

## Prédiction des incidences par la méthode des analogues

### Syndromes grippaux et diarrhées aiguës

Afin de prédire les incidences des syndromes grippaux et des diarrhées aiguës à court terme, une méthode - mise au point à l'origine pour application dans le champ de la météorologie, a été implémentée. Cette méthode, appelée méthode des analogues consiste à prédire le futur à l'aide des données passées en recherchant les comportements similaires.

La méthodologie est présentée succinctement dans ce document en prenant exemple sur la prédiction de l'incidence des syndromes grippaux. Un article scientifique publié en 2006 détaille l'ensemble de la méthode : [Viboud C, Boëlle PY, Carrat F, Valleron AJ, Flahault A. Prediction of the spread of influenza epidemics by the method of analogues. Am J Epidemiol. 2003. 158\(10\):996-1006.](#)

Le principe de la méthode des analogues (illustré dans la Figure 1) est de sélectionner dans les données historiques des sections d'incidences de longueur  $l$  semaines qui ressemblent le plus à la section d'incidences de même longueur observées au temps  $T$  actuel (illustrée dans la figure 1). Cette section d'incidence au temps  $T$  s'exprime de la façon suivante :  $X(T) = (I(T), I(T-1), \dots, I(T-l))$ , où  $I(T)$  est l'incidence au temps  $T$ . La sélection des sections d'incidences passées  $X(t)$  est réalisée grâce à un critère de distance s'exprimant de la façon suivante :

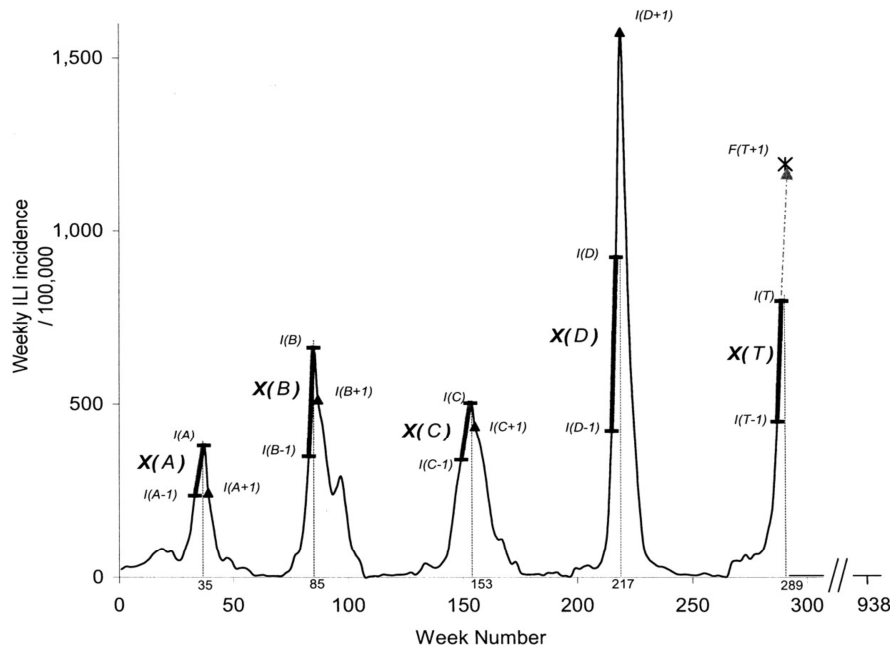
$$\text{dist}(X(T), X(t)) = \sum_{j=0}^l (I(T-j) - I(t-j))^2$$

La prédiction des futures incidences est une moyenne pondérée des valeurs d'incidences observées dans les semaines suivant les sections historiques sélectionnées, selon la formule suivante :

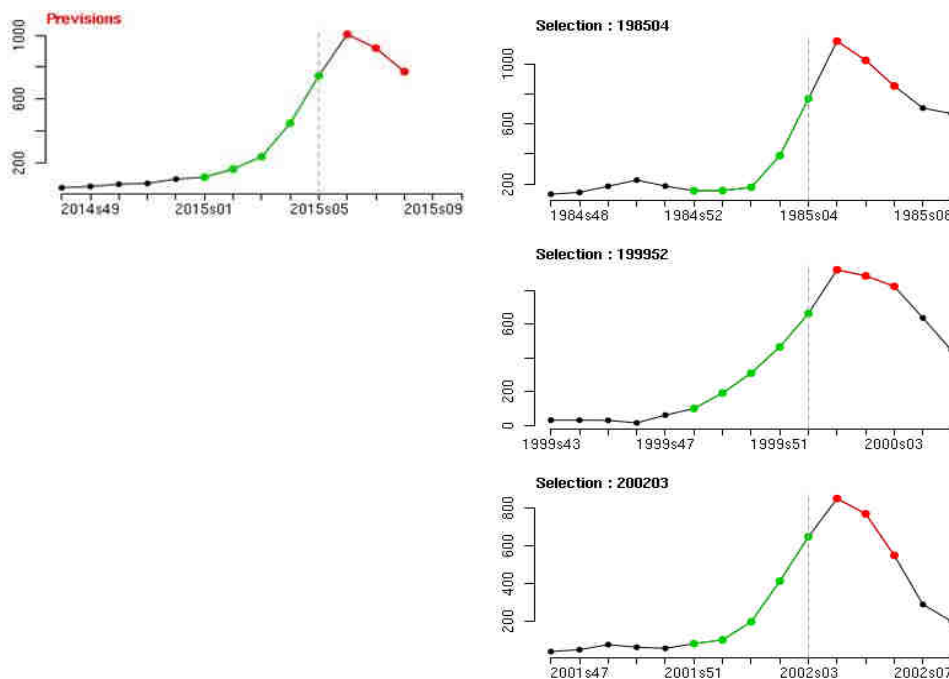
$$F(T+h) = \sum_{1 \leq i \leq v} w^i \cdot I(T^i + h)$$

Où  $F(T+h)$  est la prédiction pour la semaine  $T+h$ ,  $w^i$  est le poids assigné à la  $i$ -ème section sélectionnée (inverse de la distance à  $X(T)$ ),  $T^i$  est la semaine de fin de la  $i$ -ème section sélectionnée.

Dans le cadre de la prédiction des incidences des syndromes grippaux, les sections d'incidences recherchées dans les données passées sont de 5 semaines consécutives. La prédiction est réalisée pour les trois semaines suivantes. Un exemple est proposé dans la Figure 2, pour la semaine  $T=201505$ .



**Figure 1. Principe de la méthode des analogues pour la prédiction de l'incidence des syndromes grippaux, France, 1984-2002.** La série chronologique des incidences observées est représentée par une ligne fine ; les vecteurs sont des lignes épaisses. Le vecteur d'incidence mesuré au cours des deux dernières semaines,  $X(T) = (I(T), I(T - 1))$ , représente l'activité grippale actuelle. Les quatre vecteurs historiques des mesures d'incidence ( $X(A)$ - $X(D)$ ) correspondent le mieux à  $X(T)$ . La prévision  $F(T + 1)$  (étoile) est calculée comme une moyenne pondérée des incidences observées suivant les vecteurs historiques (triangles). Dans cette figure, les deux voisins les plus proches,  $X(B)$  et  $X(D)$ , sont utilisés, et la prévision repose sur  $I(B + 1)$  et  $I(D + 1)$ .



**Figure 2. Exemple d'utilisation de la méthode des analogues pour la prédiction de l'incidence des syndromes grippaux, France, semaine 2015s05.** La figure présente les données observées les plus récentes (à gauche en noir et vert) ainsi que leur prévisions (points rouges). A droite sont représentées les sections d'incidences sélectionnées par la méthode des analogues selon leur distance aux données observées actuelles (en vert), les points rouges sont les données utilisées pour la prédiction.