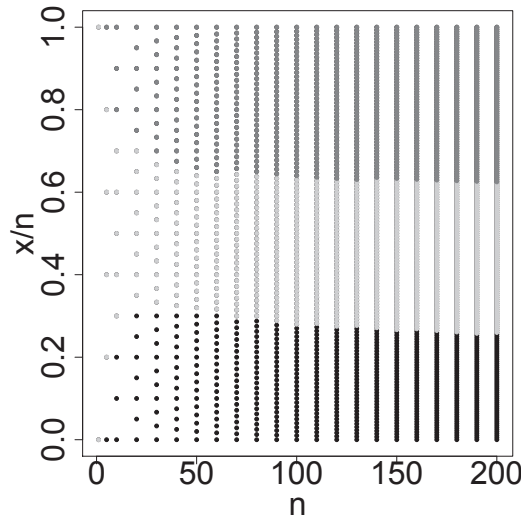


FIGURE 6.2 – Décision prise en fonction de la taille n des échantillons prélevés et de la prévalence estimée x/n grâce aux analyses microbiologiques. Le point est noir lorsque la décision d_0 est prise, gris clair lorsque d_1 est prise et gris foncé lorsque d_2 est prise (avec $[\psi] = \text{Beta}(2; 3)$, $\psi_0 = 0,2$ et $\psi_1 = 0,6$).

présente quelques valeurs de c_1 et de c_2 en fonction de n .

n	c_1	c_2
10	2	7
100	27	63
1000	225	609
10000	2081	6026

TABLE 6.8 – Valeurs des seuils c_1 et c_2 en fonction de la taille n de l'échantillon, avec $[\psi] = \text{Beta}(2; 3)$, $\psi_0 = 0,2$ et $\psi_1 = 0,6$.

Le risque bayésien $\underline{r}(\delta_n)$ est calculé pour n variant de 1 à 50 (cf. figure 6.3). Il est minimal pour $n = 14$, avec $c_1 = 4$ et $c_2 = 10$. Même si le coût d'analyses est le même dans tous les cas de figure (i.e. tous les couples (ψ, d) possibles), il est indispensable de l'ajouter. Certes, ce coût ne modifie pas la fonction de décision δ_n minimisant la fonction \underline{r} à n fixé, en revanche, il modifie la valeur de n pour laquelle \underline{r} est minimal. Le risque

bayésien sans coût d'échantillonnage diminue quand n augmente. Prendre en compte le coût d'échantillonnage permet de mettre en balance

- l'ajout d'information (donc de moins craindre les pénalités du client) ;
- le coût de cet ajout.

Dans le cas présent, l'augmentation du coût d'échantillonnage finit par dépasser la diminution du coût lié à l'augmentation d'information.

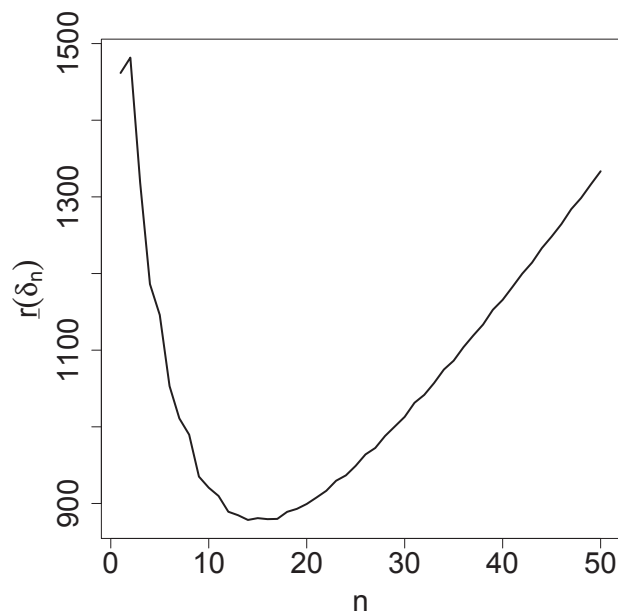


FIGURE 6.3 – Risque bayésien $\underline{r}(\delta_n)$ en fonction de n . Le minimum est atteint pour $n = 14$, avec $a = 2$, $b = 3$, $\psi_0 = 0,2$, $\psi_1 = 0,6$ et $C = 20$ €.

6.2.6 Variation des valeurs de a , b , ψ_0 et ψ_1

La valeur de n qui minimise le risque bayésien dépend, entre autres, de a , b , ψ_0 et ψ_1 . En soi, les seuils pour ψ peuvent être fixés de manière indépendante de a et b , par exemple sur des considérations réglementaires (cf. le règlement européen Regulation (CE) No 2073/2005 (2005)) ou parce que le client ou l'entreprise ne souhaite pas dépasser une prévalence donnée. Cependant, en fonction des limites choisies, toutes les décisions pourront ne pas être prises. C'est le cas si, par exemple, la distribution de prévalence est quasiment entièrement située sous ψ_0 : la décision d_0 sera toujours prise ou presque.

Prenons un exemple numérique en fixant b à 20, les valeurs des autres constantes restant inchangées. Dans ce cas, le risque bayésien est minimal pour $n = 1$ (cf. figure 6.4a) et il est inutile de procéder à des analyses. Si les limites ψ_0 et ψ_1 sont abaissées respectivement à 0,05 et 0,1, alors le risque bayésien est minimal pour $n = 48$ ($c_1 = 1$ et $c_2 = 6$), cf. figure 6.4b. La baisse du risque à partir de $n = 25$ est due au fait qu'à partir de cette taille d'échantillon, la décision d_0 peut être prise alors que pour $n < 25$ et

les valeurs numériques utilisées, seules les décisions d_1 et d_2 peuvent l'être. Une fois que la décision d_0 peut être prise, comme ψ est généralement dans Ψ_0 ou Ψ_1 (car ψ suit une distribution Beta de paramètres 2 et 20) et que la fonction de regret est peu élevée quand d_0 est prise, le risque bayésien est moins élevé à partir du moment où d_0 est une décision possible.

Dans le dernier exemple, les valeurs de ψ_0 et de ψ_1 sont rapprochées et la valeur de n pour laquelle \underline{r} est minimal est plus élevée que lorsque $\psi_0 = 0, 2$ et $\psi_1 = 0, 6$. Ce résultat est logique dans la mesure où plus d'analyses sont nécessaires pour déterminer dans quelle zone ψ se situe : l'intervalle Ψ_1 étant petit et la probabilité que ψ soit dans Ψ_1 étant non négligeable, de nombreuses analyses sont nécessaires pour déterminer si ψ est bien dans cette zone.

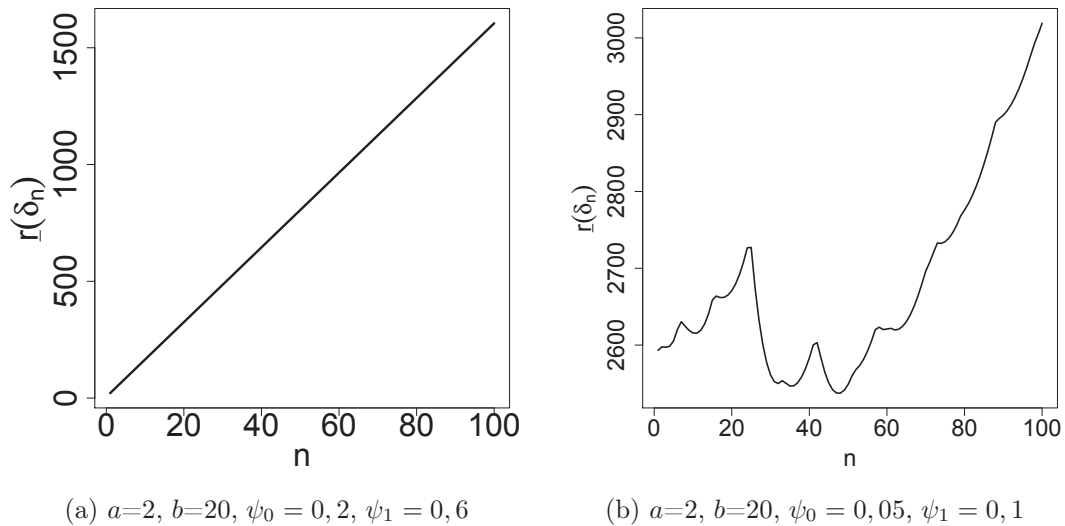


FIGURE 6.4 – Risque bayésien $\underline{r}(\delta_n)$ pour différentes valeurs de limites de prévalence ψ_0 et ψ_1 pour une taille d'échantillon n variant de 1 à 100.

6.2.7 Discussion sur les hypothèses du modèle

Dans ce modèle, j'ai fait l'hypothèse que la prévalence ψ d'une période suit une loi de distribution Beta de paramètres a et b . Comme je dispose de très peu d'information sur les valeurs possibles de ψ , j'ai choisi des valeurs arbitraires pour a et b . Or, ces valeurs ont une grande influence sur la taille optimale d'échantillon. J'aurais pu ajouter un niveau en rendant a et b aléatoires. Comme déterminer des lois pour les paramètres d'une loi Beta n'est pas intuitif, j'aurais d'abord eu recours à une reparamétrisation de façon à ce que les nouveaux paramètres aient une signification biologique. Une reparamétrisation possible est la moyenne $\mu = \frac{a}{a+b}$ et un terme $\nu = a + b + 1$ qui s'apparente à un nombre d'analyses pratiquées. Une autre reparamétrisation intéressante est la moyenne $\mu = \frac{a}{a+b}$ et la variance $\sigma^2 = \frac{ab}{(a+b)^2(a+b+1)}$.

Introduire un niveau supplémentaire transforme ψ en une variable latente. Or, c'est toujours la prévalence qui dicte la conduite du client. Pour mener le calcul du risque bayésien à bien, il faut donc intégrer selon μ et σ^2 (si cette reparamétrisation est choisie) les lois de $\psi|x, \mu, \sigma^2$ et la loi marginale de x . Je ne l'ai pas fait dans cette thèse car je souhaitais me concentrer sur la méthode à élaborer pour appliquer la théorie bayésienne de la décision et sur la construction de la fonction de coût.

Que signifie "pour minimiser les coûts, l'entreprise doit faire 14 analyses"? Cette question revient à se demander à quelle fréquence analyser. En effet, faire 14 prélèvements par semaine ou par an ne représente pas du tout la même pression de contrôle. Pour répondre à cette question, revenons aux hypothèses du modèle. La période de temps durant laquelle analyser n'est pas définie explicitement. Par référence à la notion de lot, c'est une période unitaire de fabrication homogène. Durant une période, l'entreprise fait des analyses puis en tire une conséquence (ne rien faire, mener une action corrective mineure ou majeure). Le client procède également à des contrôles afin de déterminer la prévalence de la production. Ainsi, la période d'analyse est-elle définie comme l'intervalle de temps au bout duquel l'entreprise étudie les résultats de ses prélèvements mais elle est également présentée ici comme l'intervalle de temps entre deux contrôles du client. Cela sous-entend que l'entreprise aligne ses contrôles sur ceux du client.

Il est également possible de définir la période de temps comme la durée durant laquelle la prévalence reste inchangée. La distribution *a priori* de la prévalence est un pari fait sur les valeurs possibles, i.e. la distribution représentant notre croyance de la variabilité de la prévalence inter-périodes.

Ces deux définitions (fréquence des contrôles de l'entreprise et du client et conservation de la distribution de prévalence) sont liées car si l'entreprise entreprend une action corrective, il est raisonnable de penser que cela va avoir un impact sur la prévalence. Inversement, si l'entreprise laisse son système de production inchangé, il est possible que la prévalence augmente au fil du temps. Le modèle comporte une hypothèse osée qui, elle, n'a pas d'incidence sur la définition de la période de temps. On suppose en effet que le client détecte systématiquement la "vraie" prévalence alors que l'entreprise peut se tromper. Or, le client a souvent une pression de contrôle plus faible que celle de l'entreprise. Cette hypothèse n'est donc que peu réaliste. Une modification possible du modèle est de supposer que le client prend une décision sur l'idée qu'il se fait de la prévalence à travers les résultats de ses contrôles mais qu'il peut également se tromper. L'entreprise subit donc des coûts en fonction de cette croyance. Quelle stratégie doit-elle adopter dans ces conditions? Cette question est discutée à la partie suivante.

6.3 Le client ne connaît pas exactement la prévalence

De même que l'entreprise, le client n'a pas les moyens de connaître la prévalence : il ne peut que l'estimer. Pour ce faire, il procède à des analyses de détections. Dans cette partie, on suppose que le client effectue N analyses durant une période. En fonction des résultats obtenus, il prend trois décisions possibles :

1. pas d'action particulière car la prévalence est estimée être faible ;
2. des pénalités d'un montant pb car la prévalence est considérée comme moyenne ;

3. des pénalités d'un montant PB car la prévalence est jugée élevée.

Comme dans le cas précédent où le client connaissait exactement la prévalence, ces pénalités sont modulées en fonction de la décision (d_0 , d_1 , ou d_2) prise par l'entreprise. Les trois décisions possibles de la part du client sont prises en fonction du nombre x^c de résultats positifs sur une période. Le cas suivant est envisagé :

1. si $x^c \leq \lfloor N\psi_0 \rfloor$, pas d'action particulière ;
2. si $\lfloor N\psi_0 \rfloor < x^c \leq \lfloor N\psi_1 \rfloor$, pénalité de coût pb ;
3. si $x^c > \lfloor N\psi_1 \rfloor$, pénalité de coût PB ,

avec $\lfloor z \rfloor$ l'entier le plus proche de z . Ici, la règle de décision du client est fixée, seule la règle de décision de l'entreprise est optimisée. Pour tenir compte de certaines conditions particulières, j'ai choisi de modéliser les décisions du client de la manière suivante :

- Si $\lfloor N\psi_0 \rfloor = 0$, le client ne juge jamais la prévalence comme faible ;
- si $\lfloor N\psi_1 \rfloor = 0$, le client juge la prévalence toujours élevée ;
- si $\lfloor N\psi_0 \rfloor = \lfloor N\psi_1 \rfloor$, le client ne juge jamais la prévalence comme moyenne ;
- si $\lfloor N\psi_1 \rfloor = N$, le client ne juge jamais la prévalence élevée.

Ainsi, en fonction des valeurs de ψ_0 , ψ_1 et N , le client peut ne classer la prévalence que dans une ou deux classes au lieu des trois définies. Ces cas de figure se rencontrent lorsque N est petit, ce qui est courant dans l'industrie agro-alimentaire car le client fait peu de contrôles.

La fonction de regret ne dépend plus explicitement de la prévalence ψ mais du résultat x^c obtenu par l'entreprise. Le tableau 6.9 présente la nouvelle fonction de coût et le tableau 6.10 la fonction de regret.

Résultat client	Décision entreprise		
	d_0	d_1	d_2
$x^c \leq \lfloor N\psi_0 \rfloor$	0	r	R
$\lfloor N\psi_0 \rfloor < x^c \leq \lfloor N\psi_1 \rfloor$	pb	$r + \alpha pb$	$R + \beta pb$
$x^c > \lfloor N\psi_1 \rfloor$	PB	$r + \alpha PB$	$R + \beta PB$

TABLE 6.9 – Coûts en fonction de la décision prise par l'entreprise et des résultats des analyses menés par le client durant une période. Chaque décision engendre un coût (0, r ou R). Les résultats du client occasionnent un coût pour l'entreprise (0, pb ou PB) mais ce dernier est modulé par la décision prise (réduction de 0 pour d_0 , de $1 - \alpha$ pour d_1 et de $1 - \beta$ pour d_2 , avec $0 < \beta < \alpha < 1$).

La fonction de regret $\underline{L}(x^c, \delta_n(x))$ s'écrit alors :

$$\begin{aligned} \underline{L}(x^c, \delta_n(x)) = & Cn + \mathbf{1}_{\delta_n(x)=d_0} \left(K(1, 0)\mathbf{1}_{x^c \leq \lfloor N\psi_0 \rfloor} + K(2, 0)\mathbf{1}_{\lfloor N\psi_0 \rfloor < x^c \leq \lfloor N\psi_1 \rfloor} \right) \\ & + \mathbf{1}_{\delta_n(x)=d_1} \left(K(0, 1)\mathbf{1}_{x^c \leq \lfloor N\psi_0 \rfloor} + K(2, 1)\mathbf{1}_{x^c > \lfloor N\psi_1 \rfloor} \right) \\ & + \mathbf{1}_{\delta_n(x)=d_2} \left(K(0, 2)\mathbf{1}_{x^c \leq \lfloor N\psi_0 \rfloor} + K(1, 2)\mathbf{1}_{\lfloor N\psi_0 \rfloor < x^c \leq \lfloor N\psi_1 \rfloor} \right). \end{aligned}$$

Résultat client	Décision entreprise		
	d_0	d_1	d_2
$x^c \leq \lfloor N\psi_0 \rfloor$	0	$K(0, 1)$	$K(0, 2)$
$\lfloor N\psi_0 \rfloor < x^c \leq \lfloor N\psi_1 \rfloor$	$K(1, 0)$	0	$K(1, 2)$
$x^c > \lfloor N\psi_1 \rfloor$	$K(2, 0)$	$K(2, 1)$	0

TABLE 6.10 – Fonction de regret dépendant de la décision prise par l'entreprise et des résultats des analyses menés par le client durant une période.

Le cas où le client estime la prévalence se traite de la même manière que le cas où le client connaît la prévalence en remplaçant ψ par x^c . Avant de calculer le risque bayésien, déterminons la loi de $x^c|x$. Cette loi se calcule en appliquant le théorème de Bayes :

$$\begin{aligned}
[x^c|x, n] &= \frac{[x, x^c|n]}{[x|n]} \\
&= \frac{\int [x|\psi, n][x^c|\psi, n][\psi]d\psi}{\int [x|\psi, n][\psi]d\psi} \\
&= \frac{\int C_n^x C_N^{x^c} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \psi^{x+x^c+a-1} (1-\psi)^{n-x+N-x^c+b-1} d\psi}{\int C_n^x \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \psi^{x+a-1} (1-\psi)^{n-x+b-1} d\psi} \\
&= \frac{\Gamma(N+1)}{\Gamma(N-x^c+1)\Gamma(x^c+1)} \frac{\Gamma(a+x+x^c)\Gamma(b+n+N-x-x^c)}{\Gamma(a+b+n+N)} \frac{\Gamma(a+b+n)}{\Gamma(a+x)\Gamma(n+b-x)}.
\end{aligned}$$

La fonction de regret intégrée selon x^c s'écrit donc :

$$\begin{aligned}
\int \underline{L}(x^c, \delta_n(x))[x, x^c]dx^c &= Cn \\
&+ \mathbf{1}_{\delta_n(x)=d_0} \left(K(1, 0) \sum_{x^c=\lfloor N\psi_0 \rfloor+1}^{\lfloor N\psi_1 \rfloor} [x^c|x, n] + K(2, 0) \sum_{x^c=\lfloor N\psi_1 \rfloor+1}^N [x^c|x, n] \right) [x|n] \\
&+ \mathbf{1}_{\delta_n(x)=d_1} \left(K(0, 1) \sum_{i=0}^{\lfloor N\psi_0 \rfloor} [x^c|x, n] + K(2, 1) \sum_{x^c=\lfloor N\psi_1 \rfloor+1}^N [x^c|x, n] \right) [x|n] \\
&+ \mathbf{1}_{\delta_n(x)=d_2} \left(K(0, 2) \sum_{x^c=0}^{\lfloor N\psi_0 \rfloor} [x^c|x, n] + K(1, 2) \sum_{x^c=\lfloor N\psi_0 \rfloor+1}^{\lfloor N\psi_1 \rfloor} [x^c|x, n] \right) [x|n].
\end{aligned}$$

De même que dans le cas où le client est omniscient, la décision de l'entreprise se fait en fonction de la valeur de x :

$$\begin{cases} \delta_n(x) = d_0 & \Leftrightarrow x \leq c_1 \\ \delta_n(x) = d_1 & \Leftrightarrow c_1 < x \leq c_2 \\ \delta_n(x) = d_2 & \Leftrightarrow x > c_2, \end{cases}$$

là encore, c_1 et c_2 dépendent de n et $c_1 < c_2$. Un ou plusieurs seuils peuvent ne pas être définis.

L'intégrale de la fonction de regret selon x et x^c (i.e. le risque bayésien) s'écrit :

$$\begin{aligned} \underline{r}(\delta_n) &= Cn + \int \int \underline{L}(\delta(x), x^c, n) [x^c|x, n] [x|n] dx^c dx \\ &= Cn + \sum_{x=0}^{c_1} \left(K(1, 0) \sum_{x^c=\lfloor N\psi_0 \rfloor + 1}^{\lfloor N\psi_1 \rfloor} [x^c|x, n] + K(2, 0) \sum_{x^c=\lfloor N\psi_1 \rfloor + 1}^N [x^c|x, n] \right) [x|n] \\ &+ \sum_{x=c_1+1}^{c_2} \left(K(0, 1) \sum_{x^c=0}^{\lfloor N\psi_0 \rfloor} [x^c|x, n] + K(2, 1) \sum_{i=\lfloor N\psi_1 \rfloor + 1}^N [x^c|x, n] \right) [x|n] \\ &+ \sum_{j=c_2+1}^n \left(K(0, 2) \sum_{x^c=0}^{\lfloor N\psi_0 \rfloor} [x^c|x, n] + K(1, 2) \sum_{i=\lfloor N\psi_0 \rfloor + 1}^{\lfloor N\psi_1 \rfloor} [x^c|x, n] \right) [x|n]. \end{aligned}$$

Le risque bayésien ne dépend plus que de n . Il est donc possible de déterminer pour quelle valeur de n le risque est minimal.

N	n	c_1	c_2
Tout le lot	14	4	10
300	14	4	10
100	15	4	11
30	15	4	11
10	16	4	13
5	14	3	12
4	15	2	12
3	14	4	14
2	1	-	1
1	1	-	1

TABLE 6.11 – Valeurs de la taille n optimale de l'échantillon analysé par l'entreprise et des seuils c_1 et c_2 en fonction de la taille N de l'échantillon prélevé par le client, avec la distribution de prévalence $[\psi] = \mathcal{Beta}(2; 3)$, $\psi_0 = 0, 2$ et $\psi_1 = 0, 6$. Le signe “-” signifie que la grandeur n'est pas définie.

En reprenant les valeurs numériques de la partie 6.2.5 et en faisant varier la taille de N (cf. tableau 6.11), il apparaît que :

- pour N assez grand (en l'occurrence supérieur à 300), la valeur de n est la même que pour un client omniscient ;
- pour $N = 1$, n vaut 1 parce que c'est la valeur minimale que j'ai définie pour le plan d'échantillonnage mais comme la décision d_0 est toujours prise par l'entreprise, il n'est pas nécessaire d'échantillonner ;
- lorsque N varie entre 1 et 300, n varie. La valeur prise par n peut être plus grande que la valeur de n lorsque la prévalence est connue avec certitude, d'après moi parce que la décision du client est entachée d'une grande incertitude donc l'entreprise doit se prémunir contre un classement de la prévalence de la part du

client dans une classe de prévalence plus élevée que ce qu'elle est réellement. La valeur de n peut également être plus petite sûrement parce que l'effet "le client fait peu d'analyses donc il a peu de chance de trouver un échantillon positif" l'emporte sur l'effet "le client procède à peu d'analyses donc il faut que l'entreprise en fasse plus pour être plus confiant dans la décision qu'il prend car le client peut se tromper de classe de prévalence".

La valeur de N pour laquelle le plan d'échantillonnage est le même que si tout le lot est analysé peut être très faible. Par exemple, en abaissant la valeur de a à 0,1, le plan ne varie plus dès que $N = 5$ (cf. tableau 6.12). Dans ce cas, n n'augmente pas lorsque N diminue. Lorsque $n = 1$, le seuil c_1 n'est pas défini, ce qui veut dire qu'il n'existe pas de valeur telle que $\underline{r}(d_0) \leq \underline{r}(d_1)$, donc que la décision d_1 est toujours prise et il n'est pas nécessaire pour l'entreprise d'échantillonner.

N	n	c_1	c_2
Tout le lot	3	1	3
5	3	1	3
1	1	-	1

TABLE 6.12 – Valeurs de la taille n optimale de l'échantillon analysé par l'entreprise et des seuils c_1 et c_2 en fonction de la taille N de l'échantillon prélevé par le client, avec la distribution de prévalence $[\psi] = \text{Beta}(0, 1; 3)$, $\psi_0 = 0, 2$ et $\psi_1 = 0, 6$.

Il est à noter que dans le cadre de modélisation choisi, si ψ_0 , ψ_1 et N sont tels que $\lfloor N\psi_1 \rfloor = \lfloor N\psi_0 \rfloor = 0$, alors il n'est pas nécessaire pour le client de procéder à des analyses puisque la prévalence est toujours jugée élevée (cf. l'ensemble des décisions définies ci-dessus).

Si les valeurs des seuils ψ_0 et ψ_1 sont proches, la taille de l'échantillon à analyser pour l'entreprise augmente à N fixé. Par exemple, le tableau 6.13 présente les résultats pour $\psi_0 = 0, 2$ et $\psi_1 = 0, 3$. Pour $N = 100$, n vaut 27 alors qu'il vaut 15 lorsque $\psi_1 = 0, 6$ (cf. tableau 6.11). Par ailleurs, si $N = 1$, l'entreprise n'a pas besoin d'échantillonner car la meilleure décision à prendre est toujours d_2 car les seuils ψ_0 et ψ_1 sont très bas par rapport à la distribution de ψ ($\mathbb{P}(\psi \leq \psi_1) = 0, 35$).

Que le client connaisse ou non exactement la prévalence sur une période ne change pas le principe de l'optimisation du plan d'échantillonnage ni les équations (il suffit de remplacer $[\psi|x, n]$ par $[x^c|x, n]$). En revanche, cela influe sur la valeur de n minimisant le coût espéré pour l'entreprise, qui peut beaucoup varier en fonction de N . Dans les exemples numériques présentés, n peut varier du simple au double. Dans ces mêmes exemples, la valeur de N pour laquelle le plan d'échantillonnage est le même que si le client était omniscient est très petite devant la taille du lot. Néanmoins, elle peut s'avérer très élevée compte-tenu de la taille des échantillons analysés par la grande distribution. A titre d'exemple, le distributeur de l'usine que nous avons contactée ne pratiquait qu'une analyse de détection par trimestre. Dans ce cas, il faut étudier la valeur de n pour les toutes petites valeurs de N et c'est justement dans cette zone que les écarts avec le plan d'échantillonnage optimal pour un client omniscient sont les plus élevés. Parfois même, l'entreprise n'a pas besoin d'échantillonner.

N	n	c_1	c_2
Tout le lot	31	4	10
300	31	4	10
100	27	3	9
50	23	2	8
10	15	0	7
5	15	1	10
1	1	-	-

TABLE 6.13 – Valeurs de la taille n optimale de l'échantillon analysé par l'entreprise et des seuils c_1 et c_2 en fonction de la taille N de l'échantillon prélevé par le client, avec la distribution de prévalence $[\psi] = \mathcal{Beta}(2; 3)$, $\psi_0 = 0, 2$ et $\psi_1 = 0, 3$. Le signe “-” signifie que la grandeur n'est pas définie.

Dans cette application, nous avons réparti la prévalence en trois parties. En soi, nous aurions pu définir plus de parties ou même considérer la prévalence comme continue. Dans ce cas, nous aurions pu

- modéliser les coûts des pénalités imposées par la client par une valeur pb égale à la pénalité lorsque ψ vaut 1 et que la décision d_0 a été prise par l'entreprise ;
- moduler pb par une fonction $h(\psi)$ valant 0 lorsque $\psi = 0$ et 1 lorsque $\psi = 1$ pour modéliser la baisse de la pénalité lorsque la prévalence devient faible.

Les coefficients α et β diminuant la pénalité lorsque les décisions d_1 et d_2 sont prises seraient conservés. La fonction de coût s'écrirait alors :

$$L(\psi, \delta(x)) = Cn + \mathbf{1}_{\delta(x)=d_0}(pb \times h(\psi)) + \mathbf{1}_{\delta(x)=d_1}(r + \alpha pb \times h(\psi)) + \mathbf{1}_{\delta(x)=d_2}(R + \beta pb \times h(\psi)).$$

Cependant, considérer la prévalence comme continue ne me semble pas adapté à cette application. En effet, il manque beaucoup trop d'information pour construire une fonction h correcte, surtout si l'on considère que le client estime la prévalence avec très peu d'analyses. Dans ce cas, construire une fonction h de modulation de la pénalité en fonction de la prévalence, donc une variation fine de la pénalité en fonction de la prévalence, n'a aucun sens dans la mesure où la prévalence est estimée avec très peu d'analyses. On peut donc presque la considérer comme discrète (si le client procède à 4 analyses par exemple, l'estimation de la prévalence vaudra alors 0, 0,2, 0,4, 0,6, 0,8 ou 1). Il me semble donc plus logique de répartir la prévalence en classes même si, mathématiquement, cette répartition n'est pas obligatoire.

Chapitre 7

Cas saumon fumé : plan d'échantillonnage séquentiel

La modélisation du plan d'échantillonnage dans l'application "saumon fumé" est basée elle aussi sur une pratique d'une entreprise. Les deux décisions possibles sont la livraison ou non du lot. La différence avec le cas d'application "lardons" est que plusieurs plans d'échantillonnage sont définis selon l'état de contamination de l'usine. Le plan varie (taille de l'échantillon et seuil) en fonction des résultats d'analyse de la semaine précédente.

La pratique de l'entreprise avec laquelle nous avons travaillé est la suivante : l'entreprise attribue un état discret à la contamination (bon, moyen et mauvais). Pour chaque état, un plan d'échantillonnage est appliqué : plus la contamination est jugée élevée et plus le NQL^a est petit (c'est-à-dire que plus la taille de l'échantillon est grand et plus le seuil est petit). Un plan d'échantillonnage est défini pour une semaine : chaque jour, n analyses sont menées. La décision de passer d'un état à l'autre se fait pour une semaine entière et est basé sur les résultats des analyses de la semaine précédente.

La modélisation adoptée dans le cadre de ce travail est tirée de cet exemple mais le cas a été un peu simplifié.

7.1 Principe : deux états et une fonction de coût de type "0-K"

Première simplification : seuls deux états sont considérés. Seconde simplification : l'entreprise prend une décision pour le jour suivant en fonction des résultats du jour courant. L'état de contamination supposé par l'entreprise pour le jour t est noté S_t , il peut prendre deux valeurs : 0 (contamination faible) et 1 (contamination élevée). Par ailleurs, la fonction de coût est du même type que celle du cas "lardons", i.e. du type "0-K" (cf. partie 1.3.2). Enfin, il reste à définir l'état réel de contamination pour le jour t , noté E_t , qui lui aussi est binaire.

De même que pour le cas "lardons", nous travaillerons avec la prévalence ψ_t , qui est la prévalence du jour de production t , c'est-à-dire le nombre de portions de 25 g

a. NQL signifie niveau de qualité limite. Le terme est défini partie 1.2.3.a.

	d_0 : livraison	d_1 : non livraison
$\psi_t \in \Psi_0$	0	K_2
$\psi_t \in \Psi_1$	K_1	0

TABLE 7.1 – Valeurs de $\underline{L}(\psi_t, \delta_n(x_t), E_t, S_t)$ sans tenir compte du coût d’analyse.

contaminées sur le nombre de portions de 25 g contenues dans un lot. La loi *a priori* de ψ_t est une loi Beta de paramètres $a(E_t)$ et $b(E_t)$. Ses paramètres dépendent de l’état réel de contamination. Les résultats positifs de détection d’une journée de production sont notés x_t , la vraisemblance de x_t sachant ψ_t est une loi binomiale de paramètres $n(S_t)$ et ψ_t . La taille de l’échantillon n dépend de l’état de contamination supposé par l’entreprise puisqu’en fonction de celui-ci, le plan adopté par l’entreprise diffère. Le lot est défini comme la production journalière de l’entreprise. Cette hypothèse est cohérente avec le résultat trouvé au chapitre 3 à propos de la structure de contamination du saumon fumé : le modèle B-test portion est un modèle convenable pour modéliser la contamination du saumon fumé en *L. monocytogenes*, ce qui signifie que la contamination ne dépend pas du lot de production tel que défini par l’entreprise (il y a plusieurs lots de production par jour).

Le contrôle est libératoire, l’ensemble \mathcal{D} est composé de deux décisions : d_0 (livrer le lot) et d_1 (ne pas livrer le lot). La fonction de regret \underline{L} dépend de la prévalence ψ_t , de la décision $\delta_n(x)$ associée au nombre x de présences, de l’état réel de contamination E_t et de l’état supposé de contamination S_t . Les valeurs prises par \underline{L} sont présentées au tableau 7.1. L’ensemble Ψ_0 est le segment $[0; \psi_0]$ et l’ensemble Ψ_1 le segment $]\psi_0; 1]$, avec ψ_0 une prévalence limite au-delà de laquelle la prévalence est considérée comme élevée. Le principe de cette fonction est le même que la fonction de regret de l’application “lardons” : si la bonne décision est prise (d_0 lorsque $\psi \in \Psi_0$ et d_1 lorsque $\psi \in \Psi_1$), alors cela ne coûte rien à l’entreprise. En revanche, si une mauvaise décision est prise, cela coûte K_1 ou K_2 suivant le cas. Il est raisonnable de penser que $K_1 > K_2$ car si l’entreprise met sur le marché un lot fortement contaminé, les conséquences sanitaires peuvent être graves, alors que rejeter un lot peu contaminé est certes un gaspillage pour l’entreprise mais le coût d’une telle action est tout de même limité. A ces coûts, il faut bien évidemment ajouter le coût d’analyse (le coût unitaire est noté C).

Il reste à définir le passage d’un jour à l’autre. Le passage de E_t à E_{t+1} se fait selon une matrice de passage P qui représente les probabilités des états possibles de E_{t+1} en fonctions des états de E_t . En fonction des résultats du jour t , je suppose que l’entreprise fixe la valeur de S_{t+1} de la manière suivante :

- Si $x_t \leq \lfloor \tau n(S_t) \rfloor$, alors, $S_{t+1} = 0$;
- Si $x_t > \lfloor \tau n(S_t) \rfloor$, alors, $S_{t+1} = 1$ ^b.

Ainsi, la fonction de regret pour le jour t s’écrit :

$$\underline{L}(\psi_t, \delta_n(x_t), E_t, S_t) = \mathbf{1}_{\delta_n(x_t)=d_0} \mathbf{1}_{\psi_t > \psi_0} K_1 + \mathbf{1}_{\delta_n(x_t)=d_1} \mathbf{1}_{\psi_t \leq \psi_0} K_2.$$

La loi *a posteriori* de ψ_t dépend de x_t , de S_t et de E_t et $[\psi_t | x_t, E_t, S_t] = \mathcal{Beta}(a(E_t) + x_t; b(E_t) + n(S_t) - x_t)$. La loi marginale de x_t dépend de S_t et de E_t , avec $[x_t] =$

b. Le symbole $\lfloor \]$ signifie “partie entière”, τ est une constante comprise entre 0 et 1.

$\mathcal{P}o(a(E_t), b(E_t), n(S_t))$. L'intégration de la fonction \underline{L} selon ψ_t et x_t s'écrit :

$$\begin{aligned} U(E_t, S_t) &= Cn(S_t) + \int \int \underline{L}(\psi_t, \delta_n(x_t), E_t, S_t)[x_t|S_t, \psi_t][\psi_t|E_t] dx_t d\psi_t \\ &= Cn(S_t) + \int \int \underline{L}(\psi_t, \delta_n(x_t), E_t, S_t)[\psi_t|x_t, E_t, S_t][x_t|E_t, S_t] d\psi_t dx_t \\ &= Cn(S_t) + \int \mathbf{1}_{\delta_n(x_t)=d_0} K_1[x_t|E_t, S_t](1 - I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t)) dx_t \\ &\quad + \int \mathbf{1}_{\delta_n(x_t)=d_1} K_2[x_t|E_t, S_t] I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t) dx_t, \end{aligned}$$

avec $I_{\psi_0}(a(E_t) + l, b(E_t) + n(S_t) - l)$ la fonction de répartition d'une loi beta de paramètres $a(E_t) + l$ et $b(E_t) + n(S_t) - l$ calculée en ψ_0 . La décision d_0 est choisie si :

$$\begin{aligned} K_1(1 - I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t)) &\leq K_2 I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t) \\ \Leftrightarrow I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t) &\geq \frac{K_1}{K_1 + K_2}. \end{aligned}$$

Lorsque $K_1, K_2, a(E_t), b(E_t), \psi_0$ et $n(S_t)$ sont fixés, $I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t)$ diminue lorsque x_t augmente de 0 à $n(S_t)$. Ainsi, $I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t) \geq \frac{K_1}{K_1 + K_2}$ est équivalent à $x_t \leq c$. Il est possible que même pour $x = 0$, $I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t) < \frac{K_1}{K_1 + K_2}$. Dans ce cas, la décision d_1 est toujours prise : l'entreprise ne vend jamais sa production. Cela signifie que la prévalence seuil ψ_0 n'est pas en accord avec la distribution de la prévalence : si ψ_0 est trop bas par rapport à la distribution de la prévalence quand $E_t = 0$ par exemple, $I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t)$ sera faible. Il est donc peu probable qu'il soit supérieur à $\frac{K_1}{K_1 + K_2}$, surtout si K_1 est très élevé devant K_2 . Par la suite, les valeurs de a, b et ψ_0 seront fixées de telle manière que le seuil c existe ou au moins que la valeur minimale de n pour laquelle le seuil est défini soit plus petite que la taille de l'échantillon minimisant le risque bayésien. Si jamais l'entreprise souhaite maintenir un seuil ψ_0 de prévalence trop faible par rapport à la distribution de la prévalence, cela veut dire qu'elle ne vendra rien tant qu'elle n'aura pas mis en place des mesures destinées à diminuer la prévalence et que ces dernières se seront avérées efficaces.

La valeur de c dépend de a et de b donc de E_t . Or la valeur de E_t est inconnue : on prend donc la valeur de c correspondant à $E_t = 1$ car cette valeur est plus petite que pour $E_t = 0$ à $n(S_t)$ fixé : on se place ainsi dans une hypothèse sécuritaire. Le seuil dépend de E_t et de $n(S_t)$, il faudrait donc noter $c(E_t = 1, n(S_t))$ mais on le notera $c(S_t)$ pour simplifier les notations. Calculer $c(S_t)$ pour $E_t = 1$ uniquement fait que l'entreprise ne choisit pas toujours la solution optimale. En effet, si x_t est supérieur à $c(E_t = 1, n(S_t))$ mais inférieur à $c(E_t = 0, n(S_t))$, que pour le jour t considéré $E_t = 0$ et que $I_{\psi_0}(a(0) + x_t; b(0) + n(S_t) - x_t) \geq \frac{K_1}{K_1 + K_2}$, il faudrait choisir d_0 , or l'entreprise choisit d_1 , précisément parce qu'elle n'a pas accès à la valeur de E_t . L'entreprise jette donc plus de lots qu'une connaissance parfaite de l'état réel de contamination ne le permettrait.

Avec la règle de décision telle que définie, la fonction de risque $U(E_t, S_t)$ s'écrit alors :

$$\begin{aligned} U(E_t, S_t) &= Cn(S_t) + \sum_{l=0}^{l=c(S_t)} K_1[l|E_t, S_t](1 - I_{\psi_0}(a(E_t) + l, b(E_t) + n(S_t) - l)) \quad (7.1) \\ &\quad + \sum_{l=c(S_t)+1}^{n(S_t)} K_2[l|E_t, S_t] I_{\psi_0}(a(E_t) + l, b(E_t) + n(S_t) - l). \end{aligned}$$

7.2 Calcul du risque actualisé sur plusieurs périodes

Dans ce cas d'application, on s'intéresse à la décision durant la période t mais également à l'impact de cette décision sur les périodes futures. Il reste donc à calculer le risque actualisé sur les périodes futures en tenant compte d'un taux d'actualisation noté ρ^c . Il faut donc tenir compte des quatre cas de figures possibles au temps $t + 1$:

- $E_{t+1} = 1 | E_t$ et $S_{t+1} = 1 | S_t, x_t$;
- $E_{t+1} = 1 | E_t$ et $S_{t+1} = 0 | S_t, x_t$;
- $E_{t+1} = 0 | E_t$ et $S_{t+1} = 1 | S_t, x_t$;
- $E_{t+1} = 0 | E_t$ et $S_{t+1} = 0 | S_t, x_t$;

La fonction de risque calculée en t et actualisée en tenant compte de T périodes devient :

$$\begin{aligned}
 V(E_t, S_t) &= U(E_t, S_t) + \rho \int (U(E_{t+1}, S_{t+1}) [E_{t+1}, S_{t+1} | E_t, S_t] dE_{t+1} dS_{t+1} + \dots \\
 &\quad + \rho^T \int U(E_{t+T}, S_{t+T}) [E_{t+T}, S_{t+T} | E_{t+T-1}, S_{t+T-1}] \dots [E_{t+1}, S_{t+1} | E_t, S_t] \\
 &\quad dE_{t+T} dS_{t+T} \dots dE_{t+1} dS_{t+1} \\
 &= U(E_t, S_t) + \rho \int \left(U(E_{t+1}, S_{t+1}) + \dots + \rho^{T-1} \int U(E_{t+T}, S_{t+T}) \right. \\
 &\quad \left. [E_{t+T}, S_{t+T} | E_{t+T-1}, S_{t+T-1}] \dots [E_{t+2}, S_{t+2} | E_{t+1}, S_{t+1}] dE_{t+T} dS_{t+T} \dots dE_{t+2} dS_{t+2} \right) \\
 &\quad [E_{t+1}, S_{t+1} | E_t, S_t] dE_{t+1} dS_{t+1} \\
 &= U(E_t, S_t) + \rho \int V_{t+1}(E_{t+1}, S_{t+1}) [E_{t+1}, S_{t+1} | E_t, S_t] dE_{t+1} dS_{t+1}
 \end{aligned}$$

Par conséquent, si les valeurs de $V(E_{t+1}, S_{t+1})$ sont connues par hypothèse pour les quatre couples de valeurs possibles pour (E_{t+1}, S_{t+1}) , alors $V(E_t, S_t)$ se calcule aisément en prenant l'espérance de $V(E_{t+1}, S_{t+1})$ en E_t et S_t à laquelle on ajoute $U(E_t, S_t)$. Il faut bien sûr connaître la probabilité de passage $[E_{t+1}, S_{t+1} | E_t, S_t]$ pour les seize cas possibles (deux valeurs pour chaque état E_t, S_t, E_{t+1} et S_{t+1}). La formule générale de $V(E_t, S_t)$ du temps t au temps T s'écrit alors :

$$\begin{aligned}
 V(E_t, S_t) &= U(E_t, S_t) + \rho \mathbb{E}_{|E_t, S_t} \{ V_{t+1}(E_{t+1}, S_{t+1}) \} \\
 &= \mathbb{E}_{\psi_t, \dots, \psi_T, y_t, \dots, y_T} \left\{ \sum_{k=t}^{t+T} \rho^{k-t} L(\psi_k, \delta_n(x_k), E_k, S_k) \right\} \\
 &= U(E_t, S_t) + \rho \mathbb{E}_{\psi_{t+1}, \dots, \psi_T, y_{t+1}, \dots, y_T} \left\{ \sum_{k=t+1}^{t+T} \rho^{k-t+1} L(\psi_k, \delta_n(x_k), E_k, S_k) \right\}.
 \end{aligned}$$

C'est une formule itérative. Pour le calcul, un algorithme est construit avec trois routines :

c. L'actualisation est un terme utilisé en économie. Elle consiste à ramener sur une même base des coût monétaires qui se produisent à des dates différentes en donnant une valeur présente aux coûts futurs.

1. Une calculant $U(E_t = i, S_t = j)$, avec $(i, j) = (0, 0), \dots, (1, 1)$;
2. Une calculant les probabilités de transition d'un couple (état réel, état supposé) de la période t à la période $t + 1$: $P(E_{t+1} = i, S_{t+1} = m | E_t = j, S_t = n)$, avec $(i, m, j, n) = (0, 0, 0, 0), \dots, (1, 1, 1, 1)$;
3. A partir de l'espérance du risque à partir de la période $t + 1$, la troisième routine calcule l'espérance du risque à la période t .

L'algorithme est initialisé à la période T avec $V_T(E_T, S_T) = 0$ quelles que soient les valeurs prises par E_T et S_T . Cela revient à ne pas tenir compte des coûts à partir de la période T et au-delà car elle est trop éloignée dans le temps. La valeur de T dépend de celle du taux d'actualisation ρ : plus ρ est proche de 1 et plus le nombre de périodes futures considérées est grand.

Pour la deuxième routine, il reste à calculer la probabilité de transition des états dans tous les cas de figure. A titre d'exemple :

$$\begin{aligned}
P(E_{t+1} = 1, S_{t+1} = 1 | E_t = 1, S_t = 1) &= [E_{t+1} = 1 | E_t = 1] \int \int [\psi_t | x_t, E_t = 1, S_t = 1] d\psi_t \\
&\quad [S_{t+1} = 1 | S_t = 1, E_t = 1, x_t] [x_t | E_t = 1, S_t = 1] dx_t \\
&= [E_{t+1} = 1 | E_t = 1] \int 1_{x_t > \lfloor \tau n(1) \rfloor} [x_t | E_t = 1, S_t = 1] dx_t \\
&= [E_{t+1} = 1 | E_t = 1] \sum_{l=\lfloor \tau n(1) \rfloor + 1}^{n(1)} [x_t = l | E_t = 1, S_t = 1],
\end{aligned}$$

car S_{t+1} ne vaut 1 que si $x_t > \lfloor \tau n(S_t = 1) \rfloor$. La valeur de E_{t+1} ne dépend que de celle de E_t alors que la valeur de S_{t+1} dépend de x_t (car l'état décrété par l'entreprise dépend du résultat des analyses de la semaine précédente) et donc de ψ_t et de $n(S_t)$.

7.3 Intégrons les éléments précédents selon E_t

Tous les risques calculés jusqu'à présent $(U(E_t, S_t), V(E_t, S_t))$ dépendent de l'état réel de la contamination E_t . Or ce dernier n'est pas connu. Il est donc utile de pouvoir intégrer les résultats selon la distribution de E_t . On connaît la forme de la matrice de passage de l'état E_t à l'état E_{t+1} , c'est-à-dire le noyau P de la chaîne de Markov :

$$P = \begin{pmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{pmatrix}$$

avec α et β des réels compris entre 0 et 1 exclus. Dans ce cas, on peut représenter la distribution de E_t en fonction de E_{t-1} par : $[E_t] = [E_{t-1}] \times P$, avec $[E_t]$ et $[E_{t-1}]$ des vecteurs ligne. De même, la loi de E_t en fonction de l'état initial E_0 s'écrit : $[E_t] = [E_0] \times P^t$. La matrice P est diagonalisable, ses valeurs propres sont 1 et $1 - \alpha - \beta$ et peut s'écrire $P = O^{-1}DO$ avec $D = \begin{pmatrix} 1 & 0 \\ 0 & 1 - \alpha - \beta \end{pmatrix}$ et O la matrice de passage entre P et D . Chaque vecteur ligne de O est un vecteur propre de P . Le vecteur X est vecteur propre de P pour

la valeur propre 1 si $XP = X$, avec $X = (x_1 \ x_2)$. Le système d'équations à résoudre est le suivant :

$$\begin{cases} (1 - \alpha)x_1 + \beta x_2 = x_1 \\ \alpha x_1 + (1 - \beta)x_2 = x_2 \end{cases}$$

ce qui est équivalent à $x_1 = \frac{\beta}{\alpha}x_2$. Pour la valeur propre $1 - \alpha - \beta$, on trouve comme vecteur propre : $x_1 = -x_2$. Ainsi, la matrice O s'écrit :

$$O = \begin{pmatrix} \frac{\beta}{\alpha}x & x \\ \lambda & -\lambda \end{pmatrix}$$

Si on impose en outre que le déterminant de O soit égal à 1 et $\lambda = -1$, alors :

$$O = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ -1 & 1 \end{pmatrix}.$$

La matrice élevée au rang t s'écrit donc :

$$\begin{aligned} P^t &= O^{-1}D^tO \\ &= \begin{pmatrix} 1 & -\frac{\alpha}{\alpha+\beta} \\ 1 & \frac{\beta}{\alpha+\beta} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (1 - \alpha - \beta)^t \end{pmatrix} \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1 - \alpha - \beta)^t & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1 - \alpha - \beta)^t \\ \frac{\beta}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}(1 - \alpha - \beta)^t & \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}(1 - \alpha - \beta)^t \end{pmatrix} \end{aligned}$$

Si on part d'un certain état initial et que $|1 - \alpha - \beta| < 1$, lorsque t tend vers l'infini, l'état initial est oublié et la répartition des états futurs se rapproche de plus en plus de la répartition π_∞ donnée par $\pi_\infty(0) = \pi = \frac{\beta}{\alpha+\beta}$ et $\pi_\infty(1) = 1 - \pi = \frac{\alpha}{\alpha+\beta}$.

7.3.1 Risque au temps t intégrée selon E_t

Le risque intégré selon E_t s'écrit, en reprenant l'équation (7.1) :

$$\begin{aligned} U(S_t) &= \int U(E_t, S_t)[E_t]dE_t \\ &= U(0, S_t)\pi + U(1, S_t)(1 - \pi) \\ &= \left(\sum_{l=0}^{l=c(S_t)} K_1[l|0, S_t](1 - I_{\psi_0}(a(0) + l, b(0) + n(S_t) - l)) \right. \\ &\quad \left. + \sum_{l=c(S_t)+1}^{n(S_t)} K_2[l|0, S_t]I_{\psi_0}(a(0) + l, b(0) + n(S_t) - l) \right) \pi \\ &+ \left(\sum_{l=0}^{l=c(S_t)} K_1[l|1, S_t](1 - I_{\psi_0}(a(1) + l, b(1) + n(S_t) - l)) \right. \\ &\quad \left. + \sum_{l=c(S_t)+1}^{n(S_t)} K_2[l|1, S_t]I_{\psi_0}(a(1) + l, b(1) + n(S_t) - l) \right) (1 - \pi). \end{aligned}$$

a	b	ψ_0	$n(S_t)$	$c(S_t)$
0,1 ($E_t = 0$), 2($E_t = 1$)	2 ($E_t = 0$), 0,1($E_t = 1$)	0,9	3	0
1 ($E_t = 0$), 10($E_t = 1$)	10 ($E_t = 0$), 1($E_t = 1$)	0,6	20	2
2 ($E_t = 0$), 20($E_t = 1$)	20 ($E_t = 0$), 2($E_t = 1$)	0,6	31	3
2 ($E_t = 0$), 6($E_t = 1$)	6 ($E_t = 0$), 2($E_t = 1$)	0,6	15	2
2 ($E_t = 0$), 3($E_t = 1$)	2 ($E_t = 0$), 3($E_t = 1$)	0,9	9	6

TABLE 7.2 – Valeurs minimales de l'utilité $\underline{V}(S_t)$ actualisée en t en tenant compte des plans d'échantillonnage pour différentes valeurs de a , b et ψ_0 ($\rho = 0$).

7.3.2 Risque au temps t actualisé et intégré selon E_t

Le risque actualisé en t intégré selon E_t s'écrit :

$$\begin{aligned} V_t(S_t) &= \int V(E_t, S_t)[E_t]dE_t \\ &= \pi V(0, S_t) + (1 - \pi)V(1, S_t) \end{aligned}$$

L'algorithme de calcul de $V(E_t, S_t)$ présenté à la partie 7.2 reste toujours valable pour le calcul de $V(S_t)$ à ceci près qu'il faut maintenant intégrer selon E_t .

7.4 Applications numériques

7.4.1 Cas où $\rho = 0$

Si le taux d'actualisation ρ vaut 0, cela veut dire que le futur n'est pas pris en compte. Le risque actualisé au temps t est donc égal au risque "créé" au temps t , c'est-à-dire : $V(S_t) = U(S_t)$. En calculant $V(S_t)$ pour plusieurs valeurs de $n(S_t)$, les valeurs du risque sont les mêmes quelle que soit la valeur de S_t . Sous forme mathématiques, cela revient à dire : $(V(S_t = 0)|n(0) = A) = (V(S_t = 1)|n(1) = A)$. Si $V(S_t = 0)$ est minimal pour une certaine valeur de n , $V(S_t = 1)$ sera minimal pour cette même valeur de n . Le tableau 7.2 présente les valeurs de n pour lesquelles $V(S_t)$ est minimal pour plusieurs valeurs de a , b et ψ_0 . Les valeurs numériques des autres constantes sont :

$$\begin{aligned} K_1 &= 10^5 \\ K_2 &= 10^3 \\ C &= 20 \\ \tau &= 0,2 \\ P &= \begin{pmatrix} 0,7 & 0,3 \\ 0,2 & 0,8 \end{pmatrix} \end{aligned}$$

Les valeurs de a et b ont été choisies de manière à discriminer plus ou moins la prévalence lorsque $E_t = 0$ et lorsque $E_t = 1$, c'est-à-dire que les distributions de $\psi_t(E_t = 0)$ et de $\psi_t(E_t = 1)$ sont très différentes et se recouvrent peu. Dans le tableau 7.2, les distributions pour $E_t = 0$ ont été classées dans l'ordre discriminant décroissant. La valeur

$n(S_t)$	$c(S_t)$	NQA	NQL
3	0	0,015	0,635
20	2	0,040	0,285
31	3	0,045	0,235
15	2	0,055	0,365
9	6	0,450	0,905

TABLE 7.3 – NQA et NQL pour différents plans à deux classes avec un risque fournisseur et un risque client égaux à 5%.

de ψ_0 est modifiée lorsque la valeur de n minimisant la fonction de coût correspond à un cas où une seule décision est prise. Par exemple, pour le cas $a = 0, 1$ ($E_t = 0$), 2 ($E_t = 1$) et $b = 2$ ($E_t = 0$), $0, 1$ ($E_t = 1$), le seuil ψ_0 est augmenté jusqu'à 0,9 pour trouver une valeur de n minimale telle que les décisions d_0 et d_1 sont prises en fonction de la valeur de x_t .

Les valeurs obtenues de $n(S_t)$ et de $c(S_t)$ pour les différents cas exposés au tableau 7.2 ne semblent pas liées à la discrimination entre les distributions de prévalences lorsque $E_t = 0$ et lorsque $E_t = 1$. Cependant, si l'on considère ces plans comme des plans à deux classes classiques (cf. partie 1.2.2) et que l'on calcule leur niveau de qualité acceptable (NQA) pour un risque fournisseur de 5% et leur niveau de qualité limite (NQL) pour un risque client de 5%, on constate que les distributions de prévalences les plus "séparées" correspondent à des plans dont les NQA sont les plus petits. Le tableau 7.3 présente les différents NQA et NQL pour les plans d'échantillonnages déterminés dans le tableau 7.2. Ainsi, plus les distributions de prévalence sont proches pour les deux états de contamination réelle ($E_t = 0$ et $E_t = 1$) et plus la prévalence du lot rejeté à tort dans 5% des cas est élevée. Là encore, ce résultat dépend de ψ_0 : il est tout à fait possible de trouver un plan avec un NQA élevé alors que les distributions de prévalence entre les états réels de contamination sont très différentes. Les courbes des différents plans sont tracés à la figure 7.1.

7.4.2 Cas où $\rho \neq 0$

Si $\rho \neq 0$, on tient compte du futur. L'utilité actualisée en t de $V(S_t)$ dépend à la fois des valeurs de $n(0)$ et de $n(1)$ à cause de l'actualisation.

Commençons tout d'abord par calculer les valeurs de $n(0)$ et de $n(1)$ minimisant l'utilité actualisée pour les mêmes valeurs numériques de a , b et ψ_0 de la partie précédente (cf. Tableau 7.4). Cette fois-ci, $\rho = 0, 9$ et l'actualisation porte sur 40 semaines.

La période de 40 semaines est choisie parce que $V(E_t, S_t)$ se stabilise : le risque augmente avec le nombre de semaines puis devient constant à partir de 40 semaines.

Les plans d'échantillonnage pour $S_t = 0$ et $S_t = 1$ sont très proches : la taille de l'échantillon augmente de 1 ou de 2 quand la contamination estimée par l'entreprise devient élevée. Cela est peut-être dû au fait que les valeurs de $V(E_t, S_t)$ sont proches lorsque E_t et S_t varient : en effet, ces valeurs ne diffèrent que pour le risque $U(E_t, S_t)$ calculé au temps t et l'intégration de $V(E_{t+1}, S_{t+1})$ selon $S_{t+1}|E_t, S_t$ et $E_{t+1}|E_t$. Les intégrations des

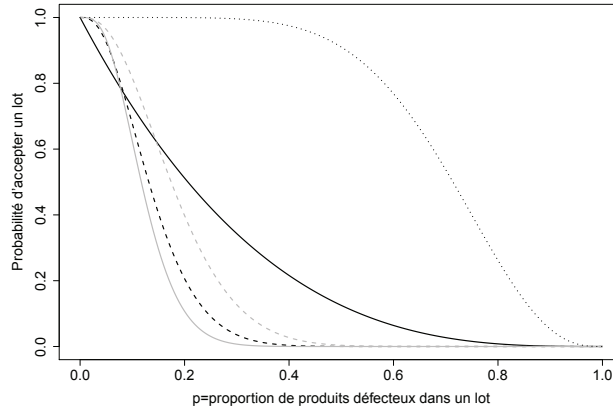


FIGURE 7.1 – Probabilité d’acceptation en fonction de la prévalence pour différents plans d’échantillonnage.

Courbe noire continue : $n=3$ et $c=0$, courbe noire discontinue : $n=20$ et $c=2$, courbe grise continue : $n=31$ et $c=3$, courbe grise discontinue : $n=15$ et $c=2$, courbe en pointillés : $n=9$ et $c=6$

a	b	ψ_0	$n(0)$	$c(0)$	$n(1)$	$c(1)$
0,1 ($E_t = 0$), 2($E_t = 1$)	2 ($E_t = 0$), 0,1($E_t = 1$)	0,99	3	1	4	2
1 ($E_t = 0$), 10($E_t = 1$)	10 ($E_t = 0$), 1($E_t = 1$)	0,6	20	2	22	3
2 ($E_t = 0$), 20($E_t = 1$)	20 ($E_t = 0$), 2($E_t = 1$)	0,6	31	3	33	4
2 ($E_t = 0$), 6($E_t = 1$)	6 ($E_t = 0$), 2($E_t = 1$)	0,6	13	1	15	2
2 ($E_t = 0$), 3($E_t = 1$)	2 ($E_t = 0$), 3($E_t = 1$)	0,9	8	5	9	6

TABLE 7.4 – Valeurs minimales du risque $V(S_t)$ actualisé en t en tenant compte des plans d’échantillonnage pour différentes valeurs de a , b et ψ_0 ($\rho = 0,9$).

K_1	K_2	$\frac{K_1}{K_1+K_2}$	P	π	$n(0)$	$c(0)$	$n(1)$	$c(1)$
10^7	10^3	0,9999	$\begin{pmatrix} 0,7 & 0,3 \\ 0,2 & 0,8 \end{pmatrix}$	0,4	41	3	43	4
10^7	10^6	0,9091	"	0,4	39	11	41	12
10^7	10^5	0,9901	"	0,4	40	8	41	8
10^5	$5 \cdot 10^4$	0,6667	"	0,4	24	6	26	7
10^5	10^3	0,9901	$\begin{pmatrix} 0,99 & 0,01 \\ 0,01 & 0,99 \end{pmatrix}$	0,5	33	4	34	4
10^5	10^3	0,9901	$\begin{pmatrix} 0,5 & 0,5 \\ 0,5 & 0,5 \end{pmatrix}$	0,5	31	3	33	4
10^5	10^3	0,9901	$\begin{pmatrix} 0,9 & 0,1 \\ 0,1 & 0,9 \end{pmatrix}$	0,8	38	7	39	7

TABLE 7.5 – Valeurs minimales du risque $V(S_t)$ actualisé en t en tenant compte des plans d'échantillonnage pour différentes valeurs de K_1 , K_2 et P , avec $\rho = 0,9$, $K_1 = 10^5$, $K_2 = 10^3$, $\psi_0 = 0,6$, $a = 2$ ($E_t = 0$), 20 ($E_t = 1$) et $b = 20$ ($E_t = 0$), 2 ($E_t = 1$).

risques selon $S_{t+i}|E_{t+i-1}, S_{t+i-1}$ et $E_{t+i}|E_{t+i-1}$ pour $i > 1$ sont les mêmes quelles que soient les valeurs de E_t et S_t puisque tous les cas de figure des couples E_{t+i}, S_{t+i} sont envisagés.

Le tableau 7.5 présente d'autres plans d'échantillonnage lorsque les coûts K_1 et K_2 , la matrice de transition P entre les différents états réels varient, avec $a = 2$ ($E_t = 0$), 20 ($E_t = 1$), $b = 20$ ($E_t = 0$), 2 ($E_t = 1$) et $\psi_0 = 0,6$. La probabilité π a une certaine influence sur le plan puisque le passage de 0,5 à 0,8 pour π augmente $n(0)$ de 7 unités et $n(1)$ de 6 unités. Enfin, les coûts influencent le plan via le rapport $\frac{K_1}{K_1+K_2}$ qui est la limite que $I_{\psi_0}(a(E_t) + x_t, b(E_t) + n(S_t) - x_t)$ dépasse si la décision d_0 est prise. La taille des échantillons augmente lorsque le rapport augmente. Cela s'explique simplement : pour $n(S_t)$ fixé, $c(S_t)$ diminue quand le rapport augmente. Ainsi, lorsque x_t varie de 0 à $n(S_t)$, la décision d_0 est moins souvent prise que d_1 donc la valeur de $U(E_t, S_t)$ est moins élevée puisque $K_2 < K_1$. Pour que l'augmentation due au coût d'échantillonnage contrebalance suffisamment la diminution due au coût de la décision, la valeur de $n(S_t)$ minimisant le risque doit être plus élevée quand le rapport augmente.

Ce cas d'application est plus difficile à appréhender que le cas "lardons", notamment en raison de l'actualisation. Il comporte de très nombreuses constantes dont il faut déterminer les valeurs. Par ailleurs, les plans d'échantillonnage ne sont pas très différents selon l'opinion de l'entreprise à propos de la contamination. Par conséquent, dans le cadre de cette modélisation, il n'est pas pertinent de procéder à des plans d'échantillonnage différents selon la contamination.

Dans le cadre de cette thèse nous n'avons pas eu la possibilité d'interroger une entreprise sur ses coûts de manière à construire une fonction de coût satisfaisante, nous avons seulement pu obtenir des informations partielles ne permettant pas de traiter ce sujet comme nous l'aurions souhaité.

7.5 Perspectives

Une extension possible de ce travail serait de travailler sur des coûts plus réalistes en rendant le coût de livraison du lot dépendant de la prévalence ou mieux de la distribution de concentration de celui-ci afin de tenir compte de la probabilité pour une personne ingérant une portion type d'être malade, via un modèle dose-réponse^d. Dans ce cas, il faudrait tenir compte de la croissance de *L. monocytogenes* entre la sortie de l'usine et la consommation pour les portions contaminées. Enfin, il reste à déterminer le coût d'une personne malade. Frenzen et al. (2005) ont par exemple estimé les coûts dus aux infections par *E. coli* O157 aux Etats-Unis en distinguant les coûts de consultation, d'hospitalisation, les coûts médicaux dus à des séquelles irréversibles de certains malades, la perte de productivité ainsi que les coûts de la mort prématurée due à ce pathogène. Scharff (2012) a mené le même travail mais pour plusieurs bactéries, parasites et virus aux Etats-Unis pour l'année 2010.

Les études d'estimation des coûts que j'ai pu lire portent sur l'impact causé par un pathogène. Si on souhaite intégrer les coûts de maladie dans la fonction de coût pour déterminer le meilleur plan d'échantillonnage dans le cas du saumon fumé, il faut estimer les coûts pour un couple pathogène/aliment. Une manière d'évaluer ces coûts serait de multiplier le coût global dû au pathogène par la fraction de malades dus à l'aliment considéré, c'est-à-dire de procéder à une analyse quantitative des risques pour le couple aliment/pathogène. Pouillot et al. (2009) ont estimé le nombre de cas de listérioses dus à la consommation de saumon fumé en France pour différentes catégories de population. Comme l'incertitude et la variabilité des différents paramètres entrant en jeu dans l'estimation du nombre de malades ont été prises en compte, le nombre de cas suit lui-même une distribution. Ainsi, le nombre de cas est estimé à 307 par an dans la population générale avec un intervalle de crédibilité à 95% très large de [10 ; 12 453] en raison de la forte incertitude de certains paramètres. En comparant la consommation de saumon fumé en France avec les données épidémiologiques de listériose, Pouillot et al. (2009) estiment que le nombre de listérioses éventuellement attribuables au saumon fumé est compris entre 0 (car le lien entre listériose et consommation de saumon fumé n'a jamais été clairement démontré) et 100, soit entre 0 et 33% des listérioses annuelles en France.

Les coûts tels qu'estimés par Frenzen et al. (2005) ou Scharff (2012) sont essentiellement supportés par la société. Or, la fonction de coût telle que présentée dans ce chapitre porte sur des coûts subis par l'entreprise. Ces coûts sont divers, ils recouvrent entre autres le retrait, le rappel, la perte d'image de marque qui peut être importante s'il y a plusieurs malades et que la couverture médiatique de l'événement est nationale. L'estimation de cette perte d'image peut s'avérer extrêmement difficile car l'entreprise peut n'avoir jamais été confrontée à ce type de situation donc elle n'a pas d'historique. De plus, un scénario catastrophe se produit, la perte peut être très élevée et être couplée à un procès donnant lieu à l'indemnisation des victimes.

Néanmoins, ce cas peut se produire et il faut bien indemniser les victimes si l'entreprise est condamnée. Cela nécessite donc de quantifier le coût d'une maladie subie par

d. Un modèle dose-réponse est un modèle reliant la dose ingérée à la probabilité d'être malade. Cette probabilité peut dépendre de la catégorie de personne (en bonne santé, immunodéprimée, âgée, femme enceinte...). Pour plus de détails sur les modèles dose-réponse, voir Haas et al. (1999).

une personne en tenant compte de la gravité et même le coût d'un mort. La "monétisation" de la maladie et même de la mort me semble être plus répandue dans les pays anglo-saxons. Ainsi, j'ai trouvé de tels coûts dans une affaire au Canada. Il y a 4 ans, des cas de listérioses se sont produits dans plusieurs provinces du Canada à cause d'aliments produits par l'entreprise canadienne Maple Leaf. Une "class-action"^e s'est constituée dans plusieurs provinces du pays. L'entreprise a procédé à un rappel de 243 produits prêts à consommés fabriqués entre le 1er janvier et le 31 août 2008. L'épidémie a causé la mort de 23 personnes et entraîné 60 malades. Le principe même de la "class-action" est de fixer des indemnités de manière précise selon l'atteinte dont une personne a été victime. Pour cette épidémie, Maple Leaf a dû provisionner 25 millions de dollars canadiens, soit 19 millions d'euros environ. L'indemnisation s'élève à 750 dollars pour une personne ayant eu des symptômes pendant 48h maximum jusqu'à 120 000 dollars pour une personne décédée avec 35 000 dollars pour le conjoint, 30 000 par enfant, 20 000 pour chaque parent et 5 000 pour les frères et sœurs ou les petits-enfants et les frais de funérailles jusqu'à 13 500 dollars (MapleLeaf (2011)).

Un moyen d'évaluer la perte d'image de marque est de consulter les assurances car certaines proposent des couvertures pour ce type de risque. Cependant, le montant de la couverture ne donne qu'une indication des coûts puisque les modalités des polices d'assurance sont variables et que la totalité des situations possibles n'est pas forcément couverte.

Pour compliquer encore un peu le problème, on peut penser qu'un retrait ou un rappel de lot pourrait permettre à l'entreprise incriminée de faire de la publicité sur le fait qu'elle est très vigilante à propos de la sécurité sanitaire des aliments et c'est pourquoi elle préfère retirer le lot. Dans ce cas, il faudrait évaluer le bénéfice pour l'entreprise s'il existe ou la modération de la perte d'image de marque du fait de cette communication.

e. La *class action* ou action collective est une action en justice qui permet à un grand nombre de personnes, souvent des consommateurs, de se regrouper et de poursuivre une personne, souvent une entreprise ou une institution publique pour faire valoir leur droit ou se faire indemniser leur préjudice. Cette procédure n'existe pas en France.

Chapitre 8

Cas plat cuisiné en restauration collective : un exemple de surveillance de la température

Cette partie présente une application de plan d'échantillonnage sur la mesure de températures dans des plats cuisinés en restauration collective. Ce travail a fait l'objet d'un article soumis à la revue *Food Control* et qui est reproduit en fin de chapitre.

8.1 Préparation d'un plat en restauration collective

Un des sous-projets du projet Quant'HACCP était intitulé "*C. perfringens* dans les plats préparés de restauration collective". Plus précisément, les plats choisis étaient les préparations de bœuf en sauce élaborés dans la cuisine centrale d'un hôpital à Paris. Des données de température ont été collectées par l'unité MAHY^a du laboratoire de sécurité sanitaire des aliments de l'ANSES dans de la viande et de la sauce depuis la cuisson jusqu'à la consommation par les patients.

Brièvement, le mode de préparation est le suivant : après une cuisson d'une heure environ à 100°C dans une grande sauteuse, la préparation est répartie dans des bacs en plastique qui sont ensuite operculés puis placés dans un chariot. Ce dernier est mis dans une cellule de refroidissement rapide. Cette cellule comporte des ventilateurs pour propager le froid de manière rapide. Une sonde mesurant la température est placée dans un des bacs. Lorsque cette température passe sous le seuil de 10°C, la sonde émet un signal sonore. Le seuil est fixé à 10°C car l'Arrêté du 21 décembre 2009 (2009) stipule à l'annexe IV que le refroidissement rapide des préparations culinaires doit être tel que la température à cœur ne doit pas demeurer plus de 2 h entre 63 °C et 10°C. A ce moment, un opérateur place une seconde sonde dans un autre endroit d'un autre bac pour effectuer une autre mesure de température. Si cette dernière est inférieure à 10°C, le chariot est retiré de la cellule de refroidissement, sinon, le chariot est maintenu en cellule et l'opérateur mesure à nouveau la température un peu plus tard. Une fois sorti, le chariot est placé dans une chambre froide. Peu avant la consommation, les bacs sont réchauffés.

a. MAHY signifie MAîtrise de l'HYgiène.

Le contrôle du procédé étant uniquement basé sur le suivi de la température, nous avons simplement tenu compte de ce dernier et nous n'avons pas étudié la contamination par *C. perfringens*. Pour une modélisation de la contamination microbiologique, voir les travaux de Jaloustre (2011), Jaloustre et al. (2011) et Jaloustre et al. (2012).

8.2 Plan d'échantillonnage thermique

La gestion de sortie de la cellule de refroidissement peut être vue comme un plan d'échantillonnage à deux classes avec $n = 2$, $c = 0$ et $m = 10^\circ\text{C}$. Deux décisions sont possibles : sortir ou ne pas sortir le chariot de la cellule de refroidissement. La différence avec le contrôle par attribut à deux classes défini à la partie 1.2.2.a est que l'échantillonnage peut avoir lieu plusieurs fois dans le temps et que de toute manière, le chariot finira par sortir de la cellule. C'est pourquoi, nous avons décidé de modéliser la température dans un chariot au cours du temps afin de déterminer la probabilité de sortie du chariot en fonction du temps.

Une fois encore, la définition du lot se pose. Ici, nous avons choisi que le lot serait constitué des préparations contenues dans un chariot puisque les bacs d'un chariot sont soumis aux mêmes conditions thermiques jusqu'au réchauffement. En outre, les chariots différents ne sont pas soumis aux mêmes conditions puisqu'ils peuvent ne pas être placés dans la même cellule de refroidissement ou parce que le temps d'attente entre la constitution des bacs et l'entrée dans la cellule de refroidissement n'est pas le même.

Dans la stratégie actuelle de l'hôpital, la température de la préparation n'est mesurée qu'en deux points du chariot. Or, d'après les relevés de températures collectés pour le projet Quant'HACCP, il apparaît que la température des produit dans un chariot n'est pas uniforme. Poumeyrol et al. (2012) ont élaboré un modèle de la température dans un chariot afin d'étudier la proportions de points dont la température est supérieure à 10°C dans le chariot alors que celui-ci a été sorti de la cellule de refroidissement. Leur modèle sépare les températures de centre géométrique d'un bac des températures des autres points du bac car les données ont montré que les variabilités de température inter et intra-bacs étaient à prendre en compte. Nous avons donc proposé plusieurs stratégies possibles afin de comparer les probabilités de sortie de la cellule de refroidissement à différents temps et d'étudier l'impact sur la probabilité que la concentration en *C. perfringens* dépasse 1 log ufc/g durant toute la phase de refroidissement (cellule de refroidissement rapide et chambre froide). Au total, quatre stratégies ont été étudiées. Toutes commencent de la même manière par la mise en place d'une sonde n'importe où dans un bac du chariot pour suivre le refroidissement, puis, lorsque l'opérateur constate que la sonde indique une température inférieure à 10°C :

1. mesure de la température en un point quelconque d'un bac. C'est la stratégie actuellement utilisée par l'hôpital ;
2. mesure de la température en un centre géométrique d'un bac ;
3. mesure de la température en deux points quelconques de deux bacs différents ;
4. mesure de la température en deux centres géométriques de deux bacs différents.

Ainsi, nous avons fait varier la taille de n par rapport à la stratégie en vigueur mais nous avons également modifié l'endroit de la mesure en imposant pour certaines stratégies que la température soit mesurée dans un centre géométrique.

En fonction de la température dans les bacs, la croissance de *C. perfringens* peut être plus ou moins élevée entre la sortie de la cellule et la consommation. Nous avons relié la température à la probabilité que la croissance dépasse 1 log ufc/g afin de relier le plan d'échantillonnage à un risque. Nous avons tracé la courbe de la probabilité de sortie de la cellule de refroidissement en fonction de la probabilité d'une croissance supérieure à 1 log ufc/g pour faire le pendant de la courbe d'efficacité des contrôles par attribut (probabilité que le lot soit jugé bon en fonction de la prévalence du lot). Les stratégies ont été comparées en étudiant le comportement des courbes obtenues.

Dans ce cas d'étude, nous n'avons pas cherché à appliquer la théorie bayésienne de la décision car nous ne disposons d'aucun élément pour établir la liste des décisions et la fonction de coût. Nous avons donc comparé des stratégies qui nous semblaient applicables dans la cuisine centrale d'un hôpital.

8.3 Modélisation de la température

Dans le modèle que nous avons élaboré, nous avons conservé la distinction entre le centre géométrique et les autres points d'un bac (ainsi que préconisé par Poumeyrol et al. (2012)) mais nous avons ajouté la dimension temporelle car, d'après les données, la température varie beaucoup entre le début et le fin du refroidissement rapide (cf. figure 2 de l'article en préparation).

En plus des données de temps-température, nous disposons de plusieurs covariables comme le jour de la mesure, la taux de remplissage du chariot, la cellule utilisée (l'hôpital disposait de trois cellules) ou le taux de remplissage de la cellule (une cellule peut accueillir trois chariots). Nous disposons de deux jeux de données, le premier portant uniquement sur des profils de temps-température de centres géométriques de bacs et l'autre de profils de points quelconque dans un bac. Nous avons donc décidé d'élaborer deux modèles, l'un pour la température de centre géométrique (variabilité inter-bacs) et l'autre modélisant la différence de température entre un centre géométrique et un point quelconque d'un même bac (variabilité intra-bac). Pour chaque type de variabilité, plusieurs modèles mixtes contenant différentes covariables ont été élaborés et nous avons retenu celui dont le BIC^b était le plus petit.

Les paramètres des deux modèles ont été estimés puis des températures de centre géométrique et de points quelconques ont été simulés de manière à estimer la fraction de points du chariot dont la température est inférieure à 10°C et la probabilité de sortir de la cellule de refroidissement pour chaque stratégie et pour des durées passées dans la cellule de 50, 60, 80, 100, 120 et 150 min. Ces résultats nous ont permis de tracer, pour chaque stratégie, la courbe de probabilité de sortie en fonction de la proportion de points inférieurs à 10°C (figure 4 de l'article en préparation).

Ensuite, nous avons simulé des profils de temps-température en tenant compte de la durée passée dans une cellule de refroidissement puis dans la chambre froide. Ces profils

b. BIC signifie *Bayesian informative criterion*.

ont été utilisés pour estimer la croissance de la population de *C. perfringens* en utilisant le modèle de germination-croissance décrit dans Jaloustre et al. (2011) afin de tracer la probabilité de sortie d'une cellule en fonction de la probabilité de croissance supérieure à $1 \log \text{ ufc/g}$ pour chaque stratégie (figure 5 de l'article en préparation).

8.4 Quel plan choisir ?

Les probabilités de sortie sont proches de 0 pour les quatre stratégies lorsque le chariot a passé moins de 60 min en cellule parce que les points ont presque tous une température supérieure à 10°C . A l'inverse, les probabilités de sortie sont proches de 1 pour une durée d'au moins 150 min en cellule. Entre 60 et 150 min, les stratégies comportant au moins une mesure de la température dans un centre géométrique (stratégies 2 et 4) ont des probabilités de sortie moindre que les autres (stratégies 1 et 3) car le centre géométrique fait partie des points les plus chauds d'un bac. En revanche, le fait de mesurer la température en un (stratégies 1 et 2) ou deux points (stratégies 3 et 4) en plus de la sonde placée dans un bac lors de l'entrée en cellule de refroidissement ne modifie pas beaucoup les probabilités de sortie puisque ces probabilités sont pratiquement les mêmes pour les stratégies 1 et 3 d'une part et les stratégies 2 et 4 d'autre part.

Quelle stratégie (et donc quel plan d'échantillonnage de température) choisir ? Tout dépend du risque que l'opérateur est prêt à prendre (cf. figure 5 de l'article en préparation). S'il considère qu'une probabilité d'une croissance supérieure à $1 \log \text{ ufc/g}$ égale à 0,15 (en termes mathématiques : $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g}) = 0,15$), alors toutes les stratégies étudiées conviennent puisque même lorsqu'un chariot n'a passé que 50 min dans une cellule (cette durée a été choisie parce que c'était la plus courte observée dans le jeu de données dont nous disposons), cette probabilité n'excède pas 0,14. La stratégie 1 est néanmoins la plus adaptée car c'est celle qui nécessite le moins de manipulations pour l'opérateur. En revanche, si l'opérateur ne veut pas que $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g})$ soit supérieure à 0,04, il a intérêt de choisir une stratégie comportant une mesure de température de centre géométrique car cette stratégie entraîne une probabilité de sortie faible lorsque $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g})$ est supérieure à 0,04. Ainsi, le chariot restera plus longtemps que 80 min dans 95% des cas, ce qui réduira $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g})$. La stratégie 2 semble la plus adaptée dans ce cas. Elle est plus contraignante que la stratégie 1 car il faut mesurer la température en un point précis (le centre géométrique d'un bac) qui n'est pas forcément très facile d'accès. Cependant, elle permet que $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g})$ ne soit pas supérieure à 0,04 dans plus de 95% des cas et elle est moins difficile à mettre en œuvre que la stratégie 4, qui nécessite la mesure de la température de deux centres géométriques de bacs.

Cette application illustre le fait que différentes stratégies peuvent être comparées en tenant compte du risque pour le décideur (ici $\mathbb{P}(\text{croissance} > 1 \log \text{ ufc/g})$) et de la difficulté de mise en œuvre d'une stratégie. Nous avons volontairement choisi de rester dans un cadre simple et de ne pas envisager trop de stratégies pour montrer ce qu'il est possible de faire lorsqu'il faut se passer d'une fonction de coût (que ce soit par manque de temps pour l'élaborer ou par manque d'information) tout en conservant un caractère opérationnel aux recommandations.

8.5 Article

L'article reproduit dans cette partie a été soumis à la revue *Food Control*.

Impact of temperature sampling strategy on the risk of *Clostridium* growth: application to rapid cooling of food in institutional food service facilities

Natalie Commeau^{a,b,c*}, Séverine Jaloustre^a

^a ANSES French Agency for Food, Environmental and Occupational Health & Safety, 23 avenue du Général de Gaulle, 94706 Maisons-Alfort, France

^b AgroParisTech/INRA, UMR MIA 518, 16 rue Claude Bernard, 75231 Paris, France

^c AgroParisTech ENGREF, 19 avenue du Maine, 75732 Paris, France

* Corresponding author. Tel: +33 1 44 08 16 74 Fax: +33 1 44 08 16 66 E-mail address: natalie.commeau@m4x.org (Natalie Commeau)

Abstract

In French institutional food services, cooked food must be cooled within two hours of cooking and the temperature of the cooked food must then be lower than 10°C to avoid significant growth of *C. perfringens*. However, food placed on trolleys and put in a rapid cooling unit (a blast-chiller) shows great variation in temperature. What criterion should then be used to decide when the trolley can be removed from the blast-chiller? We studied four different strategies of when to remove the trolley by fitting 217 time-temperature profiles collected in a central hospital kitchen to mixed-effect models. We modelled the temperatures between and within containers holding meals on trolleys. We compared the strategies as to the probability that *C. perfringens* can proliferate to at least 1 log CFU.g⁻¹. For each strategy, this probability was plotted against the probability of the trolley leaving the blast-chiller. The probability of leaving the blast-chiller varied depending on the number of measurements taken on the trolley and on the points where temperature was measured. The best strategy depends on the risk of *C. perfringens* growth exceeding 1 log₁₀ CFU.g⁻¹ the food service operator accepts and the temperature monitoring protocol that is implemented.

Keywords

Chill chain, thermal model, sampling strategy, control measures

1. Introduction

Food safety is crucial in institutional food services given the large number of meals served every year (3 10⁹ meals per year in France, Anonymous 2011). With chill-chain procedures, many meals (thousands per day and per kitchen) are cooked, cooled soon after cooking, kept cold, transported to catering facilities, and reheated just before serving. Cooling, reheating and distribution operations have been identified as the main causative factors for *Clostridium perfringens* outbreaks (de Jong, A.E.I., Rombouts, F.M., & Beumer, R.R. 2004; de Jong, A., Beumer, R.R., & Zwietering, M., 2005), which represented more than 20% of foodborne outbreaks in catering in 2009 (INVS, 2011). It is thus important for food service operators to define efficient control measures while accounting for inherent variability due to processing. Variability refers to observed differences attributed to true heterogeneity or diversity in a population and usually cannot be reduced by further measurement or study (EPA, 1997). It reflects all the possible outcomes, which have to be taken into account when defining control measures.

Regarding the cooling process, French regulations require cooling from 63°C down to 10°C within 2 h (Anonymous, 2009). To estimate the efficiency of this requirement, the present study investigated the impact of food preparation operations on the risk of *C. perfringens* growth exceeding 1 log₁₀ CFU/g in a French central hospital kitchen. More specifically, our study focused on the impact of the cooling protocol on *C. perfringens* growth during cooling period. In the French hospital, after a 1 h boiling step, food products are distributed into containers, which are then placed on a trolley. This trolley is then placed in a blast-chiller and a randomly selected container is equipped with a thermal probe for real-time temperature measurements in food products. As soon as the measured temperature reaches

10°C, a beep alerts the food service operator, who measures the temperature in a different container. If the temperature measured is at 10°C or lower, the trolley is removed from the blast-chiller and placed in refrigerated storage. Thus, in this hospital, the decision for moving food products out of the blast-chiller is only based on a double temperature measurement sampling strategy.

However, temperature is not uniform throughout the trolley, so when an operator measures the temperature of a piece of meat in a container, the result cannot necessarily be extrapolated to the rest of the trolley. In other words, is the temperature variability throughout the trolley small enough so that one measurement gives a correct overview of all the existing temperatures in the trolley? Using data acquired in a central kitchen of a French hospital (Poumeyrol, G., Morelli, E., Noel, V., & Cornu, M., 2012) estimated not only that there is a large variability on a trolley but also within a given container because the difference between the temperatures measured at random in a container and at its geometric centre can be more than 16.7°C. Thus, it is likely that when a food service operator measures a temperature of 10°C or less, the temperature at other points of the container may be actually higher and favourable for *C. perfringens* growth, which can occur between 11 and 53°C, with an optimum growth temperature around 43°C (Jaloustre, S., Cornu, M., Morelli, E., Noël, V., & Delignette-Muller, M.L., 2011). Therefore, temperature variability should be taken into account in the decision to move the trolley out of the blast-chiller.

In the present study, we used time-temperature profiles acquired in one of the central kitchens described in Poumeyrol, G., Morelli, E., Noel, V., & Cornu, M., 2012 to compare different strategies for moving food out of the blast-chiller, including the current one. The comparison included the impact of each strategy on the growth of *C. perfringens* from the entrance in the blast-chiller to consumption of the meals. Before comparing these strategies, thermal models of temperature were developed for when the trolley is in the blast-chiller and when it is in a cold room.

2. Materials and methods

2.1. Risk assessment model

To estimate the impact of the temperature sampling strategy on the length of time a trolley spends in a blast-chiller, and therefore on *C. perfringens* growth, the present risk assessment model describes the cooling process, which can be divided into two steps: an initial rapid cooling in a blast-chiller, and a final cooling in cold storage. The decision for moving food products out of the blast-chiller is based on food temperature sampling. Figure 1 shows a representation of the whole risk assessment model. This model describes the thermal changes in food products and the corresponding change in *C. perfringens* vegetative cell concentration starting from an initial spore concentration. The model was used to predict the proportion of food on a trolley that has reached a temperature of 10°C or lower as well as of *C. perfringens* growth as a function of the temperature sampling strategy. All model variables are presented in Table 1.

2.1.1. Initial concentration of *C. perfringens* spores

C. perfringens growth, defined as the difference between final and initial concentrations on a decimal logarithm (\log_{10}) scale, was the only biological parameter estimated in the present model. The initial concentration of *C. perfringens* spores was set to 0 \log_{10} CFU.g⁻¹ for simplicity.

2.1.2. Blast-chiller thermal model

The cooling operations for a beef-in-sauce meal carried out in one hospital kitchen has been continuously investigated using probes placed directly in the food product (Rabi *et al.*, submitted). According to the collected data, temperatures observed on a given trolley change:

- from day to day;
- among containers and
- within containers.

Therefore, two models were constructed to determine the kinetic of temperature of the food product anywhere on a trolley:

- 1) one was dedicated to the temperature kinetics at the geometric centre of a container; this model is called the centre temperature model (Model CT);

- 2) the other was designed to estimate the difference between the temperature anywhere within a given container and the temperature at its geometric centre, because temperature in a container appears to be linked to the temperature of its geometric centre. This model is called within-container temperature model (Model WCT) because it is used to describe the within-container variability of temperature.

Both models are mixed-effect models. A mixed-effect model is a model in which the outcome depends on fixed and random factors. The studied levels of a fixed factor are the levels of interest. A factor is considered random when the levels under study are a sample from the population of interest (Jaloustre, S., Guillier, L., Morelli, E., Noël, V., & Delignette-Muller, M.L. 2012).

Model CT

This model describes the temperature at the geometric centre of a container on a trolley. The rapid cooling step can be modelled with an exponential model: the temperature of a geometric centre at time t depends on t via the exponential function. The rapid cooling step is modelled as follows, in natural log (ln) format:

$$\ln T_{cg} = \ln T_{cg,0} - kt \quad (1)$$

where T_{cg} (resp. T_{cg0}) is the temperature of the geometric centre of a container at time t (resp. at $t=0$). We added two explanatory variables to this model: ‘filling rate’ and ‘day’. The filling rate of the trolley is important because it can explain part of the variability observed. This variable describes to what extent the trolley is filled and ranges from 0 (the trolley is empty) to 1 (the trolley is full).

Temperature of the geometric centre varies from day to day. The variable ‘day’ thus covers several factors such as:

- the interval of time between the end of the cooking step and entrance into the blast-chiller;
- the amount of sauce and meat and the spatial localisation of the meat in containers;
- the blast-chiller itself, because there were three different devices in the hospital. Each blast-chiller can contain up to three trolleys at the same time. On a given day, not all the blast-chillers were used and they were not always full;
- the level of the container on the trolley. A trolley has 20 levels. Depending on the level, the meal may cool differently.

There were not enough data to assess properly all these factors. Therefore, they were grouped together in the factor ‘day’. ‘Filling rate’ (*fil_rate*) is a fixed-effect regressor, whereas ‘time’ is both a fixed- and a random-effect regressor for the temperature of container i on day j . The equations of Model CT for temperature $T_{cg,ij}$ at the geometric centre of container i on day j are the following:

$$\ln(T_{cg,ij}) = \beta_{0j} + \beta_{1j}t + \beta_{2j}fil_rate + \varepsilon_{ij} \quad (2)$$

where $\beta_{0j} = \ln T_{cg0} + u_{0j}$ and $u_{0j} \sim N(0, \sigma_0)$

$$\beta_{1j} = \gamma_1 + u_{1j} \text{ and } u_{1j} \sim N(0, \sigma_1)$$

$$\beta_{2j} = \gamma_2$$

$$\varepsilon_{ij} \sim N(0, \sigma_d)$$

The effect of the factor ‘day’ is described by

- u_{0j} , the among-day variability of temperature at the entrance to the blast-chiller, and
- u_{1j} , the among-day variability of the time slope.

The effect of ‘filling rate’ is described via parameter γ_2

The temperature is expressed in degrees Celsius (°C) and the time is expressed in minutes (min).

Model WCT

This model describes the difference between the temperature anywhere in a container and the temperature at the geometric centre of the same container. According to previous data (Poumeyrol, G., Morelli, E., Noël, V., & Cornu, M, 2012), the temperature at the geometric centre is among the highest in a container. There are many factors (level of the container, position of the thermometer, type of blast-chiller used, number of trolleys in a blast chiller) that could account for temperature variability

within a container. However, given the size and configuration of our data set, we considered only one factor: the position where the temperature is measured in a container.

In Model WCT, 'time' is a fixed-effect regressor. The equations of Model WCT for the temperature $T_{cont,ik}$ measured at position k in container i are as follows:

$$T_{cont,ik} - T_{cg,i} = \delta_{0i} + \varphi_1 t + e_{ik}, \quad (3)$$

where $\delta_{0i} = \varphi_0 + \nu_{0i}$ and $\nu_{0i} \sim N(0, \tau_0)$

$$e_{ik} \sim N(0, \tau_e)$$

The effect of time is described via parameter φ_1 and the random effect of the probe position is described by ν_{0i} .

The coefficients of the models were estimated using the package *lme4* in R (Bates, 2010).

To check whether the models fit the data well, 5000 time-temperature profiles were simulated using model CT and the 95% confidence interval obtained from these simulations was compared to the observed time-temperature profiles.

2.1.3. Durations of blast-chiller treatment

The probability that the trolley is moved out of the blast-chiller was calculated for six different blast-chilling durations. The shortest duration spent in the blast-chiller was 50 min, because some trolleys leave the blast-chiller 50 min after entrance. We also chose 60, 80 and 100 min, because these durations correspond to those that are often used when food temperatures are monitored using Thermo Button temperature loggers. The duration 120 min was also selected because it is the maximum duration a trolley is to spend in a blast-chiller (according to Anonymous 2009). Finally, we also chose 150 min to study the effect of the trolley being left in the blast-chiller more than the legally authorized duration in France.

2.1.4. Probability of leaving the blast-chiller

For each strategy of temperature monitoring (see 2.2) and each chosen blast-chilling duration, every coefficient of both models was simulated 10,000 times and temperatures calculated at time $t=0$. Because we did not have enough information to explain all the variability with the identified covariates, residual standard deviations were high. Thus, simulated temperatures ranged from very low to very high values, which is not representative of reality. To correct for this bias, the simulated temperatures at $t=0$ above 85°C or below 40°C were removed, because these values do not fall within the range of observed temperatures. Finally, the probability of leaving the blast-chiller was calculated by counting the number of occurrences when simulated temperatures were below 10°C and by dividing this by the number of simulations. This program was coded in R (R, 2011, version 2.14.0).

This simulation pattern was chosen because it assesses the probability of leaving the blast-chiller on any given day of production, which is interesting for the food service provider.

2.1.5. Cold storage thermal model

For this model, 217 collected temperature profiles (Rabi *et al.*, submitted) were fitted to an exponential model, as follows:

$$T(t) = T_a + (T_0 - T_a) \times \exp(-k' t) \quad (4)$$

where $T(t)$ is the temperature (°C) in the food at time t (hours), T_0 and T_a initial and asymptotical temperatures (°C) and k' the exponential temperature decrease rate (hours⁻¹). In the present risk assessment model, T_0 corresponds to the temperature at the end of the rapid cooling, provided by the blast-chiller thermal model, and T_a is assumed to equal 0°C. Therefore, only k' values were estimated using collected data. Observed variability on the 217 estimated k' values was described by a Gamma distribution.

2.1.6. Growth during the whole cooling process

The model published by Jaloustre, S., Cornu, M., Morelli, E., Noël, V., & Delignette-Muller, & M.L., 2011 was used to predict *C. perfringens* growth from spores under dynamic thermal conditions. This model describes separately the uncertainty of parameters and variability on the 'work-to-be-done' (term defined by Baranyi, J. & Roberts, T.A., 1995; Robinson, T.P., Ocio, M.J., Kaloti, A., & Mackey,

B.M., 1998). to describe what *C. perfringens* cells must do during germination to reach the exponential growth phase), outgrowth and lag phase. Since our study focused on variability, only the variability of the parameters described in the published model was taken into account, using median values of uncertain parameters.

2.2. Studied strategies

The currently used protocol is the following: a blast-chiller is put in operation; an operator places a probe in a container to monitor the temperature. When the probe indicates a temperature below 10°C, it beeps. The operator then measures the temperature in another container with another thermometer. If the second temperature is also below 10°C, the trolley is removed from the blast-chiller and placed in cold storage. If the temperature is greater than 10°C, the trolley stays in the blast-chiller and the operator measures the temperature again several minutes after the previous measurement.

Four strategies were compared to study their impact on the growth of *C. perfringens*. In every strategy, the temperature is measured anywhere on the trolley and the trolley is not removed from the blast-chiller if this temperature is greater than 10°C. In addition, the trolley is removed only as soon as:

- 1) the temperature is below 10°C when measured anywhere in any other container on the trolley. (this is the current strategy);
- 2) the temperature is below 10°C when measured at the centre of another container on the trolley;
- 3) temperatures are below 10°C when measured anywhere in two other containers on the trolley;
- 4) temperatures below 10°C are measured at the centres of two other containers on the trolley.

Strategy 2 was studied because the temperature at the centre of a container is usually higher than the temperature outside of the centre. Strategy 3 (resp. 4) was studied because it seemed to be a more conservative than strategy 1 (resp. 2), because two measurements are taken instead of just one. Therefore, the probability of leaving the blast-chiller when almost all the temperatures are higher than 10°C should be lower for Strategy 3 (resp. 4) than for Strategy 1 (resp. 2).

The impact of each strategy was studied for six blast-chilling durations: 50, 60, 80, 100, 120 and 150 minutes.

To estimate the impact of the temperature sampling strategy on *C. perfringens* growth, a criterion was defined as the probability that *C. perfringens* growth exceeds 1 log₁₀ CFU.g⁻¹. Combining blast-chiller and cold-storage models made it possible to simulate 10,000 temperatures profiles describing the whole cooling-down process for each studied duration as well as the associated 10,000 final concentrations of *C. perfringens* vegetative cells. Because the initial *C. perfringens* concentration was set to 0 log₁₀ CFU.g⁻¹, the probability that *C. perfringens* growth exceeds 1 log₁₀ CFU.g⁻¹ was then estimated by calculating the proportion of simulated final concentrations exceeding 1 log₁₀ CFU.g⁻¹.

3. Results

3.1. Parameters of Models CT and WCT

The parameters of Models CT and WCT are described in Table 1. For model CT, the intercept is high compared to the fixed and the random effects of factor 'day'. For both models, the parameters linked to time (fixed and random effect of factor 'day' for model CT and constant for WCT model) had the smallest values. However, when multiplied by high values of time (e.g. $\gamma_1 \times 120 = -2.03$ in model CT), the covariate 'time' was not negligible.

Overall, as shown in Figure 1, Model CT correctly simulates temperature profiles: the 95% confidence interval of the simulated profiles contains most of the observed profiles. Although Model CT appears to slightly overestimate temperature at the geometric centre, this overestimation predicts temperatures that are more favourable for *C. perfringens* growth than those observed in the hospital kitchen. Thus Model CT can be considered as a conservative model. Model WCT also simulates temperature differences similar to the observed ones: the 95% confidence interval of the simulated temperature differences, shown in Figure 2, contains almost all the observed temperature differences. As for Model CT, simulated temperature differences are slightly higher than real ones, leading to slight overestimates of temperatures anywhere in a container. However, this was considered as a conservative alternative. The profiles of observed temperature differences were not exactly linear in

time, but the between-distance variability in time was very low, so we decided to remove this effect from the model (results not shown). The intra-class correlation coefficient was $1.6/(1.6+5.61)=0.22$, which indicates that 22% of the variance was explained by the position of the temperature logger in the container.

3.2. Probability of leaving the blast chiller according the chosen strategy

Probabilities of leaving the blast-chiller according to the strategy used and considered length of blast-chilling duration are shown in Figure 4. As expected, the highest probability of leaving the blast-chiller at each studied duration is obtained with Strategy 1, because this strategy considers only two measurements of temperature anywhere in two containers (one reading as in each strategy and one additional reading). Probability of leaving is lower for Strategy 3 than for Strategy 1, because the probability of measuring two additional temperatures below 10°C anywhere in two containers is lower than the probability of measuring one additional temperature below 10 °C anywhere in two containers. Temperature at the geometric centre of a container is usually higher than temperatures elsewhere; therefore, the probability of measuring a temperature below 10°C at the centre of a container is lower than the probability of measuring a temperature below 10°C anywhere else. Thus, probability of leaving the blast-chiller is lower for Strategy 2 than for Strategy 1. Nevertheless, there is a large difference between the probability of leaving for Strategies 1 and 3 and the probability of leaving for Strategies 2 and 4. We believe that this is due to the difference between temperatures at the geometric centre of a container and temperatures elsewhere in the container. This difference can be large, as shown in Figure 3, where the difference between central and non-central temperatures can be as high as 20°C.

At a duration of 50 min, the probability of leaving is very low (0.05 for Strategy 1 and almost 0 for the other strategies). In fact, at 50 min, the temperature is greater than 10°C at almost all the points in the containers. Therefore, the probability of leaving the blast-chiller is very low, even for Strategy 1. At a duration of 60 min, the same result is observed although the difference is less marked. The probability of leaving increases with duration, but the probability ranking of the strategies never changes and the amount of increase depends on the strategy (Figure 4). The difference between probabilities increases with duration, e.g. at 50 min, the probability of leaving is 0.02 for Strategy 1, and $0.03 \cdot 10^{-3}$ for Strategy 4, whereas at 100 min, it is 0.60 for Strategy 1 and 0.25 for Strategy 4. This difference between probabilities decreases for durations greater than 100 min. This increase is due to the within-container variability, which is highest for blast-chilling duration of 100 minutes. Due to this variability, the temperature at the geometric centre of a container and anywhere else in the container can be very different; therefore, the probability of leaving the blast-chiller is very different for 100 min. At 150 min, the probabilities of leaving for all the strategies are rather similar, indicating that most points in the containers have a temperature below 10°C so the probability of leaving the blast-chiller is high, even for Strategy 4.

3.3. Probability of *C. perfringens* growth

The probability of leaving the blast-chiller as a function of the probability of *C. perfringens* growth to exceed $1 \log_{10} \text{CFU.g}^{-1}$ is shown in Figure 5. The impact of blast-chilling duration on the probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$ clearly appears up to 100 minutes. Adding 10 more minutes to an initial blast-chilling duration of 50 minutes helps reduce this probability from $13.5 \cdot 10^{-2}$ to $8.2 \cdot 10^{-2}$. Moreover, from 60 to 100 minutes, adding 20 minutes halves the probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$. When time spent in the blast-chiller exceeds 100 minutes, the probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$ becomes relatively low and does not much change with blast-chilling duration.

An efficient temperature sampling strategy is expected to lower the probability of leaving the blast-chiller when *C. perfringens* growth may exceed $1 \log_{10} \text{CFU.g}^{-1}$. According to this requirement, for extreme blast-chilling durations (50, 120 and 150 minutes in figure 5) and corresponding extreme probabilities of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$, the four proposed strategies have equivalent efficiency. In contrast, the impact of these strategies, applied for blast-chilling durations

ranging from 60 to 100 minutes, and for probabilities of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$ is more variable. For a given probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$, Strategy 1 appears to be the least efficient strategy, whereas Strategy 4 clearly helps ensure that container temperatures have reached 10°C or less. For example, for a probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$ at $8.2 \cdot 10^{-2}$, Strategy 1, which is the currently applied protocol, leads to a probability of leaving blast-chiller that is 60-fold higher than the probability observed with Strategy 4.

4. Discussion

Because the probability that temperature has reached 10°C or less by 50 min is very low and because the only strategy with a non-negligible probability of the trolley being moved out of the blast-chiller at this time is Strategy 1, this strategy should not be used for low blast-chilling durations. For the same reason, Strategies 1 and 3 should not be used when the batch has spent only 60 min in the blast-chiller. Four strategies were tested in this study. Other strategies could have been investigated, e.g. one in which trolley is removed from the blast-chiller as soon as a minimum of four points anywhere on the trolley show temperatures below 10°C . However, we chose not to study too many strategies, because we wanted to demonstrate how strategies can be compared so that our method can be used to compare other strategies by those interested. In addition, the food service operator will not measure the temperature in many different places on a trolley because it is too time-consuming.

In this study, all the data were used to assess the parameters of the models, regardless of the time containers spent on the trolley before entering the blast-chiller, after the end of the cooking period. In the dataset we used, this waiting time ranged from 29 min to 86 min. Nevertheless, we did not have enough data to characterise this waiting time. This step could be studied to check if two different strategies of removing the trolley from the blast-chiller could be applied depending on the waiting time, because the temperature of the meal on the trolley is generally higher when the waiting time is short and lower when the waiting time is long.

In this study, the risk was defined as the probability of *C. perfringens* growth exceeding $1 \log_{10} \text{CFU.g}^{-1}$. According to the concepts developed by *Codex Alimentarius*, 2004, this criterion can be defined as a performance criterion. It would be interesting to evaluate the impact of a temperature sampling strategy on a food safety objective or a diarrheal risk for consumers. This type of evaluation would require modelling the whole process, including cold storage, reheating and distribution steps as well as the consumption of food products by patients and a dose-response model. Once all potential sources of variability have been described, it would then be possible to estimate the impact of each step on the final risk for consumers. However, the lack of data correctly describing some of these steps, from preparation to consumption, made it impossible to perform such an analysis.

Even after 2.5 h spent in the blast-chiller, not all the meat-in-sauce meals on the trolley have a temperature lower than 10°C , because the probability of leaving the blast-chiller was not equal 1 for any strategy (see Figure 5). One way to ensure that all the containers on the trolley have a temperature below 10°C is to change the French regulations and allow the trolley to spend more time in the blast-chiller or to reinforce the cooling system so that the cooling of the dish is faster than it is currently.

Many control measures applied by manufacturers of food processing or recommendations given to consumers for the preparation of their food, based on temperature monitoring. Yet sources of variation (microbiological, technological) are numerous throughout the process or preparation. Thus, these control measures or recommendations taken must be validated, that is to say, there is a need to provide evidence they are effective. The risk assessment can then help to provide such evidence (Jaloustre, S., Guillier, L., Morelli, E., Noël, V., & Delinegnette-Muller, M.L., 2012). The work we have presented here is an illustration of the use of that quantitative risk analysis can help to enable decisions to be taken on a more quantitative basis (Buchanan, R.L., 1995; Mayes, T., 1998; Popov Janevska, D., Gospavic, R., & Popov, V., 2009).

Acknowledgement

We thank Laurent Guillier for useful comments and the financial support of the ANR.

References

Anonymous (2009). Arrêté du 21 décembre 2009 relatif aux règles sanitaires applicables aux activités de commerce de détail, d'entreposage et de transport de produits d'origine animale et denrées alimentaires en contenant. JORF n°0303 du 31 décembre 2009 Texte n°241

Anonymous (2011). Le Programme national pour l'alimentation. Programme interministériel piloté par le Ministère de l'agriculture, de l'alimentation, de la pêche, de la ruralité et de l'aménagement du territoire. Février 2011. Available at <http://agriculture.gouv.fr/IMG/pdf/PNA-09022011.pdf>

Bates, D.M., (2005). Fitting linear mixed models in R using the lme4 package. R News, The Newsletter of the R Project, 5, 27–30.

Baranyi, J. & Roberts, T.A., (1995). Mathematics of predictive microbiology. *International Journal of Food Microbiology*, 26, 199-218.

Buchanan, R.L. (1995). The role of microbiological criteria and risk assessment in HACCP. *Food Microbiology* 12, 421-424.

Codex Alimentarius, (2004). Report of the twentieth session of the Codex committee on general principles.

EPA, (1997). Guiding principles for Monte-Carlo analysis. Technical report, EPA.

INVS, (2011). Surveillance des toxi-infections alimentaires collectives, données de la déclaration obligatoire, 2009.

Jaloustre, S., Cornu, M., Morelli, E., Noël, V., & Delignette-Muller, M.L., (2011). Bayesian modeling of *Clostridium perfringens* growth in beef-in-sauce products. *Food Microbiology*, 28, 311-320.

Jaloustre, S., Guillier, L., Morelli, E., Noël, V., & Delignette-Muller, M.L. (2012) Modeling of *Clostridium perfringens* vegetative cell inactivation in beef-in-sauce products: A meta-analysis using mixed linear models. *International Journal of Food Microbiology*, 154, 44-51.

de Jong, A.E.I., Rombouts, F.M., & Beumer, R.R. 2004. Behavior of *Clostridium perfringens* at low temperatures. *International Journal of Food Microbiology* 97, 71-80.

de Jong, A., Beumer, R.R., & Zwietering, M. (2005). Modeling growth of *Clostridium perfringens* in pea soup during cooling. *Risk Analysis*, 25: 61-73.

Mayes, T. (1998). Risk analysis in HACCP: Burden or benefit? *Food Control* 9, 171-176.

Popov Janevska, D., Gospavic, R., & Popov, V. (2009). A novel HACCP system supported by QMRA for increased food safety, *WIT Transactions on Biomedicine and Health*, vol. 14. 297-306.

Poumeyrol, G., Morelli, E., Noel, V., & Cornu, M, (2012). Impact of the method chosen for measuring temperatures on the efficacy of rapid cooling of foods in catering facilities. *Food control*, 23, 345-350.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing.

Rabi, J., Guilpart, J., Morelli, E., & Derens, E. Submitted. Heat transfer inside an blast-cooling cell: simulation results from a zero-order model for beef-in-sauce catering meals. *Journal of Food Engineering*.

Robinson, T.P., Ocio, M.J., Kaloti, A., & Mackey, B.M., (1998). The effect of the growth environment on the lag phase of *Listeria monocytogenes*. *International Journal of Food Microbiology*, 44, 83-92.

Captions:

Table 1: Description of the variables in the risk assessment model.

Table 2: Values of the parameters of the two mixed-effect models, CT and WCT

Figure 1: Representation of the risk assessment model.

Figure 2: Observed time-temperature profiles at the geometric centre of containers (grey) and 95% confidence interval of the profiles simulated with model CT (black).

Figure 3: Difference between the temperatures at the geometric centre of a container and at any other point in the same container. Observed differences are indicated in grey and the 95% confidence interval of the differences simulated with model WCT are indicated in black.

Figure 4: Probability of leaving the blast-chiller for four strategies: a temperature anywhere in a container of the trolley is monitored and (1) a second temperature is read anywhere in any other container (squares); (2) a second temperature is read at the centre of any other container (circles); (3) two temperatures are read anywhere in two different containers (diamonds); (4) two temperatures are read at the centre of two different containers (triangles), according to the probability that the distribution of temperatures in the blast-chiller is lower than 10°C. For each strategy, the trolley leaves the blast-chiller only if all the measured temperatures are below 10°C. Six durations were studied: 50, 60, 80, 100, 120 and 150 min.

Figure 5: Probability of leaving the blast-chiller as a function of the probability that *C. perfringens* growth exceeds 1 log₁₀ CFU.g⁻¹ for the four proposed strategies: the temperature anywhere in a container of the trolley is monitored and (1) a second temperature is measured anywhere in any other container (squares); (2) a second temperature is measured at the centre of any other container (circles); (3) two temperatures are measured anywhere in any two different containers (diamonds); (4) two temperatures are measured at the centres of two different containers (triangles), according to the probability that the distribution of temperatures in the blast-chiller is lower than 10°C. For each strategy, the trolley leaves the blast-chiller only if all the measured temperatures are below 10°C. Six durations were studied: 50, 60, 80, 100, 120 and 150 min.

Table 1

Model	Parameter	Definition	Description
	$conc_0$	Initial concentration of <i>C. perfringens</i> spores	1 CFU.g ⁻¹ (0 log ₁₀ CFU.g ⁻¹ .)
Model CT	$T_{cg,ij}$ β_{0j} β_{1j} β_{2j} ε_{ij}	Temperature at the geometric centre of container <i>i</i> on day <i>j</i> Fixed and random intercept for day <i>j</i> Fixed and random effect for day <i>j</i> linked to the variable 'time' Fixed coefficient for filling rate Random error	$\ln(T_{cg,ij}) = \beta_{0j} + \beta_{1j}time + \beta_{2j}fil_rate_j + \varepsilon_{ij}$, $\beta_{0j} = \ln T_{cg0} + u_{0j}$ and $u_{0j} \sim N(0, \sigma_0)$ $\beta_{1j} = \gamma_1 + u_{1j}$ and $u_{1j} \sim N(0, \sigma_1)$ $\varepsilon_{ij} \sim N(0, \sigma_d)$
Model WCT	$T_{cont,ik}$ $T_{cg,i}$ δ_{0i} e_{ik}	Temperature at point <i>k</i> in container <i>i</i> Temperature at the centre of container <i>i</i> Fixed and random intercept for the centre of container <i>i</i> Random error	$T_{cont,ik} - T_{cg,i} = \delta_{0i} + \varphi_1time + e_{ik}$, °C °C $\delta_{0i} = \varphi_0 + v_{0i}$ and $v_{0i} \sim N(0, \tau_0)$ $e_{ik} \sim N(0, \tau_e)$
Cold storage thermal model	$T(t)$ k' T_a T_0	Temperature in food at time <i>t</i> (°C) Exponential temperature decrease rate (hours ⁻¹) Asymptotic temperature (°C) Initial temperature (units: °C)	$T(t) = T_a + (T_0 - T_a) \times \exp(-k't)$ $k' \sim Gamma(\kappa = 2.68, \theta^{-1} = 9.51)$ 0 °C Simulated using blast-chiller thermal model
Growth model	h_0 T_{min} T_{opt} T_{max} μ_{opt}	'work-to-be-done' Minimum temperature for <i>C. perfringens</i> growth (°C) Optimum temperature for <i>C. perfringens</i> growth (°C) Maximum temperature for <i>C. perfringens</i> growth (°C) Optimum <i>C. perfringens</i> growth rate (h ⁻¹)	$\ln(h_0) \sim N(1.5, 0.26)$ 12.4 45.5 51.9 4.2

Table 2

Model	Parameter	Value
CT	$\ln T_{cg,0}$	4.43
	σ_0	$9.95 \cdot 10^{-2}$
	γ_1	$-1.69 \cdot 10^{-2}$
	σ_1	$5.0 \cdot 10^{-3}$
	γ_2	-0.38
	σ_d	0.31
WCT	ψ_0	-5.67
	τ_0	1.60
	φ_1	$-2.48 \cdot 10^{-2}$
	τ_e	5.61

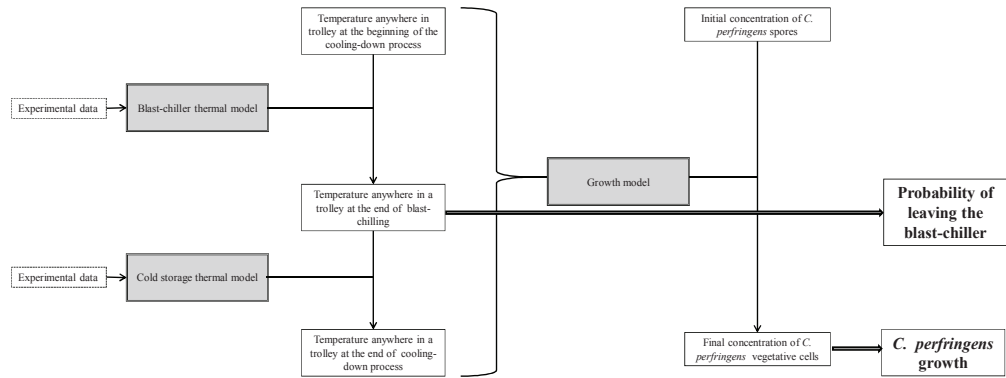


Figure 1

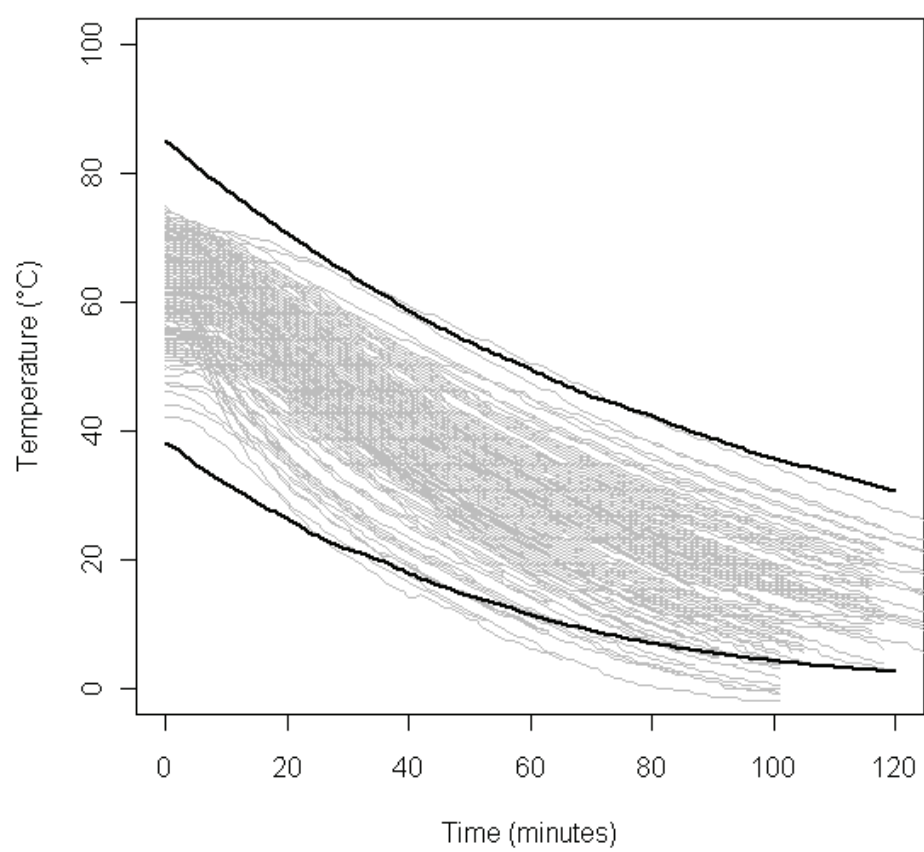


Figure 2

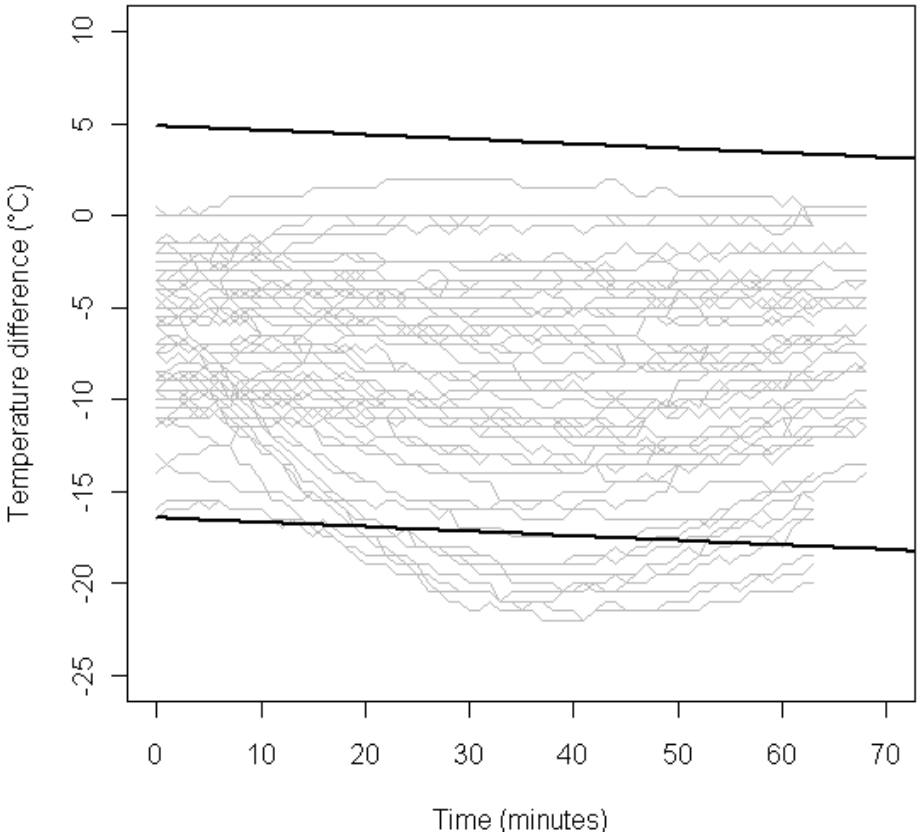


Figure 3

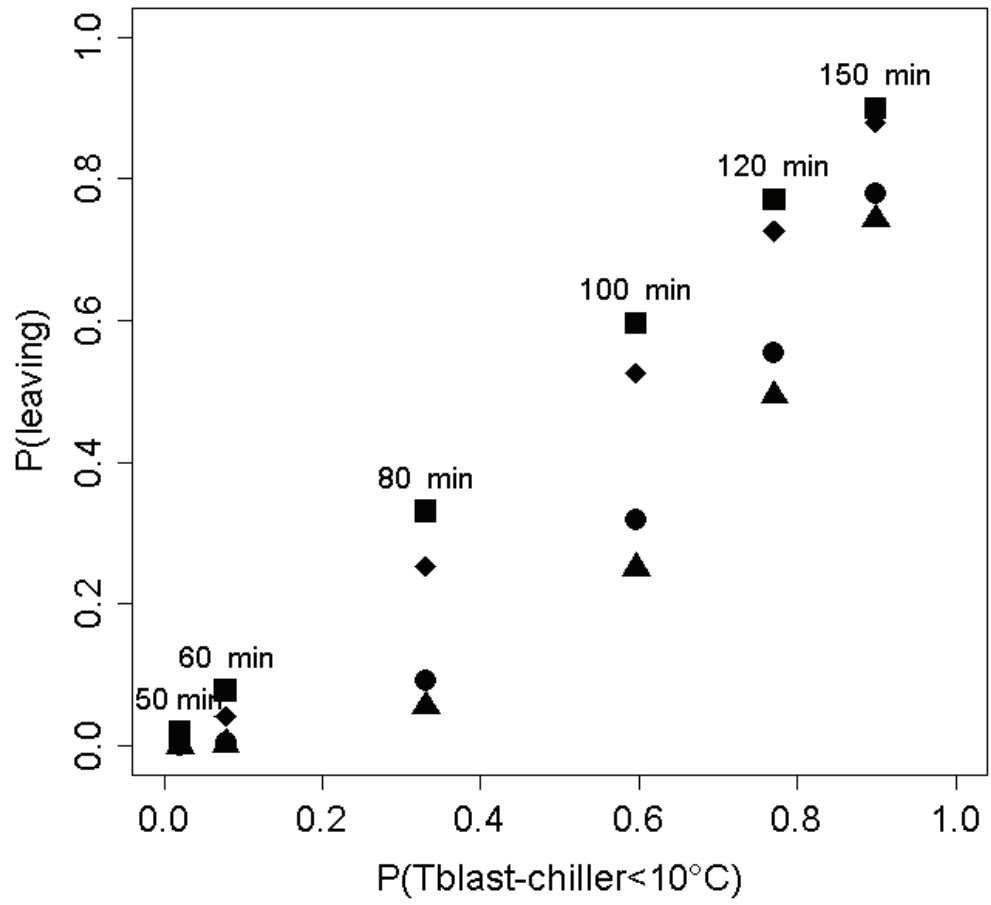


Figure 4

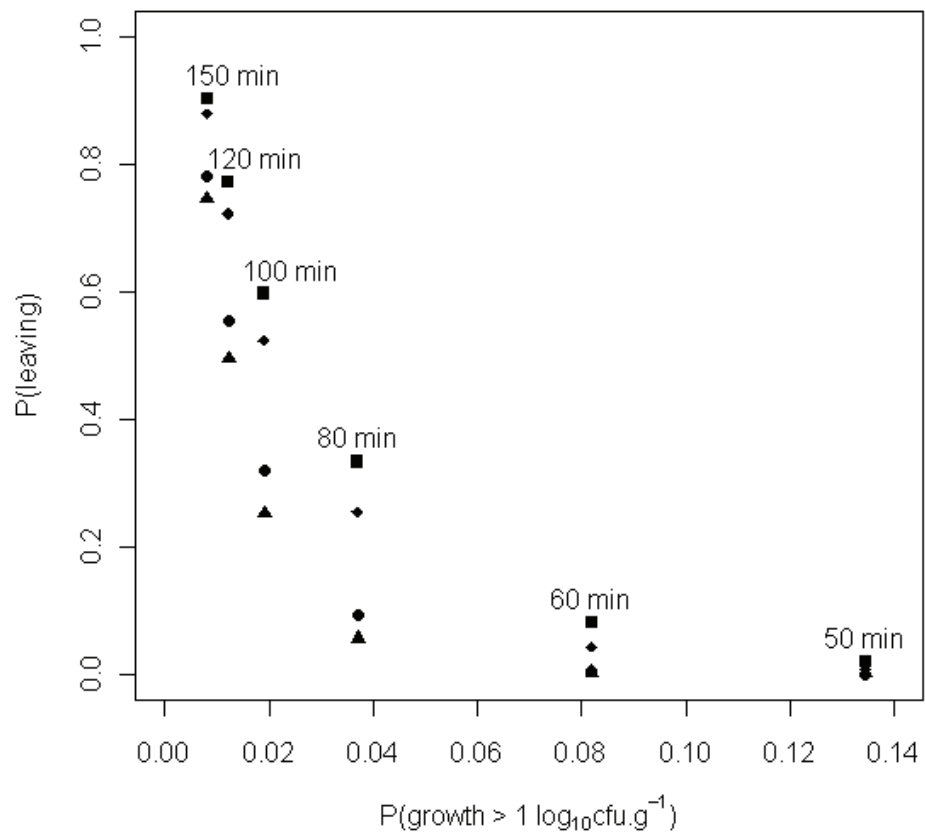


Figure 5

Dans cette partie, trois plans d'échantillonnage ont été étudiés :

- un plan d'échantillonnage microbiologique portant sur une période de production et visant à vérifier le système de production de l'usine. En fonction des résultats d'analyses, des actions correctives peuvent être menées ;
- un plan d'échantillonnage microbiologique portant sur un lot de production. Ce plan varie en fonction des résultats passés mais tient compte également de l'impact du choix de plan sur les résultats futurs ;
- un plan d'échantillonnage thermique portant sur un lot de fabrication. En fonction des températures relevées, le produit préparé passe à l'étape suivante ou reste encore à l'étape de production dans laquelle il se trouve.

Lorsqu'il met en place un plan d'échantillonnage, l'industriel doit se demander à quel endroit il est pertinent de le placer. En général, les matières premières sont analysées afin d'évaluer les fournisseurs et les produits sont analysés en fin de procédé. Cependant, selon le procédé étudié, il est possible que la sortie usine ne soit pas forcément l'étape la plus pertinente pour l'analyse. L'échantillonnage pourrait avoir lieu à une étape permettant de mieux détecter les produits contaminés ou alors d'assurer une meilleure traçabilité.

Par exemple, dans l'application "lardons" en reprenant la définition du lot "industriel" tel que défini par l'entreprise (cf. partie 2.2.2), il me semble intéressant d'échantillonner lors du raidissage. Cela permet d'une part d'obtenir les résultats alors que les produits sont encore dans l'usine puisque cette étape permet le stockage des poitrines de porc en attendant les commandes des clients qui déclenchent le tranchage. Ainsi, cela évite le retrait ou le rappel de lots. Même si les résultats étaient disponible une fois que les barquettes de lardons ont quitté l'usine, le retrait ou le rappel concernerait uniquement les lots "industriels" de barquettes constituant le lot "industriel" incriminé de poitrines de porc. Avec l'échantillonnage actuel pratiqué par l'usine, le lot jugé contaminé est un lot "industriel" de barquettes. Or ce dernier est composé de fractions de plusieurs lots "industriels" de poitrines de porc. Ainsi, il faut retirer ou rappeler tous les lots "industriels" de lardons qui ont été fabriqués à partir des lots "industriels" de poitrines qui compose le lot "industriel" de barquettes jugé contaminé après analyse des résultats du plan d'échantillonnage. Le nombre de lots "industriels" de barquettes à retirer ou à rappeler est donc plus élevé lorsque le contrôle est mené au niveau de la fin de procédé au lieu du raidissage.

Si on considère l'échantillonnage au niveau d'une période de production comme c'est le cas au chapitre 6, l'échantillonnage au niveau du raidissage n'est pas forcément plus intéressant pour l'entreprise car les résultats ne sont examinés qu'une fois que les produits ont quitté l'usine et parce que les retraits ou les rappels sont déclenchés par le client à partir des lots "industriels" de barquettes (cf. partie 6.2.3).

Bien évidemment, cette étude du lieu d'échantillonnage doit tenir compte de la distribution de contamination aux différentes étapes du procédé (en utilisant les modèles présentés aux chapitres 3 ou 5 par exemple) mais également des obligations réglementaires.

Troisième partie

CONCLUSION

Chapitre 9

Conclusion

9.1 Bilan

9.1.1 Travaux accomplis

Mon travail de thèse s’est très largement inscrit dans le projet Quant’HACCP, dont l’objectif était de proposer une voie quantitative pour adapter les mesures de maîtrises prises par les industriels. J’ai été impliquée dans les sous-projets 1 (“développements méthodologiques”), 2 (“*L. monocytogenes* dans les lardons”) et 3 (“*C. perfringens* dans les plats préparés de la restauration collective”).

L’une des premières tâches dans laquelle je me suis engagée a été la modélisation de la contamination par *L. monocytogenes* lors de la fabrication de lardons, depuis les poitrines de porc arrivant dans l’usine jusqu’aux barquettes. Ce travail a donné lieu à une publication (Billoir et al. (2011)) puis à une évolution traitée au chapitre 5.

En parallèle, nous avons décidé d’acquérir des données de contamination par *L. monocytogenes* et par la flore lactique. Pour cela, nous avons déterminé les étapes pertinentes où prélever (le choix s’est porté sur la fin du malaxage et le produit en sortie de chaîne de production) ainsi que la manière de prélever les produits (échantillonnage stratifié). Nous avons ensuite analysé les résultats de manière à pouvoir en tirer les premiers enseignements sans recourir à une quelconque modélisation. Ce travail a été explicité au chapitre 2 de cette thèse.

Par la suite, nous avons réfléchi à une autre manière de modéliser la contamination : au lieu d’étudier le devenir de la contamination d’une étape à l’autre de la fabrication et les liens entre les différentes étapes, nous nous sommes penchés sur des modèles décrivant la contamination à une étape donnée. J’ai étudié différents modèles afin d’explorer la hiérarchisation de la variabilité de la contamination : variabilité inter-lots, inter-unités ou inter-prises d’essai. L’inférence des paramètres a été menée selon l’approche bayésienne. Le chapitre 3 et l’article Commeau et al. (2012a) présentent le travail mené sur ce sujet. Ce type de modèle a l’avantage d’être plus simple que le modèle développé par Billoir et al. (2011). En particulier, ses paramètres sont beaucoup plus faciles à estimer. En revanche, cette estimation nécessite de disposer de données à l’étape du procédé considéré. En outre, aucun lien ne peut être établi entre les paramètres des modèles construits à deux étapes du procédé. Le modèle de Billoir et al. (2011) peut, lui, incorporer des données à tout

endroit de la chaîne, bien que l'opération puisse être ardue, et il est à même d'établir un lien entre les distributions de contamination observées à deux étapes données.

Lors de l'élaboration des modèles hiérarchiques, je me suis interrogée sur la nature d'une donnée microbiologique. En particulier, je me suis demandée quel impact sur l'estimation des paramètres pouvait avoir la forme du jeu de données. Ce questionnement a donné lieu à un travail de comparaison de la qualité d'inférence de paramètres sur deux modèles en utilisant les mêmes jeux de données mais exprimés différemment selon les modèles. Cette comparaison est présentée au chapitre 4 et a donné lieu à une publication (Commeau et al. (2012b)).

Les travaux sur la modélisation et sur les données microbiologiques ont très vite débouché sur la question du *lot*. Celle-ci a été assez difficile à exprimer parce que tous nos interlocuteurs (et nous-mêmes) parlions du lot sans comprendre de prime abord que ce terme n'avait pas la même signification pour l'autre. Ainsi, dans nos premiers échanges avec des industriels, nous ne nous sommes pas demandés si les lots "industriel" avaient un comportement répétitif et si leur contamination pouvait être modélisée par une même loi de probabilité. Cependant, nous butions toujours sur la définition du lot, car il est apparu au fil des mois que le lot "industriel" ne correspondait pas toujours au lot "statistique". Nous nous sommes donc penchés plus précisément sur le sujet et nous avons défini le lot "statistique" pour toutes les applications de cette thèse.

Cette réflexion sur le lot nous a été très utile pour mener à bien la seconde partie de cette thèse sur les plans d'échantillonnage. En effet, la théorie bayésienne de la décision suppose que la population dans laquelle est menée l'échantillonnage a les bonnes propriétés statistiques, mais elle ne dit rien sur les difficultés à définir une telle population.

La définition d'un lot "statistique" n'a pas été immédiate pour l'application "lardons" : il fallait élaborer des décisions pertinentes pour les lots définis, alors même que le lot "industriel" pour les barquettes n'était pas adapté. S'il était clair que le contrôle par attribut avec acceptation ou rejet de lot n'était pas le bon dès lors que l'entreprise ne le pratiquait pas et que ce type de plan ne nous semblait pas pertinent pour cette application, il a fallu définir un lot "statistique" et un ensemble de décisions possibles pour ce type de lots. Je n'aurais pas été aussi sensibilisée à cette question si je ne m'étais pas penchée sur la notion de lot durant la première partie de cette thèse. L'application de la théorie bayésienne de la décision pour le cas "lardons" est présentée au chapitre 6.

Une fois le lot défini, nous avons étudié la variabilité intra et inter-lots. Si la variabilité au sein d'un lot est plus élevée que la variabilité entre les lots, alors le contrôle libérateur semble inutile puisque le petit nombre d'analyses pratiquées par lot ne suffit pas à caractériser suffisamment sa contamination pour décider si le lot peut être livré ou non.

La théorie de la décision peut être appliquée à différents cas. Le cas "lardons" est particulier, dans la mesure où les décisions ne sont pas celles que l'on rencontre habituellement dans les exemples d'applications. En outre, trois décisions ont été considérées au lieu de deux. Le cas "saumon fumé" est plus courant dans la définition du lot "statistique" et dans les décisions prises mais il est original dans la mesure où il prend en compte les résultats futurs. Ce cas d'étude fait l'objet du chapitre 7.

Dans ces deux cas, j'ai travaillé sur la prévalence pour deux raisons : les autocon-

trôles des entreprises sont souvent constitués de données de détection qui permettent d'estimer la prévalence de la production et la loi modélisant la prévalence (la loi Beta) est conjuguée si la loi binomiale est utilisée pour modéliser les résultats de données de détection. Il est possible d'utiliser des données de dénombrement et de travailler avec la concentration. C'est un peu plus difficile à manipuler parce que la loi des paramètres de concentration n'est pas conjuguée. Le calcul analytique du risque bayésien étant impossible analytiquement, nous ne pouvions établir simplement sa valeur en faisant varier la taille de l'échantillon. Un moyen de déterminer le minimum de ce risque est de recourir à des algorithmes particuliers, comme celui présenté par Amzal et al. (2006) par exemple, permettant de déterminer le maximum d'une fonction (il s'agit en l'occurrence de la fonction d'utilité, égale à l'opposé de la fonction de coût). J'avais commencé à m'intéresser à ces algorithmes en début de thèse, puis j'ai abandonné cette piste pour me consacrer à la modélisation de la contamination des aliments par un micro-organisme et je n'ai pas eu le temps de la reprendre avant la fin de ma thèse.

Le projet Quant'HACCP comportait également un volet échantillonnage dans le sous-projet "*C. perfringens* dans les plats préparés de la restauration collective" dont j'étais responsable. Cette application me semblait intéressante, car elle comportait un plan d'échantillonnage de la température et non plus de microorganismes. Une décision était associée à l'observation de la température. Contrairement aux deux applications ci-dessus, nous n'avons pas appliqué la théorie bayésienne de la décision, qui nous semblait inadaptée à une simple comparaison de quelques stratégies très opérationnelles de mesure. Nous nous sommes donc contentés de comparer plusieurs stratégies en fonction de la probabilité de croissance de *C. perfringens* supérieure à 1 log ufc/g. Nous avons tracé, pour chaque stratégie, la probabilité de sortir d'une cellule de refroidissement en fonction de la probabilité d'une croissance supérieure à 1 log ufc/g pour obtenir des courbes semblables aux courbes d'efficacité décrites pour les plans d'échantillonnage microbiologiques afin d'indiquer aux opérateurs de cuisines centrales les risques auxquels les consommateurs sont confrontés. Cette partie du travail a été présentée au chapitre 8.

9.1.2 Application de concepts théoriques

Une large partie de mon travail de thèse a été d'utiliser des concepts de statistique et de les appliquer essentiellement à deux cas pratiques : contamination par *L. monocytogenes* et fabrication de lardons et de saumon fumé. Cela m'a permis de prendre conscience de l'importance et parfois aussi de la difficulté à adapter des concepts théoriques à des phénomènes concrets : en effet, certains concepts peuvent très bien fonctionner avec certains types de modèles ou de distributions et beaucoup moins bien avec d'autres ; il faut donc savoir ajuster un concept théorique à la question concrète que l'on se pose. Dans cette partie, j'expose les différentes adaptations que j'ai été amenée à effectuer.

Modèles hiérarchiques Ce type de modèles est apparu naturellement lorsque je me suis penchée sur la modélisation de la contamination de *L. monocytogenes*, aussi bien dans les poitrines de porc que dans le saumon fumé. Il est en effet apparu assez vite qu'il valait mieux distinguer la contamination aux niveaux de la fraction, de la prise d'essai, de l'unité et du lot. Ainsi, j'ai construit des modèles hiérarchiques à deux niveaux (A- et

B-test portion^a), à trois niveaux (A- et B-Batch, A- et B-Unit) et à quatre niveaux (A- et B-Reference). Ce travail m'a permis de réfléchir à la notion de modèles, puisque j'en ai décrit plusieurs, plus ou moins complexes, pour décrire une même réalité. Bien que les modèles soient plus simples que la réalité, il faut trouver un juste milieu entre la description précise de chaque phénomène - au prix d'un nombre excessif de paramètres à estimer au vu des données dont on dispose - et une approche comportant peu de paramètres, mais simpliste.

Modélisation mécaniste Lorsque je me suis intéressée à la modélisation de la contamination, j'ai cherché d'emblée à déterminer les raisons d'une telle contamination et de sa répartition dans la production d'une usine. J'ai donc essayé de décomposer le phénomène, puis de choisir une distribution appropriée pour chaque étape (d'où la modélisation hiérarchique citée plus haut). Je considère en effet que l'on comprend mieux le modèle probabiliste que l'on construit si celui-ci décrit les mécanismes d'un phénomène physique plutôt que s'il s'agit d'une "boîte noire" qui, comme par miracle, est capable de simuler des données proches des observations, mais dont les paramètres n'ont aucune signification physique. En outre, cette modélisation que j'appelle "mécaniste" présente l'avantage d'ouvrir la discussion avec des microbiologistes non statisticiens, car ils comprennent la signification des paramètres. Les modèles hiérarchiques sont d'ailleurs très utiles car ils permettent non seulement d'utiliser des paramètres ayant une signification physique mais aussi d'introduire des variables latentes ayant elles-mêmes une signification physique (par exemple les log concentrations moyennes d'une unité dans les modèles A- et B-Unit), dont on peut étudier le comportement pour le confronter aux connaissances des microbiologistes.

Construction de lois *a priori* : une information difficile à rassembler L'étude des différents modèles présentés dans Commeau et al. (2012a) avait pour but, non seulement de déterminer le(s) modèle(s) pouvant simuler les données les plus proches possibles des observations, mais aussi de déterminer le(s) mécanisme(s) expliquant la répartition de la contamination (par exemple, il semble raisonnable de penser que le lot de malaxage de poitrines de porc joue un rôle déterminant, puisque lors du malaxage les poitrines se frottent les unes aux autres et échangent donc des micro-organismes). Pour mettre toutes les chances de mon côté et parce que les jeux de données dont je disposais comportaient de nombreux zéros (absence pour les données de détection et zéro colonie pour les données de dénombrement), la loi *a priori* devait être choisie avec soin et être informative. Comment l'obtenir ? Disposer d'autres jeux de données ou pouvoir discuter avec des experts. Malheureusement pour le statisticien (mais heureusement pour le consommateur), le saumon fumé et les lardons en sortie usine sont peu contaminés (du moins dans les données que j'ai pu rassembler), ce qui rend difficile l'ajustement d'une loi *a priori*. C'est ce qu'ont essayé de faire Billoir et al. (2011) avec des données d'autocontrôles pour disposer d'une loi pour simuler la contamination des poitrines de porc lors de l'arrivée dans l'usine, mais cette tâche a été difficile à accomplir et la loi obtenue n'est pas classique.

Les meilleurs connaisseurs de la contamination d'un produit en sortie usine sont, à

a. Les notations sont celles de Commeau et al. (2012a)

mon avis, les responsables qualité de l'établissement, mais soit ils sont réticents à communiquer leur savoir (dans le cas "saumon fumé", nous avons réussi à obtenir des prévalences mensuelles d'autocontrôles mais pas plus), soit ils sont disposés à coopérer mais la difficulté que j'ai rencontrée alors était que les autocontrôles comportaient de nombreux zéros (dans le cas "lardons", la LOD était de 100 ufc/g donc les résultats de dénombrements étaient tous nuls et comme la fréquence d'autocontrôles était très faible, il y avait peu de données, même en les accumulant sur plusieurs années). Le fait que la LOD soit élevée empêche le responsable qualité de pouvoir estimer la contamination réelle de ses produits : il sait qu'elle est inférieure à une certaine concentration, mais il en ignore la distribution. Certes, il dispose de données de détection avec une LOD de 0,04 ufc/g mais il y avait peu de présences parmi ces résultats. Pour augmenter encore la difficulté, les paramètres pour lesquels j'avais besoin d'information étaient la plupart du temps des variances (par exemple, le modèle A-reference comporte trois paramètres de variances : σ , la variance inter-lots, λ , la variance inter-unités et ρ , la variance inter-prises d'essai). Il est malaisé de demander à un expert les quantiles à 5% et à 95% de la moyenne de la log concentration d'un lot, je vous laisse imaginer la difficulté pour obtenir qu'il réponde à la même question concernant l'écart-type de cette grandeur !

Construction de lois *a priori* : comment malgré tout construire des lois informatives ? De quoi disposions-nous pour construire des lois *a priori* ? Commençons par les "lardons" : nous disposions des autocontrôles de poitrines de porc à l'arrivée, qui avaient été utilisés par Billoir et al. (2011). Comme je travaillais sur la modélisation de la contamination des poitrines de porc en sortie malaxage, nous avons décidé d'utiliser le modèle de Billoir et al. (2011) et de simuler des données de contamination jusqu'à la fin du malaxage. Cette approche présentait l'avantage d'utiliser les données dont nous disposions et de déboucher sur des lois *a priori* même sur les variances. Pour le cas "saumon fumé", nous ne disposions que des données de prévalences (nombre de détections positives dans 25 g sur nombre total d'analyses effectuées par mois), que nous avons converties en concentration en supposant une distribution poissonnienne de la contamination et sur laquelle j'ai ajusté une loi log-normale pour obtenir une loi *a priori* sur la moyenne de la log concentration moyenne d'un lot. Bien qu'imparfaite, cette méthode a le mérite d'aboutir à des lois *a priori* informatives, au moins pour certains paramètres, malgré le peu d'information dont nous disposions.

Construction de lois *a priori* : lois compatibles Comme les modèles que nous avons comparés sont emboîtés les uns dans les autres et que nous disposions de peu de données, il fallait veiller à ce que les lois *a priori* des différents modèles soient *compatibles*, c'est-à-dire qu'un modèle ne soit pas favorisé par rapport à un autre uniquement parce que ses lois *a priori* sont mieux calibrées et non parce que c'est le meilleur. Pour construire des lois *a priori* compatibles, nous nous sommes basés sur les travaux de Celeux et al. (2006) qui ont travaillé sur des modèles de régression logistique : ils ont déterminé des lois *a priori* pour le modèle complet (prenant en compte toutes les covariables), avant d'élaborer les lois *a priori* de tous les autres modèles possibles (donc emboîtés dans le

modèle complet) en minimisant la distance de Kullback-Leibler :

$$\int \log \left(\frac{f_1(\mathbf{y}|\sigma^2)}{f_2(\mathbf{y}|\sigma^2)} \right) f_1(\mathbf{y}|\sigma^2) d\mathbf{y},$$

avec f_1 et f_2 les vraisemblances des modèles \mathcal{M}_1 et \mathcal{M}_2 (avec \mathcal{M}_2 emboîté dans \mathcal{M}_1), \mathbf{y} les observations et σ^2 la variance résiduelle. Il se trouve que dans ce cas, les lois *a priori* des coefficients de régression β_2 de \mathcal{M}_2 se déduisent facilement à partir des coefficients β_1 de \mathcal{M}_1 et des covariables. Les modèles que j'ai utilisé ne se prêtent pas si bien aux calculs et je n'ai pas réussi à trouver de formule pour déduire les lois *a priori* des paramètres des modèles emboîtés. Nous avons donc cherché à minimiser un critère construit en nous basant sur l'article de Celeux et al. (2006). Nous souhaitons que ce critère favorise les lois *a priori* telles que les simulations prédictives des données avec le modèle complet et un modèle emboîté soient les plus proches possibles. Pour les données de dénombrement, le critère a donc été défini de la manière suivante :

$$E = \frac{|\sum_i (\ln(q_{Gi} + 1) - \ln(q_{Ni} + 1))|}{\sum_i \ln(q_{Gi} + 1)},$$

avec q_{Ak} le k^{eme} quantile de la distribution prédictive des données de dénombrement du modèle A , G le modèle complet et N un des modèles emboîtés. Les quantiles utilisés sont les 5^{eme} , \dots , 95^{eme} , par pas de 5. Pour les données de détection, le critère a été bien plus simple à construire :

$$D = |P(x_G = 0) - P(x_N = 0)|,$$

avec $P(x_A = 0)$, la proportion de données de détection nulles générées de manière prédictive par le modèle A . Ces deux critères nous ont permis d'élaborer des lois *a priori* compatibles pour tous les modèles et les deux applications.

Comparaison de modèles Une fois les lois *a priori* déterminées et les inférences menées, il reste à comparer les modèles entre eux de manière à déterminer celui ou ceux qui prédisent le mieux les données observées. J'ai utilisé plusieurs outils de comparaison :

- le DIC (deviance informative criterion) : c'est un critère utilisé couramment dans la sélection de modèle bayésien lorsque les distributions *a posteriori* ont été obtenues par des chaînes de Markov Monte-Carlo (MCMC pour *Markov chain Monte Carlo*). Il a été proposé par Spiegelhalter et al. (2002) et a pour objectif de sélectionner le modèle ayant la meilleure prédiction. Il est égal à :

$$\begin{aligned} DIC &= D(\bar{\theta}) + 2p_D \\ &= \overline{D(\theta)} + p_D \end{aligned}$$

avec $D(\bar{\theta})$ est la déviance de la moyenne *a posteriori* de θ , $\overline{D(\theta)}$ est la moyenne de la déviance *a posteriori* et p_D est le nombre effectif de paramètres. Plus le DIC est petit et meilleur est le modèle. Avec les modèles de contamination utilisés, le DIC a été calculé à partir des variables latentes situées dans la couche la plus proche des données ;

- Contrairement au DIC, le facteur de Bayes n'évalue pas l'adéquation d'un modèle aux données mais compare deux modèles \mathcal{M}_1 et \mathcal{M}_2 entre eux. Il est égal au ratio des probabilité prédictives :

$$BF_{12} = \frac{[D|\mathcal{M}_1]}{[D|\mathcal{M}_2]} = \frac{\int [D|\theta_1, \mathcal{M}_1][\theta_1]d\theta_1}{\int [D|\theta_2, \mathcal{M}_2][\theta_2]d\theta_2},$$

où D sont les données et θ le paramètre. La construction de lois *a priori* informatives était indispensable pour calculer ces facteurs : à défaut, le modèle ayant la loi *a priori* la moins informative est rejeté (Kass and Raftery (1995)). Le calcul direct de $[D|\mathcal{M}_j]$ ($j = 1, 2$) étant souvent impossible, on détermine cette valeur en utilisant le théorème de Bayes. L'inverse de la vraisemblance $[D|m_j]$ s'écrit :

$$\frac{1}{[D|m_j]} = \frac{[\theta|D, m_j]}{[D|\theta, m_j][\theta|m_j]}. \quad (9.1)$$

J'ai calculé le terme de droite de l'équation (9.1) de deux manières différentes :

- en utilisant l'*importance sampling*, comme suggéré dans Kass and Raftery (1995) ;
- en utilisant les noyaux : le facteur de Bayes est calculé pour θ lorsque ce dernier est égal à sa moyenne empirique *a posteriori* : $\theta' = \sum_{g=1}^G \frac{\theta_g}{G}$. Le terme $[D|\theta', m_j][\theta'|m_j]$ est facilement calculable et la densité $[\theta'|D, m_j]$ est estimée avec un noyau gaussien de moyenne=0 et de variance=1.
- Le *Bayesian informative criterion* (BIC) est égal à $-2 \log[D|\hat{\theta}, \mathcal{M}_j] + N_j \log(n)$, avec $\hat{\theta}$ la valeur de θ au maximum de vraisemblance et N_j le nombre de paramètres du modèle \mathcal{M}_j . Des informations plus détaillées sur le BIC peuvent être trouvées dans Raftery (1995).
- Le critère de Gelfand et Ghosh (Gelfand and Ghosh (1998)) au niveau des données :

$$D(\mathcal{M}) = \sum_{l=1}^n \sigma_l^{2(\mathcal{M})} + \frac{k}{k+1} \sum_{l=1}^n (\mu_l^{(\mathcal{M})} - y_{l,obs})^2 \quad (9.2)$$

avec $\sigma_l^{2(m)} = Var(y_{l,rep}|y_{obs}, \mathcal{M})$ et $\mu_l^{(\mathcal{M})} = E(y_{l,rep}|y_{obs}, \mathcal{M})$, $y_{l,rep}$ une donnée l simulée et $y_{l,obs}$ la donnée l observée.

Ces critères sont plus ou moins faciles à programmer et à calculer. Surtout, ils utilisent les données au niveau de la prise d'essai. Or, nous souhaitions comparer les capacités des modèles à simuler correctement des données au niveau du lot. C'est pourquoi nous avons finalement retenu le critère de Gelfand et Ghosh, modifié pour être défini au niveau du lot (cf. Commeau et al. (2012a)).

Elicitation de la fonction de coût dans le cas “lardons” L'élicitation n'a pas été simple, car l'expert devait définir des actions correctives mineures ou majeures et estimer leur coût, ainsi que le montant des pénalités plausibles infligées par le client. Ce faisant, chaque coût a été estimé séparément sans vision d'ensemble ce qui aboutit à une incohérence, la meilleure décision à prendre étant de ne rien faire quelle que soit la prévalence (cf. partie 6.2.3). Plusieurs leçons peuvent être tirées :

- les acteurs ne sont pas toujours rationnels (et dans ce cas, la théorie bayésienne de la décision ne peut pas être appliquée car la rationalité est un axiome de base),

- bien que l’expert connaisse bien les coûts, il a pu se tromper dans les estimations,
- notre modèle était trop simple : il aurait fallu modéliser plus de décisions pour mieux approcher la réalité.

Une ou plusieurs de ces conclusions peuvent être vraies. Malheureusement, nous n’avons pas pu recontacter l’expert pour lui faire part des résultats et le conduire éventuellement à modifier son raisonnement. Quoiqu’il en soit, l’éllicitation de la fonction de coût est un exercice difficile, qui suppose que le statisticien connaisse bien l’application et son vocabulaire de manière à bien se faire comprendre de l’expert. Je pense qu’il aurait été possible d’imposer des contraintes à l’expert de manière à ce que les coûts minimaux correspondent aux décisions optimales lorsque la prévalence est connue par l’industriel. Lors de la séance d’éllicitation, je ne me suis pas aperçue de cette incohérence. Cependant, cela m’a permis de mieux découvrir l’univers complexe de l’éllicitation.

9.1.3 Critiques de la théorie bayésienne de la décision

Une partie substantielle de mon travail de thèse a porté sur l’application de la théorie bayésienne de la décision. Cette théorie a connu son essor dans les années 1960. Elle n’a pas été exempte de critiques. Ainsi, Cornier et al. (1980) considère que les ambitions de cette théorie sont trop vastes puisqu’elle prétend indiquer les décisions optimales dans n’importe quelle situation.

Une des critiques soulevées est que le décideur peut avoir du mal à placer sur une même échelle de valeur des éléments aussi différents que des considérations financières ou des atteintes à l’environnement par exemple. J’ai rencontré ce type de difficulté dans mon travail puisque, si les décisions retenues sont mettre ou non le lot en vente et que les contaminations sont possibles, il peut y avoir des malades et même des morts. En quelque sorte, on demande au décideur d’estimer un coût pour un échantillonnage d’un côté et le coût d’un malade ou d’un mort de l’autre. Posée ainsi, la question semble difficile voire insurmontable à régler. Je pense néanmoins que la question ne se pose pas tout à fait de cette manière. En général, le décideur a un point de vue particulier, ce qu’il fait qu’il n’envisage qu’une partie de la question. Ainsi, un industriel peut se demander combien va lui coûter un retrait ou un rappel, une annonce dans les médias et éventuellement un procès avec dommages et intérêts à verser aux victimes. Certes, cela peut sembler trop réducteur mais je pense qu’un industriel (en tant qu’industriel et non en tant qu’individu) va plutôt raisonner de cette manière. Un décideur de l’autorité publique pensera aux coûts dépensés par les industriels pour assurer la sécurité sanitaire de leurs produits mais également aux conséquences de santé publique (nombre de malades et de morts), qui sont difficiles à exprimer entièrement de manière monétaire. Si on se place du point de vue de l’Etat, une fonction de coût exprimée en monnaie n’est peut-être pas la meilleure solution, il vaut peut-être mieux exprimer les coûts en qualité de vie (comme les *quality-adjusted life years*, voir Traill and Koenig (2010) pour la définition).

Même si tous les coûts sont de la même nature (financiers par exemple), il peut être difficile de les quantifier parce que le décideur ne sait pas combien un événement rare peut coûter. Ainsi, les coûts d’une perte d’image de marque ou d’un procès sont difficilement quantifiables. Dans le cas “lardons”, nous avons choisi de ne pas considérer le cas de la vente de ce produit au consommateur et aux éventuelles conséquences néfastes car cela ne nous

a pas semblé prioritaire et parce que l'entreprise semblait plus préoccupée par ses relations commerciales avec le distributeur. Nous avons donc opté pour des conséquences de type pénalités pour l'entreprise, ce qui présente l'avantage d'être plus facilement estimable. Dans le cas "saumon fumé", je n'ai pas eu le temps de mener à bien l'élicitation de la fonction de coût, de plus, nous n'avons pas réussi à maintenir de contacts suivis avec l'industriel.

Une autre critique fréquemment formulée (notamment dans Cornier et al. (1980)) est qu'il n'est pas toujours facile ni pertinent de résumer une conséquence par une valeur d'utilité unique. Je n'ai pas rencontré ce cas dans mes travaux, mais un moyen de s'en sortir serait peut-être d'avoir deux fonction de coût, chacune exprimée dans une unité différente, le but étant de trouver une optimisation en deux dimensions...

Pourquoi décider que la meilleure décision est celle qui minimise l'espérance de la fonction de coût? Pourquoi ne pas choisir une autre grandeur ou même pourquoi ne pas conserver la fonction de coût entièrement? Considérer la fonction de coût dans son ensemble ou n'intégrer que sur les observables x par exemple rend difficile la prise de décision. En effet, une règle de décision δ_1 peut être telle que la fonction de coût $L(\theta_a, \delta_1)$ prise en δ_1 et en θ_a , une valeur du paramètre θ est inférieure à la même fonction prise en δ_2 et θ_a alors que c'est l'inverse en un point $\theta_b : L(\theta_b, \delta_1) > L(\theta_b, \delta_2)$. D'ailleurs, si la fonction de coût est intégrée en x (on parle alors de risque fréquentiste), il n'y a pas d'ordre total l'ensemble des règles de décision (cf. partie 1.3.1.a). Dans ces conditions, pour laquelle opter? Si l'on choisit une règle quand le paramètre θ prend une certaine valeur ou dans un certain intervalle et une autre règle lorsque θ est dans un autre intervalle, cela revient soit à connaître la valeur prise par θ (dans ce cas, la théorie de la décision n'est pas nécessaire), soit à avoir des informations sur θ . Il serait regrettable de s'en priver et ne pas utiliser l'approche bayésienne. Il n'est pas obligatoire de considérer le risque bayésien comme unique échelle de comparaison des décisions. L'approche fréquentiste considère par exemple comme critère de choix de la règle de décision le critère *minimax*, qui consiste à choisir la règle minimisant le risque fréquentiste avec le pire choix possible pour θ , c'est-à-dire la valeur de θ maximisant le risque fréquentiste^b. En termes mathématiques, cela se traduit par :

$$\bar{R} = \inf_{\delta \in \mathcal{D}^*} \sup_{\theta \in \Theta} R(\theta, \delta).$$

On peut considérer ce critère comme une assurance contre le pire puisque l'on choisit la décision minimisant la fonction de coût dans le cas le moins favorable (Robert (2006)).

Une critique courante de la théorie bayésienne de la décision est son manque de robustesse (Berger (1985)). La robustesse est la sensibilité de la décision aux hypothèses incertaines du modèle. Par exemple, la modification d'un coût dans la fonction de coût peut modifier la règle de décision optimale. A mon avis, le manque de robustesse dépend des cas et seule une analyse de sensibilité permet de déterminer la robustesse d'une décision. Si une fonction de coût s'avérait par trop instable, peut-être faudrait-il songer à la modifier parce qu'elle n'est pas adaptée au problème posé.

Enfin, Cornier et al. (1980) pensent que la théorie bayésienne de la décision est

b. Attention, dans ce cas, l'espace des décisions possible est \mathcal{D}^* , l'ensemble des décisions randomisées : la décision est générée selon la distribution de densité de probabilité $\delta^*(x, \cdot)$, une fois que l'observation x a été recueillie, pour plus de détails, voir Robert (2006).

séduisante sur le plan théorique mais difficilement applicable en partie parce que les mécanismes de décision par une personne restent largement inexplorés. Cependant, Robert (2006) considère au contraire que la modélisation ne peut rendre compte de la complexité de la réalité mais cette théorie permet d'établir une vraie rationalité, que les décideurs visent mais n'arrivent pas à atteindre. J'ajouterais que cette théorie permet également au décideur de se demander pourquoi il prend une certaine décision et s'il ne pourrait pas la modifier en considérant par exemple des critères plus rationnels.

9.1.4 A côté des statistiques

Mon travail de thèse a porté en grande partie sur des questions statistiques. Néanmoins, je me suis attachée à comprendre l'application sur laquelle je travaillais ainsi que l'élaboration et la signification des données que j'ai utilisées.

Travail multidisciplinaire Dans le cadre du programme de recherche Quant'HACCP, j'ai travaillé avec de nombreux chercheurs : des statisticiens mais aussi des microbiologistes ou des thermiciens. Côté personnes d'origine diverses m'a permis d'avoir une vue d'ensemble du projet Quant'HACCP et de voir comment mon travail s'intégrait. Par ailleurs, j'ai également travaillé avec des responsables qualité de deux usines. Tout l'intérêt de rencontrer des personnes différentes réside dans le fait d'apprendre à adapter son discours pour être compris par des publics différents, mais aussi dans le fait d'être curieux pour comprendre et s'approprier le mieux possible le vocabulaire technique et le point de vue de personnes de domaines variés. Le langage mathématique est d'ailleurs un outil pour regrouper les différentes spécialités : ainsi, le modèle de Billoir et al. (2011) modélisant la contamination par *L. monocytogenes* de lardons en cours de fabrication traduit en distributions, en constantes, en variables et en paramètres des connaissances microbiologiques (températures de croissance de *L. monocytogenes*, forme des courbes de croissance, etc.), industrielles (température du four d'étuvage, durée du malaxage, déroulement d'un procédé de fabrication, etc.). De même, dans l'article Commeau et al. (2012a), j'ai traduit en modèles hiérarchiques la contamination observée dans des poitrines de porc après malaxage. Pour cela, j'ai utilisé des données microbiologiques et j'ai cherché à comprendre le procédé afin d'avoir des idées de regroupement de populations pour définir les distributions. Au chapitre 8, nous avons utilisé des courbes de profil temps-température (domaine thermique) pour construire un modèle décrivant la température dans des plats cuisinés placés en cellule de refroidissement (domaine statistique) afin d'étudier différentes stratégies de sortie de cette cellule et les conséquences sur la probabilité de croissance de *C. perfringens* dans des plats consommés ensuite par les patients d'un hôpital (domaine opérationnel). Tout n'est pas toujours allé de soi car il n'est pas nécessairement facile de trouver le vocabulaire adéquat, surtout lorsque les mêmes termes recouvrent des significations différentes selon les spécialités. En outre, même si les mathématiques sont un outil pratique, ils font parfois peur aux personnes n'ayant pas l'habitude de les manipuler et il faut donc tenter de les convaincre de l'intérêt de la démarche.

Travail dans le cadre d'un projet de recherche Etre intégrée au projet Quant'HACCP m'a permis de découvrir comment fonctionne un projet de recherche regroupant plusieurs

équipes de recherches de thématiques variées. J'ai ainsi pu confronter régulièrement mon travail à d'autres statisticiens et à des chercheurs d'autres disciplines. Par ailleurs, les échéances du projet permettent de mesurer les avancées entre deux jalons et de prendre un peu de hauteur pour déterminer les tâches restant à effectuer.

9.2 Perspectives

Les travaux présentés dans ce manuscrit peuvent être poursuivis. Dans cette partie, je présente plusieurs pistes envisageables.

Au chapitre 3, j'ai présenté un critère basé sur la proposition de Gelfand and Ghosh (1998) pour comparer la capacité de différents modèles à simuler des données proches de données observées. Or, les données sont observées et simulées à un niveau (en l'occurrence des prises d'essai) mais la capacité de simulation est jugée au niveau supérieur (du lot). Il serait intéressant de définir d'autres critères de ce type et d'étudier leurs propriétés.

Dans mon travail, je ne me suis penchée que sur les données de détection et de dénombrement de *L. monocytogenes* acquises avec les protocoles définis dans ISO11290-1 (1996a) et ISO11290-2 (1998a) ou proches de ceux-ci. Or, il existe d'autres manières d'acquérir des données de dénombrement, par exemple par PCR (Rossmannith et al. (2006)) ou par la méthode NPP^c (Gonzales-Barron et al. (2010b)) L'estimation des paramètres est-elle plus ou moins précise avec ces méthodes que celle avec des décomptes de colonies ou des données exprimées sous forme de concentration ? Concernant les données de type NPP, des travaux de ce type ont été entrepris par Pouillot et al. (2012) et Williams et al. (2001).

Les observations dans les plans d'échantillonnage étudiés aux chapitres 6 et 7 sont des données de détection, il serait également possible d'observer des données de dénombrement. Dans ce cas, le calcul du risque bayésien est un peu plus difficile car la loi de la concentration n'est plus conjuguée et que les valeurs de c_1 et c_2 ne se déduisent pas facilement de n . Dans ce cas, on peut calculer la fonction de coût en faisant varier n , c_1 et c_2 mais cela peut prendre beaucoup de temps. Une autre manière de procéder est d'utiliser un algorithme pour déterminer le minimum de cette fonction de coût. Par exemple, si on considère la fonction d'utilité et qu'on la modifie pour en faire une distribution de probabilité, on peut utiliser un algorithme particulière permettant de déterminer le mode d'une fonction de probabilité. Un tel algorithme a été présenté par Amzal et al. (2006).

Les plans d'échantillonnage présentés sont situés à une étape du procédé (après le malaxage pour le cas "lardons" et en fin de procédé pour la cas "saumon fumé"). Il peut être utile de considérer un plan d'échantillonnage établi en deux étapes du procédé, par exemple pour jeter un lot en cours de production si celui-ci est trop défectueux afin d'éviter de supporter inutilement les coûts de production entière alors que le lot sera très vraisemblablement jeté lors du contrôle en sortie usine. A mon sens, le modèle de contamination utilisé pour simuler les résultats d'un tel plan d'échantillonnage doit faire le lien entre la contamination à une étape donnée et la contamination à une autre étape puisque les produits soumis aux analyses lors du second prélèvement ne sont pas identiques aux produits soumis au premier prélèvement (des lots jugés trop contaminés ont été retirés

c. NPP signifie "nombre le plus probable" et se dit MPN (*most probable number*) en anglais.

entre les deux étapes). Dans le cas d'application "lardons", le modèle de Billoir et al. (2011) ou ce même modèle modifié (cf. chapitre 5) pourraient servir à établir les distributions jointes de contamination d'un lot aux deux étapes concernées. Il reste ensuite à définir les distributions des observations pour chacune des deux étapes, et bien sûr, à déterminer les décisions et les conséquences possibles puis de déterminer leurs coûts. Cette évolution permettrait d'illustrer la capacité d'adaptation de la théorie bayésienne de la décision à l'échantillonnage multipostes, après les deux applications présentées dans le cadre de ma thèse.

ANNEXES

Annexe A

Le bayésien à travers les âges

Cette annexe évoque les principales étapes historiques de l'approche bayésienne. La principale source utilisée pour écrire cette partie est le livre de Bertsch McGrayne (2011).

L'origine : Thomas Bayes

Le courant bayésien tire son nom du révérend Bayes qui a vécu en Angleterre au XVIII^{ème} siècle et qui était un mathématicien amateur. A l'époque, les probabilités étaient utilisées dans le jeux de cartes, par exemple pour calculer la chance d'avoir quatre as dans son jeu. Bayes s'est intéressé au problème de la probabilité inverse en se demandant comment déterminer la probabilité d'un événement futur dont on ne connaît que le passé, c'est-à-dire le nombre de fois où il s'est produit et où il ne s'est pas produit. Bayes imagina une expérience pour illustrer le problème de la probabilité inverse : il est assis dos à une table carrée parfaitement plane. Une balle est lancée par un de ses amis sur cette table et Bayes ne sait pas où elle s'est arrêtée puisqu'il ne voit pas la table. Ensuite, un de ses amis lance une deuxième balle et dit à Bayes si elle se situe à gauche ou à droite de la première balle. Si elle est à gauche, Bayes est amené à penser que la première balle est située plutôt sur la partie droite de la table. Une troisième balle est lancée et on indique à Bayes sa position par rapport à la position de la première balle et ainsi de suite. A chaque lancer, Bayes acquiert une information nouvelle sur la première balle. Bayes ne sait pas exactement où se situe la première balle mais il peut être de plus en plus confiant dans le fait qu'elle se trouve tout près du côté gauche de la table par exemple. On peut donc résumer le système de Bayes ainsi : on a une croyance initiale (notre opinion sur la position de la première balle avant que les autres balles soient lancées), on acquiert des données (grâce aux lancés des balles suivantes) qui modifient notre croyance sur la position de la balle. La définition de la probabilité selon Bayes est :

“ The probability of any event is the ratio between the value at which an expectation depending on the happening of the event ought to be computed, and the chance of the thing expected upon its happening.”(Bayes (1763))

Comme Bayes ne sait pas du tout où se trouve la première balle avant le lancer des autres balles, il considère que cette balle a une égale probabilité de se trouver en n'importe quel endroit de la table (en langage moderne, il choisit une loi *a priori* uniforme). Au cours des décennies suivantes, la loi *a priori* sera souvent uniforme, ce qui vaudra à l'approche

bayésienne de nombreuses critiques portant sur le réalisme de cette façon de procéder. Bayes a écrit un article à propos de sa découverte mais il ne le publia pas. Ce n'est qu'après sa mort que Richard Price, un autre mathématicien amateur, retrouvera cet écrit, le corrigera et le publiera en 1763 sous le titre *An essay towards solving a problem of the doctrine of chances* (Bayes (1763)).

La redécouverte et les applications

Quelques années plus tard, Pierre Simon Laplace a été le premier à utiliser la règle de Bayes (sans d'ailleurs connaître les travaux du révérend) pour quantifier l'incertitude de données d'astronomie. Il souhaitait répondre à la question qui intéressait beaucoup les scientifiques de son temps : l'univers est-il stable ? Laplace disposait de données d'origine et de précisions variées (observations de la position de planètes effectuées par des peuples divers entre 1100 av. Jésus Christ et le XVIII^{ème} siècle). Après de nombreuses recherches, Laplace pensa que la probabilité constituait un outil adapté pour l'aider à traiter l'incertitude dans le système solaire. Selon lui, la probabilité était l'expression mathématique de notre ignorance. Il publia un article en 1774 intitulé *Mémoire sur la probabilité des causes par les événements* (Laplace (1774)) dans lequel se trouve la première version de ce qui est appelé aujourd'hui le théorème de Bayes. Ce dernier est uniquement exprimé avec des mots : la probabilité d'une cause sachant un événement est proportionnel à la probabilité de l'événement sachant sa cause. Durant les 40 années suivantes, Laplace clarifia, expliqua et appliqua cette règle à ce nombreuses applications en sociologie, astronomie, dans le domaine judiciaire. Il s'intéressa à des problèmes pour lesquels les données étaient très nombreuses, ce qui l'a amené à étudier les propriétés des grand échantillons. Vers la fin de sa carrière, il se tourna systématiquement vers l'approche fréquentiste des probabilités sans y voir de rupture de paradigme qui lui fasse renier la formule de Bayes. Cependant, dans les années 1810, il trouva la formule de la probabilité des causes résumant l'apprentissage grâce aux événements :

$$P(C|E) = \frac{P(E|C)P_{a\ priori}(C)}{\sum P(E|C')P_{a\ priori}(C')},$$

avec C la cause, E l'événement, $P(C|E)$ la probabilité d'une cause (ou d'une hypothèse) sachant l'événement (ou l'information), $P_{a\ priori}(C)$ l'estimation initiale de la probabilité de la cause et $P(E|C)$ la probabilité d'un événement sachant une cause. Le numérateur est divisé par la somme des probabilités des événements en considérant toutes les causes possibles.

Un relatif oublié au XIX^{ème} siècle

Au XIX^{ème} siècle, l'approche de Laplace sur la probabilité des causes fut critiquée. Il n'était plus question de lier la probabilité au manque de connaissance, l'objectivité régnait. De plus, la loi *a priori* utilisée par Laplace était la loi uniforme. Or, si cette loi peut être utilisée quand il n'y a aucune information sur les causes, comment continuer à l'utiliser si on sait, avant d'avoir observé les données, que les différentes causes ne sont pas équiprobables ? La règle de Laplace-Bayes fut donc abandonnée sauf dans quelques

domaines comme l'artillerie française (pour déterminer le placement des canons et leur angle de tir. Dans le même domaine, elle a aussi été utilisée pour tester les lots de munitions car elle permettait de tester moins de munitions qu'avec l'approche fréquentiste), les communications téléphoniques (dans l'entreprise Bell) et l'assurance (quand les Etats-Unis rendirent obligatoire l'assurance pour couvrir les accidents du travail : en très peu de temps, il fallu établir les montants des primes alors que les Etats-Unis n'avaient que peu de données sur le sujet).

Début du XX^{ème} siècle : la montée en puissance fréquentiste...

Après la Première guerre mondiale, la règle de Laplace-Bayes ne fut pas plus populaire. En Grande-Bretagne, des statisticiens de renom étaient anti-bayésiens : Fisher, Neyman et Pearson. Ronald Fisher a commencé sa carrière par travailler dans le domaine agricole en traitant des données à propos d'engrais. Il utilisait alors la règle de Bayes. Puis il a cherché une meilleure manière de concevoir des expériences. Il a été l'un des pionniers dans le domaine des méthodes de randomisation, la théorie de l'échantillonnage, les tests d'hypothèse, l'estimation du maximum de vraisemblance et l'analyse de la variance. Grâce à lui, la statistique est entrée dans le domaine des sciences expérimentales. En 1925, il publia un manuel *Statistical methods for research workers* (Fisher (1925)) expliquant les procédures statistiques à l'usage de non statisticiens. Il s'intéressa ensuite à la génétique. Il faisait lui-même des croisements afin d'étudier les pedigrees d'animaux au fil des générations. Il devint professeur de génétique à l'université de Cambridge. Toute sa vie, il fut un anti-bayésien convaincu, ce qu'il exprima dans des articles ou dans des lettres. Jerzy Neyman et Egon Pearson développèrent la théorie du test d'hypothèse en 1933. A cette époque, les statisticiens ne formulaient qu'une seule hypothèse à la fois, qu'ils rejetaient ou non. L'idée de Pearson était que la seule raison de rejeter une hypothèse était d'en accepter une plus probable. Neyman et Pearson ont ainsi défini les risques de première et de seconde espèces. Fisher et Neyman étaient de fervents anti-bayésiens : ils se limitèrent aux événements qui pouvaient être théoriquement répétés une infinité de fois. La seule information provenait de l'échantillon. Chaque nouvel échantillon donnait lieu à une inférence indépendante des jeux de données antérieurs. Bien évidemment, les lois *a priori* subjectives étaient bannies et la loi uniforme traditionnellement utilisée à cette époque comme loi *a priori* déclarée illégitime. Les fondements de ces nouvelles méthodes s'appuyaient sur la théorie mathématique des probabilités de Kolmogorov dans. Son traité *Fondements de la théorie des probabilités* paru en 1933 (Kolmogorov (1933)), ne fait pas références à l'interprétation d'une probabilité en terme de pari.

...mais le courant bayésien résiste

Cependant, même si l'approche fréquentiste dominait à cette époque, trois mathématiciens ont fait la promotion de la règle de Bayes en considérant que la connaissance pouvait être quantifiée par des paris. Ces trois hommes sont Borel, Ramsey et de Finetti. En 1924, Emile Borel, un mathématicien français concluait que la croyance d'une personne pouvait être mesurée par la somme que cette personne était prête à parier. Deux ans plus tard, Frank Ramsey, un mathématicien anglais, fit la même suggestion que Borel. Il intro-

duisit également la fonction d'utilité et la maximisation de l'utilité espérée comme critère de comportement devant l'incertitude. Il ne put poursuivre ses travaux car il mourut à 26 ans de la jaunisse contractée à la suite d'une opération. Bruno de Finetti, un professeur de mathématique italien, a donné des bases mathématiques à la subjectivité selon Bayes (de Finetti (1937)). Il a également introduit la notion d'échangeabilité.

Durant les années 1930 et 1940, le flambeau bayésien a surtout été porté par Harold Jeffreys, un chercheur à l'université de Cambridge. Il était géophysicien et fut professeur d'astronomie. Il s'intéressa aux statistiques car il étudiait les ondes de propagation des tremblements de terre. A partir de l'arrivée des ondes dans différentes stations de mesure, il cherchait à trouver l'épicentre, ce qui l'a amené à se pencher sur la probabilité inverse des causes. Il disposait de peu de données, qui plus est difficile à interpréter car les résultats variaient beaucoup d'un sismographe à l'autre. Jeffreys considérait que les probabilités étaient un outil adapté pour mesurer l'incertitude. Jeffreys et Fisher se sont affrontés par articles interposés. En réponse à leur débat, Jeffreys a écrit *Theory of probability* en 1939 (Jeffreys (1961b)), qui fut plusieurs années durant le seul ouvrage d'explication systématique sur la manière d'appliquer la règle de Bayes à des problèmes scientifiques. Une des critiques souvent faites à l'approche bayésienne qu'il n'y a pas une façon unique de construire la loi *a priori* et que cette loi peut avoir une grande influence sur l'inférence qui en résulte. Jeffreys a proposé une manière de construire une loi *a priori* aussi peu informative que possible. La construction d'une telle loi est basée sur l'information de Fisher et est telle que l'exigence d'invariance par reparamétrisation est satisfaite. La construction de cette loi est "automatique" et dépend de la distribution des observations. Jeffreys a utilisé les statistiques pour faire des inférences et non pour guider des actions futures. C'est pourtant ce dernier point qui valorisera l'approche bayésienne durant la Seconde guerre mondiale.

Bayes au service des Alliés

Les statistiques se développèrent beaucoup pendant la Seconde guerre mondiale. Elles ont été notamment développées par Turing pour déchiffrer les messages chiffrés par les Allemands. Un problème majeur pour le Royaume-Uni durant la guerre était son approvisionnement. Comme l'Europe continentale était largement occupée par les Allemands, les Britanniques se fournissaient essentiellement en Amérique. Les provisions arrivaient par bateaux. Les U-boat^a se chargeaient de couler les ravitailleurs. Il était donc vital pour les Britanniques de déchiffrer les messages envoyés aux U-boat de manière à connaître leur position. Ces messages étaient chiffrés au moyen de la machine appelée Enigma. Cette machine ressemblait à une machine à écrire sauf que la lettre tapée par l'opérateur n'était pas celle imprimée sur la papier : la transformation entre la lettre initiale et la lettre codée se faisait par une succession de permutations au moyen de trois rotors qui changeaient de position à chaque lettre tapée. La configuration des rotors utilisée pour chiffrer le message était indiquée en tête de message. Pour déchiffrer le message, il suffisait de taper le message chiffré avec une machine Enigma en ayant pris soin de la mettre dans la bonne configuration. Les machines Enigma utilisées par la marine allemande étaient plus sophistiquées : elles contenaient quatre rotors au lieu de trois (ce qui augmente considérablement

a. U-boat signifie sous-marins en allemand.

le nombre de permutations possibles) et les listes de configurations initiales se trouvaient dans un manuel à bord des sous-marins. Chaque configuration était valable 24 h. Avant la guerre, les Polonais avaient réussi à casser en partie le chiffrement et à lire 75% des messages des armées de terre et de l'air. Les Britanniques reprirent le flambeau dans leur centre de renseignements. Parmi les personnes travaillant dans ce centre se trouvait Alan Turing. Turing commença par modifier les "bombes" qui étaient des machines électromécaniques permettant de tester différents combinaisons d'arrangements d'Enigma. Un test permettait d'éliminer les moins bonnes combinaisons, les meilleurs devant ensuite être étudiées à la main. La bombe mettait du temps à examiner une combinaison, il fallait donc trouver un moyen de réduire le nombre de tests. C'est là que la règle de Bayes intervient. Turing eut l'idée suivante : si on arrivait à identifier des groupes de lettres qui se répètent dans les messages et qu'on faisait un pari sur leur signification, cela donnerait une indication sur le positionnement des roues d'Enigma, ce qui réduirait le nombre de tests à mener sur la bombe et donc diminuerait le temps de décodage. Pour estimer la probabilité d'un pari quand l'information arrive au fur et à mesure, Turing développa ce qu'on appela ensuite l'analyse séquentielle. Après plusieurs mois d'efforts, la construction de bombes plus performantes et l'obtention d'un manuel dans un U-boat, Turing parvint enfin à déchiffrer le code. Turing a développé une manière de mettre à jour ses connaissances en utilisant les données disponibles mais connaissait-il pour autant la règle de Bayes ? Oui si l'on en croit son assistant, Irving John Good qui rapporte :

"[Turing] invented a Bayesian approach to sequential data analysis, using weights of evidence (though not under that name). A weight of evidence is the logarithm of a Bayes factor ; for a Bayesian, this is the only possible definition, and the concept has been an obsession of mine ever since. [...] On one occasion I happened to meet George Barnard during the war, in London, and I confidentially mentioned that we were using sequentially a Bayesian method in terms of what we now call "weights of evidence" (log-factors) for distinguishing between two hypotheses. Barnard said that, curiously enough, in his work for the Ministry of Supply, he was using something similar. Thus Turing and Barnard invented sequential analysis independently of Wald. [...] Turing developed the sequential probability ratio test, except that he gave it a Bayesian interpretation in terms of the odds form of Bayes' theorem. He wanted to be able to estimate the probability of a hypothesis, allowing for the prior probability, when information arrives piecemeal. [...] One morning I asked Turing 'Isn't this really Bayes' theorem ?' and he said 'I suppose so.' He hadn't mentioned Bayes previously."

Les aspects bayésiens du travail de Turing sont mentionnés dans un article de Good (Good (1979)). La même méthode développée par Turing pour déchiffrer le code Enigma a été mise à profit pour déchiffrer les messages écrits avec la machine de Lorenz, utilisée par les Allemands pour les communications de haut niveau entre le commandement suprême et les quartiers-généraux des groupes d'armées.

L'approche bayésienne a également été utilisée pour localiser précisément un U-boat une fois qu'un message envoyé par le sous-marin avait été capté par des radars alliés. Une méthode fut mise au point par le mathématicien Bernard Koopman de l'université

de Columbia pour déterminer comment les officiers devaient procéder pour planifier une opération de recherche afin d'optimiser la recherche avec des ressources limitées.

Après la guerre, quelques avancées en recherche opérationnelle

Après la guerre, la règle de Bayes ne fut pas plus utilisée qu'avant la guerre. Une des raisons est peut-être que tous les travaux faits dans le domaine de la cryptographie (et donc les travaux de Turing) ont été classifiés sur ordre Churchill. Ces travaux ne sont devenus publics qu'en 1973.

Néanmoins, la règle de Bayes ne fut pas complètement abandonnée. Elle connut un regain d'utilisation dans les années 1950 avec Good, Savage et Lindley. Comme mentionné plus haut, Good était l'assistant de Turing durant la Seconde guerre mondiale. Après la guerre, il écrivit un livre sur les méthodes statistiques utilisées par Turing. Son livre, intitulé *Probability and the weighing of evidence*, parut en 1950 (Good (1950)). Il ne pouvait pas dévoiler toutes les techniques utilisées, certaines étant toujours classifiées au moment de la parution de l'ouvrage. Il travailla quelques années en Angleterre, notamment à l'université de Cambridge avant de devenir professeur de statistique à VirginiaTech : il continua à publier des articles mais resta une voix indépendante dans la communauté bayésienne.

Savage et Lindley en furent au contraire les leaders. Lindley fit des études à Cambridge. Il souhaitait faire des statistiques une branche respectable des mathématiques. Il rencontra Savage en 1954 pour découvrir que ce dernier poursuivait le même objectif que lui. En cherchant à mettre des bases solides aux techniques statistiques, ils se rendirent compte qu'ils ne pouvaient pas faire de lien logique entre leurs axiomes et les méthodes fréquentistes, ce qui les amena progressivement sur le territoire bayésien. Lindley fonda plusieurs groupes de recherche en bayésien dans différentes universités britanniques et se battit pour développer la recherche bayésienne et que des professeurs bayésiens soient nommés dans les universités. Mathématicien de formation, Leonard Jimmie Savage a travaillé au *Statistical research group* à l'université de Columbia durant la Seconde guerre mondiale, ce qui l'a amené à délaisser les mathématiques pures pour se tourner vers les statistiques. Il fonda ensuite le département des statistiques à l'université de Chicago et découvrit les écrits de Borel, Ramsey et Finetti. Il publia en 1954 *The foundations of statistics* (Savage (1954)) dans lequel il développe les idées de Ramsey sur l'utilisation du théorème de Bayes comme outil pour la décision. Son livre contribua à répandre les idées bayésiennes dans divers domaines comme la défense, la théorie des jeux, l'histoire, etc. Savage s'est proclamé subjectiviste. Comment concilier loi *a priori* subjectives (et donc différentes pour chaque personne) et l'objectivité scientifique dans le traitement des données^b ? Selon Savage, quand il y a peu de données, les scientifiques ne sont pas d'accord entre eux et sont subjectifs. Plus le nombre de données augmente et plus les scientifiques s'accordent entre eux et deviennent objectifs. A partir des années 1960 Savage devint un bayésien convaincu et participa à la bataille qui faisait rage alors contre les fréquentistes.

L'approche bayésienne a été utilisée dans plusieurs branches durant les années 1960 pour répondre à diverses questions comme : y a-t-il un lien entre le cancer du poulmon et

b. Critique répandue de l'approche bayésienne

la cigarette (Jerome Cornfield)? Quelle est la probabilité qu'une bombe nucléaire explose par accident (Albert Madansky)? L'approche bayésienne atteignit aussi l'économie avec Arthur Zellner qui traita plusieurs problèmes économiques à la fois de manière bayésienne et non bayésienne. Deux mathématiciens de la Harvard business school, Schlaifer et Raiffa, s'attaquèrent au problème de l'aide à la décision pour les gestionnaires. Schlaifer entreprit d'étudier les statistiques lorsque la Harvard business school lui a demandé d'enseigner les statistiques à ses étudiants. Historien de formation (il avait une thèse en histoire ancienne), il s'est intéressé à la physique, plus particulièrement aux sonars et à la conception d'avions. D'après Raiffa (Fienberg (2008)), il a découvert les statistiques par les livres de Fisher, Neyman et Pearson. Cette approche ne lui semblait pas appropriée pour les gestionnaires qui doivent prendre des décisions sans attendre de disposer de suffisamment de données issues d'événements répétables. Parfois, le gestionnaire n'a même pas de données pour prendre ses décisions. Encoder les informations dont dispose le gestionnaire dans une distribution de probabilité semblait un bon point de départ pour Schlaifer. A partir de là, Schlaifer s'attaqua au problème de la taille optimale de l'échantillon à recueillir. Mettre à jour les croyances initiales grâce aux données acquises conduisit Schlaifer à utiliser la règle de Bayes. En tant que "self-made statistician", Schlaifer réinventa la théorie bayésienne de la décision sans consulter les ouvrages existants. Comme il sentait que ses compétences mathématiques ne suffisaient pas, il engagea Howard Raiffa. Raiffa était statisticien fréquentiste de formation mais commença à pencher du côté bayésien après avoir lu *Theory of games and economic behavior* de von Neumann et Morgenstern (von Neumann and Morgenstern (1944)) car la distribution de probabilité lui semblait un moyen naturel de décrire sa manière de jouer face à d'autres joueurs. Convaincu par l'approche bayésienne et recruté par la Harvard business school, il travailla avec Schlaifer sur la décision. Durant leurs travaux, Schlaifer et Raiffa s'aperçurent qu'il n'existait pas de livre sur les outils bayésiens à l'usage des non-statisticiens, ils s'attelèrent donc à cette tâche. Pour simplifier les calculs, ils introduisirent les arbres de décision (qui décrivent les décisions séquentielles auxquelles les gestionnaires sont confrontés), les lois *a priori* conjuguées. Leur livre *Applied statistical decision theory* (Raiffa (1995)) est paru en 1961. Cependant, Schlaifer et Raiffa n'arrivèrent pas à faire entrer la théorie bayésienne de la décision dans la pratique courante des gestionnaires, en partie parce que Schlaifer restait un théoricien : il ne réussit pas à expliquer son projet aux entreprises.

Les Fédéralistes et les élections américaines : deux applications importantes

Le premier cas de grande échelle traité par l'approche bayésienne est l'étude menée par Mosteller et Wallace pour déterminer qui avait écrit le Fédéraliste (*Federalist papers* en anglais). Cet ouvrage est constitué de 85 articles écrits entre 1787 et 1788 pour promouvoir la signature de la Constitution américaine. L'ouvrage a été écrit par Hamilton, Madison et Jay mais aucun article n'était signé. Les historiens s'étaient accordés pour attribuer un auteur à 73 articles mais on ne savait pas qui de Madison ou d'Hamilton avait écrit les 12 articles restants. Frederick Mosteller s'est intéressé à ce problème : il souhaitait attribuer un auteur à chacun des 12 articles en se basant sur les styles et les mots propres à chaque auteur. Pour cela, les articles dont la paternité était connue servaient d'information. Il

entraîna David Wallace dans ce projet. Ils publièrent leur résultat en 1964, soit près de 20 ans après que Mosteller a commencé à travailler sur ce sujet. Ce projet avait nécessité l'utilisation des ordinateurs et de très nombreux étudiants. Leurs travaux ont montré que l'hypothèse la plus probable était que Madison avait écrit la totalité des articles.

En 1960, le bayésien fit son entrée à la télévision lors des élections présidentielles américaines. Les principales chaînes américaines se battaient pour être la première à annoncer des résultats fiables. La chaîne NBC demanda à John Tukey de se charger de la méthodologie. Tukey était un professeur à l'université de Princeton. Il avait travaillé dans des projets classifiés pour l'armée en cryptographie et avait conseillé le gouvernement américain sur plusieurs sujets civils. Lors de la soirée électorale, les dépouillements des différents comtés formant le pays arrivaient au fur et à mesure et les projections étaient mises à jours à chaque arrivée de données. Il fallait tenir compte des caractéristiques de chaque comté afin d'attribuer un poids aux voix de ce comté dans le score final. Evidemment, les mises à jours devaient être faites rapidement pour que le vainqueur des élections soit annoncé le plus vite possible. Tukey et ses collaborateurs eurent à leur disposition des ordinateurs puissants pour l'époque. Tukey participa aux soirées électorales jusqu'en 1980 où le calcul des tendances fut remplacé par des sondages. Bien que l'approche bayésienne ait été utilisée pour une application à l'échelle du pays, Tukey ne consentit jamais à reconnaître avoir utilisé l'approche bayésienne et refusa que son équipe écrive un article à ce sujet.

Le développement de l'informatique permet le développement du bayésien

Dans les années 1970, la communauté bayésienne était encore petite et l'approche fréquentiste dominait encore largement. Les applications étaient rares car les calculs étaient difficiles et de nombreux chercheurs se cantonnaient dans la sphère théorique. L'arrivée de l'informatique allait changer la donne.

Le développement de l'informatique va de pair avec le développement des algorithmes. Après la Seconde guerre mondiale, John von Neumann, Stanislaw Ulam et Nicholas Metropolis travaillaient au laboratoire de Los Alamos aux Etats-Unis sur le cheminement des neutrons et la masse nécessaire pour provoquer une explosion thermonucléaire. Ils utilisèrent les chaînes de Markov^c. L'algorithme développé à cette occasion fut appelé Monte-Carlo en référence à l'oncle d'Ulam qui était un joueur passionné. La forme moderne de l'algorithme parut en 1953 Metropolis et al. (1953). Cet algorithme fut généralisé par Hastings en 1970 (Hastings (1970)) qui fut dès lors dénommé algorithme de Metropolis-Hastings (ou Hastings-Metropolis). En 1984, en travaillant sur la restauration de photos floues, Stuart et Donald Geman décrivirent l'algorithme de Gibbs (Geman and Geman (1984)).

Toutes les briques existaient, il suffisait de les assembler. Ce fut le travail d'Alan Gelfand de l'université du Connecticut et d'Adrian Smith de l'Imperial College. En lisant différents articles traitants d'algorithmes et de leurs applications, ils eurent une idée lumineuse : remplacer les difficiles calculs d'intégrales par des échantillonnages dans la

c. Les chaînes de Markov portent le nom de leur inventeur, Andrei Markov (1856-1922).

distribution considérée. Pour tirer dans la loi voulue, il suffisait d'utiliser les chaînes de Markov avec des algorithmes appropriés (Gelfand and Smith (1990); Gelfand et al. (1990)). Ainsi, 250 ans après les écrits de Bayes, les statisticiens pouvaient enfin construire des lois *a priori* et des vraisemblances réalistes tout en étant capable de calculer les lois *a posteriori*. Tous ces calculs devenaient possibles car les ordinateurs étaient suffisamment puissants (voir Brooks (2003)). L'ensemble des algorithmes permettant de tirer des valeurs dans une loi en utilisant une chaîne de Markov fut baptisé MCMC (pour *Markov chain Monte Carlo*). Parmi d'autres chercheurs éminents du courant bayésien, on peut citer Anthony O'Hagan de l'université de Sheffield ou Andrew Gelman de l'université de Columbia.

Les problèmes auparavant difficile à résoudre devenaient bien plus simples avec la combinaison bayésien-MCMC. Depuis lors, le nombre de publications utilisant l'approche bayésienne va croissant. Les domaines d'application sont très variés. Pour faciliter encore l'application, le logiciel BUGS^d, créé par David Spiegelhalter dans l'unité de biostatistique au Medical research center de Cambridge (Lunn et al. (2000)), est un outil libre et gratuit partagé par une nombreuse communauté d'utilisateurs de statistique bayésienne.

Aujourd'hui, la bataille entre fréquentistes et bayésiens a fortement diminué. Beaucoup de statisticiens utilisent d'ailleurs l'une ou l'autre approche, selon les cas.

Lindley avait prédit que le XXI^{ème} siècle serait celui du bayésien et que la logique supérieure de cette approche renverserait les méthodes fréquentistes (Lindley (1975)). David Balckwell a tempéré cette idée (deGroot (1986)) : si l'approche bayésienne grandit dans le monde statistique, ce n'est pas en raison de l'influence des statisticiens mais en raison des actuaires, des ingénieurs et d'autres qui apprécient l'approche bayésienne et l'utilise. C'est bien ce qui est arrivé : le pragmatisme (et les ordinateurs) ont permis à l'approche bayésienne de sortir de sa confidentialité et d'être largement reconnue.

d. BUGS signifie *Bayesian inference using Gibbs sampling*.

Annexe B

Fonction d'utilité

Dans cette annexe, on s'intéresse à la fonction d'utilité et non à la fonction de coût (égale à l'opposé de la fonction d'utilité) car la théorie est souvent décrite sur la fonction d'utilité (Kadane (2011); Robert (2006); Berger (1985)). Les économistes sont des gens optimistes : ils ont élaboré la théorie de l'utilité. Les statisticiens doivent être pessimistes : ils emploient la fonction de coût dans la théorie statistique de la décision.

Deux principales difficultés sont rencontrées pour évaluer les conséquences d'une action. La première est qu'une conséquence peut ne pas avoir d'échelle de mesure évidente. Ainsi, le prestige, la bonne réputation, l'image de marque sont importants pour les entreprises mais sont difficiles à ordonner et à quantifier. Même dans le cas où une échelle monétaire est adaptée (gain au loto par exemple), la valeur du gain peut ne pas refléter la valeur réelle aux yeux du joueur. Par exemple, si une personne a le choix entre faire une tâche qu'elle juge pénible et gagner 100 euros ou ne rien faire et ne rien gagner, elle va la faire si elle juge que 100 euros est une somme importante. En revanche, on imagine qu'un millionnaire choisira de ne rien faire car 100 euros représentent peu pour lui.

Afin de déterminer mathématiquement la "valeur réelle" pour le décideur, les conséquences des décisions doivent être ordonnées numériquement. Ces nombres définissent la fonction d'*utilité*. Les bases de la théorie de l'utilité ont été développées par von Neumann and Morgenstern (1947). Elles ont été particulièrement développées dans la théorie des jeux. Dans cette partie, seul le cadre général et les axiomes nécessaires à la construction de la fonction d'utilité sont présentés.

Avant toutes choses, toutes les décisions possibles doivent être énumérées. L'ensemble des conséquences (ou récompenses) de ces décisions est noté \mathcal{R} . Très souvent, $\mathcal{R} = \mathbb{R}$. La plupart du temps, les conséquences possibles ne sont connues qu'avec incertitude. Dans ce cas, les résultats des décisions sont donc des distribution de probabilité. Soit \mathcal{P} l'espace des distributions de probabilité dans \mathcal{R} . Une relation d'ordre doit être définie dans \mathcal{P} , c'est-à-dire qu'il faut savoir si une distribution est préférée à une autre (sinon, il est illusoire de vouloir ordonner les conséquences). Une manière de faire est de trouver une fonction d'utilité U définie de \mathcal{R} dans \mathbb{R} telle que la "valeur" d'une probabilité de distribution P de \mathcal{P} serait donnée par l'utilité espérée $\mathbb{E}^P[U(r)]$. La construction de U , pour que cette dernière existe, suppose plusieurs axiomes.

Tout d'abord, on suppose qu'il est possible d'ordonner les récompenses. Il existe donc une relation d'ordre total notée \preceq telle que pour tout r_1 et r_2 deux récompenses dans \mathcal{R} ,

$r_1 \preceq r_2$ ou $r_2 \preceq r_1$.

La relation d'ordre est étendue à l'espace \mathcal{P} . Les axiomes suivants sont donc posés : soit P_1, P_2 et P_3 dans \mathcal{P} ,

Axiome B.1 *Ordre* : $P_1 \preceq P_2$ ou $P_2 \preceq P_1$.

Axiome B.2 *Transitivité* : si $P_1 \preceq P_2$ et $P_2 \preceq P_3$, alors $P_1 \preceq P_3$.

Axiome B.3 *Composition* : si $P_1 \preceq P_2$, alors $\alpha P_1 + (1 - \alpha)P \preceq \alpha P_2 + (1 - \alpha)P$ pour tout P de \mathcal{P} et pour tout α dans $]0; 1[$.

Axiome B.4 *Continuité* : si $P_1 \preceq P_2 \preceq P_3$, alors il existe α et β dans $]0; 1[$ tels que $\alpha P_1 + (1 - \alpha)P_3 \preceq P_2 \preceq \beta P_1 + (1 - \beta)P_3$.

L'axiome B.2 est un axiome de transitivité, qui est nécessaire pour comparer les distributions afin de ne pas se retrouver dans le cas $P_1 \preceq P_2 \preceq P_3 \preceq P_1$. La notation $\alpha P_1 + (1 - \alpha)P$ est un mélange : c'est une distribution qui génère P_1 avec une probabilité α et P avec une probabilité $1 - \alpha$. L'axiome B.3 précise que si une personne peut comparer P_1 à P_2 , alors elle conserve le même classement s'il y a une probabilité $1 - \alpha$ que P_1 et P_2 soient remplacés par P . L'axiome B.4 dit qu'il n'y a pas de récompense égale à $+\infty$ ou $-\infty$. Par exemple, si P_1 était considérée comme "infiniment mauvaise", alors il serait impossible de trouver β dans $]0; 1[$ tel qu'une personne désirant P_3 accepte de risquer d'avoir P_1 avec une probabilité β pour éviter P_2 . Cet axiome implique la mort causée par une intoxication alimentaire par exemple a une utilité finie. Les axiomes précédents garantissent l'existence de la fonction d'utilité U .

Théorème B.1 *La relation d'ordre \preceq satisfait les quatre axiomes précédents si et seulement si il existe $U : \mathcal{R} \mapsto \mathbb{R}$ telle que pour tout P et Q de \mathcal{P} , $P \preceq Q$ si et seulement si $\mathbb{E}^P[U(r)] \leq \mathbb{E}^Q[U(r)]$. De plus, si une autre fonction U' satisfait la même relation d'équivalence, alors il existe $a > 0$ et b tels que $U'(r) = aU(r) + b$, c'est-à-dire que la fonction U est unique à une transformation affine positive près.*

L'existence d'une fonction d'utilité peut être prouvée avec des axiomes plus faibles mais ce thème ne sera pas développé ici. Pour plus de précisions, voir Robert (2006), Mas-Colell and Monteiro (1991) et Cho and Kreps (1987).

Même dans le cas d'une récompense uniquement financière, la construction de la fonction U ne va pas forcément de soi car U n'est pas linéaire, surtout pour les grandes récompenses. Wilkinson (2008) pense même que l'échelle monétaire n'est pas appropriée pour prendre des décisions à propos de jeux financiers. En particulier, la valeur marginale de l'argent diminue pour la plupart des gens. Ainsi, recevoir 100€ lorsque l'on ne dispose de rien a plus de valeurs que lorsque l'on possède déjà 1.000.000 €. De même, il n'est généralement pas équivalent pour une personne de gagner 10.000 € de manière sûre ou de gagner 20.000 € avec probabilité 1/2. Wilkinson (2008) distingue d'ailleurs l'espérance de gain pour un jeu et l'espérance d'utilité pour ce même jeu. Pour une construction pas à pas d'une fonction d'utilité monétaire, voir (Berger, 1985, chap.2). Les personnes averses au risque ont une fonction d'utilité monétaire concave et majorée pour les grandes valeurs

de récompense. Au contraire, les amateurs de risque ont une fonction d'utilité convexe puisqu'ils préfèrent un gain aléatoire à une espérance de gain (cours de Wilkinson (2008)).

La théorie de la décision suppose que chaque décision d prise tandis que l'état de la nature se révèle valoir θ conduit à une récompense r , avec une utilité $U(r)$. Comme la récompense dépend uniquement de la décision d et du paramètre θ , l'utilité sera dorénavant notée $U(\theta, d)$. Si d'autres facteurs aléatoires Z interviennent dans U , alors $U(\theta, d) = \mathbb{E}_{\theta, d}[U(Z)]$. Une fois la fonction U construite, la fonction de coût est définie comme :

$$L(\theta, d) = -U(\theta, d).$$

Elle est parfois appelée *desutility function* en anglais.

Annexe C

Théorie bayésienne de la décision appliquée au contrôle par attribut

L'exemple présenté dans cette annexe porte sur le contrôle par attribut comme défini dans la partie 1.2.2. Ainsi qu'il est précisé dans la partie 1.3, le cadre de la théorie bayésienne de la décision nécessite de définir :

- l'ensemble \mathcal{D} des décisions ;
- l'ensemble Θ des états de la nature ainsi que la distribution *a priori* de θ notée $[\theta]$;
- la famille des distributions pour les observations, notées $[x|\theta]$;
- la fonction de coût L associée aux décisions et aux états de la nature.

Pour le contrôle par attribut, l'ensemble \mathcal{D} est binaire et vaut 0 si le lot est rejeté ou 1 si le lot est accepté par convention.

La qualité d'un lot est quantifié par sa prévalence, donc l'ensemble Θ est l'intervalle $[0; 1]$. La loi *a priori* de θ est la loi beta de paramètres a et b . Cette loi est choisie car elle est définie sur $[0; 1]$ et a une grande variété de formes, selon les valeurs des paramètres.

Les observations sont les résultats d'analyses de recherche donc sont sous formes de présence (codé 1) ou d'absence (codé 0). Il est naturel de choisir pour vraisemblance la fonction $[x|\theta] = \mathcal{B}in(n, \theta)$, c'est-à-dire une fonction binomiale. La paramètre n est le nombre d'analyses menées par lot de production. Dans ce cas, la loi *a posteriori* de θ est conjuguée : c'est une loi beta de paramètres $a + x$ et $b + n - x$.

La fonction de coût traduit le coût induit pour l'entreprise selon la décision prise. Quelle que soit la valeur prise par la fonction δ , l'entreprise paie le coût de l'échantillonnage. Celui-ci vaut Cn en notant C le coût de la vérification d'un produit. Si le lot est rejeté, l'entreprise a fabriqué le lot pour rien, ce qui lui coûte N en considérant le coût de fabrication d'un produit comme unitaire. En revanche, si le lot est accepté, l'entreprise met sur le marché des produits non-conformes en quantité $N\theta$. En notant k le coût de la "non-qualité" pour un produit, le coût total est $Nk\theta$. Les coûts sont résumés comme suit :

$$\begin{cases} L(\theta, \delta(x)) = Cn + N \times 1 & \text{si } \delta(x) = 0 \\ L(\theta, \delta(x)) = Cn + kN\theta & \text{si } \delta(x) = 1 \end{cases}$$

On se place ici dans le cadre prédictif, c'est-à-dire que les résultats x des analyses n'ont pas encore été observés. Le cadre expérimental e est ici défini par la taille n de

l'échantillon préleva dans chaque lot. Ainsi, la fonction de coût dépend de n . De même, x dépend de n puisque x varie de 0 à n . Pour ne pas alourdir les notations, l'indice n pour x est omis car il est clair que l'observation dépend de n , ainsi que cela apparaît dans l'écriture de la fonction L . Le risque bayésien s'écrit :

$$\begin{aligned} r(\delta_n) &= \int \int_{\mathcal{X}, \Theta} L(\theta, \delta_n(x)) d\theta dx \\ &= \int_{\mathcal{X}} \left(\int_{\Theta} (\mathbf{1}_{\delta_n(x)=0}(Cn + N) + \mathbf{1}_{\delta_n(x)=1}(Cn + kN\theta)) [\theta|x] d\theta \right) [x] dx \\ &= Cn + \int_{\mathcal{X}} \left(\mathbf{1}_{\delta_n(x)=0} N + \mathbf{1}_{\delta_n(x)=1} kN \frac{a+x}{a+b+n} \right) [x] dx \end{aligned}$$

La décision $\delta_n(x) = 1$ est choisie si $1 \geq k \frac{a+x}{a+b+n}$, autrement dit si $x \leq \frac{a+b+n}{k} - a$. Comme x est entier alors que le membre droit de l'inégalité ne l'est pas forcément, on prend la partie entière de ce membre. Notons

$$c_n = \lfloor \frac{a+b+n}{k} - a \rfloor \quad (\text{C.1})$$

La règle de décision optimale pour n et x fixés est donc la suivante :

$$\begin{cases} \text{si } x \leq c_n, \delta_n(x) = 1 \\ \text{si } x > c_n, \delta_n(x) = 0 \end{cases}$$

Pour obtenir la règle optimale δ'_n , il reste à intégrer selon x . La loi de x , dite de Polya, s'écrit :

$$\begin{aligned} [X = x] &= \int_0^1 [x|\theta][\theta] d\theta \\ &= \frac{\Gamma(a+b) \Gamma(n+1) \Gamma(a+x) \Gamma(n-x+b)}{\Gamma(a) \Gamma(b) \Gamma(x+1) \Gamma(n-x+1) \Gamma(n+a+b)}. \end{aligned}$$

Revenons au calcul du risque bayésien :

$$\begin{aligned} r(\delta_n) &= Cn + \sum_{x=c_n}^n N[X = x] + \sum_{x=0}^{c_n} kN \frac{a+x}{a+b+n} [X = x] \\ &= Cn + N + N \sum_{x=0}^{c_n} \left(k \frac{a+x}{a+b+n} - 1 \right) [X = x]. \end{aligned}$$

Ce risque se calcule analytiquement. Le risque est calculé pour différentes valeurs de n afin de déterminer la valeur de n minimisant r . La figure C.1 représente r en fonction de n . Les valeurs numériques utilisées sont : $a = 2, 18$, $b = 38, 18$, $k = 17$, $C = 1$ et $N = 3000$. La courbe en forme de "vagues" s'explique car le seuil c prend la même valeur pour plusieurs valeurs consécutives de n .

Ici, avec ces valeurs, le minimum est atteint pour $n = 90$. c_n vaut alors 5, d'après l'équation C.1.

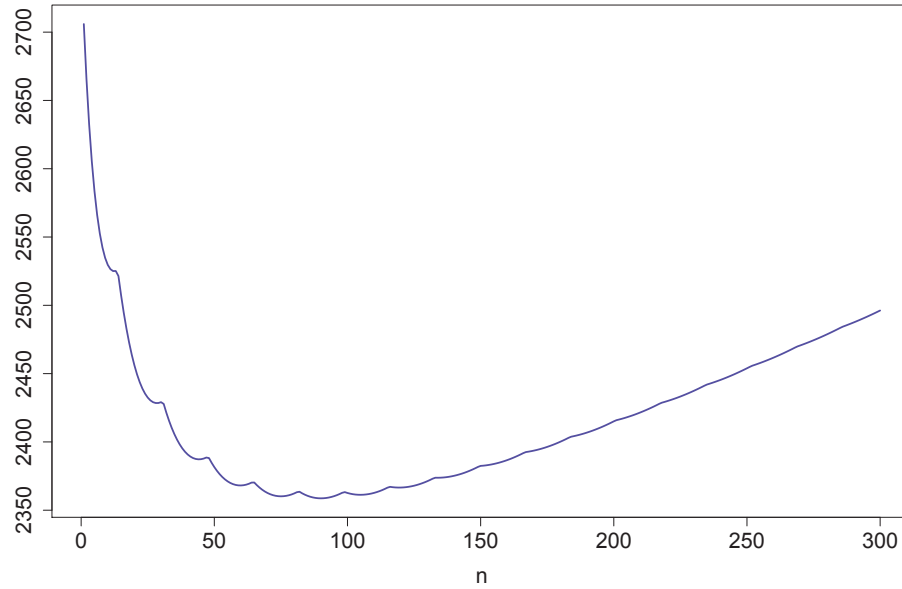


FIGURE C.1 – Graphe du risque bayésien r en fonction de la taille n de l'échantillon

Ce cas est simple car les lois sont conjuguées, la loi marginale de x est calculable et la fonction de coût est donnée. Les applications présentées aux chapitres 6 et 7 sont plus complexes, notamment parce que les valeurs des paramètres de la loi *a priori* et la fonction de coût sont à définir.

Annexe D

Inférence bayésienne et inférence classique

L'*inférence statistique* est une technique permettant de tirer des conclusions sur une population à partir d'un échantillon de cette population. L'inférence peut se réaliser en utilisant l'approche bayésienne ou l'approche classique. La manière de mener l'inférence est d'ailleurs directement liée à la définition de ces approches. Ces approches diffèrent car elles interprètent différemment la probabilité. L'approche bayésienne interprète la probabilité comme un pari. La probabilité permet d'exprimer l'incertitude que l'on peut avoir sur une grandeur (Bayes (1763)). Elle reflète des croyances personnelles sur les événements. La probabilité au sens classique est définie comme la fréquence relative limite d'événements répétables. Considérons une expérience aléatoire dont l'ensemble des résultats possibles est connu et est constitué d'un sous-ensemble A d'événements. On s'intéresse au résultat d'expérience binaire^a : soit l'événement A se produit, soit il ne se produit pas. La probabilité que A se produise est égale à la limite de la fréquence d'occurrence de A :

$$P(A) = \lim_{n \rightarrow +\infty} \frac{n_A}{n},$$

avec n_A le nombre d'occurrence de A parmi les n répétitions de l'expérience. Si on considère par exemple un modèle statistique paramétrique avec un paramètre noté θ , l'inférence sur θ consiste à estimer θ en tenant compte des observations x . Ainsi, on observe les conséquences (les données) et on en déduit une caractérisation sur les causes (paramètre)^b.

D.1 Inférence bayésienne

Dans l'approche bayésienne, le paramètre θ est caractérisé par une distribution qui reflète l'information que l'on a (ou pas) sur lui avant l'observation des données. Il peut s'agir d'une grandeur physique comme la population d'un pays à un instant donné mais comme cette grandeur est difficilement mesurable de manière exacte, on représente notre

a. Cela s'appelle également une épreuve de Bernoulli.

b. c'est pourquoi la statistique était appelée *probabilités inverses* au temps de Bayes et de Laplace (Robert (2006)).

croissance sur les valeurs possibles prises par cette grandeur par une fonction de probabilité : la loi *a priori* d'un paramètre peut être vue comme une façon de résumer l'information disponible sur le paramètre θ , que ce dernier soit réellement aléatoire ou non (Robert (2006)). L'inférence bayésienne permet de mettre à jour les croyances ou le degré de confiance que l'on a sur un paramètre après observation des données. Pour ce faire, le théorème de Bayes est utilisé : soit deux événements A et B tels que $P(B) \neq 0$, alors

$$\begin{aligned} P(A|B) &= \frac{P(B|A)P(A)}{P(B)} \\ &= \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}, \end{aligned} \quad (\text{D.1})$$

avec A^c , le complémentaire de l'événement A . On peut également interpréter ce théorème en terme d'actualisation puisque notre croyance sur l'occurrence de l'événement A ($P(A)$) a été modifiée après avoir observé l'événement B ($P(A|B)$). Ce théorème peut être vu comme une manière d'actualiser nos connaissances sur le paramètre θ puisqu'il permet de calculer la distribution *a posteriori* $[\theta|x]$ du paramètre θ en tenant compte des observations x . Le théorème de Bayes s'écrit alors (si θ est un paramètre continu) :

$$[\theta|x] = \frac{[x|\theta][\theta]}{\int [x|\theta][\theta]d\theta}.$$

D.2 Inférence fréquentiste

Dans cette approche, le paramètre θ est considéré comme fixe mais inconnu. L'inférence est faite en tenant compte des données à disposition et ne tient pas compte de croyances que l'on pourrait avoir. Le statisticien peut s'intéresser à trois types de problèmes :

- l'estimation ponctuelle : la réponse est un point de l'espace Θ du paramètre ;
- le test d'hypothèse : Θ est partitionné en deux sous-espaces Θ_0 et $\Theta_1 = \Theta \setminus \Theta_0$ et le statisticien doit déterminer dans quel sous-ensemble se situe la vraie valeur de θ ;
- l'estimation ensembliste : le statisticien s'intéresse à une estimation ponctuelle et l'accompagne d'une partie de Θ censée contenir la vraie valeur de θ . Si on travaille en une dimension, cette partie s'appelle un intervalle de confiance.

D.3 Intervalle de crédibilité et intervalle de confiance

Les deux approches fournissent des estimations sous forme d'intervalle : l'intervalle de crédibilité pour l'approche bayésienne et l'intervalle de confiance pour l'approche fréquentiste.

L'intervalle de crédibilité de niveau α du paramètre θ est un intervalle dans lequel α % des valeurs de θ se trouvent. Par exemple, si un intervalle de crédibilité de niveau α vaut $[a; b]$, alors $\mathbb{P}(a \leq \theta \leq b) = \alpha$.

L'intervalle de confiance permet d'évaluer la précision de l'estimation du paramètre θ . Considérons un intervalle de confiance de niveau α : en moyenne la vraie valeur de θ se trouve avec une probabilité α dans l'intervalle. Dit autrement, si on réalise plusieurs fois un échantillonnage de la population d'intérêt et que l'on calcule l'intervalle de confiance à chaque fois, la vraie valeur de θ sera contenue dans α % des intervalles de confiance.

Il arrive que la définition de l'intervalle de confiance soit confondue avec celle de l'intervalle de crédibilité or, en fréquentiste, comme le paramètre n'a qu'une seule valeur possible (mais inconnue) la probabilité que cette valeur se trouve dans un intervalle de confiance vaut 0 ou 1. Les intervalles de confiance et de crédibilité ne coïncident pas toujours car les intervalles de crédibilité dépendent des données et de la loi *a priori* de θ alors que l'intervalle de confiance ne dépend que des données.

D.4 Tests d'hypothèse

D.4.1 Approche fréquentiste

En fréquentiste, le test d'hypothèse est une méthode pour prendre une décision en fonction des données observées. En général, ce test repose sur le choix entre deux hypothèses : l'hypothèse nulle H_0 et l'hypothèse alternative H_1 . La question à laquelle le test répond est : en supposant que l'hypothèse H_0 est vraie, quelle est la probabilité d'observer une certaine valeur pour la statistique de test qui est au moins aussi extrême que la valeur de la statistique calculée à partir des données observées ? Une statistique de test est une fonction des données, c'est un résumé des données qui permet de réduire les données à un petit nombre de valeurs. Le test se déroule de la manière suivante :

- formulation de l'hypothèse nulle et de l'hypothèse alternative (hypothèse que l'on propose pour décrire la réalité si H_0 est rejetée) ;
- calcul de la statistique de test T retenue et détermination de sa distribution si l'hypothèse H_0 est vraie. Les valeurs possibles de la statistique sont séparées en deux régions : la région pour laquelle l'hypothèse H_0 est rejetée et la région pour laquelle cette hypothèse n'est pas rejetée ;
- calcul de la statistique de test t_{obs} à partir des données observées ;
- décision de rejeter (en faveur de l'hypothèse H_1) ou de ne pas rejeter l'hypothèse H_0 selon que t_{obs} se situe dans la région de T pour laquelle l'hypothèse H_0 est rejetée ou non.

La valeur seuil t_0 de la statistique de test séparant la région pour laquelle l'hypothèse H_0 n'est pas rejetée de la région pour laquelle cette hypothèse est rejetée dépend du niveau de significativité α du test, fixé par le décideur. Par exemple, si ce niveau est fixé à 5 %, alors la valeur seuil de la statistique de test T est égal au 95^e percentile de T en supposant que l'hypothèse H_0 est vérifiée. On peut alors calculer la *p-valeur* p qui est la probabilité d'observer des données aussi extrêmes que les données observées. En terme mathématiques : $p = \mathbb{P}(T > t_{obs})$. Si p est inférieur à 5 % alors H_0 est rejetée car on considère que la probabilité d'observer les données dont on dispose sous l'hypothèse H_0 est trop faible. La *p-valeur* donne un poids à la force de conviction dans le rejet ou non de H_0 . Plus p est petit (tout en étant inférieur à 5 %) et plus on est confiant dans le rejet de H_0 .

		Décision	
		Rejet de H_0	Non rejet de H_0
Réalité	H_0 vraie	Risque de 1ère espèce α Faux positif	$1 - \alpha$ Vrai négatif
	H_1 vraie	Puissance $1 - \beta$ Vrai positif	Risque de 2nde espèce β Faux négatif

TABLE D.1 – Risques de première et seconde espèce et puissance de test en fonction de la réalité et de la décision prise.

La décision prise n'est pas de choisir entre H_0 et H_1 mais bien de rejeter ou non H_0 . Soit on rejette H_0 parce que les données observées sont incompatibles avec cette hypothèse soit on ne la rejette pas parce que les données sont compatibles. Cependant, on n'accepte pas H_0 pour autant parce que la compatibilité des observations avec une hypothèse ne garantit pas que cette hypothèse soit vraie.

Quelle que soit la décision prise, il est possible que l'on se soit trompé. Il existe deux erreurs possibles :

- erreur de première espèce α : probabilité de rejeter H_0 alors que H_0 est vraie^c. Le risque de première espèce est aussi appelé risque de faux positif car l'hypothèse H_1 est considérée comme validée (positif) alors que c'est faux.
- erreur de seconde espèce β : probabilité de ne pas rejeter H_0 alors que H_1 est vraie. On ne peut pas calculer β exactement car on ne connaît pas la valeur du paramètre θ si H_1 est vraie. Pour calculer β , il faut spécifier une hypothèse H_1 particulière (donc une valeur précise pour θ). C'est ce qui est fait pour calculer le niveau de qualité limite NQL (cf. partie 1.2.3.a). Le risque de deuxième espèce est également appelé risque de faux négatif car l'hypothèse H_0 n'est pas rejetée (négatif) alors que c'est faux. La puissance d'un test est égale à $1 - \beta$. Elle caractérise la capacité de rejeter H_0 quand H_1 est vraie.

En plus des risques de faux positifs et de faux négatifs, il y a aussi la probabilité de vrai positif (H_1 est vraie et le test conduit à rejeter H_0) et la probabilité de vrai négatif (H_0 est vraie et le test conduit à ne pas rejeter H_0). Le tableau D.1 récapitule les différents risques d'erreur.

Idéalement, ou souhaiterait que les erreurs des deux types soient nulles, mais c'est impossible tant qu'on ne dispose que d'un nombre fini d'observations.

D.4.2 Approche bayésienne

Dans l'approche bayésienne, on peut comparer deux hypothèses H_0 et H_1 . Pour cela, on assigne aux hypothèses H_0 et H_1 des probabilités *a priori* et on met à jour ces probabilités en tenant compte des observations via la vraisemblance en appliquant le

c. Le risque de première espèce est communément appelé α de même que le niveau de significativité. Le lien entre les deux est qu'un test d'hypothèse qui rejette H_0 pour tous les échantillons dont p est inférieur au niveau de significativité j a un risque de première espèce égal à j .

$\log(BF_{01})$	BF_{01}	Preuve contre H_1
0 à 1/2	1 à 3,2	à mentionner simplement
1/2 à 1	3,2 à 10	substantielle
1 à 2	10 à 100	forte
>2	>100	décisive

TABLE D.2 – Preuve contre H_1 par rapport à H_0 en fonction du résultat numérique du facteur de Bayes BF_{01} .

théorème de Bayes. Par exemple, la probabilité *a posteriori* de H_0 vaut :

$$\mathbb{P}(H_0|D) = \frac{\mathbb{P}(D|H_0)\mathbb{P}(H_0)}{\mathbb{P}(D|H_0)\mathbb{P}(H_0) + \mathbb{P}(D|H_1)\mathbb{P}(H_1)},$$

avec D les données, $\mathbb{P}(D|H_i)$ la vraisemblance des données D sachant l'hypothèse i ($i = 0; 1$) et $\mathbb{P}(H_i)$ la probabilité *a priori* de H_i .

Pour comparer les deux hypothèses, on peut utiliser le facteur de Bayes qui vaut :

$$BF_{01} = \frac{\mathbb{P}(H_0|D)/\mathbb{P}(H_0)}{\mathbb{P}(H_1|D)/\mathbb{P}(H_1)} = \frac{\mathbb{P}(D|H_0)}{\mathbb{P}(D|H_1)}.$$

En fonction du résultat obtenu, on décide si H_0 est plus probable que H_1 . Jeffreys (1961a) a suggéré d'interpréter le facteur de Bayes selon l'échelle logarithmique décimale (cf. tableau D.2).

De même que dans l'approche fréquentiste, on ne décide pas si une hypothèse est vraie ou fausse. On étudie la plausibilité d'une hypothèse par rapport à une autre en tenant compte des données.

Le facteur de Bayes est aussi utilisé dans la sélection de modèles, ce qui est le même problème que le test d'hypothèse. En effet, supposer une hypothèse H_i revient à supposer l'existence d'un certain modèle (cf. partie 9.1.2).

Annexe E

Algorithme EM

Cette annexe présente l'application de l'algorithme EM pour estimer les paramètres du modèle \mathcal{M}_{CLD} décrit au chapitre 4.

E.1 Principe de l'algorithme EM

L'algorithme espérance-maximisation appelé par la suite algorithme EM, a été proposé par Dempster et al. (1977). Il permet de trouver le maximum de vraisemblance des paramètres de modèles probabilistes lorsque le modèle dépend de variables latentes. L'algorithme EM est constitué de deux étapes :

- une étape d'évaluation de l'espérance (E), où l'on calcule l'espérance de la vraisemblance en tenant compte des dernières valeurs des paramètres,
- une étape de maximisation (M), qui sert à déterminer le maximum de vraisemblance des paramètres en maximisant l'expression trouvée à l'étape (E).

Les paramètres trouvés à l'étape (M) sont ensuite utilisés comme point de départ d'une nouvelle phase d'évaluation de l'espérance, et l'on itère ainsi.

La vraisemblance du modèle \mathcal{M}_{CLD} s'écrit :

$$\begin{aligned}y &\sim \mathcal{P}(\lambda M' \frac{V_r}{V_d}) \\x &\sim \mathcal{Ber}(1 - \exp(-\lambda M)) \\ \lambda &\sim \mathcal{LN}(\mu, \sigma).\end{aligned}$$

En passant la vraisemblance sous forme logarithmique, en isolant le terme $\log[x, y|\theta]$ et en notant $\theta = (\mu, \sigma)$, il vient :

$$\log[x, y|\theta] = \log[x, y, \lambda|\theta] - \log[\lambda|\theta, x, y]$$

L'objectif est d'atteindre le maximum de vraisemblance, solution de $\frac{\partial \log[x, y|\theta]}{\partial \theta} = 0$ lorsque la vraisemblance, non explicite, est celle d'un modèle hiérarchique.

Les lois de probabilité jointes $[x, y, \lambda|\theta]$ et la conditionnelle $[\lambda|\theta, x, y]$ ont des supports indépendants de θ . Si on intègre la log-vraisemblance par rapport la densité de probabilité $[\lambda|\theta', x, y]$, il vient :

$$\begin{aligned}\log[x, y|\theta] &= \int_{\lambda} \log[x, y, \lambda|\theta][\lambda|\theta', x, y]d\lambda - \int_{\lambda} \log[\lambda|\theta, x, y][\lambda|\theta', x, y]d\lambda \\ &= Q(\theta, \theta') - H(\theta, \theta')\end{aligned}\tag{E.1}$$

On peut montrer facilement que, sous des conditions de régularité du modèle, la fonction $\theta \mapsto H(\theta, \theta')$ atteint son maximum en $\theta = \theta'$. Considérons la formule de l'équation (E.1) pour θ et pour θ'

$$\log[x, y|\theta] - \log[x, y|\theta'] = (Q(\theta, \theta') - Q(\theta', \theta')) + (H(\theta', \theta') - H(\theta, \theta'))$$

L'algorithme EM est une procédure itérative pour passer d'une valeur du paramètre θ' à une valeur θ^* telle que $Q(\theta^*, \theta') > Q(\theta', \theta')$. Le mieux que l'on puisse faire à chaque étape est de trouver

$$\theta^* = \text{Arg max}_{\theta} Q(\theta, \theta')$$

Ainsi, le passage de l'étape k qui fait passer de la valeur θ^k à la valeur θ^{k+1} :

1. **Phase E** : calculer pour tout θ , l'expression : $Q(\theta, \theta^k) = \int_{\lambda} \log[x, y, \lambda|\theta] \times [\lambda|\theta^k, x, y]d\lambda$.

Ce calcul d'espérance de $\log[x, y, \lambda|\theta]$ à x , y et θ fixés est rarement explicite. Comme lors de l'inférence bayésienne, on a ici souvent recours à des méthodes de simulation de $[\lambda|\theta^k, x, y]$ de type importance sampling ou MCMC.

2. **Phase M** : chercher θ^{k+1} qui réalise le maximum de $Q(\theta, \theta^k)$ soit :

$$\frac{\partial Q(\theta, \theta^k)}{\partial \theta} = 0 = \int_{\lambda} \frac{\partial \log[x, y, \lambda|\theta]}{\partial \theta} [\lambda|\theta^k, x, y]d\lambda.$$

Comme dans toute méthode de gradient, l'obtention d'un maximum global n'est pas garantie, l'algorithme peut donc se piéger dans un minimum local.

E.2 Application de l'algorithme EM au modèle \mathcal{M}_{RD}

E.2.1 Calcul de $Q(\theta, \theta')$

La quantité $Q(\theta, \theta')$ s'écrit dans le cas présent :

$$\begin{aligned}
Q(\theta, \theta') &= \int \ln[x, y, \lambda|\theta][\lambda|\theta', y, x]d\lambda \\
&= \sum_j \int \ln[x_j, \lambda|\theta][\lambda|\theta', x_j]d\lambda + \sum_i \int \ln[y_i, \lambda|\theta][\lambda|\theta', y_i]d\lambda \\
&= Q_j(\theta, \theta') + Q_i(\theta, \theta') \\
&= \sum_j \int \left(x_j \ln(1 - \exp(-\lambda_j m_j)) - (1 - x_j)\lambda_j m_j - 0.5 \ln(\sigma^2 2\pi) \right. \\
&\quad \left. - 0.5(\mu - \ln \lambda_j)^2 / \sigma^2 - \ln \lambda_j \right) [\lambda_j|\theta', x_j]d\lambda_j \\
&\quad + \sum_i \int \left(-\lambda_i m_i d + y_i \ln(\lambda_i m_i d) - \ln(y_i!) - 0.5 \ln(\sigma^2 2\pi) \right. \\
&\quad \left. - 0.5(\mu - \ln \lambda_i)^2 / \sigma^2 - \ln \lambda_i \right) [\lambda_i|\theta', y_i]d\lambda_i
\end{aligned}$$

La quantité $Q(\theta, \theta')$ est maximale lorsque sa dérivée par rapport à μ et à σ^{-2} est nulle :

$$\begin{aligned}
\frac{\partial Q_j}{\partial \mu} + \frac{\partial Q_i}{\partial \mu} = 0 &\Leftrightarrow \int \sum_j \left(-(\mu - \ln \lambda_j) / \sigma^2 [\lambda_j|\theta', x_j]d\lambda_j \right) \\
&\quad + \sum_i \left(-(\mu - \ln \lambda_i) / \sigma^2 [\lambda_i|\theta', y_i]d\lambda_i \right) = 0 \\
&\Leftrightarrow -N_x \sigma^{-2} \mu + \int \sigma^{-2} \sum_j \ln \lambda_j [\lambda_j|\theta', x_j]d\lambda_j - N_y \sigma^{-2} \mu \\
&\quad + \int \sigma^{-2} \sum_i \ln \lambda_i [\lambda_i|\theta', y_i]d\lambda_i = 0 \\
&\Leftrightarrow \mu = \frac{1}{N_x + N_y} \left(\sum_j \int \ln \lambda_j [\lambda_j|\theta', x_j]d\lambda_j + \sum_i \int \ln \lambda_i [\lambda_i|\theta', y_i]d\lambda_i \right)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Q_j}{\partial \sigma^2} + \frac{\partial Q_i}{\partial \sigma^2} = 0 &\Leftrightarrow \int \sum_j \left(-\frac{1}{2\sigma^2} + 0.5 \frac{(\mu - \ln \lambda_j)^2}{\sigma^4} \right) [\lambda_j | \theta', x_j] d\lambda_j \\
&+ \int \sum_i \left(-\frac{1}{2\sigma^2} + 0.5 \frac{(\mu - \ln \lambda_i)^2}{\sigma^4} \right) [\lambda_i | \theta', y_i] d\lambda_i \\
&\Leftrightarrow N_x \sigma^2 - \mu^2 N_x + 2\mu \int \sum_j \ln \lambda_j [\lambda_j | \theta', x_j] d\lambda_j - \sum_j \int (\ln \lambda_j)^2 [\lambda_j | \theta', x_j] d\lambda_j \\
&+ N_y \sigma^2 - \mu^2 N_y + 2\mu \int \sum_i \ln \lambda_i [\lambda_i | \theta', y_i] d\lambda_i - \sum_i \int (\ln \lambda_i)^2 [\lambda_i | \theta', y_i] d\lambda_i \\
&\Leftrightarrow \sigma^2 = \frac{1}{N_x + N_y} \left(\mu^2 (N_x + N_y) \right. \\
&\quad \left. - 2\mu \left(\sum_j \int \ln \lambda_j [\lambda_j | \theta', x_j] d\lambda_j + \sum_i \int \ln \lambda_i [\lambda_i | \theta', y_i] d\lambda_i \right) \right. \\
&\quad \left. \left(\sum_j \int (\ln \lambda_j)^2 [\lambda_j | \theta', x_j] d\lambda_j + \sum_i \int (\ln \lambda_i)^2 [\lambda_i | \theta', y_i] d\lambda_i \right) \right)
\end{aligned}$$

Il faut donc calculer, pour chaque itération :

$$\begin{aligned}
Ae_i &= \int \ln \lambda_i [\lambda_i | \theta', y_i] d\lambda_i \\
Ad_j &= \int \ln \lambda_j [\lambda_j | \theta', x_j] d\lambda_j \\
Be_i &= \int (\ln \lambda_i)^2 [\lambda_i | \theta', y_i] d\lambda_i \\
Bd_j &= \int (\ln \lambda_j)^2 [\lambda_j | \theta', x_j] d\lambda_j
\end{aligned}$$

E.2.2 Calcul de Ae_i , Ad_i , Be_j et Bd_j par importance sampling

Loi d'importance pour $\lambda_i | \theta', y_i$ La loi de $\lambda_i | \theta', y_i$ n'est connue qu'à une constante près ($[\lambda_i | \theta', y_i] \propto [y_i | \lambda_i] [\lambda_i | \theta']$). Pour calculer les intégrales, on utilise la technique de l'importance sampling (IS) pour calculer Ae_i et Be_i . Comme le modèle Poisson Log-Normal est assez proche du modèle Poisson Gamma, nous avons donc choisi la loi Gamma comme loi d'importance pour λ_i : $\lambda_i \sim \mathcal{G}(\alpha, \beta)$.

Comment choisir α et β ? A partir des paramètres μ' et $\sigma^{2'}$. En effet, comme le logarithme de λ_i suit une loi normale :

$$\begin{aligned}
E(\lambda_i) &= \exp\left(\mu' + \frac{\sigma^{2'}}{2}\right) & V(\lambda_i) &= \exp\left(\mu' + \frac{\sigma^{2'}}{2}\right) (\exp(\sigma^{2'}) - 1) \\
E(\lambda_i) &= \frac{\alpha}{\beta} & V(\lambda_i) &= \frac{\alpha}{\beta^2}
\end{aligned}$$

D'où :

$$\alpha = \frac{\exp\left(\mu' + \frac{\sigma^{2'}}{2}\right)}{\exp(\sigma^{2'}) - 1} \quad \beta = \frac{1}{\exp(\sigma^{2'}) - 1}$$

Pour tout $i = 1, \dots, N$, on tire G valeurs de λ_i dans une loi Gamma de paramètres a_i et b , puis on calcule les poids $\omega_i^{(g)}$:

$$\begin{aligned} \forall g = 1, \dots, G, \quad \omega_i^{(g)} &= \frac{[y_i|\lambda_i^{(g)}][\lambda_i^{(g)}|\theta']}{[\lambda_i^{(g)}]} \\ &= \frac{\mathcal{P}(y_i|\lambda_i^{(g)})\mathcal{LN}(\lambda_i^{(g)}|\mu', \sigma'^2)}{\mathcal{G}(\lambda_i^{(g)}|a_i, b)} \end{aligned}$$

Une fois les poids normalisé, on obtient Ae_i et Be_i en calculant la moyenne des poids :

$$\begin{aligned} Ae_i &= \sum_g \ln \lambda_i^{(g)} \omega_i^{(g)} \\ Be_i &= \sum_g (\ln \lambda_i^{(g)})^2 \omega_i^{(g)} \end{aligned}$$

A l'usage, je me suis rendue compte que lorsqu'il y a peu ou pas de 0, la fonction d'importance n'est pas très adaptée : le nombre de particules effectives est très faible (de l'ordre de 5% du nombre réel de particules). Cela est principalement dû au fait que la variance de la loi Gamma est assez faible et que la moyenne de $\lambda_i^{(g)} M'd$ est relativement loin de y_i . Dans ce cas, la probabilité que y_i soit tiré dans une Poisson de paramètre $\lambda_i^{(g)} M'd$ est très faible voire nulle. Pour corriger ce phénomène, la fonction d'importance a été modifiée de manière à ce que la moyenne de $\lambda_i^{(g)} M'd$ soit très proche de y_i . Désormais, $\lambda_i^{(g)}$ est tiré dans une loi gamma de paramètres $a_i = (\beta + M'd) \times y_i / (M'd) + 0, 1$ et $b_i = \beta + M'd$. Dans ce cas, $E(\lambda_i^{(g)} M'd) = M'd ((\beta + M'd) \times y_i / (M'd) + 0, 1) / (\beta + md) \simeq y_i$. Dans la formule de a_i , lorsque y_i vaut 0, les particules simulées $\lambda_i^{(g)}$ valent souvent 0, ce qui incompatible avec la loi log-normale. Ajouter 0,1 à a_i permet de contourner ce problème.

Loi d'importance pour $\lambda_j|\theta', x_j$ Comme l'information apportée par la détection n'est pas très grande, on considère que $[\lambda_j|\theta', x_j]$ est proche de $[\lambda_j|\theta']$. On choisit donc cette dernière comme loi d'importance. Dans ce cas, le poids non normalisé $\omega_j^{(g)}$ d'une particule g vaut :

$$\omega_j^{(g)} = \mathcal{Ber}(x_j|\lambda_j^{(g)})$$

E.2.3 Algorithme obtenu

Les étapes de l'algorithme, à l'itération k , sont donc :

1. Calcul de β

$$\beta = \frac{1}{\exp(\sigma^{2(k)}) - 1}$$
2. Quel que soit $i = 1, \dots, N_y$, tirage de G particule pour λ_i dans une loi Gamma de paramètres $a_i = (\beta + M'd) \times y_i / (M'd) + 0, 1$ et $b = \beta + M'd$ et calcul des poids $\omega_i^{(g)}$:

$$\omega_i^{(g)} = \frac{\mathcal{P}(y_i|\lambda_i^{(g)})\mathcal{LN}(\lambda_i^{(g)}|\mu^{(k)}, \sigma^{2(k)})}{\mathcal{G}(\lambda_i^{(g)}|a_i, b)}$$

3. Quel que soit $j = 1, \dots, N_x$, tirage de G particule pour λ_j dans une loi lognormale de paramètre θ^k et calcul des poids $\omega_j^{(g)}$:

$$\omega_j^{(g)} = \mathcal{Ber}(x_j | \lambda_j^{(g)})$$

4. Normalisation des poids : $\underline{\omega}_j^{(g)} = \frac{\omega_j^{(g)}}{\sum_g \omega_j^{(g)}}$ et $\underline{\omega}_i^{(g)} = \frac{\omega_i^{(g)}}{\sum_g \omega_i^{(g)}}$.

5. Calcul de $Ae_i = \sum_g \ln \lambda_i^{(g)} \underline{\omega}_i^{(g)}$, $Be_i = \sum_g (\ln \lambda_i^{(g)})^2 \underline{\omega}_i^{(g)}$, $Ad_j = \sum_g \ln \lambda_j^{(g)} \underline{\omega}_j^{(g)}$ et de $Bd_j = \sum_g (\ln \lambda_j^{(g)})^2 \underline{\omega}_j^{(g)}$.

6. Calcul de

$$\mu^{(k+1)} = \frac{1}{N_x + N_y} \left(\sum_i Ae_i + \sum_j Ad_j \right)$$

et de

$$\sigma^{2(k+1)} = \mu^{2(k+1)} - \frac{1}{N_x + N_y} 2\mu^{(k+1)} \left(\sum_i Ae_i + \sum_j Ad_j \right) + \frac{1}{N_x + N_y} \left(\sum_i Be_i + \sum_j Bd_j \right)$$

Ces étapes sont menées jusqu'à ce que $\mu^{(k)}$ et $\sigma^{(k)}$ convergent. Pour utiliser cet algorithme, il reste une étape lors de la première itération : le calcul de $\mu^{(1)}$ et de $\sigma^{2(1)}$, qui peut se faire par exemple à partir de l'estimation de μ et de σ obtenue par maximum de vraisemblance sur les données transformées ou alors en assignant à $\mu^{(1)}$ et à $\sigma^{2(1)}$ des valeurs constantes indépendamment du jeu de données.

Annexe F

Article Billoir et al.

Cette partie est la reproduction de l'article intitulé Probabilistic modeling of the fate of Listeria monocytogenes in diced bacon during the manufacturing process écrit par Elise Billoir en collaboration avec Jean-Baptiste Denis, Natalie Commeau, Marie Cornu et Véronique Zuliani. Il a été publié dans la revue Risk Analysis (2011, 31(2), 237-254)

Abstract To assess the impact of the manufacturing process on the fate of *Listeria monocytogenes*, we built a generic probabilistic model intended to simulate the successive steps in the process. Contamination evolution was modeled in the appropriate units (breasts, dice, and then packaging units through the successive steps in the process). To calibrate the model, parameter values were estimated from industrial data, from the literature, and based on expert opinion. By means of simulations, the model was explored using a baseline calibration and alternative scenarios, in order to assess the impact of changes in the process and of accidental events. The results are reported as contamination distributions and as the probability that the product will be acceptable with regards to the European regulatory safety criterion. Our results are consistent with data provided by industrial partners and highlight that tumbling is a key step for the distribution of the contamination at the end of the process. Process chain models could provide an important added value for risk assessment models that basically consider only the outputs of the process in their risk mitigation strategies. Moreover, a model calibrated to correspond to a specific plant could be used to optimize surveillance.

Keywords : Process chain modeling ; Bayesian network ; Microbial risk assessment ; *Listeria monocytogenes* ; diced bacon.

F.1 Introduction

The use of modeling as a tool to investigate the fate of microbiological hazards during food processing chains is an emerging approach. Only a few process chain models have been published. Nauta⁽¹⁾ first introduced the "Modular Process Risk Model" (MPRM), where the food pathway is divided into smaller steps. Since then, the concept has been further investigated^(2,3), and the methodology has been applied to *Staphylococcus aureus* in unripened cheese made from raw milk⁽⁴⁾, to poultry-processing⁽⁵⁾, to cross-contamination

during the smearing of cheese⁽⁶⁾, to *Campylobacter* in broiler meat⁽⁷⁾, to *Bacillus cereus* in courgette purée⁽⁸⁾, and to *Escherichia coli* O157 in milk⁽⁹⁾ or in beef^(10,11).

In this paper, we focus on the fate of *L. monocytogenes* during the processing of diced bacon. In France, 50% of pork breasts are transformed into diced bacon^(12,13), and this production is increasing every year. Diced bacon consists of processed pork breasts tumbled with brine, steamed (but not cooked), diced and packed under a modified atmosphere (between 20%CO₂/80%N₂ and 50%CO₂/50%N₂). The initial pH of this product is between 5.6 and 5.8, its water activity is around 0.96, and its shelf life is between 35 and 52 days. It is not unusual for *L. monocytogenes* to be detected in pork breasts, and the manufacture of diced bacon does not include any physical or chemical treatment sufficiently drastic to eradicate any *L. monocytogenes* present. Moreover, given the physico-chemical conditions of this product during the post-processing chill chain, *L. monocytogenes* growth is theoretically possible. Diced bacon is expected to be cooked before consumption, but in fact it is occasionally consumed raw by 14% of French people (from several times a week to less than once a month)⁽¹⁴⁾, so that diced bacon can be considered to be a ready-to-eat (RTE) product.

L. monocytogenes is responsible for listeriosis, a severe foodborne disease that may lead to gastroenteritis, meningitis, septicemia, spontaneous miscarriage, and perinatal infections with high rates of lethality in immuno-compromized individuals. Its ubiquitous occurrence and ability to grow or survive in a chilled environment makes *L. monocytogenes* a significant challenge in food production, particularly for RTE products. For these reasons, criteria for *L. monocytogenes* in RTE products have been imposed by many legislations, in most cases absence in 25 g and/or 100 CFU.g⁻¹ at the end of the shelf-life.

In this study, we propose a basic model of the diced bacon manufacturing process chain, from incoming breasts to outgoing packaging units, which can be used to evaluate the impact of each step on the contamination. Following the principles of Nauta's MPRM⁽¹⁾, we adapted the model to the process under investigation, and to the specificity of the product, *i.e.* its solidity and heterogeneity characteristics. In particular, bacteria were assumed to be present only on the surface of the breasts, and to have higher affinity (and therefore higher likelihood of being present) for the lean areas than for the fatty ones. This meant that specific mixing and partitioning models had to be developed. Both the masses and areas of the items were considered, and we switched between them when necessary. To calibrate the model, parameter estimation was carried out from data provided by partner business operators, from data provided by IFIP (French pork institute), from the literature, and based on expert opinion. By means of simulations, we demonstrated the potential of the model and its flexibility. Simulations were performed using both a baseline calibration and alternative scenarios, in order to assess the impact of changes in the process and of accidental events. Results are reported as contamination distributions, and as the probability of acceptance with regards to the regulatory safety criterion.

We start by giving an overview of the model, and then detail how each step is modeled. The simulation design is then described, and some results are presented. In the last part, the simulation results, the assumptions underlying the model, and the prospects for the use of the model are discussed.

F.2 The model

F.2.1 Overview of the model

The model built to study the behavior of bacteria during the process chain of the manufacture of diced bacon has been devised as a sequence of several steps : (1) *arrival* (A) corresponds to the delivery of the raw material, which consists of boned and pared pork breasts, (2) *brining-and-tumbling* (BT) (around 8 hours), during which the breasts are put in a tumbler where they absorb some brine, (3) *steaming* (S) (around 2.5 hours) which contributes to producing the sensorial qualities of diced bacon, (4) *dicing* (D) during which the processed product is shaped, and (5) *packaging* (P) when an appropriate number of dice are combined in units suitable for retailing. In the real process chain, storage steps occur between these five steps. However, based on reasonable assumptions concerning the physico-chemical conditions and duration of the different storage steps, the bacterial contamination is not expected to change significantly during these steps (temperature below 4 ° C). Consequently we have neglected the storage steps in this study, and focused on steps (1) to (5).

The observational unit changes during the process chain, so we successively followed the contamination of the whole breasts, the diced meat, and then the packaging units. For practical purposes, we determined the contamination in Colony Forming Units (CFU), but we also modeled the masses of the units, so that the results could also be expressed as a concentration.

In this manuscript, \log stands for natural logarithm, in contrast to \log_{10} . We denoted the contamination c (in CFU), and the mass m (in g) of the units at the end of each step using the first letter of the step name as a superscript (c^A and m^A for the breasts when they arrived, c^{BT} and m^{BT} for the breasts during the *brining-and-tumbling*, etc). When needed, we also introduced the contamination levels and masses at the beginning of the step by adding a b before the initial (for instance, c^{bBT} for the concentration at the beginning of *brining-and-tumbling*).

These variates (c and m) were assumed to depend only on the state of the units at the end of the previous step, and on the parameters modeled. We built our model as a Bayesian network⁽¹⁵⁾, based on conditional independence assumptions. A schematic diagram depicting the whole process is shown in Figure 1. The quantities involved in the model are represented as nodes. The links between nodes show direct dependencies, and can be either deterministic or stochastic. The nodes without incoming arrows are the parameters of the model. The different parts of the Bayesian network are detailed below.

For the contamination, we considered both *L. monocytogenes* (*Lm*) and lactic acid bacteria (LAB). Except for the contamination at arrival, we adopted similar modeling for *L. monocytogenes* and LAB, using same or different parameter values depending on the step concerned (see Calibration). Below, we use the general term *bacteria* to refer to either *L. monocytogenes* or LAB.

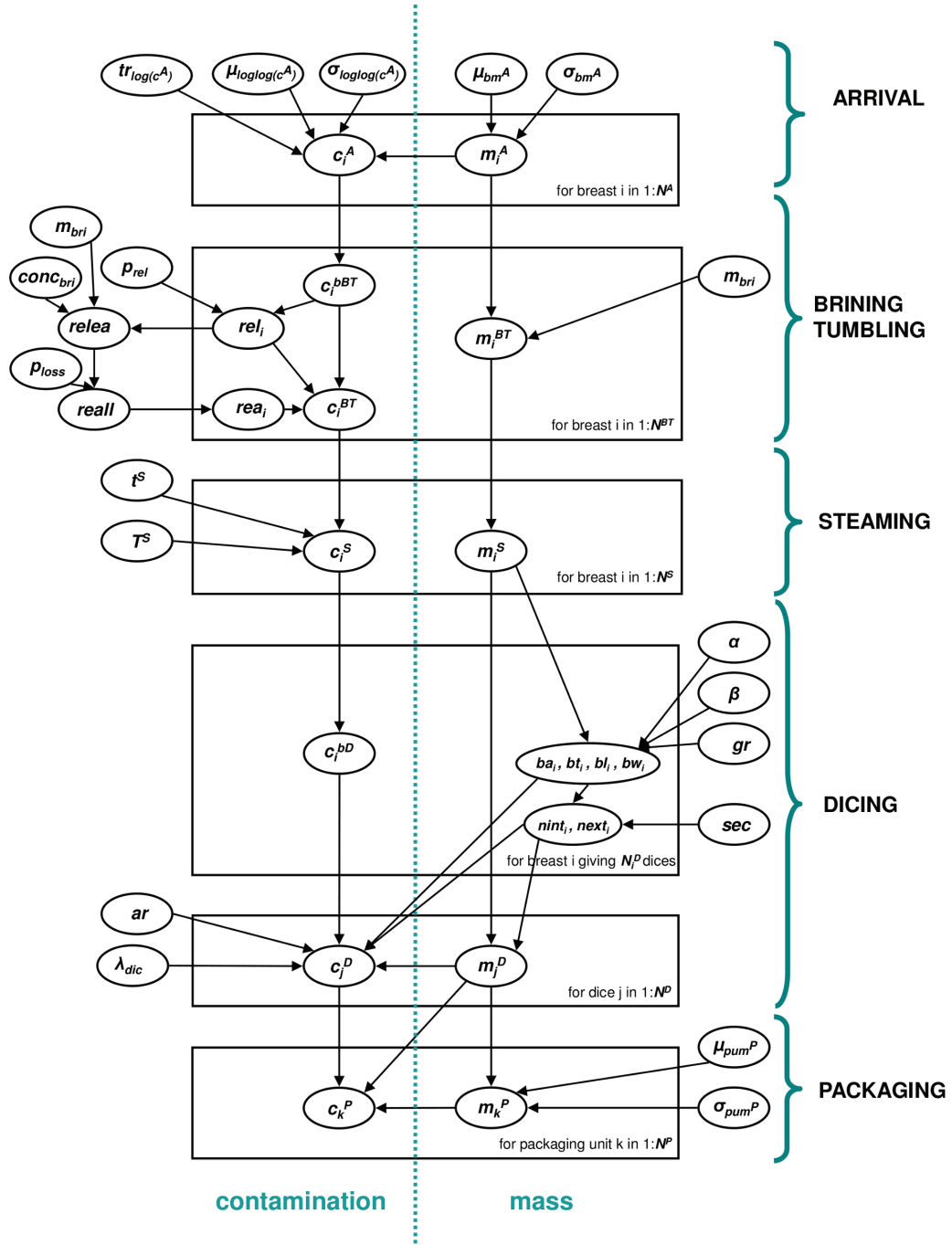


FIGURE F.1 – Diagram of the model describing the diced bacon manufacturing process chain. For the purpose of clarity, microbiological inactivation parameters are not shown.

F.2.2 Step combination

In the diced bacon processing chain, our informal definition of a batch varies at each step : at arrival, a batch consists of the entire set of pork breasts that arrive on a given day. For the brining-and-tumbling step, it is the set of breasts contained in a tumbler ; for steaming, it is the set of breasts contained in one oven ; for dicing it is the set of dice processed during a certain period ; finally for packaging, it is the set of packaging units produced from a dicing batch. The batches for the different steps do not always match ; a batch may sometimes be divided between one step and the next one, and they can also be combined. For instance, a tumbling batch can be divided into several steaming batches, and a dicing batch can include several steaming batches. The step during which the strongest interactions occur between breasts is the brining-and-tumbling step, where the breasts are mechanically mixed with the brine for several hours. During the other steps, slight contact may occur between the breasts (or dice), but we have ignored the corresponding potential for cross-contamination (except *via* the dicing machine). Consequently, we considered the brining-and-tumbling batches to be the most important in terms of a homogenized set, and so we decided to simulate the history and the subsequent fate of these basic batches, without taking into account any combinations that could occur later in the process. In this study, we therefore assumed that the arrival, brining-and-tumbling, and steaming batches did in fact correspond, so that the same number of breasts was included in each step ($N^A = N^{BT} = N^S$). We then deduced the resulting number of dice and packaging units, as described in paragraphs 2.6 and 2.7. Nevertheless, we describe step modeling as a general process, so that the successive batch sizes could be dealt with in a different way ($N^A \neq N^{BT} \neq N^S$).

F.2.3 Arrival

The breasts are considered to be already boned and pared when they are delivered for the manufacture of diced bacon. In this section, we consider a number of breasts in an arrival batch N^A .

F.2.3.a Mass m^A

The weights of the breasts were assumed to be independent and identically-distributed :

$$m_i^A |_{i=1..N^A} \sim \mathcal{Norm}(\mu_{bm}^A, \sigma_{bm}^A) \quad (\text{F.1})$$

where μ_{bm}^A and σ_{bm}^A (g) are the mean and the standard deviation of the breast mass at arrival, respectively.

F.2.3.b Contamination c^A_{Lm}, c^A_{LAB}

Two different modeling approaches were chosen for the contamination at arrival, because the level of *L. monocytogenes* contamination is much lower than that of LAB.

We considered a probability distribution for concentration of *L. monocytogenes*, and we derived a contamination level, *i.e.* a number of CFU for each breast which could be

null or not. The *L. monocytogenes* contamination (CFU) of every breast was assumed to follow a conditional Poisson distribution :

$$c_{i\ Lm}^A|_{i=1..N^A} \sim \mathcal{Pois}(\lambda = \text{conc}_{i\ Lm}^A m_i^A) \quad (\text{F.2})$$

where $\text{conc}_{i\ Lm}^A$ and m_i^A stands for the *L. monocytogenes* concentration (CFU.g⁻¹) and the mass (g) of the breast number i . In this study, the *L. monocytogenes* concentration conc_{Lm}^A (CFU.g⁻¹) was modeled using a log-lognormal distribution. Such a distribution is rarely used to describe contamination, however, it allowed us to have a parametric distribution to simulate the contamination at arrival. This choice was driven by an empirical reasoning in order to adapt the distribution to the results obtained during calibration (see paragraph 3.1).

$$\log(\log(\text{conc}_{i\ Lm}^A) + \text{tr}_{\log(\text{conc}^A)_{Lm}})|_{i=1..N^A} \sim \mathcal{Norm}(m = \mu_{\log \log(\text{conc}^A)_{Lm}}, \text{sd} = \sigma_{\log \log(\text{conc}^A)_{Lm}}) \quad (\text{F.3})$$

where $\mu_{\log \log(\text{conc}^A)_{Lm}}$ and $\sigma_{\log \log(\text{conc}^A)_{Lm}}$ are the mean and the standard deviation of the log-lognormal distribution, and $\text{tr}_{\log(\text{conc}^A)_{Lm}}$ is a translation term necessary to remain in the definition interval of the log function, that is $]0, +\infty[$.

The bacterial concentration of LAB, conc_{LAB}^A (CFU.g⁻¹), was assumed to follow a lognormal distribution :

$$\log_{10}(\text{conc}_{i\ LAB}^A)|_{i=1..N^A} \sim \mathcal{Norm}(m = \mu_{\log_{10}(\text{conc}^A)_{LAB}}, \text{sd} = \sigma_{\log_{10}(\text{conc}^A)_{LAB}}) \quad (\text{F.4})$$

where $\mu_{\log_{10}(\text{conc}^A)_{LAB}}$ and $\sigma_{\log_{10}(\text{conc}^A)_{LAB}}$ are the mean and the standard deviation of the \log_{10} of the LAB concentration. Finally, we deduced the contamination (CFU) by multiplying the concentrations (CFU.g⁻¹) by the masses (g).

$$c_{i\ LAB}^A|_{i=1..N^A} = \text{round}(\text{conc}_{i\ LAB}^A m_i^A) \quad (\text{F.5})$$

where $\text{round}()$ stands for the closest integer. Note that Eq (5) is only applicable to LAB.

F.2.4 Brining-and-tumbling

In this section, we consider a number of breasts in a tumbling batch, N^{BT} .

F.2.4.a Mass m^{BT}

During the brining-and-tumbling step, the breasts absorb brine, so their mass increases slightly.

$$m_i^{BT}|_{i=1..N^{BT}} = m_i^A + m_{bri} m_i^A \quad (\text{F.6})$$

where m_{bri} corresponds to the addition of brine (g.g⁻¹ of meat), which is assumed to be the same for all the breasts.

F.2.4.b Contamination c^{BT}

The breasts in a given tumbling batch are assumed to exchange bacteria during the tumbling step, so that contamination is partly homogenized. In this step, we describe the behavior of the bacteria on the N^{BT} breasts that make up a tumbling batch, and we broke up the process as follows. We successively considered the number of bacteria released from each breast, rel , the total number of bacteria released in a tumbler, $relea$, the total number of reallocated bacteria, $reall$, and the number of bacteria reallocated to each breast, rea .

Let c_i^{bBT} be the number of CFU on breast i at the beginning of the brining-and-tumbling step. A part rel_i (CFU) of the bacteria is assumed to be released from breast i during the brining-and-tumbling step, following :

$$rel_{i|i=1..N^{BT}} \sim \mathcal{Binom}(size = c_i^{bBT}, prob = p_{rel}) \quad (\text{F.7})$$

where p_{rel} is the probability that each bacterium will be released.

Also we took into account a potential contamination of the brine added, denoted c_{bri} (CFU).

$$c_{bri} = \sum_{i=1..N^{BT}} conc_{bri} m_{bri} m_i^A \quad (\text{F.8})$$

where $conc_{bri}$ is the bacterial concentration in the brine (CFU.g^{-1}), m_{bri} is the brine added (g.g^{-1} of meat) and m_i^A the mass (g) of breast number i at the beginning of the brining-and-tumbling step.

In the tumbler, the total number of released bacteria, denoted $relea$ (CFU), equals :

$$relea = \sum_{i=1..N^{BT}} rel_i + c_{bri} \quad (\text{F.9})$$

Some bacteria are also assumed to be lost as a result of remaining on the tumbler surface, whereas it is assumed that the others are reallocated onto the breasts that make up the tumbling batch. Let $reall$ (CFU) be the total number of bacteria reallocated during the tumbling step :

$$reall \sim \mathcal{Binom}(size = relea, prob = 1 - p_{loss}) \quad (\text{F.10})$$

where p_{loss} is the probability that each bacterium will remain on the tumbler surface.

Finally, the total number of reallocated bacteria $reall$ was distributed to the breasts using a multinomial distribution, in order to ensure the conservation of bacteria. Every breast was assumed to be an equivalent receptor, with equal allocation probability. Let rea_i (CFU) be the number of bacteria reallocated to breast i :

$$(rea_1, rea_2, \dots, rea_i, \dots, rea_{N^{BT}}) \sim \mathcal{Multinom}(size = reall, prob = rep(\frac{1}{N^{BT}}, N^{BT})) \quad (\text{F.11})$$

where $rep(\frac{1}{N^{BT}}, N^{BT})$ is a vector of N^{BT} elements, all of which being equal to $\frac{1}{N^{BT}}$.

The contamination c^{BT} (CFU) at the end of the brining-and-tumbling step finally corresponds to the initial contamination minus the CFU released plus the CFU reallocated.

$$c_i^{BT}|_{i=1..N^{BT}} = c_i^{bBT} - rel_i + rea_i \quad (\text{F.12})$$

F.2.5 Steaming

In this section, we consider a number of breasts of a steaming batch, N^S .

F.2.5.a Mass m^S

During the steaming step, the air humidity is controlled so that the breasts do not lose weight. The breast masses were therefore assumed to remain unchanged during the steaming step.

$$m_i^S|_{i=1..N^S} = m_i^{BT} \quad (\text{F.13})$$

F.2.5.b Contamination c^S

The steaming step can sometimes reduce the contamination of the meat ; nevertheless this is not a high-temperature decontamination step. As a primary model, we used the deterministic Chick's model⁽¹⁶⁾, with first order kinetics.

$$c_i^S|_{i=1..N^S} = c_i^{bS} \exp(-k_i t^S) \quad (\text{F.14})$$

where c_i^{bS} and c_i^S are the contamination (CFU) on breast i at the beginning and end of the steaming step, respectively, t^S is the duration of steaming (h), and k_i is the reaction speed constant (h^{-1}) for breast i . To link k to the steaming temperature T^S that we assumed to be constant, we used the Bigelow's secondary model⁽¹⁷⁾ :

$$k_i(T^S)|_{i=1..N^S} = \frac{2.303}{D_{ref_i}} \exp\left(\frac{2.303}{z_i}(T^S - T_{ref})\right) \quad (\text{F.15})$$

where D_{ref_i} is the time for a decimal reduction at the temperature T_{ref} for the breast i , and z_i is the increase in temperature resulting in a 10-fold reduction in the time to first decimal reduction D_{ref_i} , with :

$$D_{ref_i}|_{i=1..N^S} \sim \text{Norm}(mean = \mu_{D_{ref}}, sd = \sigma_{D_{ref}}) \quad (\text{F.16})$$

where $\mu_{D_{ref}}$ and $\sigma_{D_{ref}}$ are the mean and standard deviation of D_{ref} (min), and :

$$z_i|_{i=1..N^S} \sim \text{Norm}(mean = \mu_z, sd = \sigma_z) \quad (\text{F.17})$$

where μ_z and σ_z are the mean and standard deviation of z ($^{\circ}\text{C}$).

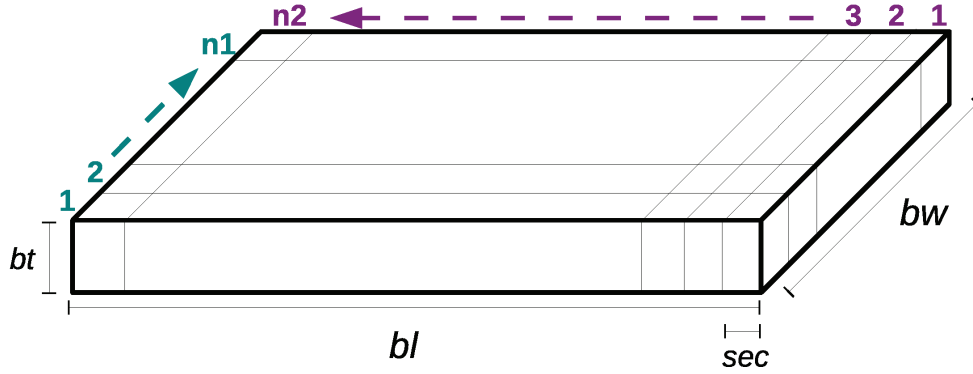


FIGURE F.2 – Breast geometry and dicing scheme.

F.2.6 Dicing

In this section, we consider a number of dice in a dicing batch, N^D .

For the dicing step, we had to estimate how many dice would result from each breast. Consequently we began with geometrical considerations, and adopted the following approximations : (1) the breasts are cuboid in shape (see Figure 2) with proportional dimensions (they all have the same scale factor between length, width and thickness), (2) their mass is proportional to their volume. For every breast i at the dicing step, we deduced its area ba_i^D (only one main side of the cuboid), and its thickness bt_i^D from its mass bm_i^D , using :

$$ba_i^D = \alpha (bm_i^D)^{2/3} \quad \text{and} \quad bt_i^D = \beta (bm_i^D)^{1/3} \quad (\text{F.18})$$

where α ($\text{cm}^2 \cdot \text{g}^{-2/3}$) is a parameter required to go from mass to area and β ($\text{cm} \cdot \text{g}^{-1/3}$) a parameter required to go from mass to thickness.

We also deduced the breast length bl_i^D (cm) and its width bw_i^D (cm) from its area ba_i^D (cm^2), using :

$$bl_i^D = \sqrt{\frac{ba_i^D}{gr}} \quad \text{and} \quad bw_i^D = gr bl_i^D \quad (\text{F.19})$$

where gr is the geometrical ratio between the lengths and the widths of the breasts.

Given these dimensions, we calculated the number of dice N_i^D manufactured from each breast :

$$N_i^D = n_{1i} n_{2i} = \text{round}\left(\frac{bl_i^D}{sec}\right) \text{round}\left(\frac{bw_i^D}{sec}\right) \quad (\text{F.20})$$

where n_1 and n_2 are the rounded number of dice cut in the length and in the width of the breast (see Figure 2), and sec is the dice cross-section (cm). $N^D = \sum_{i=1..Ns} N_i^D$. The thickness is not sliced and remains equal (see Figure 2).

F.2.6.a Mass m^D

Following the assumption of the geometrical regularity of the breasts, we deduced the mass (g) of dice j from that of its original breast i :

$$m_{ji}^D = \frac{m_i^S}{N_i^D} \quad (\text{F.21})$$

where the subscript ji specifies dice j coming from breast i . All the dice originating from a given breast are assumed to have the same mass.

F.2.6.b Contamination c^D

Because the bacteria are more likely to be located on the surface of the breasts, we distinguished two kinds of dice : exterior dice, which are more contaminated because they have a larger exterior surface area, and interior dice, which are less contaminated because they have a smaller exterior area exposed to the environment. The number of exterior and interior dice are $next_i = 2n_{1i} + 2n_{2i} - 4$ and $nint_i = (n_{1i} - 2)(n_{2i} - 2)$, respectively.

The bacteria were assumed to be independently distributed on the breast surface. The bacteria on each breast were shared between the resulting dice by a multinomial drawing, with the allocation probability proportional to the exterior area, weighted by an affinity ratio, ar . Indeed, bacteria are known to have a lower affinity for fatty areas than for lean areas, and we considered that one of the main sides of the breast was fatty. Under these assumptions, the allocation probability was $sec_i^2(1 + ar) + sec_i bt_i$ for the $next_i$ exterior dice, and $sec_i^2(1 + ar)$ for the $nint_i$ interior dice, respectively. Thus, the contamination of the dice (CFU) is :

$$c_1^D, \dots, c_{ji}^D, \dots, c_{ND}^D \sim \text{Multinom}(\text{size} = c_i^{bD}, \text{prob} = \text{rep}(sec_i^2(1 + ar) + sec_i bt_i, next_i), \text{rep}(sec_i^2(1 + ar), nint_i)) \quad (\text{F.22})$$

where c_i^{bD} (CFU) is the contamination of breast i at the beginning of the dicing step (in our study, this is assumed to be equal to the contamination at the end of the steaming step). $\text{rep}(sec_i^2(1 + ar) + sec_i bt_i, next_i)$ is a vector of $next_i$ elements all equal to $sec_i^2(1 + ar) + sec_i bt_i$ and $\text{rep}(sec_i^2(1 + ar), nint_i)$ is a vector of $nint_i$ elements all equal to $sec_i^2(1 + ar)$. The subscript ji specifies a dice j from breast i .

We also assumed that additional *L. monocytogenes* contamination could come from the dicing machine. To allow for this, we considered a stock of *L. monocytogenes* (assumed to be permanently renewed by growth) on the blades or other places in the dicing machine. We also assumed that each dice would pick up c_{dic} bacteria (CFU), with :

$$c_{dic} \sim \text{Poisson}(\lambda = \lambda_{dic}) \quad (\text{F.23})$$

where λ_{dic} (CFU) is the parameter of the Poisson distribution. Finally we added c_{dic} to the dice contamination c^D (CFU). We ignored the possibility of a transfer in the opposite direction, *i.e.* from the dice to the dicing machine.

F.2.7 Packaging

In this section, we consider a number of packaging units in a packaging batch, N^P .

F.2.7.a Mass m^P

The masses (g) of the packaging units were assumed to follow a normal distribution :

$$m_k^P |_{k=1..N^P} \sim \mathcal{Norm}(\text{mean} = \mu_{pum}^P, \text{sd} = \sigma_{pum}^P) \quad (\text{F.24})$$

where μ_{pum}^P and σ_{pum}^P are the mean and the standard deviation of the packaging unit mass (g).

F.2.7.b Contamination c^P

For every packaging unit, we added dice until its requisite mass was reached, and we cumulated the contaminations of the dice (CFU). We then filled another packaging unit with other dice, and so on until no dice was left. After the last piece of diced bacon from a given breast has been packed, the next breast is used to continue to fill the same packaging unit. The dice were sorted by breast of origin but were randomly considered within each breast of origin. So we included n_k dice (deduced from the algorithm) in each packaging unit k , and we obtained N^P packaging units.

$$c_k^P |_{k=1..N^P} = \sum_{jk=1..n_k} c_{jk}^D \quad (\text{F.25})$$

where the subscript jk specifies a dice j going into packaging unit k .

F.3 Baseline calibration

The baseline parameter values are summarized in Table 1.

F.3.1 Arrival (μ_{bm}^A , σ_{bm}^A , $\mu_{\log \log(\text{conc}^A)}_{Lm}$, $\sigma_{\log \log(\text{conc}^A)}_{Lm}$, $tr_{\log(\text{conc}^A)}_{Lm}$, $\mu_{\log_{10}(\text{conc}^A)}_{LAB}$ and $\sigma_{\log_{10}(\text{conc}^A)}_{LAB}$)

According to the data provided by IFIP (industry data), the mean and the standard deviation of the mass of boned and pared breast are $\mu_{bm}^A = 3033$ g and $\sigma_{bm}^A = 274$ g. These values correspond to industrial requirements.

To estimate the *L. monocytogenes* contamination at the beginning of the process, we pooled the surveillance data from four partner business operators (own-checks). This gave us $n = 1994$ detection data dd , corresponding to 25 g test portions. The details of our reasoning are as follows.

The detailed sampling protocols were not available, so a random test portion area sa (cm²) of lean surface was attributed to every 25 g portion. We reasoned in terms of areas, because bacteria were assumed to occur only on the surface. On the basis of the geometrical dimension means of the breasts, we estimated the total breast area equivalent

TABLE F.1 – Baseline calibration of the parameters. *Lm* and *LAB* stands for *L. monocytogenes* and lactic acid bacteria, respectively. The symbol - means dimensionless.

Symbol	Dimension	Interpretation	Value	Source of information
ARRIVAL				
μ_{bm}^A	g	breast mass mean	3,033	IFIP
σ_{bm}^A	g	breast mass std. dev.	274	IFIP
$\mu_{\log \log(\text{conc}^A)_{Lm}}$	$\log \log(\text{CFU.g}^{-1})$	<i>Lm</i> loglog concentration mean	-0.614	Company own data
$\sigma_{\log \log(\text{conc}^A)_{Lm}}$	$\log \log(\text{CFU.g}^{-1})$	<i>Lm</i> loglog concentration std. dev.	0.451	Company own data
$tr_{\log(\text{conc}^A)_{Lm}}$	$\log(\text{CFU.g}^{-1})$	<i>Lm</i> log translation	7.567	Company own data
$\mu_{\log_{10}(\text{conc}^A)_{LAB}}$	$\log_{10}(\text{CFU.g}^{-1})$	<i>LAB</i> log concentration mean	2	Expert's opinions
$\sigma_{\log_{10}(\text{conc}^A)_{LAB}}$	$\log_{10}(\text{CFU.g}^{-1})$	<i>LAB</i> log concentration std. dev.	0.25	Expert's opinions
BRINING AND TUMBLING				
N^{BT}	#	number of breasts per batch	1000	Company own data
m_{bri}	g.g^{-1} (of meat)	brine mass added	0.150	Company own data
conc_{bri}	CFU.g^{-1}	brine contamination	0	Expert's opinions
p_{rel}	-	release probability	0.5	Expert's opinions
p_{loss}	-	loss probability	0.01	Expert's opinions
STEAMING				
T^S	$^{\circ}\text{C}$	steaming temperature	45	Company own data
t^S	h	steaming duration	1.5	Company own data
$\mu_{D_{ref}}$	min	mean of the time to the first decimal reduction at $T_{ref} = 60^{\circ}\text{C}$	3.779	ICMSF
$\sigma_{D_{ref}}$	min	std. dev. of the time to the first decimal reduction at $T_{ref} = 60^{\circ}\text{C}$	0.319	ICMSF
μ_z	$^{\circ}\text{C}$	mean of the increase in temperature resulting in a 10-fold reduction in D_{ref}	7.702	ICMSF
σ_z	$^{\circ}\text{C}$	std. dev. of the increase in temperature resulting in a 10-fold reduction in D_{ref}	0.416	ICMSF
DICING				
α	$\text{cm}^2.\text{g}^{-2/3}$	from mass to area	5.531	IFIP
β	$\text{cm.g}^{-1/3}$	from mass to thickness	0.198	IFIP
gr	-	width/length geometrical ratio	0.4	IFIP
ar	-	fat/lean affinity ratio	0.5	Expert's opinions
sec	cm	dice cross-section	0.8	Company own data
λ_{dic}	CFU	cross-contamination parameter	0.0001	Expert's opinions
PACKAGING				
μ_{pum}^P	g	packaging unit mass mean	98	Specific data
σ_{pum}^P	g	packaging unit mass std. dev.	4	Specific data

to lean area as 2382 cm². Actually one of the two main sides was considered to be fatty and consequently its area was weighted by the fat/lean affinity ratio, ar (see the dicing baseline calibration for details). We thus deduced the breast proportion $bprop$ corresponding to every test portion by :

$$bprop_l | sa_l |_{l=1..n} = \frac{sa_l}{2382} \quad (\text{F.26})$$

Let $sprev$ denote the prevalence in 25 g test portions, and $bprev$ denote the prevalence in the breasts.

Assuming a Poisson distribution of the bacteria on the breast surface, the probability that a test portion of sa cm² will be uncontaminated is equal to $1 - spreval = \exp(-\gamma * sa)$, with γ the mean concentration in 1 cm². The probability that a breast of 2382 cm² will be uncontaminated is $1 - bprev = \exp(-\gamma * 2382)$. So the two latter probabilities are related as follows :

$$(1 - spreval_l) | bprev_l, bprop_l |_{l=1..n} = (1 - bprev_l)^{bprop_l} \quad (\text{F.27})$$

Then, we linked the observed detection data, dd , to the prevalence in 25 g test portions, $sprev$:

$$dd_l | spreval_l |_{l=1..n} \sim \text{Bern}(prob = spreval_l) \quad (\text{F.28})$$

Finally, as the *L. monocytogenes* contamination at arrival, c^A , is assumed to follow a Poisson distribution of parameter $\lambda = conc^A m^A$, the probability of obtaining 0 CFU is $\exp(-conc^A m^A)$. We identified the probability of obtaining 0 CFU with $1 - bprev$ and m to $\mu_{bm}^A = 3033$ g to deduce the concentration, $conc^A$ (in CFU.g⁻¹) :

$$conc^A | bprev_l |_{l=1..n} = -\frac{\log(1 - bprev_l)}{3033} \quad (\text{F.29})$$

The following distributions were used as priors. The test portion area, sa (cm²), was based on expert opinion and specific tests.

$$bprev \sim \text{Unif}(a = 0, b = 1) \quad (\text{F.30})$$

$$sa \sim \text{Unif}(a = 3, b = 40) \quad (\text{F.31})$$

where a and b are the distribution bounds.

By means of Monte Carlo Markov Chains (MCMC) we performed Bayesian inference using WinBugs⁽¹⁸⁾, and we thus obtained a $conc^A$ simulated posterior distribution without convergence difficulty. A double log transformation of $conc^A$ simulated posterior distribution was necessary to correctly fit a closed distribution, as shown in Figure 3. The parameter estimates are $\mu_{\log \log(conc^A)_{Lm}} = -0.614$, $\sigma_{\log \log(conc^A)_{Lm}} = 0.541$ and $tr_{\log(conc^A)_{Lm}} = 7.567$.

In the case of LAB, on the basis of expert opinion, we fixed $\mu_{\log_{10}(conc^A)_{LAB}} = 2$ and $\sigma_{\log_{10}(conc^A)_{LAB}} = 0.25$.

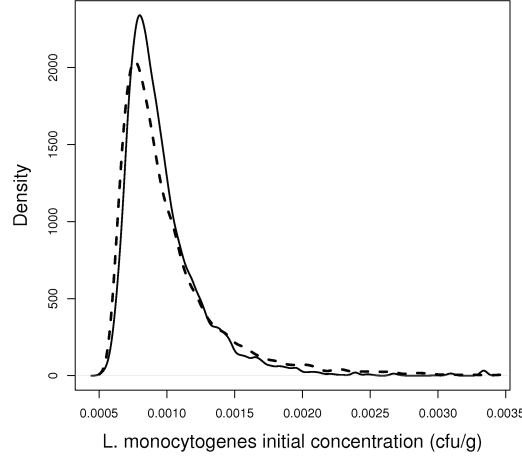


FIGURE F.3 – $conc^A_{Lm}$ (CFU.g⁻¹) simulated posterior distribution (solid line) and fitted posterior lognormal distribution (long-dashed line).

F.3.2 Brining and tumbling (N^{BT} , m_{bri} , $conc_{bri}$, p_{rel} and p_{loss})

According to the practices of our partner business operator, the carrying capacities of the tumblers N^{BT} , and the amount of brine added m_{bri} , are around 1000 breasts (3 tons) and 0.150 g.g⁻¹ of meat, respectively. We used these values as baseline levels, and we assumed that the brine was not contaminated : $conc_{bri} = 0$ CFU.g⁻¹.

With regards to the tumbler, the probability of release p_{rel} and of loss p_{loss} are very difficult to calibrate, and so far no relevant data is available. According to expert opinion, the release of bacteria from the breast can be expected to be a frequent event, because the breasts are mechanically shaken for several hours. In contrast, the loss of bacteria that remain on the tumbler surface is expected to be a rare event, because the bacteria are expected to have a low affinity with the fatty slime that remains in the tumbler. We took $p_{rel} = 0.7$ and $p_{loss} = 0.01$ as the baseline values, but also tried assigning other values to these two parameters (see Sensitivity analysis).

F.3.3 Steaming (t^S , T^S , $\mu_{D_{ref}}$, $\sigma_{D_{ref}}$, μ_z and σ_z)

According to data collected from partner business operators, the duration and air temperature during the steaming step are around 2.5 hours and 50 °C, respectively. However, to account for the thermal inertia of the breasts, we discounted the first hour of steaming, which corresponded to the heat-up phase. Moreover, the outer surface of the breast was expected to be cooler than the surrounding air. We therefore finally assumed as baseline that the classical breast steaming step corresponded to exposing the bacteria to 45 °C for 1.5 hours.

To estimate the inactivation parameters D_{ref} (min) and z (°C), we made use of data pooled by ICMFS⁽¹⁹⁾, corresponding to the inactivation of *L. monocytogenes* on various meat matrices. We used 63 pairs of observed D_{ref} values (denoted D_{obs}) at observed T_{ref} values (denoted T_{obs}) ($m = 14$ different T_{obs}). We chose $T_{ref} = 60$ °C as the reference

temperature to estimate D_{ref} and z , and we carried out a Bayesian statistical analysis of these data, using the following conditional links :

$$\log(\mu_D)_m | T_{ref}, D_{ref}, z, T_{obsm} = \log(D_{ref}) + \frac{T_{ref} - T_{obsm}}{z} \quad (\text{F.32})$$

$$\log(D_{obs})_m | \mu_{D_m}, \sigma_D \sim \mathcal{N}orm(\text{mean} = \mu_{D_m}, \text{sd} = \sigma_D) \quad (\text{F.33})$$

where μ_{D_m} corresponds to a theoretical D value at the temperature T_{obsm} . We based on the following prior distribution⁽²⁰⁾ :

$$\log(D_{ref}) \sim \mathcal{N}orm(\text{mean} = 3, \text{sd} = 1) \quad (\text{F.34})$$

$$z \sim \mathcal{N}orm(\text{mean} = 7.55, \text{sd} = 1) \quad (\text{F.35})$$

$$\sigma_D^{-2} \sim \mathcal{G}amma(\text{scale} = 0.001, \text{shape} = 0.001) \quad (\text{F.36})$$

By means of Monte Carlo Markov Chains (MCMC) we performed Bayesian inference using WinBugs⁽¹⁸⁾, and we obtained D_{ref} and z simulated posterior distributions without convergence difficulty. We then fitted normal distributions to these simulated posterior distributions in order to get closed distributions for D_{ref} and z . Figure 4 shows the superimposition of prior, simulated posterior and fitted posterior distributions.

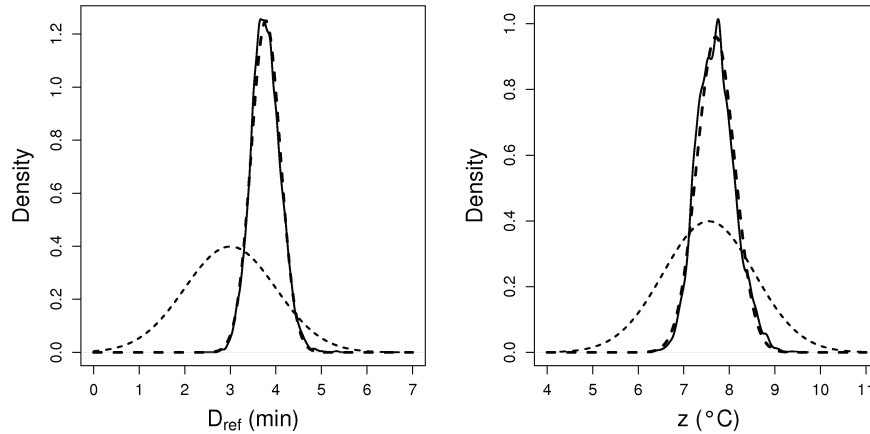


FIGURE F.4 – D_{ref} (left) and z (right) prior distributions (short-dashed line), simulated posterior distributions (solid line), and fitted posterior normal distributions (long-dashed line).

We also simulated probability distributions of predicted values D_{pred} for each T_{obs} , and we compared them to the observed values D_{obs} . The predictions and the observations were in agreement, and consequently we validated this Bayesian analysis. The parameter estimates are : $\mu_{D_{ref}} = 3.779$ min, $\sigma_{D_{ref}} = 0.319$ min, $\mu_z = 7.702$ ° C and $\sigma_z = 0.416$ ° C.

On the basis of expert opinion, we used the same values for LAB.

F.3.4 Dicing (α , β , gr , ar , sec , λ_{dic})

The mean area of one main side of the breast μ_{ba}^D equals 1,272 cm², the mean breast thickness μ_{bd}^D equals 3 cm, and the mean breast mass μ_{bm}^D equals 3488 g after it has absorbed 0.150 g of brine per g of meat (industry data provided by IFIP). On the basis of these means, we estimated α (in cm² g^{-2/3}) and β (in cm g^{-1/3}) as :

$$\alpha = \mu_{ba}^D \mu_{bm}^D^{-2/3} \simeq 5.531 \quad \beta = \mu_{bd}^D \mu_{bm}^D^{-1/3} \simeq 0.198$$

We fixed the width/length geometrical ratio to its mean value $gr = \frac{2}{5}$ (industry data provided by IFIP), and we fixed the fat/lean affinity ratio as $ar = \frac{1}{2}$, on the basis of expert opinion. The dice cross-section sec is a process parameter that we fixed at 0.8 cm for the baseline calibration.

The contamination of the dice by the dicing machine is not thought to be an unusual event. Indeed, *L. monocytogenes* is known readily to form biofilm structures, which tend to facilitate the detachment of microbial cells^(21,22). We fixed $\lambda_{dic} = 0.0001$ CFU as the baseline value, but we also tried other values for this parameter (see Model exploration).

F.3.5 Packaging (μ_{pum}^P and σ_{pum}^P)

Like the dice cross-section, the packaging can vary depending on customer specifications. In this study we chose 100 g packaging units as the baseline. $\mu_{pum}^P = 98$ g and $\sigma_{pum}^P = 4$ g were estimated from specific data acquired for this study. Notice that our packaging algorithm by definition simulated packaging units that exceeded the packaging unit mass m^P .

F.4 Design of simulations

All the simulations were performed using R⁽²³⁾.

F.4.1 Sensitivity analysis towards uncontrolled process parameters

Among all the parameters related to the manufacturing process of diced bacon, we identified three parameters that we considered as uncontrolled by the manufacturer : the probability of release and loss in the tumbler, p_{rel} and p_{loss} , and cross-contamination from the dicing machine, λ_{dic} . When we describe these parameters as being uncontrolled or beyond the manufacturer's control we mean that they are very difficult to measure or to estimate, and so they seem to be impossible for the manufacturer to control, and their calibration was only based on expert opinion.

In this sensitivity analysis, our objective was to investigate the impact of these three uncontrolled parameters, and to check whether the lack of knowledge about their value was critical or not for a sensible use of the model. We therefore tested three values (low, baseline and high) of each parameter ; these values are shown in Table 2. To do this, we carried out a complete sampling plan ($3*3*3=27$ parameter sets). The other parameters

TABLE F.2 – Values tested in the sensitivity analysis towards uncontrolled process parameters.

Parameter	Low	Baseline	High
p_{rel}	0.5	0.7	0.9
p_{loss}	0.001	0.01	0.1
λ_{dic}	0.00001	0.0001	0.001

were fixed at their baseline value (see Table I). For each parameter set $(p_{rel}, p_{loss}, \lambda_{dic})$, 2,500 simulations were performed.

F.4.2 Adapting the model to various processes

In contrast to the uncontrolled parameters, we considered that all the other process parameters were directly controllable by the manufacturer. For these controllable process parameters, we tested possible values other than the ones fixed as baseline, in order to exemplify the adaptability of our model to various diced bacon processes. The following cases were tested. In every case, 2,500 simulations were performed.

1. (P1) 5 mm dice cross-section : $sec = 0.5$ cm (in this case the product is called *allumettes*)
2. (P2) 200 g packaging unit : $\mu_{pum}^P = 198$ g
3. (P3) steaming for 1.5h at 50 ° C : $T^S = 50$ ° C
4. (P4) steaming such that no inactivation can occur

F.4.3 Alternative scenarios

In addition to the changes in the process mentioned above, we also investigated alternative scenarios corresponding to accidental events. The following scenarios were tested. For every scenario, 2,500 simulations were performed.

1. (S1) initial *L. monocytogenes* contamination ten times higher than baseline^(*) : $tr_{\log(conc^A)_{Lm}} = 5.264 \log(\text{CFU.g}^{-1})$
2. (S2) initial *L. monocytogenes* contamination one hundred times higher than baseline^(*) : $tr_{\log(conc^A)_{Lm}} = 2.962 \log(\text{CFU.g}^{-1})$
3. (S3) initial *L. monocytogenes* contamination ten times higher than baseline for a part of the tumbling batch (250 breasts over 1000)
4. (S4) initial *L. monocytogenes* contamination one hundred times higher than baseline for a part of the tumbling batch (250 breasts over 1000)
5. (S5) Brine contaminated by *L. monocytogenes* at $conc_{bri} = 0.1 \text{ CFU.g}^{-1}$

Scenarios (S3) and (S4) correspond to the situation when two arrival batches are combined to form a single tumbling batch, as is often the case.

(*) To multiply the initial *L. monocytogenes* contamination, we used the property :

$$\log(\log(\gamma \text{conc}^A_{Lm}) + tr_{\log(\text{conc}^A)_{Lm}}) = \log(\log(\text{conc}^A_{Lm}) + \log(\gamma) + tr_{\log(\text{conc}^A)_{Lm}}) \quad (\text{F.37})$$

Thus we adjusted the $tr_{\log(\text{conc}^A)_{Lm}}$ value as the baseline value plus $\log(\gamma)$, γ being the multiplying factor required.

F.5 Results

To exemplify the results obtained from one simulation (follow-up of one brining-and-tumbling batch), the change in the level of contamination during the process is shown by means of histograms. In contrast, the results of 2,500 simulations are represented by cumulative probability distributions. Notice that in the first three steps (arrival, brining-and-tumbling, steaming), the items considered consisted of 1000 breasts, then the resulting dice (dicing) and finally packaging units (packaging). For *L. monocytogenes*, the contamination is expressed both in CFU.g⁻¹ and in number of CFU per item.

F.5.1 Baseline calibration

As an example, Figure 5 shows the results of a simulation performed using the baseline calibration. The contamination distribution after brining-and-tumbling was narrower than after arrival, because during the brining-and-tumbling step, the contamination was homogenized within the breasts. Steaming then reduced the contamination. This simulation predicted that a large proportion of the dice would not be contaminated by *L. monocytogenes*, but that the load of some dice was 1 CFU.g⁻¹. In the packaging units, the prevalence of *L. monocytogenes* was not negligible, however the level of contamination remained low, with a maximum of 0.05 CFU.g⁻¹ in this simulation.

The results of one simulation allows us to comment qualitatively on every step. However, because our model is probabilistic, every simulation is different, and numerous simulations are necessary to describe the baseline model outputs. Every tumbling batch of 1000 breasts yielded 1,986,000 [1,973,000-2,000,000] dice and then 35,270 [34,940-35,640] packaging units.

Figure 6 shows the results of 2,500 simulations in terms of the bacterial load. For each process step, the median simulation results are plotted (solid lines), as well as the 95% extreme simulation results (short-dashed lines). At every step the simulation results were quite similar, indicating a low variability between the simulations. According to the median simulation results, at arrival about 5% of the breasts were uncontaminated by *L. monocytogenes*, 50% of the breasts had no more than 0.001 CFU.g⁻¹ (3 CFU per breast), and 100% of the breasts had no more than about 0.004 CFU.g⁻¹ (12 CFU per breast). After brining-and-tumbling, 100% of the breasts had no more than about 0.003 CFU.g⁻¹ (9 CFU per breast). The steaming step reduced *L. monocytogenes* contamination, so that 50% of the breasts had no more than 0.0005 CFU.g⁻¹ (2 CFU per breast), and 100% of the breasts had no more than about 0.002 CFU.g⁻¹ (6 CFU per breast). On the dice, using the baseline calibration, the maximum contamination observed was 1.2 CFU.g⁻¹ (2 CFU per dice), and 99% of the dice were predicted to be uncontaminated by *L. monocytogenes*.

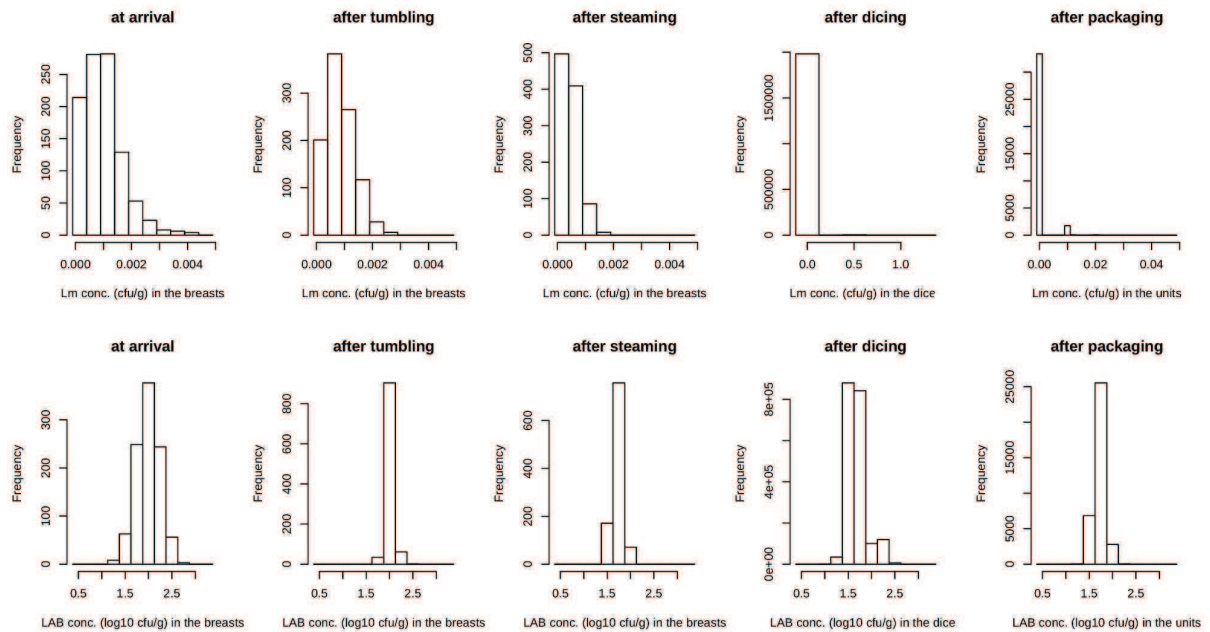


FIGURE F.5 – Example of the results of one simulation. Bacterial concentration after each of the steps considered. First row : *L. monocytogenes*, second row : LAB. First three columns : on breasts, fourth column : on dice, fifth column : in packaging units.

After the packaging step, the maximum contamination observed in the median simulation was 0.02 CFU.g^{-1} (2 CFU per packaging unit), and 95% of the packaging units were predicted to be uncontaminated by *L. monocytogenes*.

In the case of LAB, Figure 6 (second row) shows the same pattern : the overall concentration range was from $[1-3] \log_{10}(\text{CFU.g}^{-1})$ after arrival, $[1.8-2.6] \log_{10}(\text{CFU.g}^{-1})$ after brining-and-tumbling, $[1.2-2.3] \log_{10}(\text{CFU.g}^{-1})$ after steaming, $[0.8-2.8] \log_{10}(\text{CFU.g}^{-1})$ after dicing, and finally $[1.2-2.4] \log_{10}(\text{CFU.g}^{-1})$ in the packaging units. The cumulative distributions of LAB are smoother than those of *L. monocytogenes* because there is no small-number effect in this case. We can suppose that the dip in the cumulative distribution after dicing highlights the distinction between interior and exterior dice.

From the simulations, it is also possible to examine the change in the total number of *L. monocytogenes* bacteria in the 1000 breasts constituting one batch. The results of the 2,500 simulations are shown as boxplots in Figure 7. At arrival, the number of bacteria per batch is in the range $[2,785-3,294] \text{ CFU}$, this number then falls slightly because of loss in the tumbler. The steaming step almost halves the number of bacteria as a result of thermal inactivation. The total contamination of each batch then increases by about 10% as a result of contamination from the dicing machine. Finally the packaging step does not modify the total contamination. At the end of the process, the number of bacteria per batch is in the range $[1,496-1,789] \text{ CFU}$.

Finally, to make our results meaningful in terms of industrial practices, we expressed the results with regards to European regulation (EC) N° 2073/2005 and its safety criterion at the end of the process : no bacteria in 5 out of 5 25-g samples. Thus, for each batch, we simulated 10,000 samplings of five 25-g samples after the last process step, and we

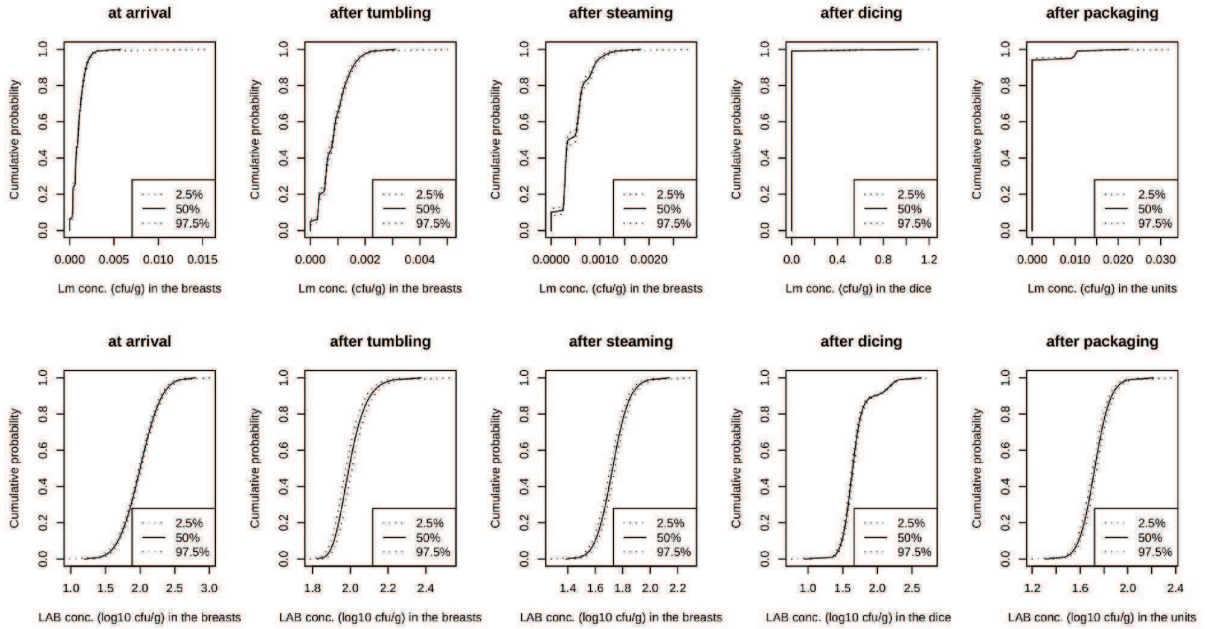


FIGURE F.6 – Cumulative distributions of the bacterial concentration after each step considered. First row : *L. monocytogenes*, second row : LAB. First three columns : on breasts, fourth column : on dice, fifth column : in packaging units. Solid lines correspond to simulation median, dashed lines correspond 95% interval of the simulations. The cumulative distributions are associated with the distribution of items during a given step; the two quantiles assess the variability of the process.

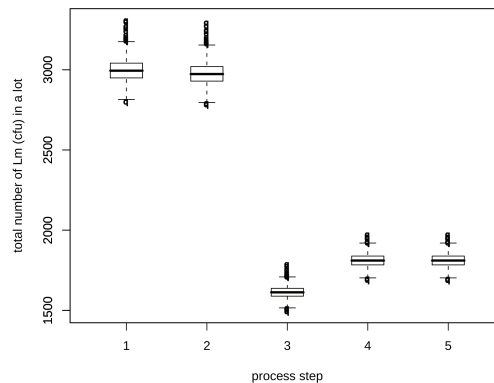


FIGURE F.7 – Total number of *L. monocytogenes* CFU per batch after each considered step : (1) at arrival, (2) after tumbling, (3) after steaming, (4) after dicing, and (5) after packaging. The boxplots display the simulation variability.

calculated the probability of acceptance of a lot as the frequency over 10,000 replicates to obtain 0 CFU in all 5 samples (*i.e.* absence in 25 g). Our modeling and its baseline calibration led to a probability of acceptance in the range [0.9209-0.9404].

F.5.2 Sensitivity analysis towards uncontrolled process parameters

The results were analyzed using regression analysis⁽²⁴⁾. The complete output of a simulation is a probability distribution at each step. Here, we chose to synthesize it by the 99% percentile of the packaging unit contamination that we thought to be one of the most interesting statistics to consider. In order not to be biased by a particular simulation, we used 2,500 independent simulations (each comprising 1000 breasts), and we considered the mean of the corresponding 2,500 values in order to quasi eliminate the uncertainty.

We tested the linear and the quadratic effects of the parameters, as well as their interaction. We distinguished the most significant effects in comparison with a residual variance in which we included the effects of all the terms with a degree of at least 4. The most significant factors (p-value lower than 0.05) were : λ_{dic} (p-value<1E-16), p_{loss} (p-value=4E-13), λ_{dic}^2 (p-value=3E-08), p_{rel} (p-value=2E-07), $p_{rel} * p_{loss}$ (p-value=9E-07) and $p_{loss} * \lambda_{dic}$ (p-value=0.0002). The quadratic effects of the model parameters were not significant (except λ_{dic}^2). The only significant interactions were $p_{loss} * \lambda_{dic}$ and $p_{rel} * p_{loss}$, the two parameters involved in the tumbling modeling. These results indicate that our model is not excessively sensitive to complex effects or to interactions between the uncontrolled parameters in the process.

As expected, the 99% percentile of the packaging unit contamination :

- decreases when the probability of bacterial release in the tumbler p_{rel} increases, since it homogenizes the bacterial load,
- decreases when the probability of bacterial loss in the tumbler p_{loss} increases, since it reduces the bacterial load of the product,
- increases when the contamination from the dicing machine λ_{dic} increases.

The parameter sets leading to the highest and lowest contamination loads are ($p_{rel} = 0.5$, $p_{loss} = 0.001$, $\lambda_{dic} = 0.001$) and ($p_{rel} = 0.9$, $p_{loss} = 0.1$, $\lambda_{dic} = 0.00001$), respectively. The response ratio between these two extreme parameter sets is only 1.036.

F.5.3 Adapting the model to various processes

Some results obtained for the four different processes are summarized in Table 3. In these four cases, we tested changes in the process during the steaming, dicing or packaging steps. This meant that there was no reason for the contamination of the incoming breasts to be modified from the baseline calibration. Regarding the outgoing packaging units, the reduction in dice cross-section (P1) and the doubling of packaging unit weight (P2) reduced the variability of the 95th percentile of the contamination. Indeed, in these cases, the packaging units included more dice than in the baseline calibration, so the contamination was homogenized within the packaging units of a batch. For the baseline and (P1), the 95th percentile of the concentration was about 0.01 CFU.g⁻¹, which corresponded to 1 CFU in 100 g. When the packaging weight was doubled (P2), the bacterial load

was diluted, the 95th percentile of the concentration was about $0.005 \text{ CFU}\cdot\text{g}^{-1}$, which corresponded to 1 CFU in 200 g.

TABLE F.3 – For the baseline calibration and the alternatives tested, *L. monocytogenes* concentration in 10^{-3} CFU.g $^{-1}$: 5th, 50th and 95th percentiles (median [95% extreme] simulation) for the breasts at arrival (process chain inputs), and 90th, 95th and 99th percentiles (median [95% extreme] simulation) for the packaging units (process chain outputs). P1 to P4 correspond to the four different processes that have been tested (see paragraph 4.2). S1 to S5 correspond to the five accidental scenarios that have been tested (see paragraph 4.3).

	INPUTS (BREASTS)			OUTPUTS (PACKAGING UNITS)		
	5th per.	median	95th per.	90th per.	95th per.	99th per.
Baseline	0 [0-0]	0.9142 [0.8547-0.9521]	2.198 [2.068-2.337]	0 [0-0]	9.140 [0-9.460]	10.49 [10.45-10.52]
P1	0 [0-0]	0.9142 [0.8547-0.9521]	2.198 [2.068-2.337]	0 [0-0]	9.735 [9.665-9.794]	10.60 [10.56-10.64]
P2					5.030 [5.022-5.039]	5.189 [5.175-5.204]
P3					0 [0-0]	0 [0-0]
P4					10.04 [10.00-10.07]	10.73 [10.69-10.77]
S1	5.464 [5.222-5.697]	9.053 [8.833-9.275]	16.75 [15.78-17.92]	10.74 [10.69-10.80]	20.03 [19.93-20.14]	29.18 [28.46-29.65]
S2	65.30 [64.02-66.55]	89.15 [87.42-90.86]	161.3 [151.2-173.2]	77.67 [75.19-79.62]	88.65 [86.22-90.92]	111.5 [108.3-115.5]
S3	0 [0-0.2983]	1.165 [1.085-1.240]	11.94 [11.21-12.71]	9.906 [9.867-9.944]	10.32 [10.29-10.35]	19.29 [13.55-19.63]
S4	0 [0-0.2983]	1.165 [1.085-1.240]	114.4 [107.8-121.7]	30.01 [29.63-30.38]	38.69 [33.07-39.66]	53.42 [51.67-58.88]
S5	0 [0-0]	0.9142 [0.8547-0.9521]	2.198 [2.068-2.337]	20.26 [20.22-20.31]	21.43 [21.33-21.55]	32.09 [31.84-32.46]

With scenario (P3) the 99th percentile of the packaging unit contamination was reduced to 0. So an increase of 5°C in the steaming temperature seemed to drastically reduce the prevalence in packaging units. In contrast, for the process in which no inactivation occurs during steaming (P4), the results were close to the baseline level. However, the prevalence in the packaging units was 0.08 [0.08-0.08], *i.e.* higher than the baseline 0.05 [0.04-0.05].

For these four processes, the probabilities of batch acceptance with regard to European regulation (absence for 5 out of 5 25-g samples) are in the ranges : [0.912-0.933] for (P1), [0.920-0.940] for (P2), [0.985-0.994] for (P3) and [0.865-0.895] for (P4).

F.5.4 Alternative scenarios

Some results obtained with the five alternative scenarios are also summarized in Table 3. In scenarios (S1) to (S4), the initial contamination of some or all the incoming breasts was higher than baseline. The median breast contamination was, as expected, increased tenfold for (S1), and 100-fold for (S2). When only some of the breasts were overcontaminated (scenarios (S3) and (S4)), the median of the contamination was not far from the baseline, whereas the 99th percentiles were markedly increased. These scenarios logically led to an increase in the contamination of the outgoing packaging units, as reported in Table 3. The prevalence in the packaging units were : 0.36 [0.35-0.37] for (S1), 0.98 [0.98-0.98] for (S2), 0.14 [0.13-0.14] for (S3) and 0.63 [0.62-0.65] for (S4).

With scenario (S5), where the brine was contaminated at 0.1 CFU.g⁻¹, the contamination of the incoming breasts was not modified. Regarding the high percentiles of the packaging units, the results are of the same order of magnitude as those obtained with scenario (S1). In this case, the prevalence in packaging units is 0.51 [0.52-0.51].

For these five scenarios, all of which are overcontaminated compared to baseline, the probabilities of batch acceptance in accordance with the European regulation (absence for 5 out of 5 25-g samples) are in the ranges : [0.505-0.550] for (S1), [0.0005-0.0035] for (S2), [0.825-0.860] for (S3), [0.28-0.34] for (S4) and [0.335-0.380] for (S5).

F.6 Discussion

F.6.1 What we can learn from the simulations

The sensitivity analysis shows that in the ranges tested, the proposed model is not very sensitive to the probabilities of release and loss in the tumbler, or to cross-contamination from the dicing machine. For all the possible combinations of these three parameters, we observed a difference of less than 4% for the 99th percentile of the packaging unit contamination. As we tested parameter values that are quite representative of the expert's uncertainty, the findings of the sensitivity analysis suggest that this lack of knowledge is not an issue for the proposed modeling of the tumbling and dicing steps.

We already knew that the tumbling-and-brining step is an essential one, since adding brine decreases the water activity of the product, and therefore makes the conditions less suitable for bacterial growth. Our results highlight another phenomenon ; the homogenization of the tumbler content, which appears to be important for the distribution of

the contamination at the end of the process. An interesting point to investigate experimentally would be the breast contamination level after tumbling⁽²⁵⁾. This could lead to a recommendation that the manufacturers should carry out input checking after this step rather than on the raw material, since this would increase the probability of detecting contamination.

From detection data obtained by analyzing packaging units provided by our business partner operators, we estimated the prevalence in 100 g within the 95% confidence interval [0-0.215]. The results obtained with the baseline calibration are consistent with this estimation. Even through the estimated confidence interval is wide, it allows us to make a first attempt to compare the outputs of our model with the reality of the packaging unit contamination. Further investigations are necessary to validate the baseline model with certainty.

F.6.2 Assumptions underlying the model

In the present study, the *L. monocytogenes* initial contamination $conc^A_{Lm}$ is simulated with a log-lognormal distribution, which is unusual. We did that because of the posterior distribution of $conc^A_{Lm}$ we obtained. Since then, we collected both new detection and enumeration data after the tumbling step and we intend, using these data, to build a new posterior distribution of $conc^A_{Lm}$. We hope that only one log transformation will be necessary to correctly fit a normal distribution.

To model the dicing step, every breast was assumed to be cuboid, which is not actually the case. Also every dice resulting from a same breast was assumed to have a regular exterior surface area, and to have the same weight. In contrast, observation of the dice produced reveals variability, both in shape and weight, because of the irregular shape of the breasts, and probably also because some of the dice are broken after dicing. Each dice of bacon is modeled separately in our model. However, with regard to potential human exposure, several dice are combined in a single food portion for consumption, and several dice are analyzed together as a test portion for microbiological analysis. Consequently, we can consider that the consequence of our assumptions would be balanced at these portion scales.

In this study, we did not consider bacterial growth, arguing that growth would be negligible during the process, given the surrounding physico-chemical conditions, and the duration of the successive steps. However, growth can occur during the steaming step if the air temperature is around 40°C for example. Besides, growth may occur if a storage step is longer than intended and/or carried out at an accidentally high temperature. Of course, growth can occur at subsequent steps not considered in this paper, such as retail, transport or home storage. The addition of growth and of further steps in the model will be done in future work.

The brining-and-tumbling is a step involving strong cross-contamination, during which the breasts are mechanically tumbled for several hours. In contrast, we assumed that the effects due to contacts between the breasts that can occur during the other steps are negligible.

With regard to the steaming conditions, we made the rough assumption of a constant temperature. Further experiments could help to elaborate more complex models for com-

parison to the simpler assumption. However, even if temperature follow-up could be performed both in the surrounding air and in the center of the breasts, the temperature to which the bacteria on the surface of the breasts are really exposed remains difficult to determine.

F.6.3 Step combination

We simplified the step combination, since we followed tumbling batches without considering the possibility that they would overlap during the process. In reality, the step chain is complicated by practical constraints, both in terms of time and in terms of batch overlaps. The duration of storage after tumbling can vary between the sub-batches of a given tumbling batch. In addition, these sub-batches are steamed in cells where the temperature could be slightly different. And so on for every step. In order to refine the step combination, it would be necessary to develop a process chain manager model, taking open stocks and carrying capacities into account at every step.

Besides, the overlap of several brining-and-tumbling batches later in the process can only result in a lower level of contamination than that in the most contaminated brining-and-tumbling batch, so our approach (following-up of brining-and-tumbling batches) is safe. Moreover, only a small minority of packaging units do actually contain dice originating from several brining-and-tumbling batches.

F.6.4 Prospects for the use of the model

The model presented in this paper is a generic method applicable to any production site producing diced bacon, at least any of the French sites of which we are aware. We intend next to refine this process model in order to focus on one specific plant, using additional data we have acquired there in the meantime. We have been monitoring the temperatures during the steaming step, *L. monocytogenes* has been detected and counted and LAB have been counted on breasts after brining-and-tumbling, and on packaging units after packaging. Using a standard Monte Carlo approach, it would be difficult to integrate this contamination data as inputs at a later step than the initialization of the model. Nevertheless, the formalism of a Bayesian network such as that chosen here looks promising from this perspective⁽¹⁵⁾.

As illustrated in paragraphs 5.2, 5.3 and 5.4, such process models are primarily intended to be used for assessing the impact on the output of the input parameters and their uncertainty and/or variability (using sensitivity analysis in the broader sense) and for mitigation (using what-if scenarios). In the present case, the output is not directly the risk but the contamination of the product at the end of the processing step. It is important to question most of the issues raised by the operators, but it is not sufficient to assess the effects on the risk, in view of recent studies that have highlighted the importance of the output chosen for the conclusion of a sensitivity analysis^(26,27). Nevertheless, this could provide an important added value in risk assessment models which either model the process itself, like a black box, or initialize the contamination at a later point than manufacturing process and then consider in their risk mitigation strategies only the outputs of the process (prevalence and/or initial contamination)^(26,28,29,30,31).

The model presented and/or the plant-specific model are also intended to optimize surveillance, *i.e.* to use a risk-based approach to determine how microbial contamination can best be monitored by the producers. We could therefore use the process model to discuss the impact of sampling raw materials at arrival, and/or intermediary products, and/or final products, in various numbers, and then to study the effectiveness of different sampling plans.

F.7 Acknowledgments

This work was supported by a grant from the Agence National de la Recherche (ANR) (France), as a part of the Quant'HACCP project. We are grateful to the business operators which answered our enquiry.

References

1. Nauta MJ. A modular process risk model structure for quantitative microbiological risk assessment and its application in an exposure assessment of *Bacillus cereus* in a REPFED. Report nr. 149106007 : RIVM, Bilthoven. 2001.
2. Nauta MJ. Modelling bacterial growth in quantitative microbiological risk assessment : is it possible? *International Journal of Food Microbiology* 2002 ;73 :297-304.
3. Nauta MJ. Microbiological risk assessment models for partitioning and mixing during food handling. *International Journal of Food Microbiology* 2005 ;100 :311-322.
4. Lindqvist R, Sylven S, Vagsholm I. Quantitative microbial risk assessment exemplified by *Staphylococcus aureus* in unripened cheese made from raw milk. *International Journal of Food Microbiology* 2002 ;78 :155-170.
5. Nauta MJ, van der Fels-Klerx I, Havelaar A. A poultry-processing model for quantitative microbiological risk assessment. *Risk analysis* 2005 ;25 :85-98.
6. Aziza F, Mettler E, Daudin J-J, Saana M. Stochastic, compartmental and dynamic modeling of cross-contamination during mechanical smearing of cheeses. *Risk analysis* 2006 ;26 :731-745.
7. Nauta MJ, Jacobs-Reitsma WF, Havelaar AH. A risk assessment model for *Campylobacter* in broiler meat. *Risk analysis* 2007 ;27 :845-861.
8. Afchain AL, Carlin F, Nguyen-the C, Albert I. Improving quantitative exposure assessment by considering genetic diversity of *B. cereus* in cooked, pasteurised and chilled foods. *International Journal of Food Microbiology* 2008 ;128 :165-173.
9. Clough HE, Clancy D, French NP. Quantifying exposure to Vero-cytotoxigenic *Escherichia coli* O157 in milk sold as pasteurized : A model-based approach. *International Journal of Food Microbiology* 2009 ;131 :95-105.
10. FSIS-USDA. Draft risk assessment of the public health impact of *Escherichia coli* O157 :H7 in ground beef. Report 00-023N, 2001.
11. RIVM. Risk assessment of Shiga-toxin producing *Escherichia coli* O157 in steak tartare in the Netherlands. Report 257851003, 2001.

12. IFIP. Le porc par les chiffres. Paris, France : IFIP Editions, 2009.
13. FICT. Statistiques de production industrielle française, 2008. Available at : <http://www.fict.fr>, Accessed on May 5, 2010.
14. AFSSA. Etude individuelle nationale des consommations alimentaires, 2009. Available at : <http://www.afssa.fr/Documents/PASER-Ra-INCA2.pdf>, Accessed on May 5, 2010.
15. Albert I, Grenier E, Denis J-B, Rousseau J. Quantitative risk assessment from farm to fork and beyond : A global bayesian approach concerning food-borne diseases. *Risk analysis* 2008 ;28 :557-571.
16. Chick H. An investigation of the laws of disinfection. *Journal of Hygiene of Cambridge* 1908 ;8 :92-158.
17. Bigelow WD. The logarithmic nature of thermal death time curves. *Journal of Infectious Diseases* 1921 ;29 :528.
18. Lunn DJ, Thomas A, Best N, Spiegelhalter D. Winbugs - a bayesian modelling framework : concepts, structure, and extensibility. *Statistics and Computing* 2000 ;10 :325-337.
19. ICMFS. Microorganisms in Foods 5 : Microbiological Specifications of Food Pathogens. London, UK : Kluwer Academic / Plenum Publishers, 1996.
20. Zuliani V. Prédiction de la contamination bactérienne lors de la fabrication et de la conservation d'un aliment. Université d'Auvergne, France, PhD Thesis, 2005.
21. Midelet G, Carpentier B. Transfer of microorganisms, including *Listeria monocytogenes*, from various materials to beef. *Applied and environmental microbiology* 2002 ;68 :4015-4024.
22. Keskinen LA, Todd ECD, Ryser ET. Impact of bacterial stress and biofilm-forming ability on transfer of surface-dried *Listeria monocytogenes* during slicing of delicatessen meats. *International Journal of Food Microbiology* 2008 ;127 :298-304.
23. R Development Core Team. R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2007.
24. Frey HC, Mokhtari A, Danish T. Evaluation of Selected Sensitivity Analysis Methods Based Upon Applications to Two Food Safety Process Risk Models. Washington, DC : Office of Risk Assessment and Cost-Benefit Analysis, U.S. Department of Agriculture, 2003. Available at : <http://www.ce.ncsu.edu/risk/Phase2Final.pdf>, Accessed on May 5, 2010.
25. Commeau N, Parent E, Billoir E, Zuliani V, Cornu M. Modelling contamination to build a sampling plan : application to French diced bacon industry and *Listeria monocytogenes*. Proceedings of ICPMF6, Washington DC, USA, 2009.
26. Pouillot R, Goulet V, Delignette-Muller ML, Mahé A, Cornu M. Quantitative Risk Assessment of *Listeria monocytogenes* in French Cold-Smoked Salmon : II. Risk Characterization. *Risk Analysis* 2009 ;29 :806-819.
27. Ellouze M, Gauchi JP, Augustin JC. Global Sensitivity Analysis Applied to a Contamination Assessment Model of *Listeria monocytogenes* in Cold Smoked Salmon at Consumption. *Risk Analysis* 2010 ;30 :841-852.
28. Pouillot R, Miconnet N, Afchain AL, Delignette-Muller ML, Beaufort A, Rosso L, Denis JB, Cornu M. Quantitative risk assessment of *Listeria monocytogenes* in French

cold-smoked salmon : I. Quantitative exposure assessment. *Risk Analysis* 2007 ;27(3) :683-700.

29. Ross T, Rasmussen S, Fazil A, Paoli G, and Sumner J. Quantitative risk assessment of *Listeria monocytogenes* in ready-to-eat meats in Australia. *International Journal of Food Microbiology* 2009 ;131 :128-137.

30. Delignette-Muller ML, Cornu M. Quantitative risk assessment for *Escherichia coli* O157 :H7 in frozen ground beef patties consumed by young children in French households. *International Journal of Food Microbiology* 2008 ;128(1) :158-164.

31. FDA/FSIS. Quantitative Assessment of Relative Risk to Public Health from Foodborne *Listeria monocytogenes* Among Selected Categories of Ready-to-Eat Foods. Washington, DC :U.S. Food and Drug Administration (HHS)/Food Safety and Inspection Agency (USDA), 2003.

Annexe G

Modification de l'article de Billoir et al.

Ce poster a été présenté au congrès FoodMicro 2010 à Copenhague.



Adjusting the Control Parameters of the Diced Bacon Process According to Food Safety Targets for *Listeria monocytogenes*

Laurent Guillier¹, Natalie Commeau^{1,2}, Jean-Baptiste Denis³ and Pascal Garry⁴

¹ Agence nationale de sécurité sanitaire ANSES – Laboratoire de sécurité des aliments, Maisons-Alfort, FRANCE (laurent.guillier@anses.fr)
² MORSE Team - UMR 518, Mathématiques et informatique appliquées, AgroParisTech/INRA – Paris, FRANCE
³ UR 341, Mathématiques et Informatique appliquées- INRA Jouy-en-Josas, FRANCE
⁴ IFIP, Institut de la filière porcine, Pôle viandes fraîches et produits transformés, Maisons-Alfort, FRANCE



1. Introduction

In France, 11.22% of the pork's breasts is transformed into diced bacon. Diced bacon consists in processed pork breasts tumbled with brine and then steamed, diced and packed under modified atmosphere. *L. monocytogenes* can be present on pork breasts. According to the physico-chemical characteristics of diced bacon and during post-processing storage (the self-life is between 35 and 52 days), growth of the pathogen is possible. Diced bacon is a product usually used in home cooked preparations but occasionally consumed as a ready-to-eat by about 10% of the consumer population in France. In this context, a probabilistic model was developed (Billoir et al. in press) to assess the impact of the manufacturing process on the fate of *L. monocytogenes* and lactic acid bacteria (LAB). Contamination evolution was modeled in the adequate units (breasts, dices, or packaging units) through the successive process steps (arrival, brining, steaming, dicing and packaging). This model did not consider growth during process. Recently, additional data have been acquired. This concerns (i) *L. monocytogenes* (detection and enumeration) on pork breasts just after tumbling (Commeau et al. 2009), (ii) fate of inoculated *L. monocytogenes* in packed diced bacon (Cornu et al., accepted) and (iii) temperatures at the breast's surfaces and in the surrounding air during steaming.



Figure 1. Diced bacon process chain

3. Methods

Refinement of the process model

First the modelling approach of *L. monocytogenes* concentration at arrival was modified. The Bayesian model used to describe contamination at arrival is composed of two parts: the process part and the data one, the latter being composed of detection and enumeration (Commeau et al., 2009). The log-lognormal distribution previously used by Billoir et al. (in press) to describe concentration on breasts (c_i^A). The previous model presented in Billoir et al. (2010) described concentration on breasts at arrival with a log-lognormal distribution. In the new model, c_i^A depends on the logarithm the concentration in a batch (conCB). IconCB follows a normal distribution of parameter μ^A and σ^A . Flat priors were set because very few information about breast contamination at arrival was available. Bayesian estimation of the posterior distributions of the lognormal distribution was performed using the software OpenBUGS (3.0.3) and the R package BRugs. 100,000 iterations were performed. 1 value out of 1000 was taken to avoid autocorrelation. The new model of simulation of the contamination is presented in Figure 2.

Growth during heat-up phase of the steaming phase was considered in the new process model. The same primary and secondary growth models as used for post-processing were used to predict growth of LAB and *L. monocytogenes*.

Extension of the model to post-processing

The model was extended to the post-process storage until consumption, by considering further steps such as refrigerated transport and home storage. A shelf-life of 52 days was considered. Temperature was assumed to be constant at 8°C throughout it. Growth and competition between *L. monocytogenes* and LAB were considered. For primary growth model, a model that takes into account Jameson effect was used. The growth competition was considered at the dice level.

For LAB and *L. monocytogenes*, secondary growth models of Mejlholm and Dalgaard (2007) and Mejlholm et al. (2010) were considered respectively.

The following environmental factors were considered for prediction of growth rate of LAB and *L. monocytogenes*: were done according to values of physico-chemical analyses (pH, water activity, lactate, diacetate, CO₂) carried out in diced bacon of three different producers. Between batch variability was taken into account (e.g. for acid content).

Link between control parameters and food safety targets

Several process parameters are directly controllable by the manufacturer: for example temperature and duration of the steaming step. For these parameters we tested different possible values in order to exemplify the link between process control measures and food safety targets. Five scenarios were tested in their ability to comply with different food safety target levels. The scenarios combined different control measures that are listed in Table 1. For each of these scenarios, we checked the compliance of diced bacon packages at the end of the storage step with three different food safety targets (1 cfu/100 g, 1 cfu/10 g, 1 cfu/g and 100 cfu/g).

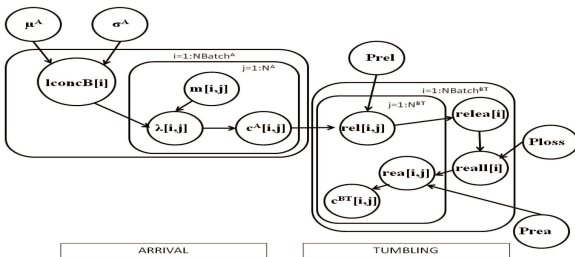


Figure 2. Diagram of the model describing the first two steps of the process. μ^A : probability for a bacterium to be released; σ^A : part of bacteria released from a breast; $releal[i]$: total number of bacteria released; $reall[i]$: total number of bacteria reallocated; $real[i]$: number of bacteria reallocated to a breast; $Prea$: probability for a bacterium to be reallocated; and c^{BT} is the contamination at the end of the tumbling step.

2. Objectives

The objectives of this work were:

- To refine the process model according to new data collected
- To extend the model for the storing step
- To link the controllable parameters of the diced bacon process according to food safety targets

5. Conclusions

- Modelling approach confirms experimental results of challenge test
- Growth is sparse during refrigerated storage of diced bacon and according to level at the end of the process step, the probability to reach high levels of *L. monocytogenes* at the end of the shelf-life is low
- Model can be used to adjust process parameters in order to comply with different food safety objectives

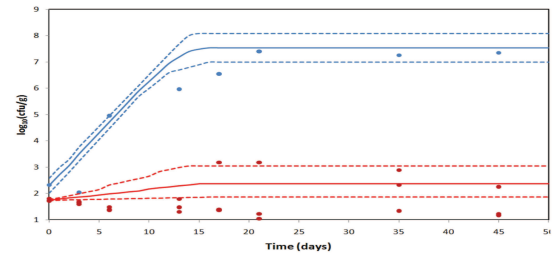


Figure 3. Observed (●) and predicted (–) growths of LAB (blue) and *L. monocytogenes* (red) in diced bacon at 8°C ($a_w=0.96$, $pH=5.8$, 1.52% of lactate, 40% of CO₂). (–) represent 95% confidence intervals of predictions.

4. Results

The means of posterior distributions for μ^A and σ^A were respectively of -2.33 and 0.93. Primary model associated with secondary growth models for LAB and *L. monocytogenes* satisfactorily predict the observed data of challenge-testing (Figure 3). These two models were thus incorporated in the storing step of the model.

For all the different sets of process parameters or for a reduced salt level, the probabilities to surpass level higher than 1 cfu of *L. monocytogenes*/g at the end of the shelf life is small (Figure 4).

Figure 5 illustrates the impact on changing process parameter or diced bacon composition on food safety objectives.

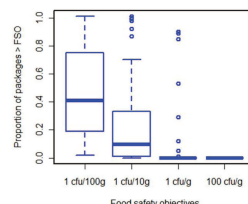


Figure 4. Probability to comply with different food safety objectives at the end of the shelf-life for scenario S5.

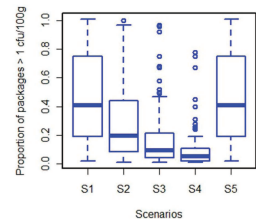


Figure 5. Probability to comply with a food safety target of 1 ufc/100g at the end of the shelf-life for the different scenarios

Table 1. Parameter values for different scenarios of diced bacon processing

Scenario	Maximum temperature after heat-up step during steaming	Duration of exposing at maximum temperature	Water activity
S1	45°C	1h30	0.96
S2	50°C	0h30	
S3	50°C	1h00	
S4	50°C	1h30	
S5	45°C	1h30	0.965

References

Billoir E, Zuliani V, Cornu M, Commeau N, and Denis J-B (in press). Probabilistic modeling of the fate of *Listeria monocytogenes* in diced bacon during the manufacturing process. Risk analysis. Accepted for publication.
 Cornu M, Billoir E, Bergis H, Beaufort A and Zuliani V (in press). Modeling microbial competition in foods. Application to the behaviour of *Listeria monocytogenes* and lactic acid flora in diced bacon. Food Microbiology. Accepted for publication.
 Commeau N, Parent E, Billoir E, Zuliani V and Cornu M (2009). Modeling contamination to build a sampling plan: application to French diced bacon industry and *Listeria monocytogenes*. Proceedings of ICPMIF, Washington DC, USA.
 Mejlholm O and Dalgaard P (2007). Modeling and predicting the growth of lactic acid bacteria in lightly preserved seafood and their inhibiting effect on *Listeria monocytogenes*. Journal of Food Protection. 70:2485-2497.
 Mejlholm O, Gunvig A, Bornggaard C, Blom-Hansen J, Mellefont L, Ross T, Leroi F, Else T, Visser D, and Dalgaard P (2010). Predicting growth rates and growth boundary of *Listeria monocytogenes* - An international validation study with focus on processed and ready-to-eat meat and seafood. International Journal of Food Microbiology. 141:137-150.

Annexe H

Rapport qualité

Ce rapport a été élaboré par l'équipe de management de la qualité du laboratoire de sécurité des aliments (ANSES).

RAPPORT QUALITE POUR LE MEMOIRE DE THESE DE N. COMMEAU

« PROJET QH-ECH »

Depuis 1990, le laboratoire de sécurité des aliments de l'Anses a mis en place une démarche qualité. Il est accrédité par le COFRAC selon le référentiel NF EN ISO 17025 (portée disponible sur www.cofrac.fr) pour un ensemble de programmes « Essais » correspondant à ses activités d'analyse de contaminants microbiologiques et physico-chimiques dans les aliments. Depuis 2001, le laboratoire a entrepris une démarche de qualité en recherche et a suivi pour cela les recommandations de la commission AFNOR sur l'assurance qualité en recherche. Un système de management de la qualité en recherche a été mis en place au sein du laboratoire qui s'applique aux travaux de thèse. Un manuel pour la qualité en recherche auquel les unités accueillant des doctorants doivent se référer est en vigueur. Des audits internes sont réalisés régulièrement auprès des unités par le service de management de la qualité du laboratoire. C'est dans ce contexte que le présent rapport qualité pour les travaux de thèse faisant l'objet de ce mémoire a été rédigé. Les audits et le rapport ci-dessous ne concernent que les travaux effectués à l'Anses, au laboratoire de sécurité des aliments de Maisons-Alfort.

Rédacteurs et auditeurs	Service de management de la Qualité : H. Lelièvre (responsable qualité des laboratoires), C. Cecconi et M. Poumeyrol (adjointes) Sous-direction en charge de la recherche au laboratoire : S. Dragacci
Audit Qualité	1 ^{er} décembre 2009 ; auditeurs : H. Lelièvre, M. Poumeyrol et S. Dragacci
Audit de traçabilité	29 mars 2011 ; auditeur : H. Lelièvre et S. Dragacci
Unité d'accueil	Unité MOB-Modélisation des comportements bactériens, chef d'Unité A. BEAUFORT - ANSES-Laboratoire de sécurité des aliments
Date début de thèse	Novembre 2008
Acronyme Thèse	QH-Ech
Sujet de Thèse	Optimisation de l'échantillonnage pour la surveillance des dangers microbiens dans les procédés agro-alimentaires <i>Ce sujet de recherche fait partie du projet de recherche ANR-Quant'HACCP</i>
Direction de thèse	M. Simon-Cornu (Anses, Unité MOB) & E. Parent (AgroParisTech/Engref)
Chef de travaux à l'Anses	M. Cornu & A. Beaufort (Anses, Unité MOB)
Ecole doctorale	ABIES Agriculture, Alimentation, Biologie, Environnement et Santé (ED n° 0435, AgroParisTech)
Comité de pilotage	B. Amzal (EFSA), JC Augustin (ENVA), K. BOQUJET (AgroParisTech-Engref), ML. Delignette-Müller (ENVL), MB. PEYRAT (DGAL), J. ROUSSEAU (Université Paris Dauphine), A. Beaufort (Anses)
Correspondant qualité de l'Unité	S. Rudelle

Extrait de la charte des thèses de l'Anses : « *Thèse et système de management de la qualité* : Les travaux de thèse doivent s'intégrer dans le système de management de la qualité en recherche en place au sein de l'entité. Le management de la qualité en recherche permet de garantir la bonne conduite de la recherche et ainsi de donner confiance dans les résultats de la recherche.

A ce titre, le doctorant accepte d'effectuer ses expériences et manipulations conformément aux méthodes contenues dans le manuel de la qualité du laboratoire dont il a pris connaissance. Le doctorant accepte qu'un

audit qualité ait lieu au cours de son travail ainsi qu'il y ait éventuellement un audit qualité à la fin de son travail, préliminaire à la rédaction du bilan qualité de la thèse par le responsable qualité de l'entité et qui devra accompagner son mémoire de thèse. D'autre part, le doctorant accepte qu'une fois par an ait lieu un bilan qualité de son travail dont les conclusions pourront comporter des suggestions d'amélioration. L'entité d'accueil du doctorant s'engage à prendre en charge et assurer la formation qualité de l'étudiant et à mettre à sa disposition toute la documentation qualité dont il a besoin. »

Initiation à l'assurance qualité

La doctorante a été formée de façon satisfaisante au système d'assurance qualité par la correspondante qualité de l'unité. Elle s'est rapidement approprié le manuel de la qualité en recherche du laboratoire et en particulier de la procédure « Gestion de projet ».

Présentation par la doctorante de son projet de recherche (objectifs, hypothèses de travail)

Le sujet de thèse a été clairement présenté par la doctorante et les hypothèses de travail ont été posées de façon pertinente. Ce sujet de recherche est issu d'un projet financé par l'ANR et coordonné par l'unité d'accueil : Quant'HACCP-Extension des méthodologies d'appréciation quantitative des risques microbiologiques aux mesures de maîtrise des dangers pour respecter des objectifs de sécurité des aliments (2008-2011). De ce fait, le reporting exigé par l'ANR et auquel aura contribué la doctorante pour sa part de projet y est particulièrement soigné (description des tâches, définition des équipes et responsabilité, délais et jalons, rapports d'avancement, etc.).

Organisation de la documentation qualité

La documentation qualité est convenablement utilisée : documents de définition et d'acceptation du projet, nomination du chef de projet et composition de l'équipe de recherche et du comité de pilotage, description des tâches, diagramme de Gantt, rapports d'avancement, rapports de réunion (notamment du comité de pilotage), valorisation, collaborations.

L'archivage de l'ensemble du dossier de recherche est effectué sur le réseau informatique de l'établissement et sur le site web interne au projet ANR « Quant'HACCP » (sites sécurisés). Il existe un document décrivant la traçabilité du stockage des dossiers.

Gestion du projet

La démarche entreprise par la doctorante pour la réalisation de ses recherches est ici excellente : la définition des objectifs, l'édification des hypothèses (parfois dénommées dans ce projet « alternatives »), les travaux de modélisation, ses résultats, la discussion et les conclusions sont effectués et détaillés dans des documents adéquats et tracés. La valorisation des résultats est d'emblée inscrite dans le déroulement des travaux (rédaction des résultats validés sous format de publication).

La doctorante s'est vue recommander et s'est attachée à étoffer les critères de sélection des données entrantes pour nourrir ses travaux de modélisation, que ces données proviennent d'autres équipes du laboratoire ou de l'extérieur de l'Agence. De même, des critères pour l'acceptation et la validation des logiciels utilisés pour la construction des programmes (ou feuilles de calcul) dans le cadre du projet ont été réfléchis.

La valorisation du projet de recherche est avérée par 3 publications publiées dont 2 en tant que premier auteur, plusieurs communications à congrès-colloques dont 2 orales et par une contribution à un rapport d'expertise pour l'EFSA-European Agency for food safety.

Expérimentations et traçabilité

La traçabilité de la démarche de recherche est exemplaire : il existe des documents expliquant la démarche de recherche entreprise pour chaque question posée. Ces documents datés sont sous forme de fichiers pdf. Ils expliquent l'état de la question et des résultats précédemment acquis, les nouvelles hypothèses pour la construction des modèles statistiques, la modélisation proprement dite (formulation mathématique), les résultats produits et une discussion éventuellement assortie de nouvelles pistes de modélisation à entreprendre. Ils sont incrémentés et figés par un enregistrement sous pdf au fur et à mesure de l'avancée des travaux de modélisation, ce qui permet de conserver une trace historique de l'évolution de la recherche menée. Concernant les données entrantes pour les travaux de modélisation : la qualité (nature, format) de ces données nourrissant la construction des modèles statistiques est constamment discutée au long du processus de recherche car elle conditionne la qualité même de construction de ces modèles et leur représentativité pour la modélisation/optimisation et pour la signification des plans d'échantillonnage. Certaines données sur la contamination naturelle par *Listeria monocytogenes* dans des aliments (saumon) sont issues d'analyses effectuées dans l'unité dont la traçabilité d'acquisition a été constatée.

Les logiciels couramment utilisés dans ces travaux sont des logiciels d'accès libres *via* internet : le logiciel R, le logiciel OpenBugs, le logiciel Jags, et de manière limitée Excel. Ces logiciels évoluant constamment, les nouvelles versions sont utilisées au fur et à mesure de leur apparition. Il est précisé dans les documents de recherche que telle modélisation a été réalisée avec telle version.

EN CONCLUSION, L'EQUIPE DE MANAGEMENT DE LA QUALITE DU LABORATOIRE DE SECURITE DES ALIMENTS ATTESTE QUE LES TRAVAUX DE THESE MENES PAR N. COMMEAU SE SONT DEROULES EN APPLIQUANT LES ELEMENTS DU SYSTEME QUALITE DU LABORATOIRE, CE QUI PERMET D'AVOIR CONFIANCE DANS LA CONDUITE DU PROJET DE THESE. N COMMEAU S'EST EGALEMENT BEAUCOUP IMPLIQUEE D'UNE PART DANS LA DEFINITION DE CRITERES DE SELECTION DES DONNEES ENTRANTES DE SES MODELES ET, D'AUTRE PART, DANS UNE REFLEXION SUR LA VALIDATION DES LOGICIELS UTILISES POUR SES TRAVAUX DE MODELISATION.

Fait à Maisons-Alfort, le

20 Avril 2012

H. Lelièvre

Responsable du service Qualité
Anses-Laboratoire de sécurité des aliments
et laboratoire de santé animale

* * * * *

Index

C. perfringens, 7, 8, 21, 59, 62, 63, 79, 197–200, 221, 223, 230

L. monocytogenes, 7, 8, 21, 23, 24, 27, 49, 50, 52, 56, 58–61, 63, 67, 69, 70, 72–74, 77, 81, 84, 86–89, 93, 125, 147–150, 154, 155, 186, 195, 221, 223, 230, 231

ALOP, 22–25

BIC, 199, 227

DIC, 93, 226, 227

Facteur de Bayes, 227, 239, 257

FSO, 23–25, 27, 49

L. monocytogenes, 161, 162

LOD, 72, 74, 81, 85, 150, 154, 164, 225

LOQ, 72–74, 81, 84, 85, 89, 129, 150, 154, 164

lot, 7, 24, 26, 27, 30–35, 37, 38, 40, 41, 50, 54–58, 61, 62, 67, 68, 74–78, 80–83, 85, 88, 90–93, 124, 126, 147, 148, 150–155, 161–165, 170, 171, 178, 181–183, 185, 186, 192, 195, 198, 199, 221–225, 227, 228, 231, 232, 249, 250

NQA, 33, 34, 38, 39, 192

NQL, 33, 34, 38, 39, 185, 192, 256

PC, 24, 25

PMS, 28, 29, 162–165

PO, 23–25, 49

Prévalence, 50, 51, 53, 59, 61, 62, 74, 86–88, 90, 162, 164, 165, 169, 171, 172, 174–183, 185–187, 191–193, 195, 199, 222, 223, 225, 227, 249

Bibliographie

- Abadias, M., Canamas, T., Asension, A., Anguera, M., and Vinas, I. (2006). Microbial quality of commercial ‘Golden Delicious’ apples throughout production and shelf-life in Lleida (Catalonia, Spain). *International Journal of Food Microbiology*, 108(3) :404–409.
- Amzal, B., Bois, F., Parent, E., and Robert, C. (2006). Bayesian optimal design via interacting particle systems. *Journal of the American Statistical Association*, 101(474) :773–785.
- Ancelet, S. (2008). *Exploiter l’approche hiérarchique bayésienne pour la modélisation statistique de structures spatiales*. PhD thesis, AgroParisTech.
- Anonymous (1997). Healthy people 2020 summary of objectives : Food safety.
- ANSES (2008). Recommandations pour l’élaboration de critères microbiologiques d’hygiène des procédés. Technical report, ANSES.
- ANSES (2009). *Enquête individuelle et nationale sur les consommations alimentaires 2 (INCA2)*, (<http://www.afssa.fr/Documents/PASER-Ra-INCA2.pdf>).
- ANSES (2010). Fiche de description de danger microbiologique transmissible par les aliments : *Clostridium perfringens*. pages 1–3. <http://www.anses.fr/Documents/MIC-Fi-Clostridiumperf.pdf>, last accessed on January, 25th 2012.
- Arrêté du 21 décembre 2009 (2009). Arrêté du 21 décembre 2009 relatif aux règles sanitaires applicables aux activités de commerce de détail, d’entreposage et de transport de produits d’origine animale et denrées alimentaires en contenant. *Journal officiel de la République française du 31 décembre 2009*, 0303 :23335.
- Arrêté du 8 juin (2006). Arrêté ministériel relatif à l’agrément ou à l’autorisation des établissements de production primaire et des établissements mettant sur le marché des produits d’origine animale ou des denrées contenant des produits d’origine animale. *Journal Officiel*.
- Augustin, J.-C. and Carlier, V. (2006). Lessons from the organization of a proficiency testing program in food microbiology by interlaboratory comparison : analytical methods in use, impact of methods on bacterial counts and measurement uncertainty of bacterial counts. *Food microbiology*, 23 :1.

- Aziza, F., Mettler, E., Daudin, J.-J., and Sanaa, M. (2006). Stochastic, compartmental, and dynamic modeling of cross-contamination during mechanical smearing of cheeses. *Risk Analysis*, 26(3) :731–745.
- Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 53 :370–418.
- Beaufort, A., Rudelle, S., Gnanou-Besse, N., Toquin, M.-T., Kerouanton, A., Bergis, H., Salvat, G., and Cornu, M. (2007). Prevalence and growth of *Listeria monocytogenes* in naturally contaminated cold-smoked salmon. *Letters in applied microbiology*, 44(4) :406–411.
- Berger, J. (1985). *Statistical decision theory and Bayesian analysis*. Springer Verlag, New York, second edition.
- Bertsch McGrayne, S., H. (2011). *The theory that would not die*. Yale University Press, New Haven and London.
- Billoir, E., Zuliani, V., Commeau, N., Cornu, M., and Denis, J.-B. (2011). Probabilistic modeling of the fate of *Listeria monocytogenes* in diced bacon during the manufacturing process. *Risk Analysis*, 31(2) :237–254.
- Boreux, J.-J., Parent, E., and Bernier, J. (2010). *Pratique du calcul bayésien*. Springer-Verlag, France.
- Brand, K. and Small, M. (1995). Updating uncertainty in an integrated risk assessment : conceptual framework and methods. *Risk Analysis*, 15(6) :719–731.
- Brooks, S. (2003). Bayesian computation : A statistical revolution. *Transactions of the Royal Society, Series A*, 361 :2681–2697.
- Brown, M., Gill, C., Hollingsworth, J., Nickelson II, R., Seward, S., Sheridan, J., Stevenson, T., Sumner, J., Theno, D., Usborne, W., and Zink, D. (2000). The role of microbiological testing in systems for assuring the safety of beef. *International Journal of Food Microbiology*, 67 :7–16.
- Burfoot, D., Archer, J., Horváth, E., Hooper, G., Allen, V., Hutchison, M., and Harrison, D. (2010). Fate of *Salmonella* spp. on broiler carcasses before and after cutting and/or deboning. Technical report, Campden BRI and University of Bristol. <http://www.efsa.europa.eu/en/scdocs/doc/45e.pdf>, last accessed on January 17, 2011.
- Busschaert, P., Geeraerd, A., Uyttendaele, M., and Van Impe, J. (2010). Estimating distributions out of qualitative and (semi)quantitative microbiological contamination data for use in risk assessment. *International journal of food microbiology*, 138 :260–269.
- Busschaert, P., Geeraerd, A., Uyttendaele, M., and Van Impe, J. (2011). Hierarchical Bayesian analysis of censored microbiological contamination data for use in risk assessment and mitigation. *Food Microbiology*, 28 :712–719.

- CDC (2010). Number of infections and incidence per 100,000 persons by site.
- CDC (2011). CDC estimates of foodborne illness in the united states.
- Celeux, G., Marin, J., and Robert, C. (2006). Sélection bayésienne de variables en régression linéaire. *Journal de la société française de statistique*, 147(1) :59–79.
- Cho, I.-K. and Kreps, D. (1987). Signalling games and stable equilibria. *Quarterly Journal of Economics*, 102 :179–121.
- Clough, H., Clancy, D., O’Neill, P., Robinson, S., and French, N. (2005). Quantifying uncertainty associated with microbial count data : a Bayesian approach. *Biometrics*, 61 :610–616.
- Cochran, W. (1977). *Sampling techniques*. John Wiley & Sons, New York.
- Codex Alimentarius (1997a). Joint FAO/WHO food standards programme Codex Alimentarius Commission. Food hygiene, Supplement to volume 1B-1997. Hazard analysis and critical control point (HACCP) system and guidelines for its application.
- Codex Alimentarius (1997b). Joint FAO/WHO food standards programme Codex Alimentarius Commission. Food hygiene, Supplement to volume 1B-1997. Recommended international code of practice, General principles of food hygiene.
- Codex Alimentarius (1997c). Principles for the establishment and application of microbiological criteria for foods. *Codex Alimentarius*.
- Codex Alimentarius (1999). Principes et directives régissant la conduite de l’évaluation des risques microbiologiques.
- Codex Alimentarius (2003). Guidelines on the judgement of equivalence of sanitary measures associated with food inspection and certification systems.
- Codex Alimentarius (2004). Report of the twentieth session of the Codex committee on general principles.
- Codex Alimentarius (2007). Guidelines on the application of general principles of food hygiene to the control of *listeria monocytogenes* in foods.
- Codex Alimentarius (2008). Directives relatives à la validation des mesures de maîtrise de la sécurité alimentaire.
- Commeau, N., Cornu, M., Albert, I., Denis, J.-B., and Parent, E. (2012a). Hierarchical bayesian models to assess between- and within-batch variability of pathogen contamination in food. *Risk Analysis*, 32(3) :395–415.
- Commeau, N., Parent, E., Delignette-Muller, M.-L., and Cornu, M. (2012b). Fitting a lognormal distribution to enumeration and absence/presence data. *International journal of food microbiology*, 155(3) :146–152.

- Cornier, C., Curioni, M., Hatchuel, A., Lefol, Y., Moisdon, J., Molet, H., Saporta, G., and Tonneau, D. (1980). Théorie de la décision, une théorie en quête d'application. Technical report.
- Cornu, M., Billoir, E., Bergis, H., Beaufort, A., and Zuliani, V. (2011). Modeling microbial competition in food : Application to the behavior of *listeria monocytogenes* and lactic acid flora in pork meat products. *Food Microbiology*, 28 :639–647.
- Crouch, E. and Golden, N. (2005). A risk assessment for *clostridium perfringens* in ready-to-eat and partially cooked meat and poultry products. Technical report, USDA.
- Crépet, A., Albert, I., Dervin, C., and Carlin, F. (2007). Estimation of microbial contamination of food from prevalence and concentration data : application to *Listeria monocytogenes* in fresh vegetables. *Applied and environmental microbiology*, 73(1) :250–258.
- Dalgaard, P. and Jørgensen (1998). Predicted and observed growth of *listeria monocytogenes* in seafood challenge tests and in naturally contaminated cold-smoked salmon. *International Journal of Food Microbiology*, 40(1-2) :105–115.
- Dauphin, G. and Ragimbeau, P. (2001). Use of PFGE typing for tracing contamination with *Listeria monocytogenes* in three cold-smoked salmon processing plants. *International Journal of Food Microbiology*, 64 :51–61.
- deFinetti, B. (1937). *La prévision : ses lois logiques, ses sources subjectives*. Institut Henri Poincaré.
- deGroot, M. (1986). A conversation with David Blackwell. *Biometrika*, 1(1) :40–53.
- Delignette-Muller, M. and Cornu, M. (2008). Quantitative risk assesement for *escherichia coli* O157 :H7 in frozen ground beef patties consumed by young children in French households. *International journal of food microbiology*, 128 :158–164.
- Delignette-Muller, M.-L., Cornu, M., Pouillot, R., and Denis, J.-B. (2006). Use of bayesian modelling in risk assessment : application to growth of *Listeria monocytogenes* and food flora in cold-smoked salmon. *International Journal of Food Microbiology*, 106 :195–208.
- Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39 :1–38.
- FDA (2003). Quantitative assessment of relative risk to public health from food-borne *listeria monocytogenes* among selected categories of ready-to-eat foods. Technical report. <http://www.fda.gov/Food/ScienceResearch/ResearchAreas/RiskAssessmentSafetyAssessment/ucm183966.htm>, last accessed on January 26th, 2012.
- Fienberg, S. (2008). The early statistical years : 1947-1967 a conversation with Howard Raiffa. *Statistical Science*, 23(1) :136–149.
- Fisher, R. (1925). *Statistical methods for research workers*. Oliver and Boyd.

- Food Inspection Agency, C. (2010). Food recall report.
- FoodNet (2009). Foodborne active disease surveillance network (foodnet) : 2009 surveillance report.
- Forster, J. (2010). Bayesian inference for Poisson and multinomial log-linear models. *Statistical Methodology*, 7 :210–224.
- Frenzen, P., Drake, A., Angulo, F., and the merging infections program FoodNet working group (2005). Economic cost of illness due to *escherichia coli* O157 infections in the United States. *Journal of Food Protection*, 68(12) :2623–2630.
- Gelfand, A. and Ghosh, S. (1998). Model choice : a minimum predictive loss approach. *Biometrika*, 85(1) :1–11.
- Gelfand, A., Hills, S., Racine-Poon, A., and Smith, A. (1990). Illustration of Bayesian inference in Normal data models using Gibbs sampling. *Journal of American Statistical Association*, 85(412) :972–985.
- Gelfand, A. and Smith, A. (1990). Sampling-based approaches to calculating marginal densities. *Journal of American Statistical Association*, 85(410) :398–409.
- Gelman, A. (2007). Comment : Bayesian checking of the second levels of hierarchical models. *Statistical science*, 22(3) :349–352.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6 :721–740.
- Gnanou Besse, N., N, A., Beaufort, A., Colin, P., Cornu, M., and Lombard, B. (2004). A contribution to the improvement of *Listeria monocytogenes* enumeration in cold-smoked salmon. *International Journal of Food Microbiology*, 91 :119–127.
- Gonzales-Barron, U. and Butler, F. (2011a). Characterisation of within-batch and between-batch variability in microbial counts in foods using poisson-gamma and poisson-lognormal regression models. *Food Control*, 22 :1268–1278.
- Gonzales-Barron, U. and Butler, F. (2011b). A comparison between the discrete poisson-gamma and poisson-lognormal distributions to characterise microbial counts in foods. *Food Control*, 22 :1279–1286.
- Gonzales-Barron, U., Kerr, M., Sheridan, J., and Butler, F. (2010a). Count data distributions and their zero-modified equivalents as a framework for modelling microbial data with a relatively high occurrence of zero counts. *International journal of food microbiology*, 136 :268–277.
- Gonzales-Barron, U., Raidmond, G., and Butler, F. (2010b). Modelling prevalence and counts from most probable number in a Bayesian framework : an application to *Salmonelle typhimurium* in fresh pork sausages. *Journal of food protection*, 73(8) :1416–1422.

- Gonzales-Barron, U., Soumpasis, I., Butler, F., Prendergast, D., Duggan, S., and Duffy, G. (2009). Estimation of prevalence of *salmonella* on pig carcasses and pork joints, using a quantitative risk assessment model aided by meta-analysis. *Journal of food protection*, 72(2) :274–285.
- Good, I. (1979). Studies in the history of probability and statistics. XXXVII A. M. Turing's statistical work in World War II. *Biometrika*, 66(2) :393–396.
- Good, I. J. (1950). *Probability and the Weighing of Evidence*. Charles Griffin, London.
- Gorton, S., Kliebenstein, J., and Beran, G. (1997). Cost of onfarm microbial testing for *Salmonella* : An application by farm size and prevalence level. Technical Report ASL-R1413, Iowa state university.
- Haas, C., Rose, J., and Gerba, C. (1999). *Quantitative microbial risk assessment*. John Wiley and sons, INC.
- Habib, I., Sampers, I., Uyttendaele, M., Berkvens, D., and De Zutter, L. (2008). Baseline data from a Belgium-wide survey of *Campylobacter* species contamination in chicken meat preparations and considerations for a reliable monitoring program. *Applied and environmental microbiology*, 74(17) :5483–5489.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chain and their applications. *Biometrika*, 57(1) :97–109.
- Helsel, D. (2006). Fabricating data : how substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere*, 65 :2434–2439.
- Hudson, J. and Mott, S. (1993). Growth of *listeria monocytogenes*, *aeromonas hydrophila* and *yersinia enterocolitica* on cold-smoked salmon under refrigeration and mild temperature abuse. *Food Microbiology*, 10 :61–68.
- ICMSF (1986). *Microorganisms in Foods 2 : sampling plan for microbiological analysis : principles and specific applications, 2nd edition*. Toronto, University of Toronto press.
- ICMSF (2002). *Microorganisms in Foods 7 : microbiological testing in food safety management*. New-York, Kluwer academic/Plenum Publisher.
- ICMSF (2011). *Microorganisms in Foods 8 : use of data for assessing process control and product acceptance*. Springer, New-York.
- ILSI (2004). Food safety objectives - role in microbiological food safety management.
- ILSI (2005). Achieving continuous improvement in reductions in foodborne listeriosis - A risk-based approach. *Journal of Food Protection*, 68(9) :1932–1994.
- ILSI (2010). Impact of microbial distribution on food safety. 1-68.
- InVS (2004). Morbidité et mortalité dues aux maladies infectieuses d'origine alimentaire en France. pages 1–192. http://opac.invs.sante.fr/doc_num.php?explnum_id=5780, last accessed on January, 25th 2012.

- InVS (2010). Nombre de cas de listérioses déclarés par an et taux d'incidence annuel de 1999 à 2010. Technical report. http://www.invs.sante.fr/surveillance/listeriose/cas_listeriose_1999_2010.pdf, last accessed on January 9th, 2012.
- ISO 16140 (2003). Protocole pour la validation des méthodes alternatives. Technical report, AFNOR.
- ISO11290-1 (1996a). Horizontal method for the detection and enumeration of *Listeria monocytogenes* - Part 1 : detection method.
- ISO11290-1, N. E. (1996b). Méthode horizontale pour la recherche et le dénombrement de *Listeria monocytogenes* - Partie 1 : Méthode de recherche.
- ISO11290-2 (1998a). Horizontal method for the detection and enumeration of *Listeria monocytogenes* - Part 2 : enumeration method.
- ISO11290-2, N. E. (1998b). Méthode horizontale pour la recherche et le dénombrement de *Listeria monocytogenes* - Partie 2 : Méthode de dénombrement.
- ISO19036 (2006). Spécification technique ISO/TS 19036 microbiologie des aliments - lignes directrices pour l'estimation de l'incertitude de mesure pour les détermination qualitatives.
- Jaloustre, S. (2011). *Appréciation quantitative des risques pour l'évaluation de mesures de maîtrise sanitaire dans une filière agro alimentaire. Application à Clostridium perfringens en restauration hospitalière*. Phd thesis, AgroParisTech.
- Jaloustre, S., Cornu, M., Morelli, E., Noël, V., and Delignette-Muller, M.-L. (2011). Bayesian modeling of *clostridium perfringens* growth in beef-in-sauce products. *Food Microbiology*, 28 :311–320.
- Jaloustre, S., Guillier, L., Morelli, E., Noël, V., and Delignette-Muller, M.-L. (2012). Modeling of *clostridium perfringens* vegetative cell inactivation in beef-in-sauce products : A meta-analysis using mixed linear models. *International Journal of Food Microbiology*, 154 :44–51.
- Jeffreys, H. (1939,1948,1961b). *Theory of probability*. Clarendon Press.
- Jeffreys, H. (1961a). *Theory of probability*. Oxford Classis texts in the physical sciences, Oxford university press.
- Jemmi, T. and Stephan, R. (2006). *Listeria monocytogenes* : food-borne pathogen and hygiene indicator. *Rev. sci. tech. Off. int. Epiz*, 25(2) :571–580.
- Jongenburger, I., Reij, M., Boer, E., Gorris, L., and Zwietering, M. (2010). Factors influencing the accuracy of the plating method used to enumerate low numbers of viable micro-organisms in food. *International Journal of Food Microbiology*, 143(1-2) :32 – 40.

- Jørgensen, L. and Huss, H. (1998). Prevalence and growth of *listeria monocytogenes* in naturally contaminated seafood. *International Journal of Food Microbiology*, 42 :127–131.
- Kadane, J. (2011). *Principles of uncertainty*. Carnegie Mellon University, Pittsburgh.
- Kass, R. and Raftery, A. (1995). Bayes factor. *Journal of American statistical association*, 430 :773–795.
- Kolmogorov, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer-Verlag.
- Kuhnert, P., Martin, T., and Griffiths, S. (2010). A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters*, 13 :900–914.
- Laplace, P.-S. (1774). Mémoire sur la probabilité des causes par les événements. *Mémoires de Mathématique et de Physique Présentés à l'Académie Royale des Sciences, Par Divers Savans, et Lûs dans ses Assemblées*, 6 :621–656.
- Legan, J., Vandeven, M., Dahms, S., and Cole, M. (2001). Determining the concentration of microorganisms controlled by attribute sampling plans. *Food control*, 12 :137–147.
- Lindley, D. (1975). The future of statistics : A Bayesian 21st century. *Advances in Applied Probability*, 7, Supplement : Proceedings of the Conference on Directions for Mathematical Statistics :106–115.
- Lorimer, M. and Kiermeier, A. (2007). Analysing microbiological data : Tobit or not tobit? *International journal of food microbiology*, 116 :313–318.
- Lund, B. M., Knox, M., and Cole, M. (1989). Destruction of *Listeria monocytogenes* during microwave cooking. *Lancet*, 28 :218.
- Lunn, D., Thomas, A., Best, N., and Spiegelhalter, D. (2000). Winbugs – a bayesian modelling framework : concepts, structure, and extensibility. *Statistics and Computing*, 10 :325–337.
- MapleLeaf (2011). Maple Leaf foods class action settlement.
- Mas-Colell, A. and Monteiro, P. (1991). Self-fulfilling equilibria : an existence theorem for a general state space. Economic Theory Discussion Paper 2, Harvard University.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21 :1087–1092.
- Moala, F. and O'Hagan, A. (2010). Elicitation of multivariate prior distributions : A non-parametric Bayesian approach. *Journal of Statistical Planning and Inference*, 140 :1635–1655.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33 :341–365.

- NFV45-008 (2009). Méthode pour le dénombrement de *Listeria monocytogenes* aux faibles niveaux de contamination dans le saumon fumé et la truite fumée. Technical report, AFNOR.
- O’Hagan, A., Buck, C., Daneshkhah, A., Eiser, R., Garthwaite, P., Jenkinson, D., Oakley, J., and Rakow, T. (2006). *Uncertain judgements : eliciting expert probabilities*.
- Parent, E. and Bernier, J. (2007). *Le raisonnement bayésien*. Springer Verlag, France.
- Philip Dawid, A. and Lauritzen, S. (2000). Compatible prior distributions. ISBA Proceedings 2000.
- Plummer, M. (2008). *JAGS : Just another Gibbs sampler. Version 1.0.3*.
- Pouillot, R., Albert, I., Cornu, M., and Denis, J. (2003). Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. *International journal of food microbiology*, 81(2) :87–104.
- Pouillot, R. and Delignette-Muller, M.-L. (2010). Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages. *International Journal of Food Microbiology*, 142 :330–340.
- Pouillot, R., Goulet, V., Delignette-Muller, M.-L., Mahé, A., and Cornu, M. (2009). Quantitative risk assessment of *Listeria monocytogenes* in French cold smoked salmon : II. risk characterization. *Risk analysis*, 29(6) :809–819.
- Pouillot, R., Miconnet, N., Afchain, A.-L., Delignette-Muller, M.-L., Beaufort, A., Rosso, L., Denis, J.-B., and Cornu, M. (2007). Quantitative risk assessment of *Listeria monocytogenes* in French cold smoked salmon : I. quantitative exposure assessment. *Risk Analysis*, 27(3) :683–700.
- Poumeyrol, G., Morelli, E., Noel, V., and Cornu, M. (2012). Impact of the method chosen for measuring temperatures on the efficacy of rapid cooling of foods in catering facilities. *Food Control*, 23 :345–350.
- R Development Core Team (2011). *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Raftery, A. (1995). Bayesian model selection in social research. *Sociological methodology*, 25 :111–163.
- Regulation (CE) No 178/2002 (2002). laying down the general principles and requirements of food law, establishing the European food safety authority and laying down procedures in matters of food safety. *Official Journal of the European Communities*.
- Regulation (CE) No 2073/2005 (2005). On microbiological criteria for foodstuffs. *Journal officiel de l’Union européenne*, L338 :1–26.

- Reinders, R., De Jonge, R., and Evers, E. (2003). A statistical method to determine whether micro-organisms are randomly distributed in a food matrix, applied to coliforms and *Escherichia coli* O157 in minced beef. *Food microbiology*, 20 :297–303.
- Réglement (CE) No 178/2002 (2002). établissant les principes généraux et les prescriptions générales de la législation alimentaire, instituant l'autorité européenne de sécurité des aliments et fixant des procédures relatives à la sécurité des denrées alimentaires. *Journal officiel de l'Union européenne*, pages 1–24.
- Réglement (CE) No 2073/2005 (2005). Concernant les critères microbiologiques applicables aux denrées alimentaires. *Official Journal of the European Union*, L338 :1–26.
- Robert, C. (2006). *Le choix bayésien*. Springer-Verlag, France.
- Robert, C. and Casella, G. (2004). *Monte-Carlo statistical methods*. Springer, 645p.
- Ross, T. and Sumner, J. (2002). A simple, spreadsheet-based, food safety risk assessment tool. *International Journal of Food Microbiology*, 77 :39–53.
- Rossmannith, P., Krassnig, M., Wagner, M., and Hein, I. (2006). Detection of *listeria monocytogenes* in food using a combined enrichment/real-time PCR method targeting the *prfA* gene. *Research in Microbiology*, 157 :763–771.
- Roverato, A. and Consonni, G. (2004). Compatible prior distributions for directed acyclic graph models. *Journal of the Royal statistic society B*, 66(1) :47–61.
- Ryser, E. and Marth, E. (2007). *Listeria, listeriosis and food safety*. Food science and technology, Boca Raton, FL :CRC Press.
- Sanaa, M., Coroller, L., and Cerf, O. (2004). Risk assessment of listeriosis linked to the consumption of two soft cheeses made from raw milk : Camembert of Normandy and Brie of Meaux. *Risk Analysis*, 24(2) :57–62.
- Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.
- Scharff, D. (2012). Economic burden from health losses due to foodborne illness in the United States. *Journal of Food Protection*, 75(1) :123–131.
- Sosa Mejia, Z., Beumer, R., and Zwietering, M. (2010). Risk evaluation and management to reaching a suggested fso in a steam meal. *Food Microbiology*.
- Spiegelhalter, D., Best, N., Carlin, B., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal statistic society, Series B*, 64(4) :583–639.
- Spiegelhalter, D., Dawid, A., Lauritzen, S., and Cowell, R. (1993). Bayesian analysis in expert systems. *Statistical science*, 8(3) :219–283.
- Spor, A., Dillman, C., Wang, S., de Vienne, D., Sicard, D., and Parent, E. (2010). Hierarchical Bayesian modelling for *Saccharomyces cerevisiae* population dynamics. *International journal of food microbiology*, 142 :25–35.

- Szabo, E. A. and Cahill, M. E. (2009). Nisin and ALTA 2341 inhibit the growth of *listeria monocytogenes* on smoked salmon packaged under vacuum or 100% CO₂. *Risk analysis*, 29(6) :809–819.
- Tebaldi, C., Smith, R., Nychka, D., and Mearns, L. (2005). Quantifying uncertainty in projections of regional climate change : A Bayesian approach to the analysis of multimodel ensembles. *Journal of Climate*, 18(10) :1524–1540.
- Teisl, M. and Roe, B. (2010). Consumer willingness-to-pay to reduce the probability of retail foodborne pathogen contamination. *Food Policy*, 35 :521–530.
- Thomas, A., O’Hara, B., Ligges, U., and Sturtz, S. (2006). Making BUGS open. *R News*, 6 :12–17.
- Traill, W. and Koenig, A. (2010). Economic assessment of food safety standards : Costs and benefits of alternative approaches. *Food Control*, 21 :1611–1619.
- Uyttendaele, M., Busschaert, P., Valero, A., Geeraerd, A., Vermeulen, A., Jaxsens, L., Goh, K., de Loy, A., Van Impe, J., and Devlieghere, F. (2009). Prevalence and challenge tests of *Listeria monocytogenes* in Belgian produced and retailed mayonnaise-based deli-salads, cooked meats products and smoked fish between 2005 and 2007. *International journal of food microbiology*, 133 :94–104.
- van der Gaag, M., Saatkamp, H., Backus, G., van Beek, P., and B.M.H., R. (2004). Cost-effectiveness of controlling *Salmonella* in the pork chain. *Food Control*, 15 :173–180.
- van Schothorst, M., Zwietering, M., Ross, T., Buchanan, R., and Cole, M. (2009). Relating microbiological criteria to food safety objectives and performance objectives. *Food Control*, 20 :967–979.
- Verdonck, F. A. M., Jaworska, J., Thas, O., and Vanrolleghem, P. A. (2001). Determining environmental standards using bootstrapping, Bayesian and maximum likelihood techniques : a comparative study. *Analytica Chimica Acta*, 446(1-2) :427–436.
- Vermeulen, A., Devlieghere, F., De Loy-Hendrickx, A., and Uyttendaele, M. (2011). Critical evaluation of the EU-technical guidance on shelf-life studies for *L. monocytogenes* on RTE-foods : A case study for smoked salmon. *International Journal of Food Microbiology*, 145 :176–185.
- Vogel, B., Huss, H., Ojeniyi, B., and Ahrens, P. (2001). Elucidation of *listeria monocytogenes* contamination routes in cold-smoked salmon processing plants detected by DNA-based typing methods. *Applied and environmental microbiology*, 67(6) :2586–2595.
- von Neumann, J. and Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press, Princeton.
- von Neumann, J. and Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton University Press, Princeton, 2nd edition.

Wilkinson, D. (2008). Decision making and bayes linear methods. Technical report.

WTO (1995). The WTO agreement on the application of sanitary and phytosanitary measures (SPS agreement).