

# Statistique bayésienne en écologie

Marie-Pierre Etienne thx to E. Rivot

<https://github.com/MarieEtienne>

Novembre 2018



# Outline

- ① Objectifs
- ② Réflexion sur la modélisation
- ③ L'approche bayésienne
- ④ Bilan

# Outline

- 1 Objectifs
- 2 Réflexion sur la modélisation
- 3 L'approche bayésienne
- 4 Bilan

# Objectifs bayésiens

- Principes fondamentaux
- Liens Ressemblances et différences par rapport à l'analyse statistique fréquentiste ("classique")
- Intérêts / Limites
- Calcul Bayésien dans la pratique (simulations MC)

# Objectifs de modélisation

*Modèles hiérarchiques* : modèles à espace d'état (state-space models), modèles intégrés ... Gérer les multiples source d'incertitude dans la modélisation

*Outils d'inférence et de décision* Simulations de Monte Carlo Propagation des incertitudes (analyses de risque)

# Outline

- ① Objectifs
- ② Réflexion sur la modélisation
- ③ L'approche bayésienne
- ④ Bilan

# Dans quel but

- Un modèle pour formaliser des connaissances ou des hypothèses
- confronter le modèles aux données
- estimation de paramètre
- comparaison de modèles
- prédiction / projection

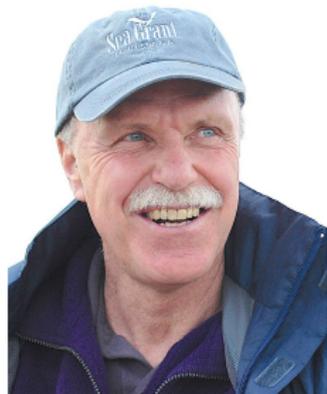
Modélisation statistique = mixer des connaissance et des données

# Difficultés

- Quantifier notre incertitude et ses conséquences
- Intégrer des données hétérogènes
- Des techniques d'estimation flexibles pour utiliser des modèles complexes (mécanistes) en prenant en compte l'incertitude

## Communiquer sur l'incertitude

*Lets go back to testable hypotheses and evidence. Avoid faith-based support for ideas, or scientists will become advocates! An advocate knows the answer and looks evidence to support it. A scientist asks nature (=data) how much support there is for competing hypotheses (Ray Hilborn, 2006 Faith-based fisheries. Fisheries, 31(11): 554-555)*



# Utiliser l'incertitude pour quantifier les risques

Le risque = probabilité qu'un évènement indésirable se produise

Si on veut gérer un système, il faut quantifier le risque et l'inclure dans la décision.

Ignorer l'incertitude c'est minimiser le risque et se montrer sur confiant dans notre capacité à gérer un système.

L'approche bayésienne permet de facilement propager l'incertitude dans un système

# Formule Bayes

- proposée par Sir Reverend Thomas BAYES (1702-1761) et indépendamment par Pierre-Simon LAPLACE (1749-1827)

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- Très simple mais conduit à des calculs difficiles voire impossibles. Peu utilisée avant l'avènement des ordinateurs (après 1970, 1980).

# Formule Bayes en pratique

- Algorithmes numériques, qui demande de grandes ressources de calculs : Monte Carlo par chaînes de Markov (MCMC), méthodes particulières, Monte Carlo séquentielle, approximation de Laplace



Source : Science et Vie, Novembre 2012

# Formule de Bayes en écologie

## Publications

- Efron B en 1986. Why isn't every one a Bayesian ? (with comments and reply). *American Statistician*, 40:1-11
- Aaron M. Ellison (2004) Bayesian inference in ecology, *Ecology letters* *The application of Bayesian inference to ecological questions has blossomed since the publication in 1996 of a series of papers on Bayesian inference for ecological research and environmental decision making (Dixon & Ellison 1996).*
- Clark, J. S. (2005). Why environmental scientists are becoming Bayesians? *Ecology Letters*, 8, 2-14

# Pourquoi un tel succès

- Flexibilité dans la modélisation
- Aide à la décision en univers incertain
- Intégration de données hétérogènes
- des algorithmes d'estimation efficaces et **surtout** des logiciels qui les rendent accessibles (WinBugs, 1989 ; Jags, 2003 ; INLA, 2009 ; Stan, 2011 ; Nimble, 2014; ... )

# Outline

- ① Objectifs
- ② Réflexion sur la modélisation
- ③ L'approche bayésienne**
- ④ Bilan

# Penser l'incertitude comme une probabilité

Utiliser les lois de probabilité pour représenter notre connaissance et non seulement une variabilité

# Exemple 1

**Exemple** on s'intéresse à la probabilité d'une pièce à tomber sur Pile

**A priori** que peut on en dire ?

**Données** on lance 5 fois une pièce on obtient la séquence *PFFF*

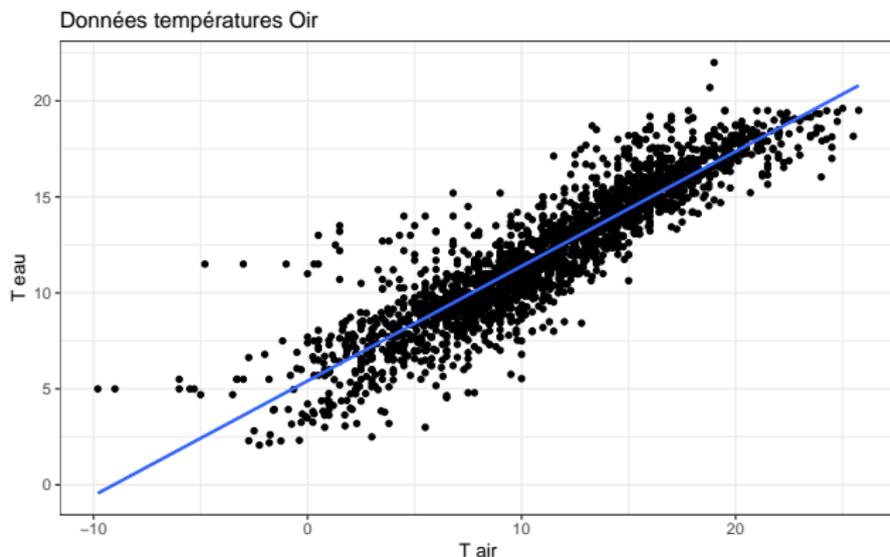
**A posteriori** comment mettre à jour notre perception de cette probabilité ?

## Exemple 2: la regression simple version bayésienne

**Exemple** Quel est le lien entre température de l'air et température de l'eau ?

**A priori** que peut on en dire ?

**Données**



**A posteriori** comment mettre à jour notre perception de cette probabilité ?

## Exemple 1 mis à jour

**Données historique** la séquence  $PFFF$

**Nouvelles Données** on lance 2 fois la même pièce on obtient la séquence  $PP$

**A posteriori** comment mettre à jour notre perception de cette probabilité ?

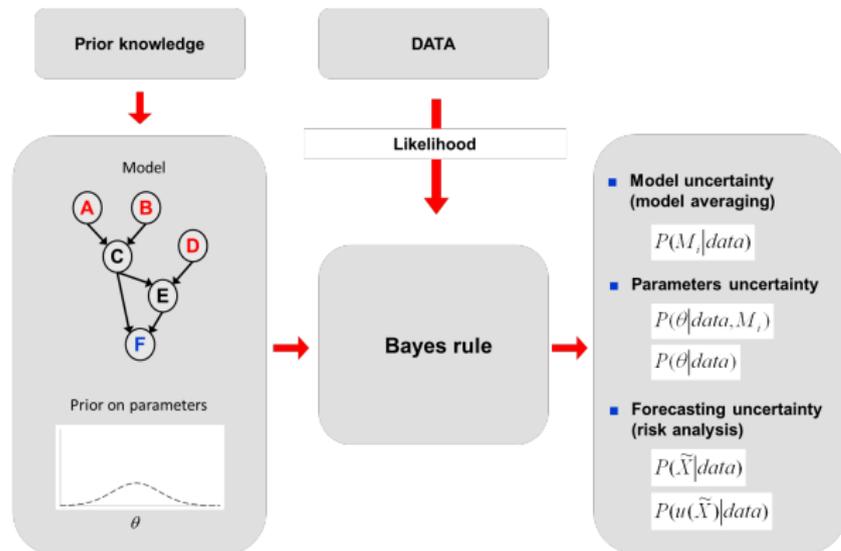
# Intervalle de confiance vs Intervalle de crédibilité

Quelles différences ?

# Prédiction en univers incertain

Exemple 1 : Quelle est la probabilité d'observer pile au prochain tirage ?

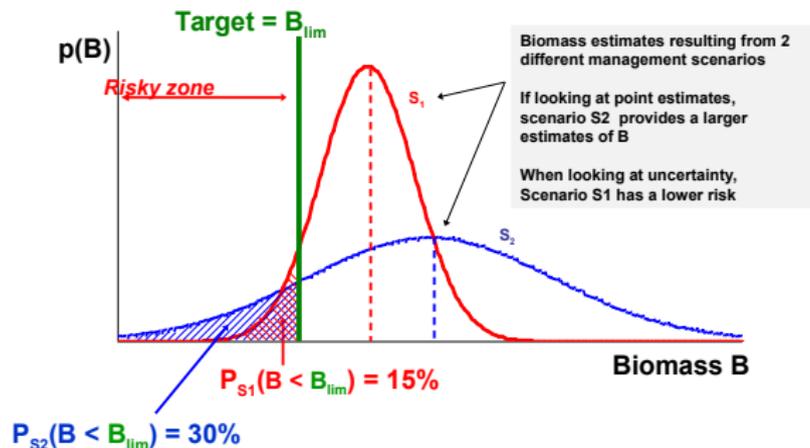
# Approche bayésienne en 1 schéma



Source : E. Rivot lecture

# Préserver une espèce dans un monde incertain

- Accounting for uncertainty may lead to « re-rank » the management options



Source : E. Rivot lecture

# Intégrer l'incertitude de modèle

Moyenner des prédictions données par plusieurs modèles.

# Outline

- ① Objectifs
- ② Réflexion sur la modélisation
- ③ L'approche bayésienne
- ④ Bilan

# Approche bayésienne

- Pour intégrer toutes les sources d'incertitudes
- pour intégrer l'information a priori
- profiter d'algorithmes d'estimation performant

# Attention

- les questions de modélisation demeurent (identifiabilité notamment)
- le choix de modèle reste un problème difficile
- les algorithmes donnent toujours un résultat, mais est-il valide ?