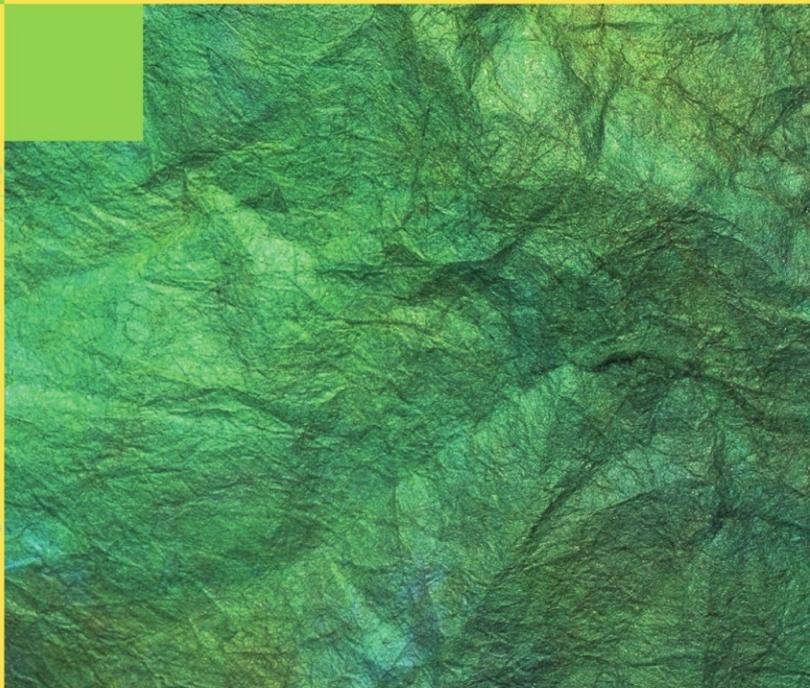


David Makowski, Hervé Monod

# Analyse statistique des risques agro-environnementaux

## Études de cas



$$\phi(n) = \left(1 - \frac{1}{12}\right) \left(1 - \frac{2}{15}\right)$$

$$P(A) = \frac{25}{216} \text{ et } P(B)$$

$$\begin{aligned} (\forall B \in \beta_{\mathbb{R}}) P_X(B) &= P(X^{-1}(B)) \\ &= P(\{\omega \in \Omega \mid X(\omega) \in B\}) \end{aligned}$$

 Springer

$$P(A_2) = 1 - P(\overline{A_2}) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0,491.$$

**Springer**

*Paris*

*Berlin*

*Heidelberg*

*New York*

*Hong Kong*

*Londres*

*Milan*

*Tokyo*

**David Makowski**  
**Hervé Monod**

**Analyse statistique des risques  
agro-environnementaux**  
**Études de cas**

 **Springer**

David Makowski  
Directeur de Recherche INRA  
UMR 211 INRA AgroParisTech  
BP 01  
78850 Thiverval-Grignon

Hervé Monod  
Directeur de Recherche INRA  
Unité MIAJ (UR341)  
78352 Jouy-en-Josas Cedex

ISBN-13 : 978-2-8178-0250-3 Springer Paris Berlin Heidelberg New York

© Springer-Verlag France, 2011

Imprimé en France

Springer-Verlag France est membre du groupe Springer Science + Business Media

Cet ouvrage est soumis au copyright. Tous droits réservés, notamment la reproduction et la représentation, la traduction, la réimpression, l'exposé, la reproduction des illustrations et des tableaux, la transmission par voie d'enregistrement sonore ou visuel, la reproduction par microfilm ou tout autre moyen ainsi que la conservation des banques de données. La loi française sur le copyright du 9 septembre 1965 dans la version en vigueur n'autorise une reproduction intégrale ou partielle que dans certains cas, et en principe moyennant les paiements des droits. Toute représentation, reproduction, contrefaçon ou conservation dans une banque de données par quelque procédé que ce soit est sanctionnée par la loi pénale sur le copyright.

L'utilisation dans cet ouvrage de désignations, dénominations commerciales, marques de fabrique, etc., même sans spécification ne signifie pas que ces termes soient libres de la législation sur les marques de fabrique et la protection des marques et qu'ils puissent être utilisés par chacun.

La maison d'édition décline toute responsabilité quant à l'exactitude des indications de dosage et des modes d'emploi. Dans chaque cas il incombe à l'utilisateur de vérifier les informations données par comparaison à la littérature existante.

*Maquette de couverture : Jean-François Montmarché*



**Collection**  
**Statistique et probabilités appliquées**

**dirigée par Yadolah Dodge**

Professeur Honoraire  
Université de Neuchâtel  
Suisse  
yadolah.dodge@unine.ch

**Comité éditorial :**

**Aurore Delaigle**

Département de mathématiques  
et de statistique  
Université de Melbourne  
Victoria 3010  
Australie

**Christian Genest**

Département de mathématiques  
et de statistique  
Université McGill  
Montréal H3A 2K6  
Canada

**Marc Hallin**

Université libre de Bruxelles  
Campus de la Plaine  
CP 210  
1050 Bruxelles  
Belgique

**Ludovic Lebart**

Télécom-ParisTech  
46, rue Barrault  
75634 Paris Cedex 13  
France

**Christian Mazza**

Département de mathématiques  
Université de Fribourg  
Chemin du Musée 23  
CH-1700 Fribourg  
Suisse

**Stephan Morgenthaler**

École Polytechnique Fédérale  
de Lausanne  
Département de Mathématiques  
1015 Lausanne  
Suisse

**Louis-Paul Rivest**

Département de mathématiques  
et de statistique  
Université Laval  
Québec G1V 0A6  
Canada

**Gilbert Saporta**

Conservatoire national  
des arts et métiers  
292, rue Saint-Martin  
75141 Paris Cedex 3  
France

## Dans la même collection :

- *Statistique. La théorie et ses applications*  
Michel Lejeune, avril 2004
- *Optimisation appliquée*  
Yadolah Dodge, octobre 2004
- *Le choix bayésien. Principes et pratique*  
Christian P. Robert, novembre 2005
- *Régression. Théorie et applications*  
Pierre-André Cornillon, Éric Matzner-Løber, janvier 2007
- *Le raisonnement bayésien. Modélisation et inférence*  
Éric Parent, Jacques Bernier, juillet 2007
- *Premiers pas en simulation*  
Yadolah Dodge, Giuseppe Melfi, juin 2008
- *Génétique statistique*  
Stephan Morgenthaler, juillet 2008
- *Maîtriser l'aléatoire. Exercices résolus de probabilités et statistique, 2<sup>e</sup> édition*  
Eva Cantoni, Philippe Huber, Elvezio Ronchetti, septembre 2009
- *Pratique du calcul bayésien*  
Jean-Jacques Boreux, Éric Parent, décembre 2009
- *Statistique. La théorie et ses applications, 2<sup>e</sup> édition*  
Michel Lejeune, septembre 2010
- *Le logiciel R*  
Pierre Lafaye de Micheaux, Rémy Drouilhet, Benoît Liquet, novembre 2010
- *Probabilités et processus stochastiques*  
Yves Caumel, avril 2011

# Avant-propos

Les ingénieurs et scientifiques sont souvent sollicités pour réaliser des analyses de risque dans le cadre d'expertises pour des agences gouvernementales. Ce type d'analyse comporte généralement trois étapes — l'évaluation du risque, sa gestion et la communication des résultats — qui nécessitent la mise en œuvre de méthodes statistiques et de modèles mathématiques. Dans cet ouvrage, nous nous intéressons à des risques d'un type particulier : les risques associés aux activités agricoles. Ces risques concernent à la fois les produits agricoles et l'environnement, et touchent aussi bien l'agriculteur que le consommateur ou le citoyen.

Cet ouvrage n'a pas la prétention d'être exhaustif mais il constitue une introduction aux principaux types de modèle et aux principales méthodes statistiques utiles pour l'analyse des risques agro-environnementaux. Le chapitre 1 est un chapitre introductif qui souligne le rôle que peuvent jouer les modèles et les méthodes statistiques pour l'analyse des risques agro-environnementaux. Le chapitre 2 présente des notions de base en modélisation et statistique. Le chapitre 3 décrit plusieurs types de modèle utiles pour l'évaluation du risque. Le chapitre 4 décrit des méthodes pour optimiser des décisions et gérer les risques, et le chapitre 5 présente une série de techniques pour l'analyse et la communication de l'incertitude.

Chaque type de modèle et chaque méthode statistique sont illustrés par une ou plusieurs applications. Les programmes informatiques utilisés pour développer les modèles et appliquer les méthodes statistiques sont présentés et commentés en détail. Ils ont tous été réalisés avec des logiciels librement téléchargeables. Les chapitres 2 à 5 incluent également des exercices.

Nous espérons que notre livre sera utile au plus grand nombre et qu'il encouragera les ingénieurs et scientifiques à s'intéresser à la modélisation et à la statistique. Dans ce but, nous avons souhaité que les applications décrites concernent des risques variés : risque de pollution de l'eau par les nitrates, invasion par des espèces nuisibles, diminution de la biodiversité, flux de gènes d'une culture OGM vers une culture non OGM, pertes de rendement et de qualité, émission de gaz à effet de serre, etc. Ces applications ont été réalisées au cours des 10 dernières années en collaboration avec des agronomes, biologistes et écologues appartenant à divers organismes, notamment à l'INRA et à AgroParisTech.

Nous les remercions très chaleureusement. Sans eux, cet ouvrage n'existerait pas.

Nous remercions tout particulièrement, par ordre alphabétique : Katarzyna Adamczyk, Bruno Andrieu, Frédérique Angevin, Aude Barbottin, Nicolas Beaudoin, Liliane Bel, Claire Cadet, Marion Casagrande, Bruno Chauvel, Jean-Jacques Daudin, Jean-Baptiste Denis, Thierry Doré, Sabah Ennaïfar, Robert Faivre, Benoît Gabrielle, Arnaud Gauffreteau, Martine Guérif, Laurence Guichard, Jonathan Hillier, Marie-Hélène Jeuffroy, Jim W. Jones, Anne Lacroix, Matieyendou Lamboni, François Laurent, Marc Lavielle, Marianne Le Bail, Chantal Loyce, Philippe Lucas, Alexandra Maltas, Antoine Messéan, Jean-Marc Meynard, Marie Morfin, Nicolas Munier-Jolain, Cédric Naud, Eric Parent, Annette Penaud, Aurore Philibert, Sophie Primot, Lorène Prost, Marie Taverne, Muriel Tichit, Laurent Ruck, Muriel Valantin-Morison, Valérie Viaud, Daniel Wallach.

Nous remercions également Yadolah Dodge, Stephan Morgenthaler et Charles Ruelle pour avoir soutenu et accompagné notre projet d'ouvrage dès ses premières versions.

Nous exprimons notre reconnaissance tout particulièrement à Stephan, dont les relectures détaillées et les nombreux conseils ont permis d'améliorer très fortement la qualité de l'ouvrage.<sup>1</sup>

David Makowski et Hervé Monod  
INRA France

---

1. Cet ouvrage a été réalisé avec le style L<sup>A</sup>T<sub>E</sub>X « iris » développé par Nicolas Puech, avec l'aide et les suggestions de Patrik Fuhrer, Michel Lejeune, Bruno Petazzoni, Laurent Decreusefond et Philippe Martins.

# Sommaire

<b>Sommaire</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Notions de base</b>	<b>5</b>
2.1 Variables aléatoires et lois de probabilité . . . . .	5
2.1.1 Variable aléatoire . . . . .	5
2.1.2 Variables aléatoires discrètes . . . . .	6
2.1.3 Variables aléatoires continues . . . . .	7
2.1.4 Valeurs caractéristiques d'une variable aléatoire . . . . .	10
2.1.5 Dépendance entre variables aléatoires . . . . .	12
2.2 La notion de modèle en statistique . . . . .	14
2.2.1 Description . . . . .	14
2.2.2 Fonction de vraisemblance d'un modèle statistique . . . . .	15
2.3 Inférence statistique . . . . .	16
2.3.1 Approche fréquentiste et approche bayésienne . . . . .	16
2.3.2 Estimateur . . . . .	17
2.3.3 Test statistique et intervalle de confiance . . . . .	18
2.3.4 Inférence bayésienne . . . . .	19
2.4 Les quatre étapes de la modélisation . . . . .	20
2.4.1 Définition des variables . . . . .	20
2.4.2 Choix des équations . . . . .	22
2.4.3 Estimation des paramètres . . . . .	22
2.4.4 Évaluation des modèles . . . . .	23
2.4.5 Importance de la planification expérimentale . . . . .	23
2.5 Exercices . . . . .	24
<b>3 Modèles statistiques et évaluation des risques</b>	<b>27</b>
3.1 Modèle linéaire . . . . .	27
3.1.1 Définition . . . . .	27
3.1.2 Généralité du modèle linéaire . . . . .	28
3.1.3 Estimation des paramètres . . . . .	29
3.1.4 Évaluation et limites du modèle linéaire . . . . .	30

3.1.5	Exemple : prédiction de la teneur en azote et de la teneur en protéines des grains de blé . . . . .	31
3.2	Modèle linéaire généralisé . . . . .	48
3.2.1	Définition . . . . .	48
3.2.2	Exemple : présence/absence d'oiseaux dans une prairie . . . . .	49
3.3	Modèle non linéaire . . . . .	58
3.3.1	Définition . . . . .	58
3.3.2	Exemple : reliquat d'azote dans le sol à la récolte . . . . .	59
3.4	Modèle hiérarchique . . . . .	69
3.4.1	Définition et intérêt . . . . .	69
3.4.2	Exemple : reliquat d'azote dans le sol à la récolte . . . . .	72
3.4.3	Exemple : variabilité intra-parcellaire des densités de mauvaises herbes . . . . .	83
3.5	Estimation de valeurs extrêmes par régression quantile . . . . .	87
3.5.1	Définition . . . . .	87
3.5.2	Exemple : risque de sclérotinia du colza . . . . .	88
3.6	Exercices . . . . .	94
<b>4</b>	<b>Optimisation des décisions et gestion des risques</b>	<b>97</b>
4.1	Les quatre étapes de l'optimisation . . . . .	97
4.1.1	Présentation . . . . .	97
4.1.2	Exemple : détermination d'une température optimale pour le traitement thermique du bois destiné à l'exportation . . . . .	99
4.2	Optimisation d'une règle de décision binaire par analyse ROC . . . . .	102
4.2.1	Introduction . . . . .	102
4.2.2	Règle de décision binaire et ses deux types d'erreur . . . . .	102
4.2.3	Estimation et évaluation par la méthode ROC . . . . .	103
4.2.4	Exemple : gestion du risque d'invasion par les mauvaises herbes . . . . .	104
4.2.5	Exemple : gestion du risque de sclérotinia du colza . . . . .	110
4.3	Optimisation d'une variable décisionnelle par simulation . . . . .	117
4.3.1	Méthode . . . . .	117
4.3.2	Exemple : calcul de doses optimales d'engrais . . . . .	118
4.4	Exercices . . . . .	123
<b>5</b>	<b>Analyse et communication de l'incertitude</b>	<b>125</b>
5.1	Les différents types d'incertitude et leurs conséquences . . . . .	125
5.2	Décrire l'incertitude par des distributions de probabilité . . . . .	126
5.2.1	Objectif . . . . .	126
5.2.2	Exemple basé sur des calculs analytiques : risque d'invasion par une espèce nuisible . . . . .	127

---

5.2.3	Exemple basé sur des simulations de Monte-Carlo : reliquat d'azote dans le sol . . . . .	129
5.2.4	Exemple combinant un modèle dynamique et des mesures en cours de saison : estimation du carbone du sol . . . . .	134
5.3	Calculer des indices de sensibilité . . . . .	138
5.3.1	Objectifs et définitions . . . . .	138
5.3.2	Exemple basé sur des simulations de Monte-Carlo : reliquat d'azote minéral dans le sol . . . . .	139
5.3.3	Exemple basé sur des simulations planifiées : influence du parcellaire sur les flux de gènes . . . . .	142
5.4	Exercices . . . . .	145
<b>6</b>	<b>Recommandations</b>	<b>149</b>
6.1	Quelques conseils . . . . .	149
6.2	Pour continuer . . . . .	150
	<b>Bibliographie</b>	<b>151</b>
	<b>Index</b>	<b>159</b>



Collection  
**Statistique  
et probabilités  
appliquées**

**Dirigée par  
Yadolah Dodge**

COMITÉ ÉDITORIAL :

**Aurore Delaigle**

Université de Melbourne, Australie

**Christian Genest**

Université McGill, Montréal

**Marc Hallin**

Université libre de Bruxelles, Belgique

**Ludovic Lebart**

Télécom-ParisTech, Paris

**Christian Mazza**

Université de Fribourg, Suisse

**Stephan Morgenthaler**

EPFL, Lausanne

**Louis-Paul Rivest**

Université Laval, Québec

**Gilbert Saporta**

CNAM, Paris

**David Makowski, Hervé Monod**

# Analyse statistique des risques agro-environnementaux

## Études de cas

*Cette collection met à la disposition du public intéressé par la statistique (étudiants, enseignants, chercheurs) des ouvrages qui concilient effort pédagogique et travail permanent de mise à jour.*

*Cette démarche implique de prendre en compte de façon sélective et critique les renouvellements des concepts, des champs d'application et des outils de traitement. Seules une compréhension profonde et une appropriation des connaissances permettront de s'adapter aux évolutions qui n'ont pas fini de bouleverser cette discipline.*

Conçu comme un véritable manuel pratique, ce livre est une introduction aux méthodes statistiques les plus couramment utilisées pour l'analyse des risques agro-environnementaux. Celles-ci peuvent être regroupées au sein de trois grandes sections :

- La modélisation des risques en fonction de facteurs environnementaux et anthropiques (modèle linéaire, modèle linéaire généralisé, modèle non linéaire, modèle hiérarchique, régression quantile) ;
- L'optimisation de décisions ou de règles de décision pour mieux gérer les risques, en intégrant des variables décisionnelles dans les modèles (optimisation de seuils de décision, optimisation par simulation, analyses ROC) ;
- L'analyse et la communication des incertitudes associées aux modèles (estimation et description de distributions de probabilité, assimilation de données, analyse de sensibilité).

L'utilisation de chaque méthode est illustrée par une ou plusieurs applications à des problèmes concrets (pollution de l'eau par les nitrates, invasion par des espèces nuisibles, flux de gènes d'une culture OGM vers une culture non OGM, etc.). Les programmes informatiques R ou WinBUGS utilisés dans les exemples sont présentés et commentés en détail. À la fin de chaque chapitre, des exercices permettront aux lecteurs de tester leur compréhension des méthodes étudiées.

30 € TTC

ISBN : 978-2-8178-0250-3



9 782817 802503

 [springer.com](http://springer.com)