



Sums of Smooth Exponentials



1 Model

Model

Sums of smooth exponentials

- **But du modèle** : modéliser des courbes de séries de dénombrement (ici les décès y_i) avec i l'indice pour l'âge. Les décès y_i ont une moyenne μ_i . On modélise cette moyenne en la divisant en trois composantes : $\gamma_k, k = \{1, 2, 3\}$. Le taux de mortalité μ_i est la somme des trois composantes pondérée par l'exposition $\mu_i = \sum_{k=1}^3 e_i \gamma_{ik}$.

- Chaque composante représente un intervalle d'âges et est modélisée par des coefficients associés aux B-Splines

$$\gamma_{ik} = \exp\left(\sum_{j=1}^{J_k} B_{jk}(x_i) \alpha_{jk}\right)$$

- Le nombre de Splines par composantes est déterminé de manière arbitraire et correspond à un découpage optimal de l'intervalle des âges.

Equation du modèle :

$$\mu_i = e_i \left(\exp\left(\sum_{j=1}^2 B_{j1}(x_i) \alpha_{j1}\right) + \exp\left(\sum_{j=1}^{25} B_{j2}(x_i) \alpha_{j2}\right) + \exp\left(\sum_{j=1}^{25} B_{j3}(x_i) \alpha_{j3}\right) \right)$$

k	1 γ_{i1}	2 γ_{i2}	3 γ_{i3}
Âges	1 – 50	1 – 110	1 – 80
Nombre de Splines J_k	2	25	25
Période modélisée	Mortalité infantile	Mortalité senescente	Accident hump

Model

B-Spline Smoothing with Penalties



FIG. 1. Illustrations of one isolated B-spline and several overlapping ones (a) degree 1; (b) degree 2.

Une B-Spline est une combinaison de morceaux polynomiaux. Les B-Splines de degré 2 consistent en 3 morceaux de fonctions quadratiques et 4 nœuds adjacents : x_1 , x_2 , x_3 et x_4 (Figure 1). Une courbe de données (x_i, y_i) est modélisée par une combinaison linéaire $\hat{y} = \sum_{j=1}^n \hat{a}_j B_j(x)$ avec n le nombre de B-Splines correspondant à $n = \text{nombre de nœuds} + \text{degré des Splines}$ et \hat{a}_j les coefficients. Le fit se fait de la manière d'une régression des points de données (x_i, y_i) sur le set de B-Splines, par minimisation d'une fonction objectif. Les pénalité λ_k sont ajoutées sur les différences des coefficients des Splines adjacentes pour permettre le lissage. La fonction objectif est la déviance du modèle. Les pénalités optimales sont déterminées en minimisant le critère BIC.

$$Q = \text{DEV}(y|\mu) + \sum_{k=1}^K \lambda_k \|D_k \alpha_k\|^2 = 2 \sum_{i=1}^m y_i \ln \left(\frac{y_i}{\mu_i} \right) + \sum_{k=1}^K \lambda_k \|D_k \alpha_k\|^2$$



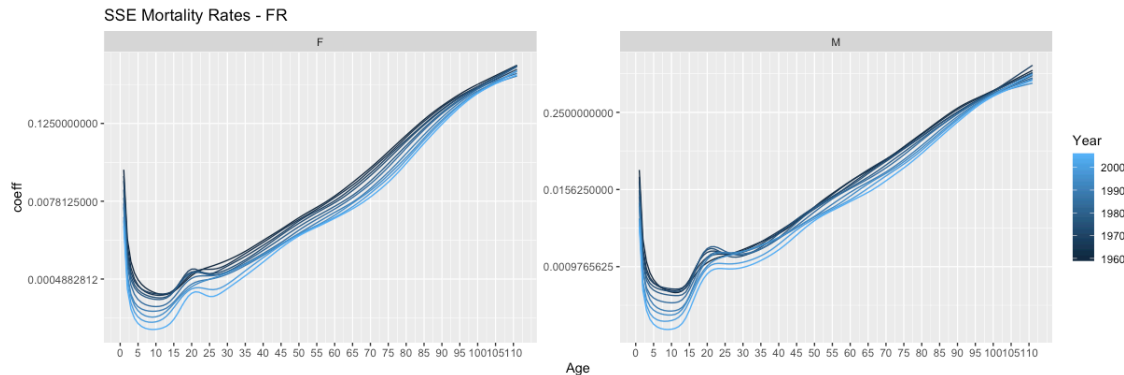
2

Fit des données

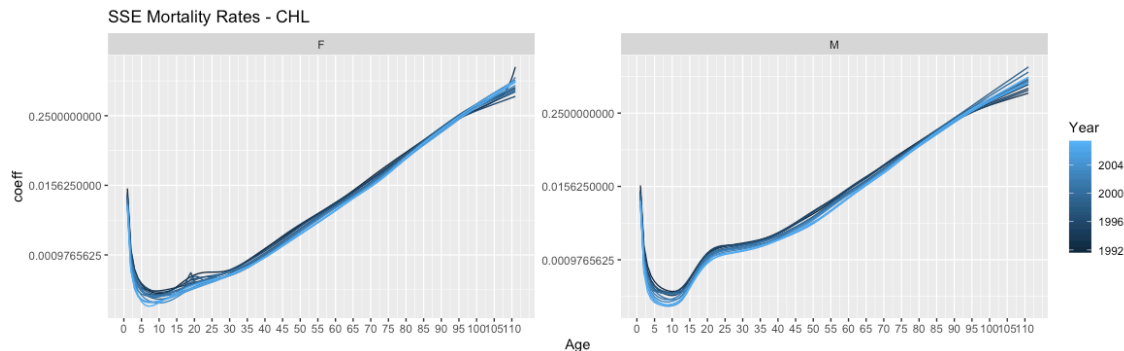
Fit des données

Taux de mortalité fittés

FRANCE
1960-1990



CHILI
1992-2007



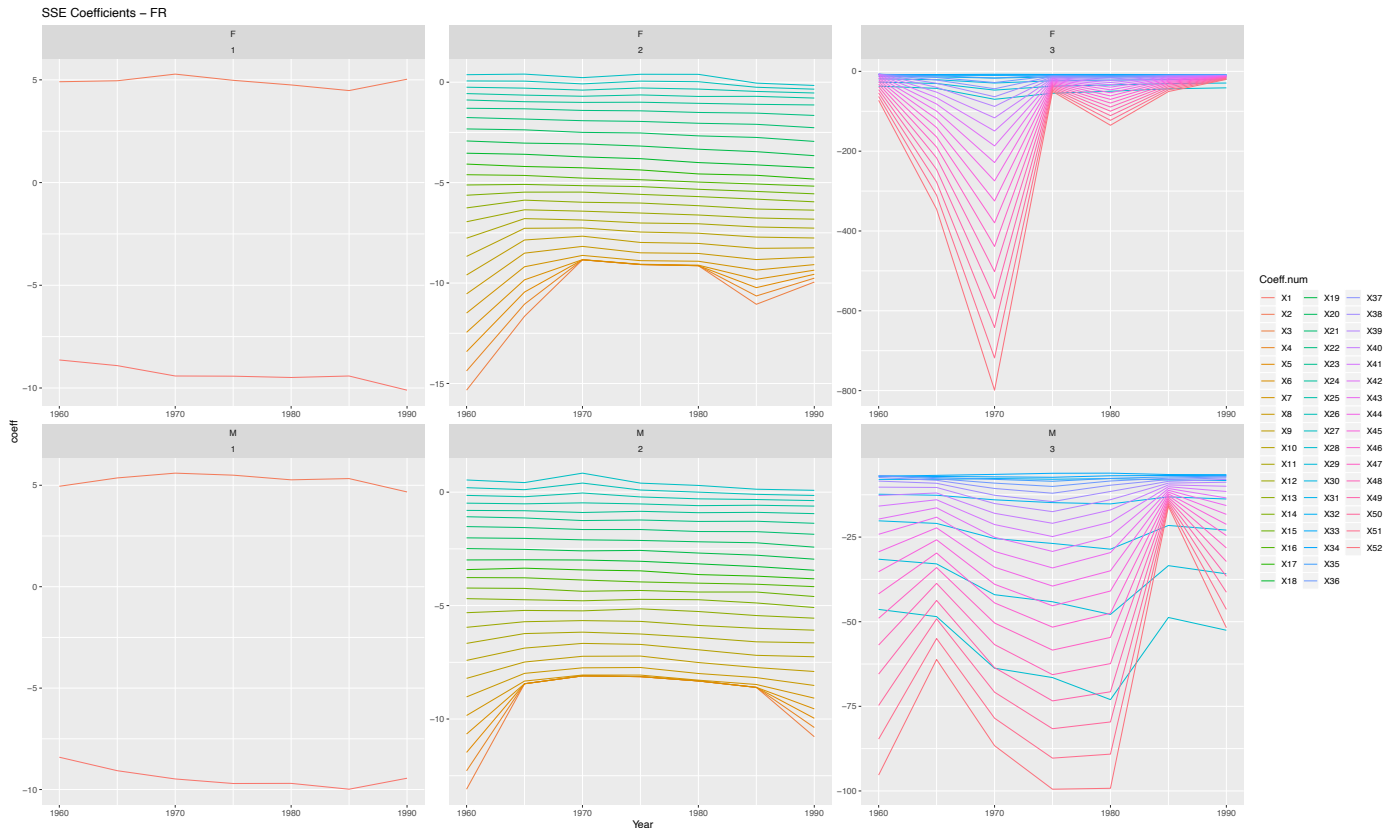
Fit des données

Taux de mortalité fittés



Coefficients α_{jk} estimés pour les Femmes (en haut) et les Hommes (en bas), groupés par composantes.

Les coefficients à droite sont représentés avec en abscisses les périodes et en ordonnées la valeur montant des coefficients. Ils varient selon la période mais ne sont pas dépendant de l'âge. Pour une composante : $\gamma_{ik} = \exp\left(\sum_{j=1}^J B_{jk}(x_i)\alpha_{jk}\right)$. Le alpha est multiplié par la spline pour tous les âges x_i .



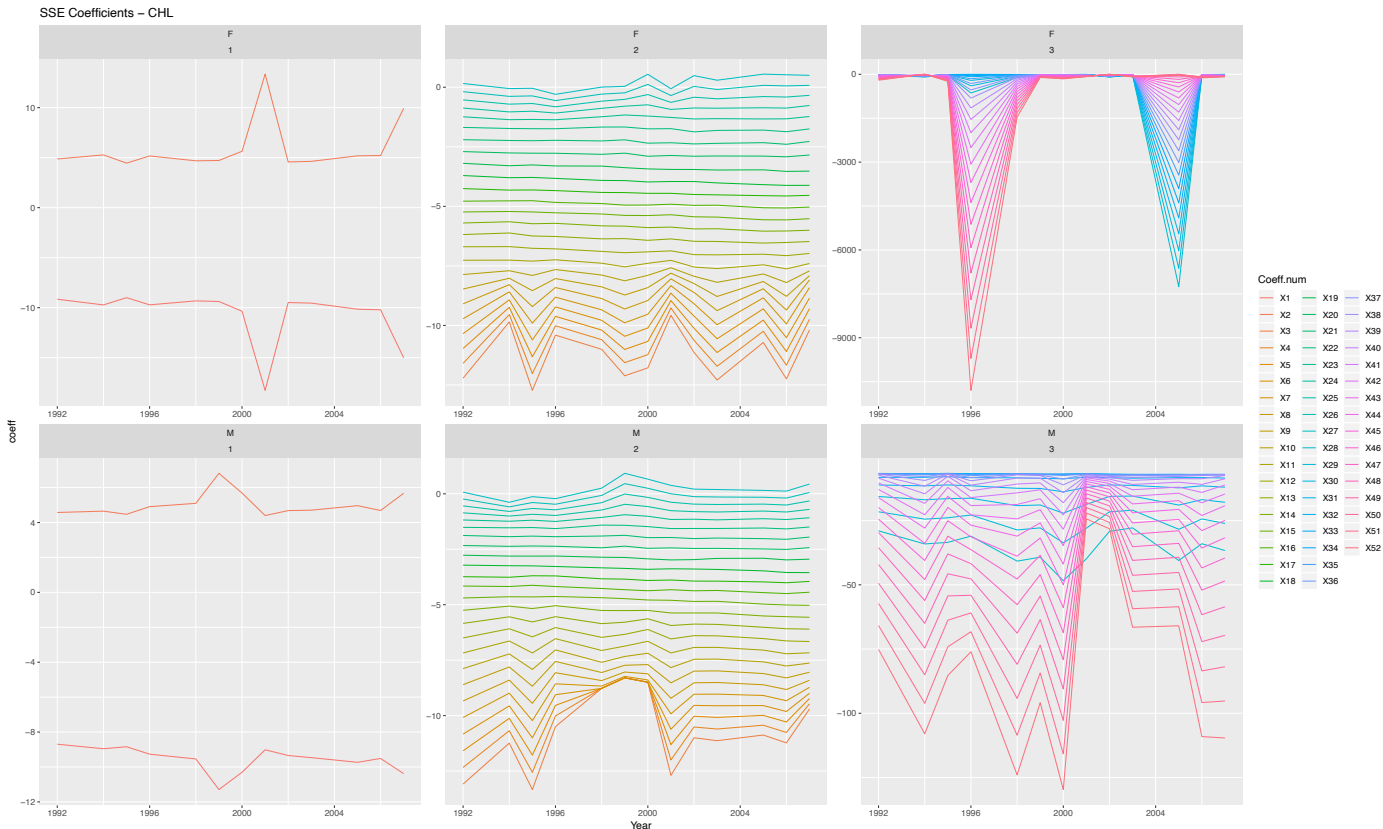
Fit des données

Taux de mortalité fittés



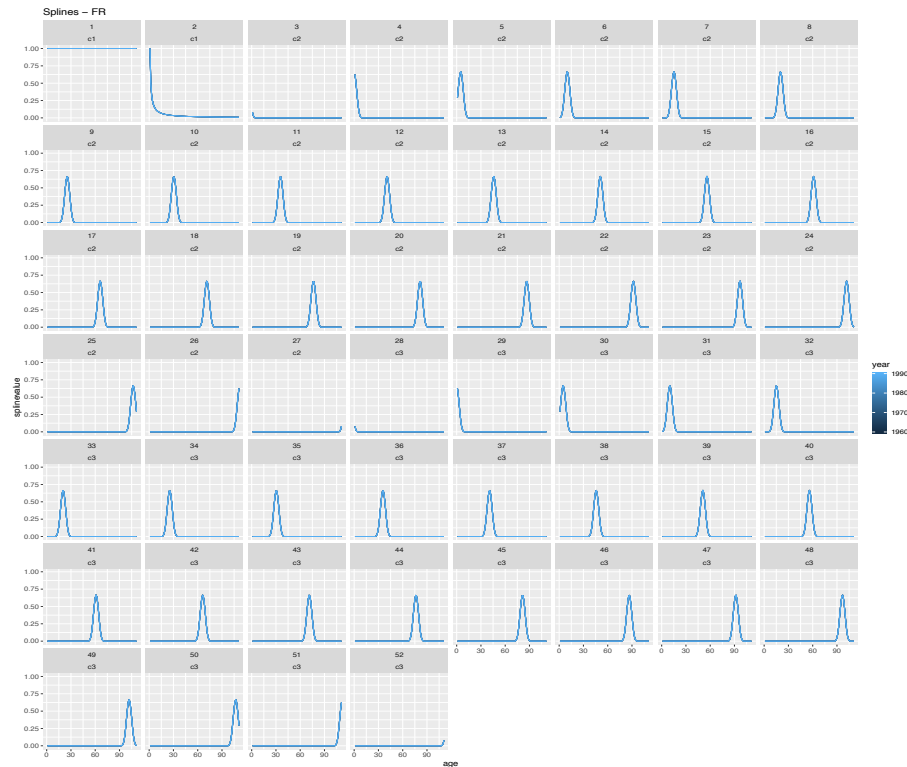
Coefficients α_{jk} estimés pour les Femmes (en haut) et les Hommes (en bas), groupés par composantes.

Les coefficients à droite sont représentés avec en abscisses les périodes et en ordonnées la valeur montant des coefficients. Ils varient selon la période mais ne sont pas dépendant de l'âge. Pour une composante : $\gamma_{ik} = \exp\left(\sum_{j=1}^J B_{jk}(x_i)\alpha_{jk}\right)$. Le alpha est multiplié par la spline pour tous les âges x_i .



Fit des données

B-Splines



Les B-Splines sont calculées pour les composantes 2 et 3 par le biais du package Mortality Smoothing qui repose sur la technique décrite par Eilers and Marx (2010). Le nombre de B-Splines est nombre de nœuds + degré polynomial. Ici 25 splines = 22 nœuds + degré polynomial 3. Les Splines de la composante 1 sont composées d'une Spline constante égale à 1 et d'une Spline de fonction $\frac{1}{x}$. Toutes les composantes sont créées sur l'intervalle des âges de 1 à 111. Le nombre de nœuds est calculés en divisant l'intervalle par 5 ($110/5=22$).

À gauche : Graphique des composantes $\beta_{jk}(x_i)$ en fonction de l'âge x_i . *Remarques* : Pas de variation en fonction de l'année. Pas de variation en fonction du genre. Pas de variation en fonction du pays.



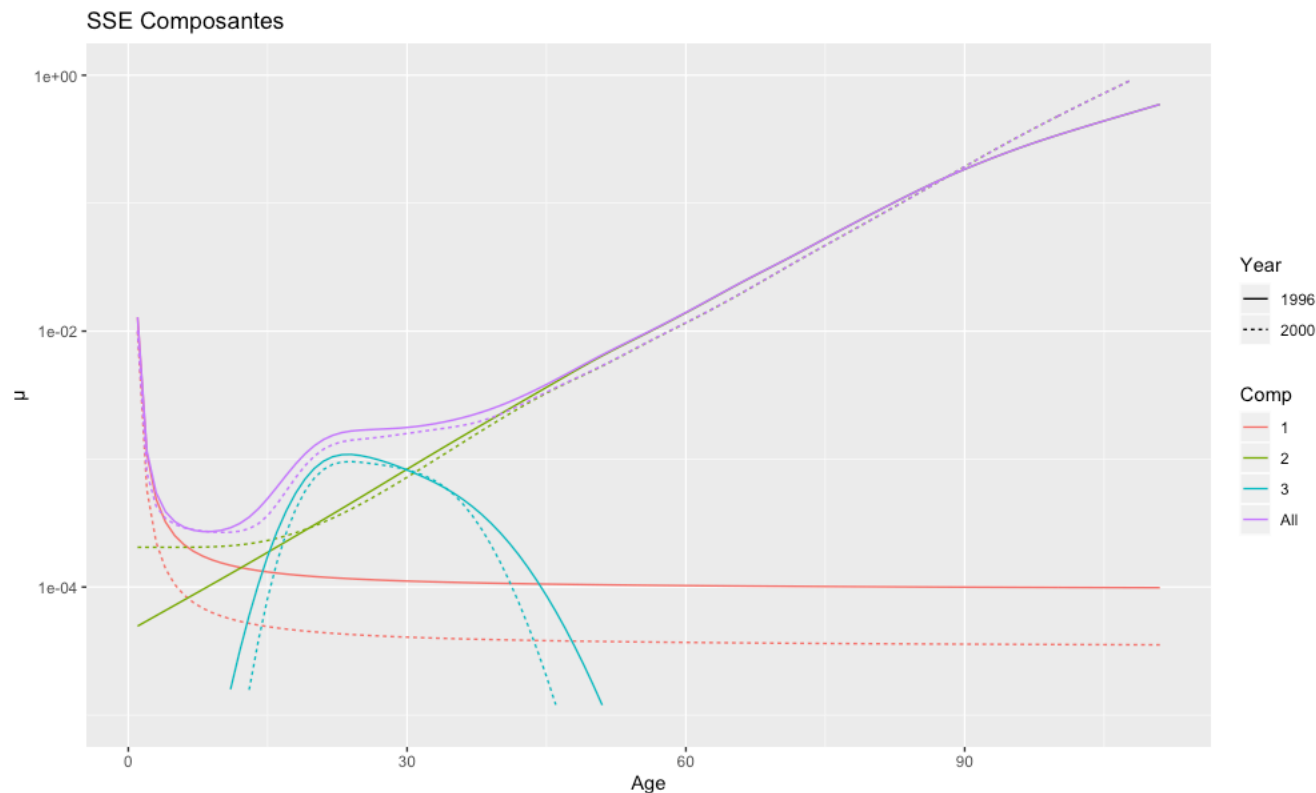
3 Décomposition

Décomposition

Sur le graphique :

Décomposition des taux de mortalité en composantes. Les trois composantes sont représentées en Bleu, Vert et Rouge. La somme des trois composantes est représentée par la courbe violette.

Les courbes sont représentées en 1996 et en 2000.

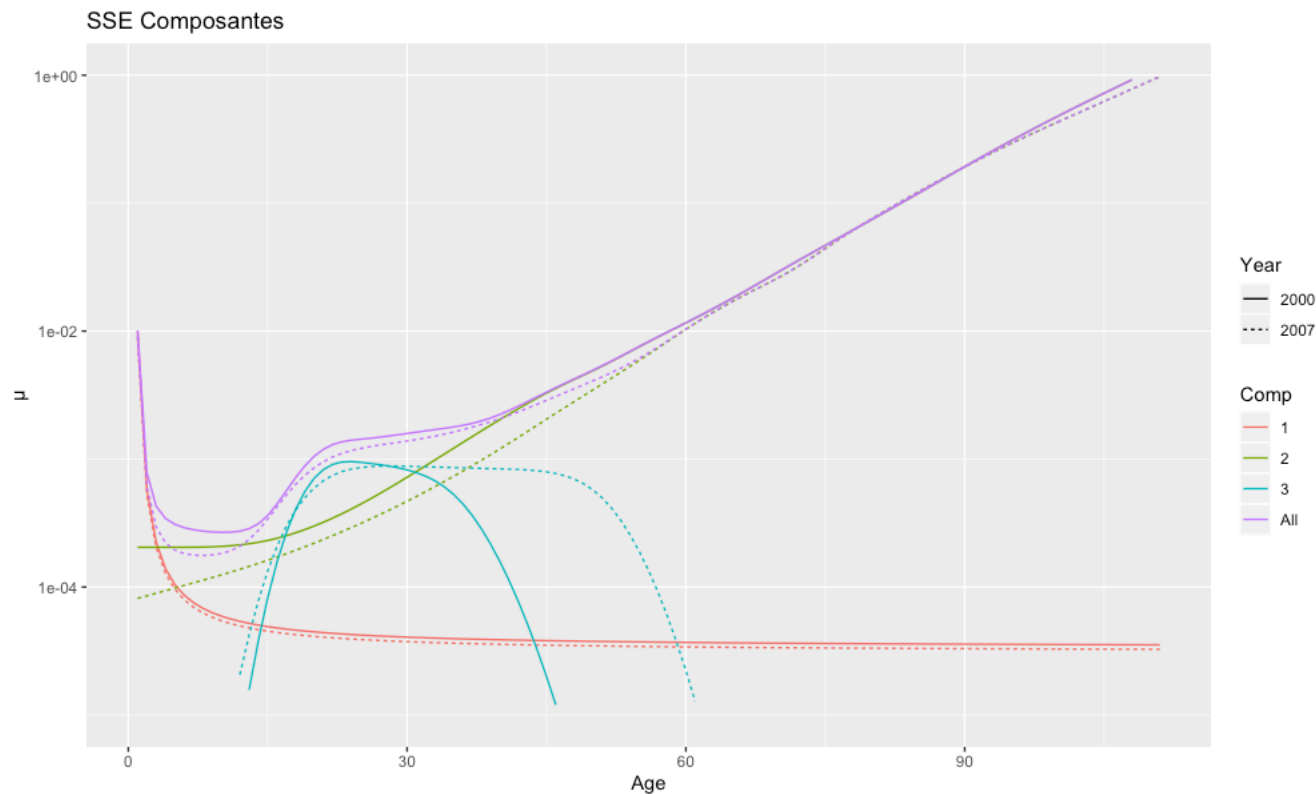


Données de la population Française (HMD), Hommes.

Décomposition

Entre 2000 et 2007 :

La forme de la bosse des accidents à changé et la variation de mortalité infantile a diminué.



Données de la population Française (HMD), Hommes.

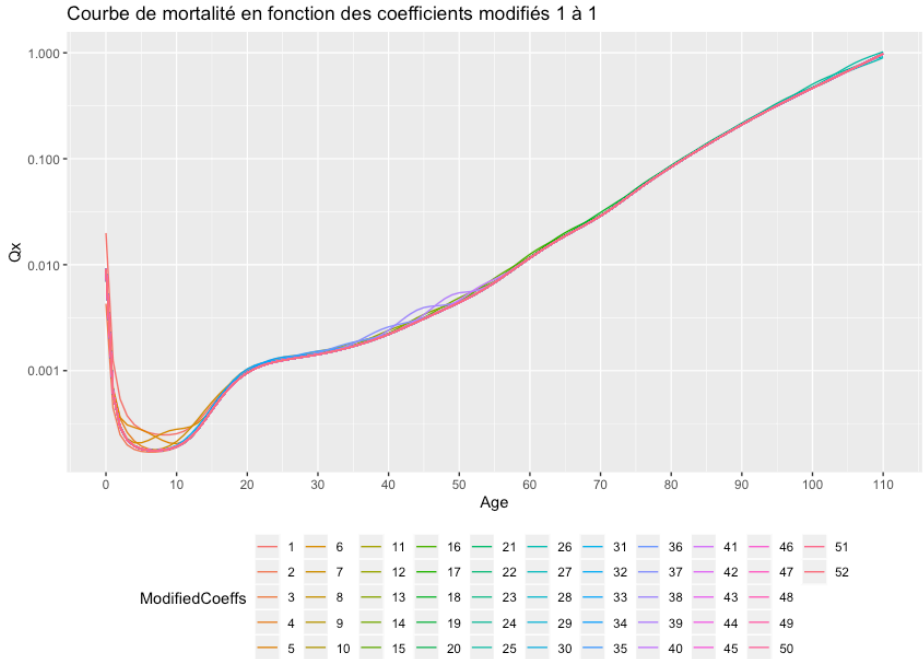
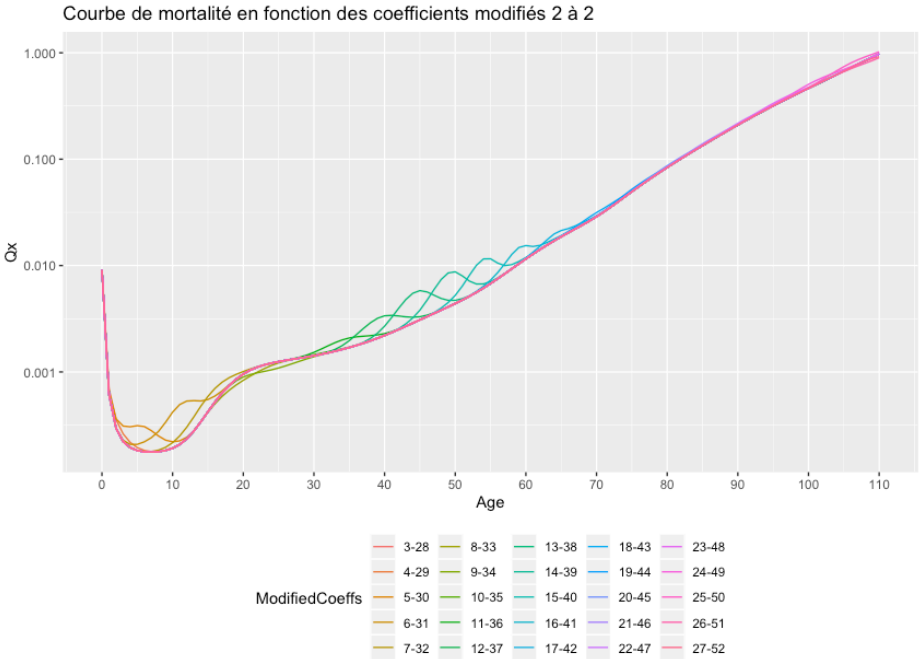


4

Sensibilité des coefficients

Sensibilité des coefficients

Etude des sensibilités afin de sélectionner les paramètres pertinents et d'analyser leur comportement. Calcul des espérances de vie suite à des variations des coefficients uns à uns puis deux par deux. Les variations des coefficients sont calculées avec un choc sur le coefficient du montant du coefficient de variation ($CV=\sigma/\mu$).



Sensibilité des coefficients

Etude des sensibilités afin de sélectionner les paramètres pertinents et d'analyser leur comportement. Calcul des espérances de vie suite à des variations des coefficients uns à uns puis deux par deux. Les variations des coefficients sont calculées avec un choc sur le coefficient du montant du coefficient de variation ($CV=\sigma/\mu$).

<i>IndiceComposante2</i>	$\Delta Ex2$	<i>IndiceComposante3</i>	$\Delta Ex3$	$\Delta ExBoth$	Ex
14	0,092	39	0,175	0,644	73,39
15	0,110	40	0,089	0,592	73,44
13	0,070	38	0,175	0,515	73,52
16	0,124	41	0,020	0,370	73,66
...
3	0,001	28	0,000	0,001	74,03
25	0,000	50	0,000	0,000	74,03
27	0,000	52	0,000	0,000	74,03
26	0,000	51	0,000	0,000	74,03

Sensibilité des coefficients

Etude des sensibilités afin de sélectionner les paramètres pertinents et d'analyser leur comportement. Calcul des espérances de vie suite à des variations des coefficients uns à uns puis deux par deux. Les variations des coefficients sont calculées avec un choc sur le coefficient du montant du coefficient de variation ($CV=\sigma/\mu$).

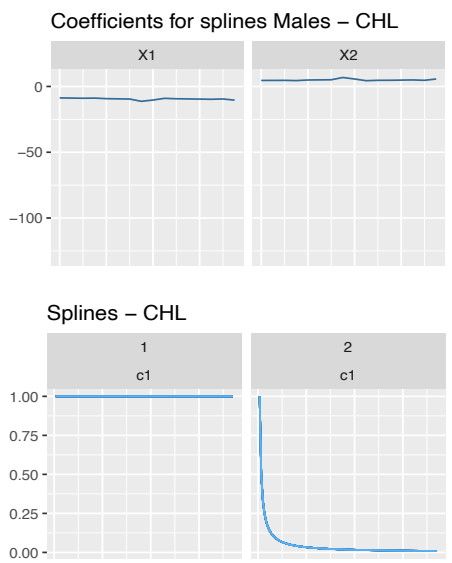
Remarque 1 : Les coefficients 1 et 2 font le plus varier l'espérance de vie :

	Ex	ΔEx
1	73,01	1,03
2	74,41	0,39

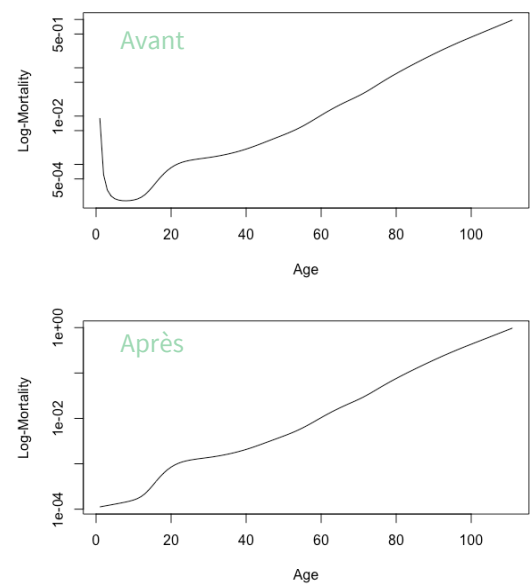
Explication 1 : Le coefficient 1 agit à tous les âges, la Spline 1 est constante et égale à 1 à tous les âges. En réduisant le coefficient 1 à 0, les taux ne sont plus compris en 0 et 1.

Explication 2 : Le coefficient 2 agit seulement sur les âges jeunes, en le rendant nul on supprime la mortalité infantile. Le coefficient 2 va servir à piloter l'espérance de vie par la mortalité infantile.

Zoom sur les coefficients 1 et 2 et les Splines associées



Impact du coefficient 2 sur la courbe de mortalité





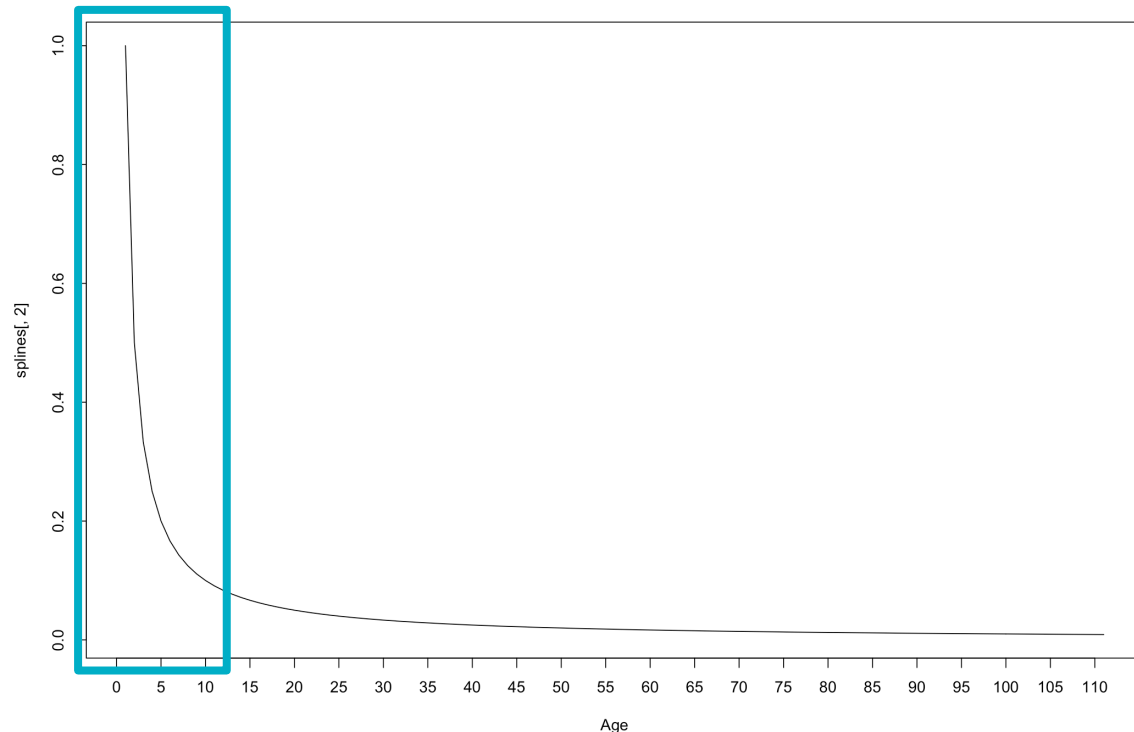
6

Pilotage de l'espérance de vie

Pilotage de l'espérance de vie

Mortalité Infantile

Spline 2 correspondant au paramètre α_2



Test sur :

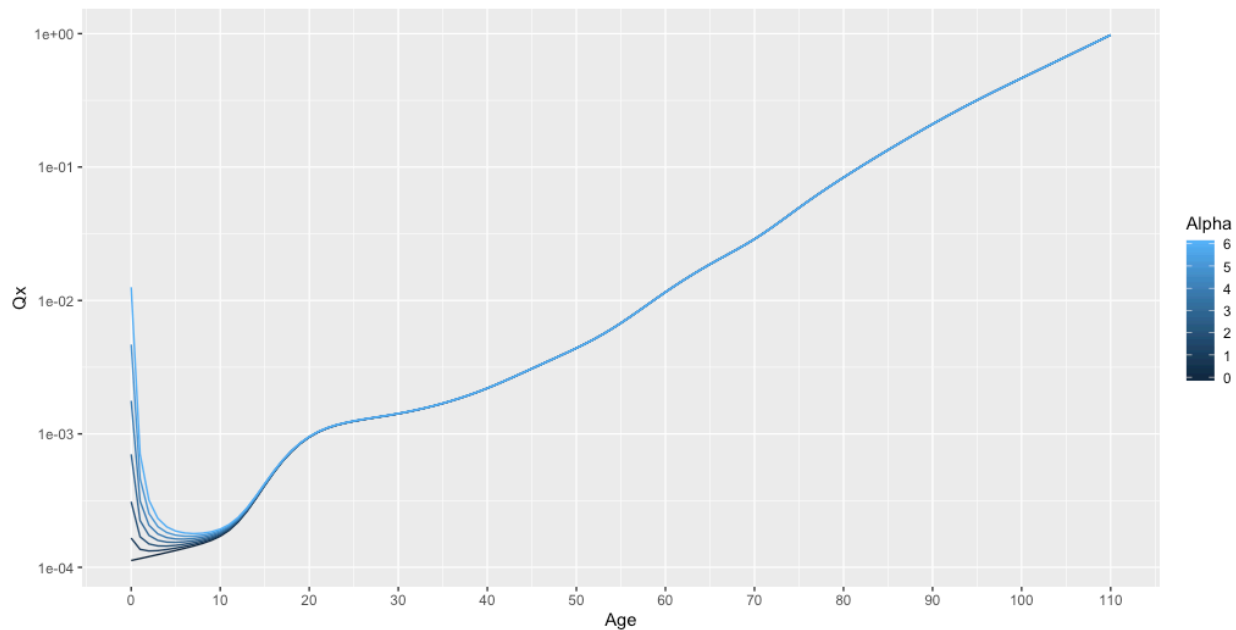
- Chili
- Année 2007
- Genre M

Dans le modèle SSE : la mortalité infantile est gouvernée par le coefficient α_2 de la composante 1 : les âges vont de 0 à ≈ 10 .

Pilotage de l'espérance de vie

Mortalité Infantile

Courbe de mortalité en fonction du paramètre alpha 2



Pilotage de l'espérance de vie

Mortalité Infantile

Espérance de vie en fonction du paramètre alpha 2

