

# Un apport de l'IA aux crises épidémiologiques Kiss the Covid

P. Mathieu<sup>a,1,\*</sup>, N. Mauhé<sup>a</sup>

<sup>a</sup>*Univ. Lille, CNRS, Centrale Lille, UMR 9189 – CRISTAL Lab, F-59000 Lille, France*

---

## Abstract

Lors d'une crise sanitaire liée à l'apparition d'une maladie infectieuse, il est impératif d'avoir des systèmes d'aide à la décision des politiques publiques (évolution de la maladie, immunité collective, nombres dates et durées de périodes de confinement).

L'évolution d'une maladie infectieuse peut-être modélisée de différentes façons (approche centrée-individus, réseaux sociaux, equations mathématiques, modèles de flux).

Toutes ces modélisations necessite l'usage de paramètres pour calibrer le modèle. Plus il y a de paramètres, plus il est facile de coller aux données (tout comme il existe toujours un polynome de degrés  $n$  qui passe par  $n$  points).

Certains modélisateurs pensent que plus les paramètres sont précis (et nombreux), plus le modèle sera fiable. C'est un course sans fin.

L'important est de trouver l'essence du problème, le modèle le plus économe en paramètres (parcimonieux) rendant compte des phénomènes observés.

L'IA est d'une grande aide pour éviter les biais et régler au mieux ces modèles.

*Keywords:* Infectious disease, SIR modelling, Decision support system, artificial intelligence

*2010 MSC:* 00-01, 99-00

---

---

\*Corresponding author

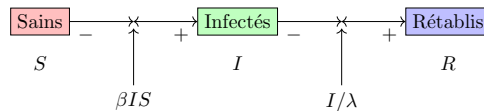
*Email addresses:* philippe.mathieu@univ-lille.fr (P. Mathieu),  
nicolas.mauhe@univ-lille.fr (N. Mauhé)

## 1. Introduction (ce qui se fait en épidémiologie)

Depuis exactement un siècle et le papier séminal de Kermack et McKendrick, la modélisation épidémiologique a pris une part prépondérante dans l'étude de ces phénomènes complexes. La crise du H1N1 en 2008 a nettement amplifié ce phénomène, faisant apparaître un nombre important de modèles plus complexes les uns que les autres. Dans ce papier nous montrons que les techniques d'intelligence artificielle Utilisées à bon escient, Permettent d'une part de concevoir des modèles plus faciles à comprendre et à utiliser par les thématiques et d'autre part de faciliter grandement la conception d'outils d'aide à la décision pour les politiques sanitaires.

Le modèle présenté s'appuie sur des taux de contagion adaptable chaque jour, dont la valeur est automatiquement calculée par apprentissage génétique.

Le modèle SIR de Kermack et McKendrick [1] a été proposé en 1927 pour expliquer la hausse et la baisse rapides du nombre de patients infectés observées lors d'épidémies telles que la peste (à Londres en 1665, à Bombay en 1906) ou le choléra (à Londres en 1865). Il a été remis à l'avant-plan à la fin des années 1970 par Anderson et May [2].



## 2. L'apport de l'informatique

L'usage de l'IA est une méthode nouvelle pour l'épidémiologie, mais ancienne pour l'informatique.

Certains pensent que si on complexifie le modèle, on s'approche de la réalité : or plus il y a de paramètres, plus c'est facile de "fitter" : Problématique de l'overfitting (mettre une belle courbe qui passe par n points alors qu'en fait c'est une droite qui généralise le mieux).

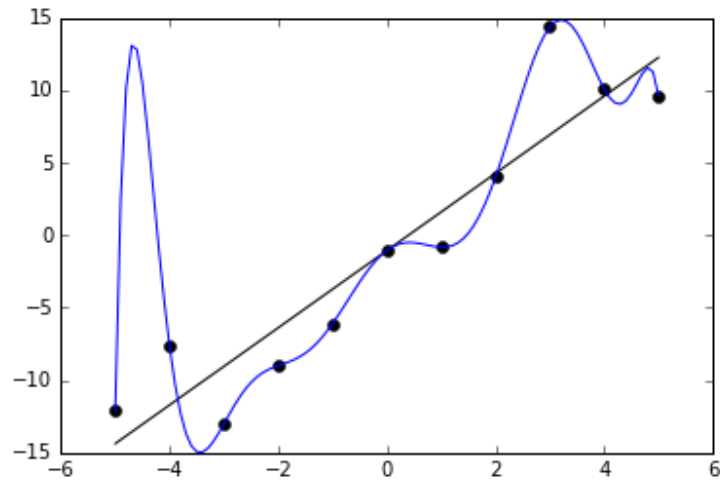


Figure 1: ©<https://en.wikipedia.org/wiki/Overfitting> Des données bruitées (approximativement linéaires) sont approximées par une fonction linéaire et une fonction polynomiale. Bien que la fonction polynomiale soit parfaitement ajustée, on constate que la fonction linéaire généralise surement bien mieux

Un modèle qui ne fonctionne pas (parce que par exemple des  $R_0$  sont délirants) peut très bien néanmoins “fitter” les données et donc continuer à prédire correctement les personnes en Réa

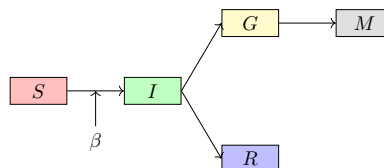
L’apport de l’informatique :

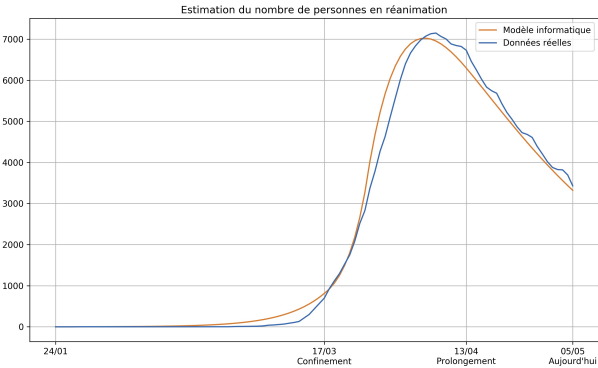
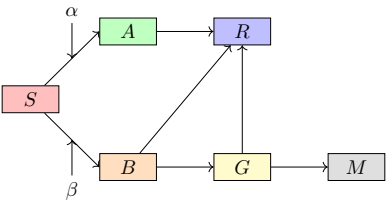
30

- des techniques et des outils : algos d’apprentissage
- s’attacher à l’essence du problème : principe de parcimonie

### 3. Hypothèse de simplicité

Avec quelque chose de très simple on “fitter” les données





References