



# Méthodes dynamiques d'extrapolation de séries démographiques

*Centre d'Economie de la Sorbonne  
UMR 8174 du CNRS - Université Paris 1 Panthéon-Sorbonne*

*Maison des sciences économiques*

*Luxi CHEN*

*13 Janvier 2017*



## I. Introduction

- A. Contexte
- B. Données utilisés

## II. Les méthodes dynamiques

- A. Équations aux Différences finies des Séries Dynamiques
- B. Retournement de tendances
- C. Utilisation de l'Extrapolateur Dynamique
- D. Application de l'extrapolateur dans un modèle d'ajustement

## III. Comparaison de résultats

## IV. Conclusion

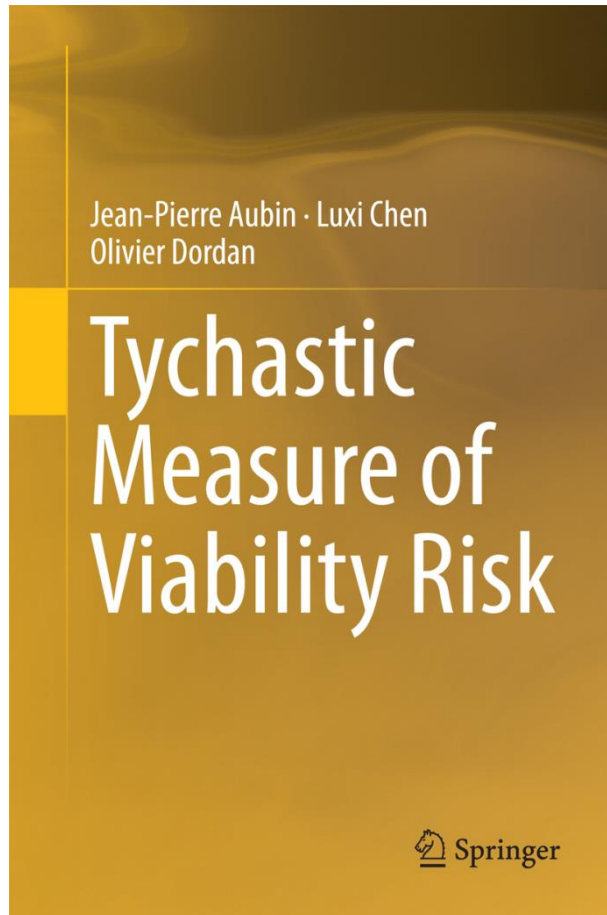


# Contexte

L'objectif de cet exposé est de présenter de nouvelles méthodes de prédiction et d'extrapolations de séries démographique,

1. En construisant des équations aux différences finies et les reproduisant,
2. En utilisant des méthodes d'extrapolation dynamiques prenant en compte les dérivées des séries démographique au lieu de techniques statistiques de régression.





***Tychastic Measure of Viability Risk***

*AUBIN J.-P., CHEN LX, DORDAN O.*

**2014**



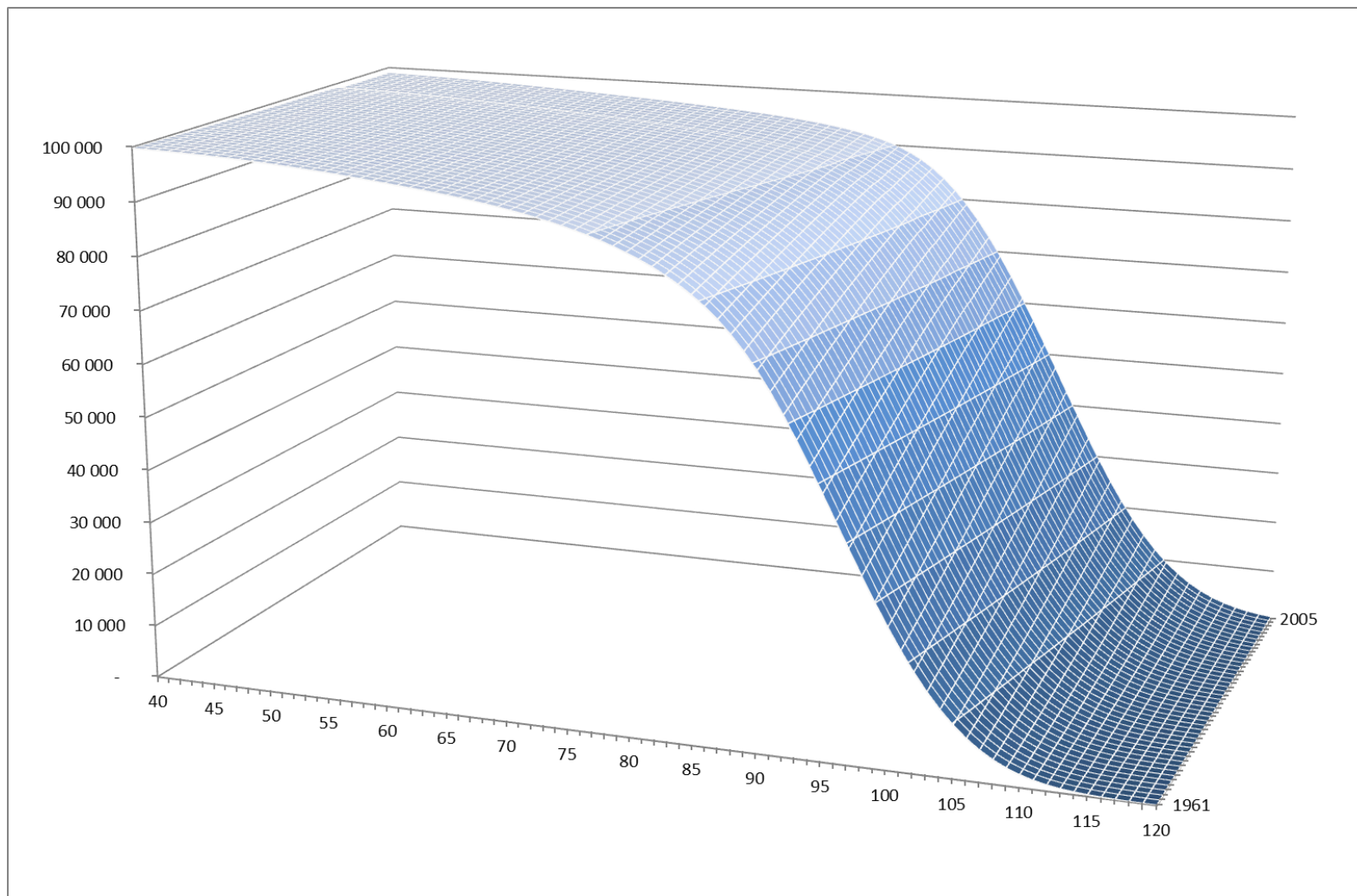
# Données Utilisés

Les tables TGF05 et TGH05 fournissent les données brutes du nombre de survivants, âge par âge, pour les hommes et pour les femmes, pour toutes les générations entre 1900 et 2005. Il s'agit de la table réglementaire pour les opérations de rentes viagères. Les données sont limitées à la tranche d'âge comprise entre 0 à 120 ans. Dans cette étude on utilise ces deux tables de mortalité sous forme de  $L_x$ , dans certains exemples, pour mieux illustrer les données sont limitées à la tranche d'âge comprise entre 40 à 120 ans et les années comprises entre 1961 à 2005.



## Données utilisés

TGH05/TGF05



# **Les Méthodes Dynamiques**



## Les méthodes dynamiques

Pour chaque date (année)  $t$  et chaque âge  $x$ , on désigne par  $L_{t,x}$  *le nombre de personnes d'âge  $x$  vivantes à la date  $t$ .*

La question : comment prévoir

1. à chaque date  $t$ , la fonction  $x \mapsto L_{t,x}$  de la population  $x \mapsto L_{t,x}$  en fonction de son âge ;
2. à chaque âge  $x$ , l'évolution de l'âge de la population  $t \mapsto L_{t,x}$  en fonction de la date ;
3. l'évolution  $(t, x) \mapsto L_{t,x}$  en fonction de la date et de l'âge de la population (évolution structurée en âge) ;

Pour simplifier l'exposé, nous considérons l'évolution de la population totale  $x \mapsto L_{t,x}$  entre les âges 0 et 120 pour chaque année  $t$  entre 1961 et 2005, et illustrer ce comportement à la date 2005 comme exemple avant de les juxtaposer.





## Les méthodes dynamiques

Nous introduisons

$$V_{t,x} := L_{t,x} - L_{t,x-1}, \quad W_{t,x} := L_{t,x} - L_{t-1,x} \text{ et } D_{t,x} := L_{t,x} - L_{t-1,x-1} = V_{t,x} + W_{t,x-1} \quad (1)$$

Pour cela, nous proposons deux types de réponses : pour chaque date  $t$ , connaissant  $L_{t,x}$

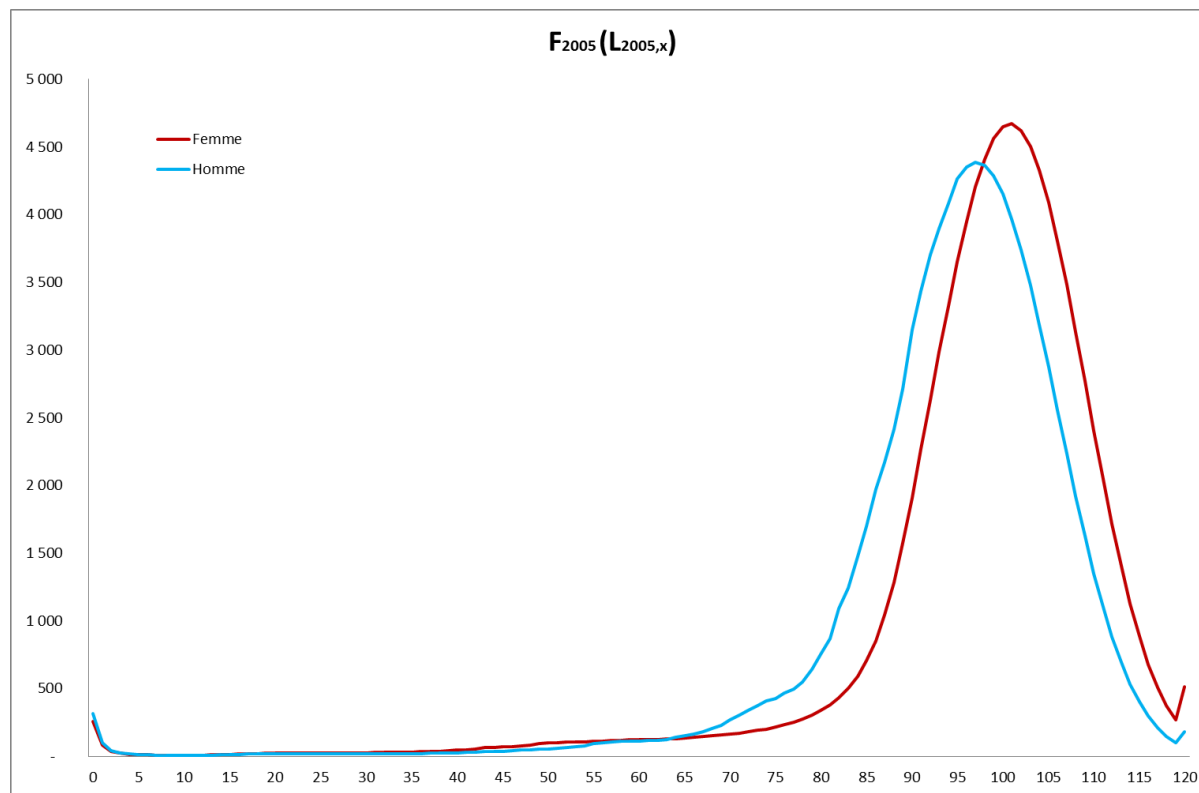
1. étudier la graphe fonction  $F_t : L_{t,x} \mapsto F_t(L_{t,x}) := V_{t,x}$  qui décrit l'équation aux différences finies de la population (connaissant les naissances  $L_{t,0}$ , on peut en déduire  $L_{t,x}$  pour tout  $x \in \{1, 120\}$ )
2. la population  $L_{t,x_i}$  pour les âges de  $x_i$  de 0 à  $n - 1$ , déduire la prévision de la population  $L_{t,x_{n+1}}$  : on pourra utiliser une méthode de *régression statistique* ou *l'extrapolateur dynamique* que nous allons présenter et les comparer;



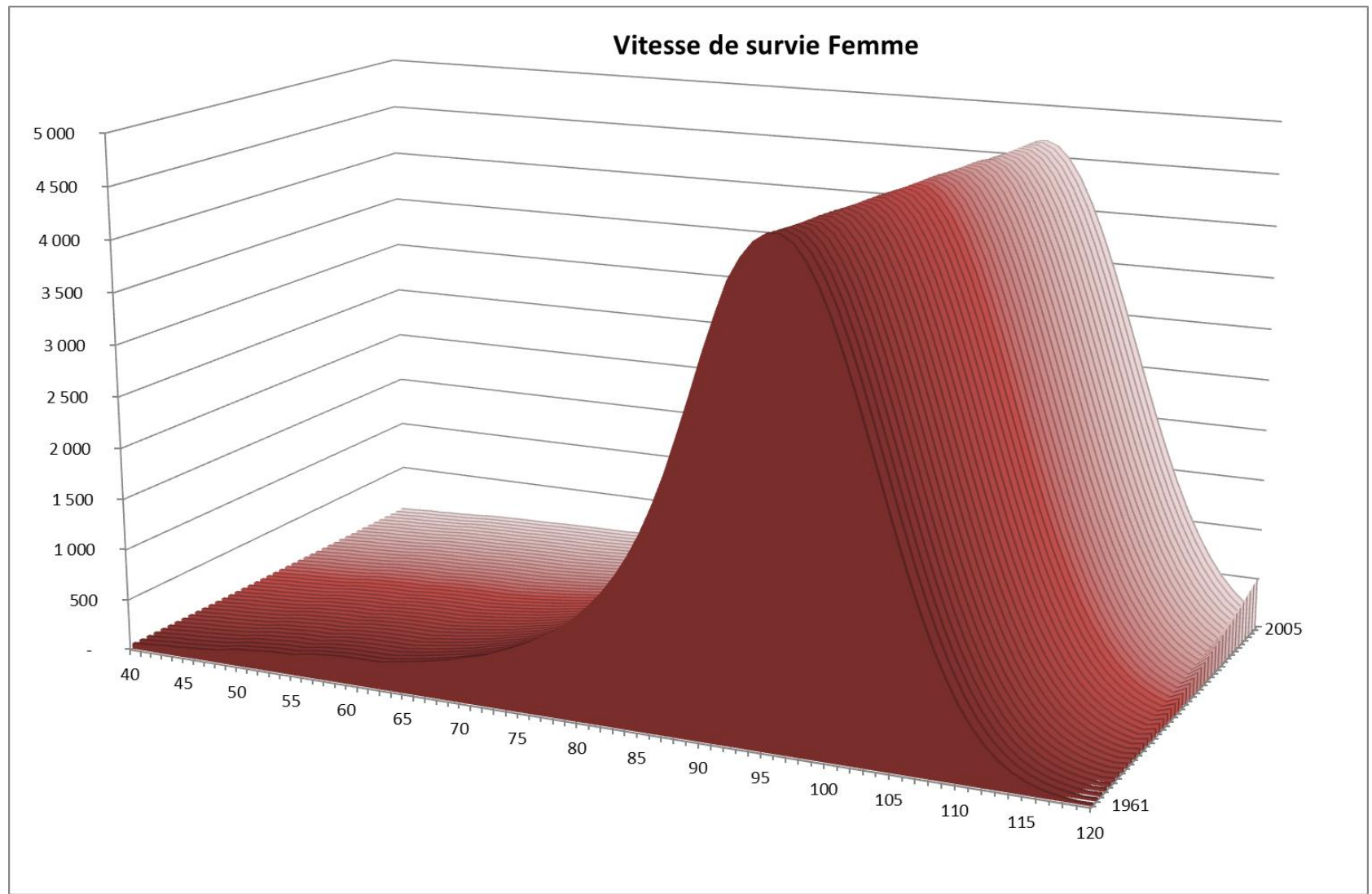
## *Relation aux différences finies de la population survie*

Pour  $t = 2005$ , on peut déduire de la fonction  $x := x \mapsto L_{2005,x}$  la fonction  $F_{2005,} : L_{2005,x} \mapsto F_{2005}(L_{2005,x}) := V_{2005,x}$  et les fonctions  $F_t(L_{t,x})$ .

Voir les graphiques de  $V_{2005,xF}$ ,  $V_{2005,xH}$  et  $V_{[1961-2005],xF}$  :



# Les méthodes dynamiques

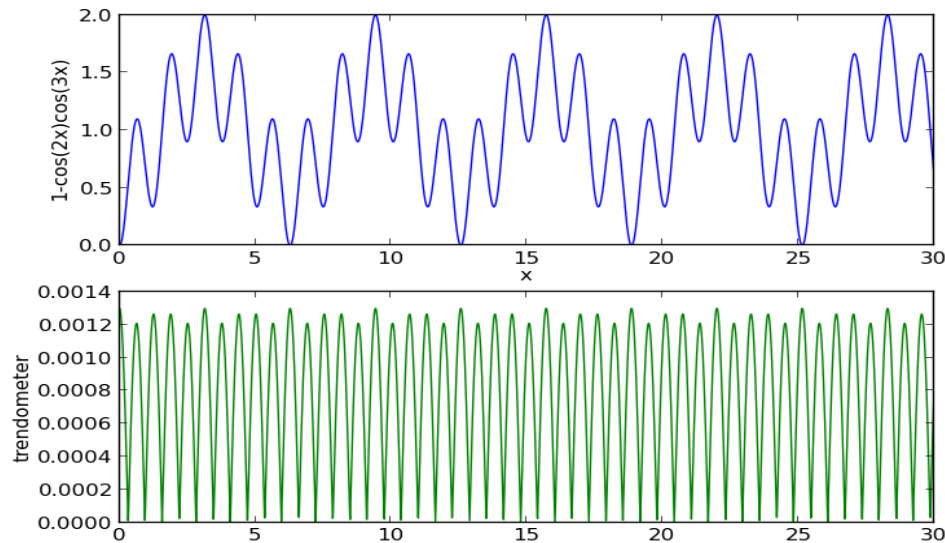


### *Le Tendomètre*

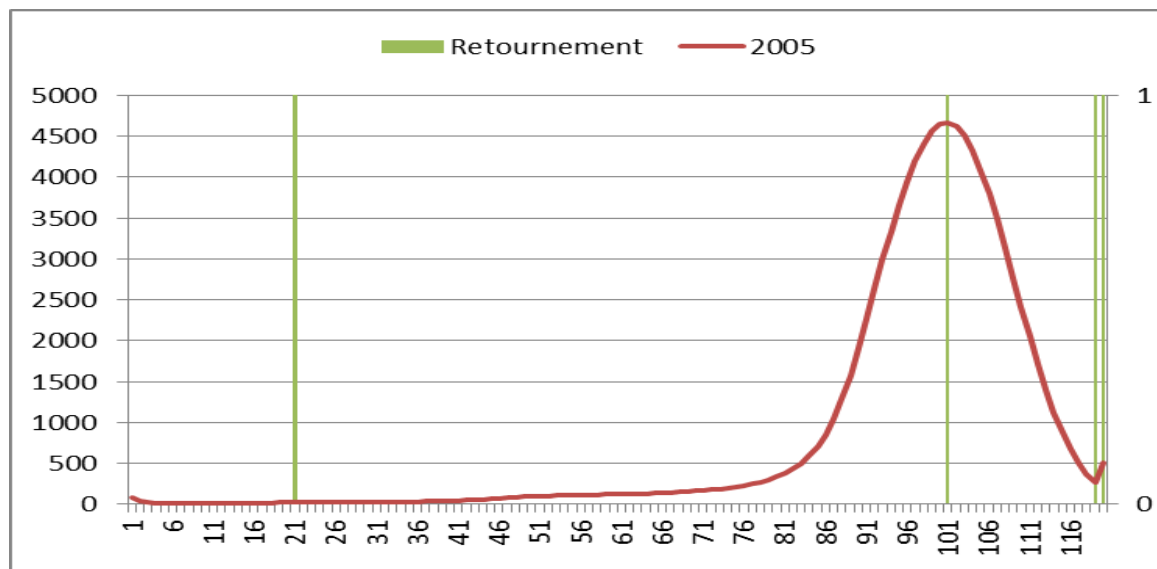
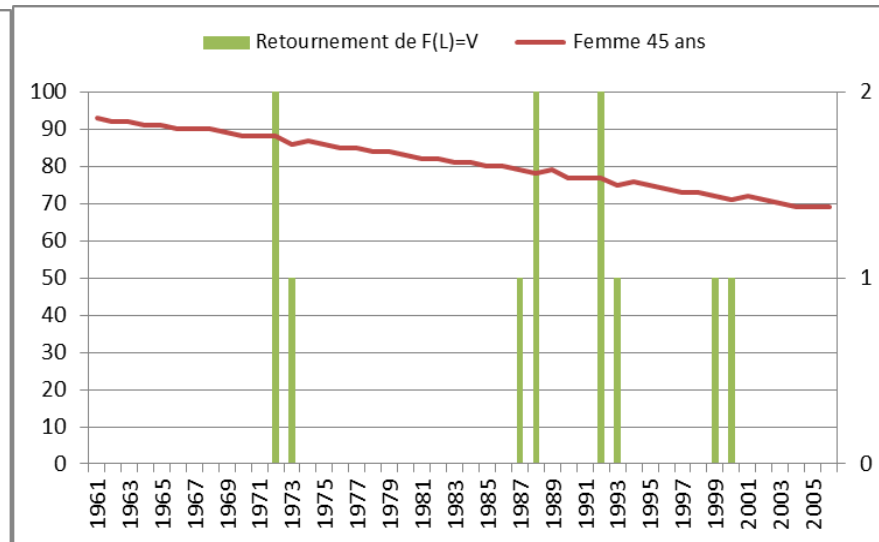
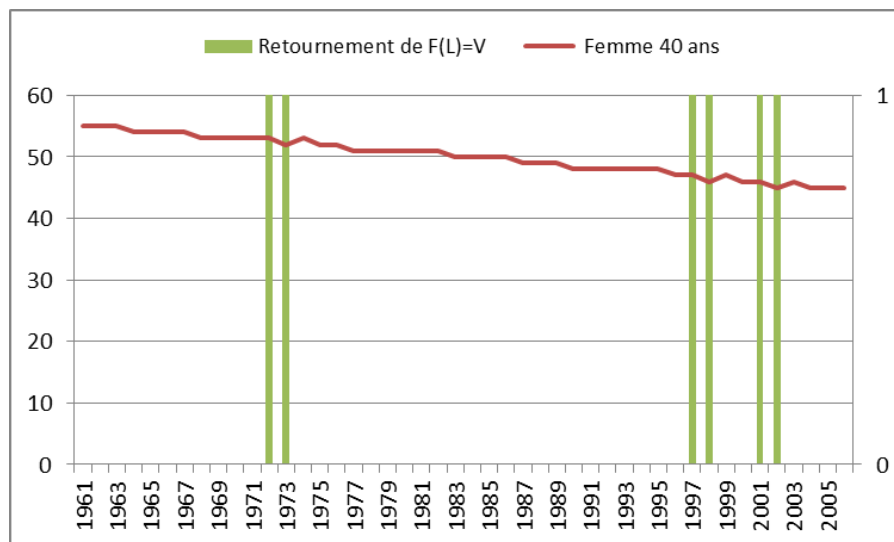
A chaque date, le tendomètre

1. Détecte automatiquement dates de *retournement de tendance* où une fonction atteint soit un minimum local, soit un maximum local;
2. calcule un indicateur (non linéaire) mesurant l'intensité retournement et un autre indicateur de fréquence des dates de retournement.

$$X \rightarrow V(x) := 1 - \cos(2x)\cos(3x)$$



# Les méthodes dynamiques



# **L'Utilisation l'Extrapolateur Dynamique**



### *L'Extrapolateur dynamique*

Considérons une série temporelle discrète  $(x_j)_{j \in \mathbb{Z}}$ . En utilisant mesures de Dirac  $\delta_j$  aux dates  $j \in \mathbb{Z}$ , nous pouvons plonger la série temporelle discrète dans l'espace des distributions en lui associant son peigne Dirac  $\sum_{j \in \mathbb{Z}} x_j \delta_j$  et les régulariser en prenant leur produit de convolution  $(\lambda \star x)(t) := \int_{-\infty}^{+\infty} \lambda(\tau) x(t\tau) d\tau$  de fois différentiable que  $\lambda$  (voir les processus de Hawkes, par exemple).

*L'extrapolateur dynamique* appartient à cette classe et est basé sur les techniques d'analyse numérique (éléments finis). En prenant  $p = 4$ , on obtient un extrapolateur d'ordre 4 qui reprend les tendances des séries chronologiques régularisées: ses valeurs, ses vitesses, ses accélérations et ses secousses. Les valeurs supérieures ne donnent pas de résultats sensiblement meilleurs.

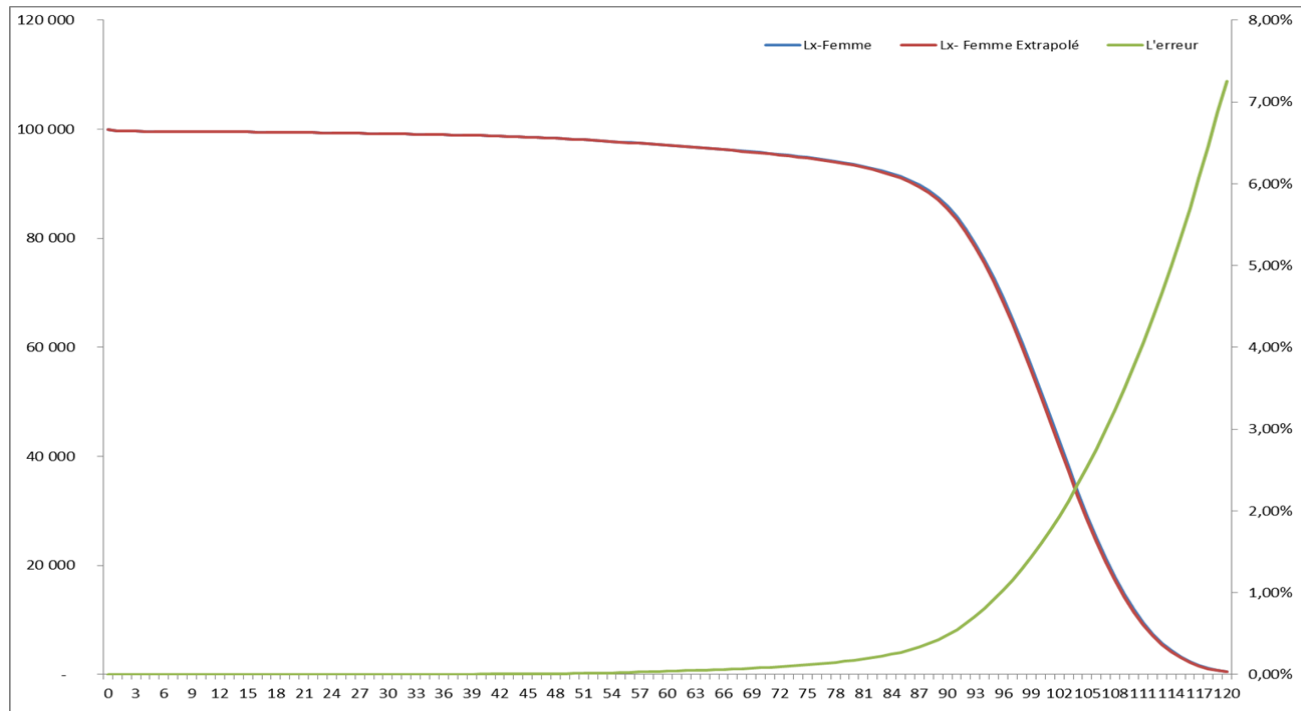


## L'Utilisation l'Extrapolateur Dynamique

En prenant les séries historiques des femmes  $L_{2001,x}$ ,  $L_{2002,x}$ ,  $L_{2003,x}$ ,  $L_{2004,x}$ , on obtient l'extrapolation  $x \mapsto Ext(L_{2005,x})$ , le graphique ci dessous représente la série  $x \mapsto L_{2005,x}$ , son extrapolation  $x \mapsto Ext(L_{2005,x})$  et des erreurs relatives  $\delta \mapsto \frac{|L_{2005,x} - Ext(L_{2005,x})|}{L_{2005,x}}$ .

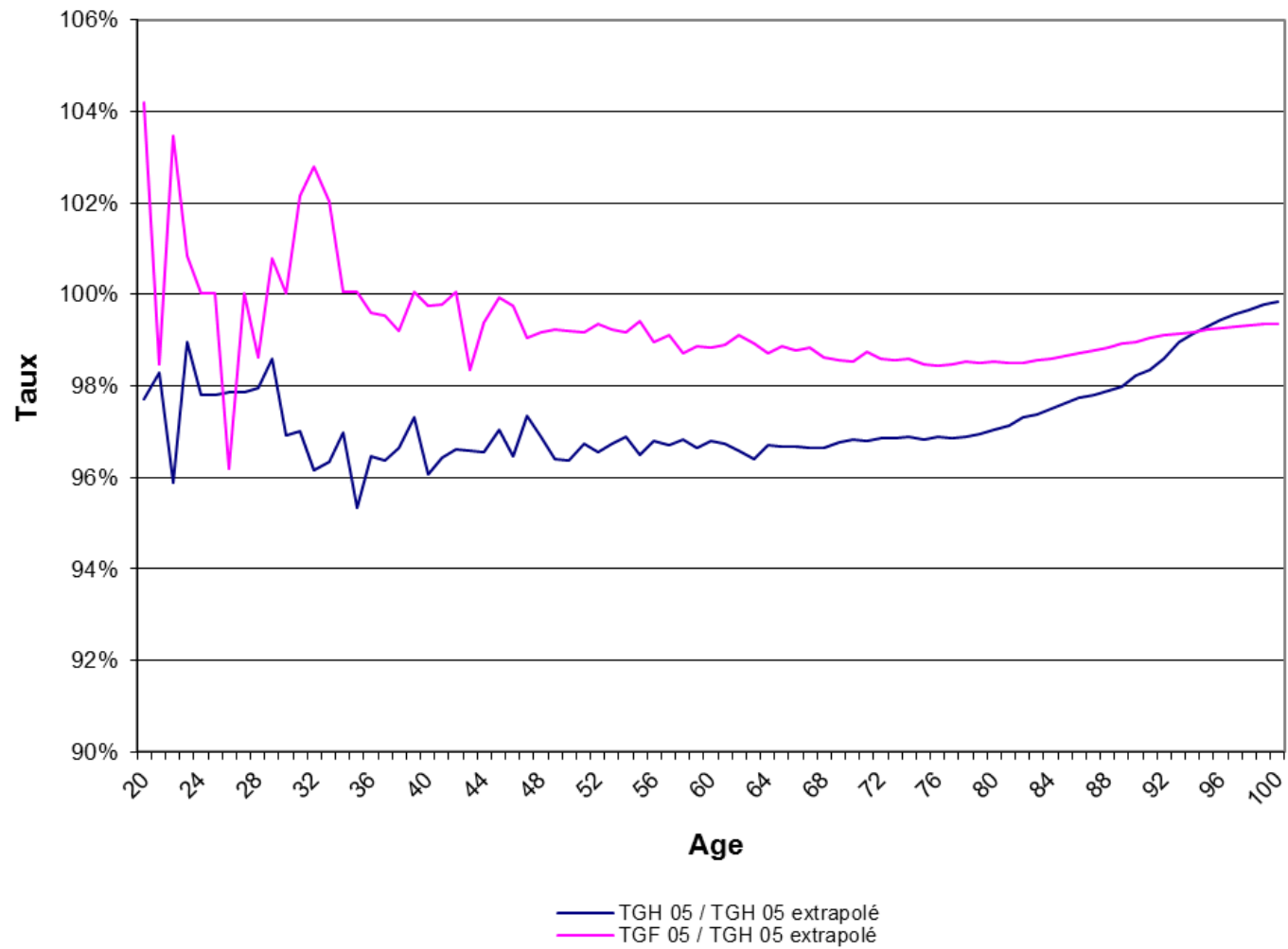
L'erreur moyen pour tout  $x \in \{1, 120\}$  est 0,81%;

L'erreur moyen pour tout  $x \in \{40, 80\}$  est 0,05%.

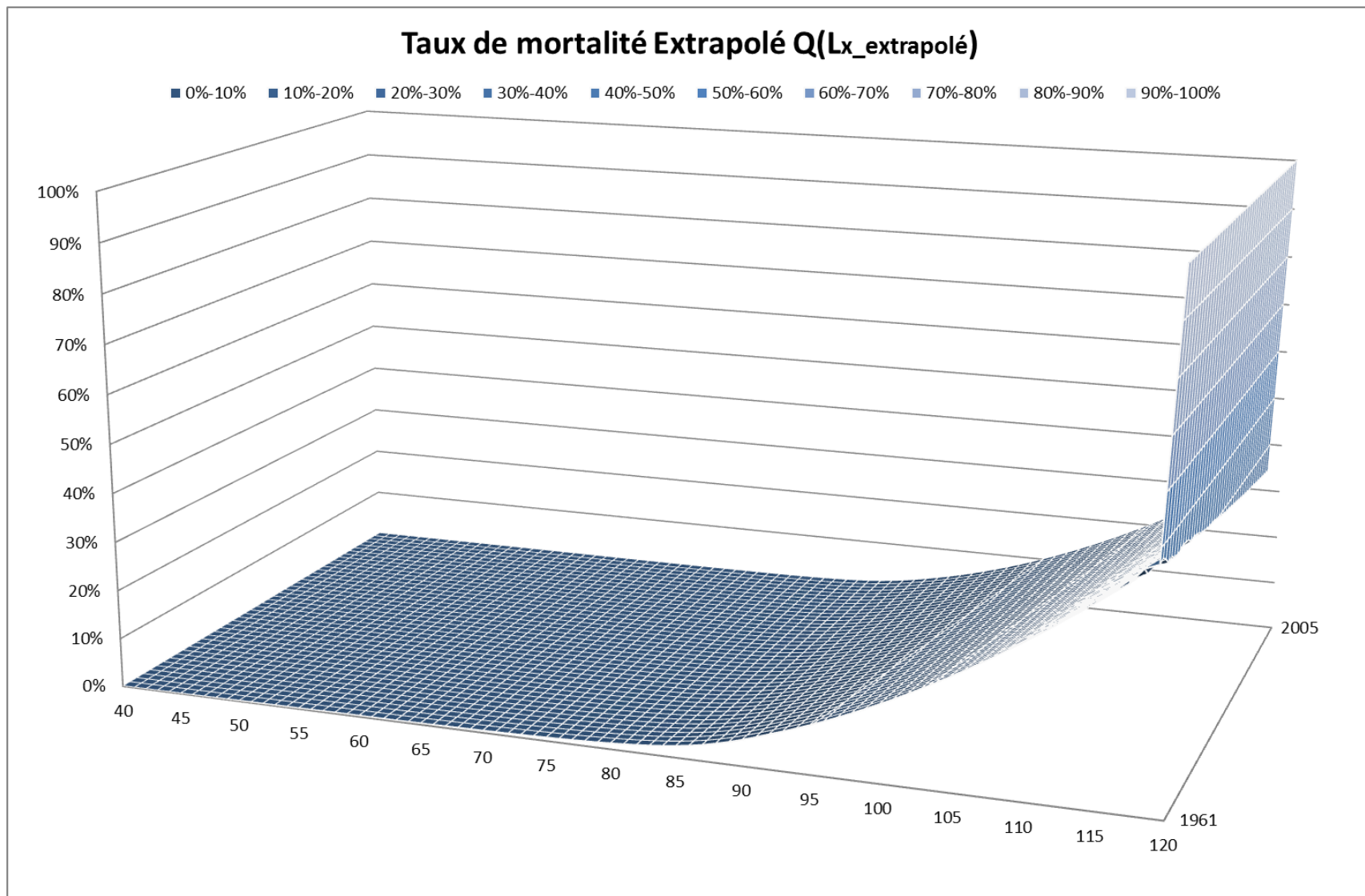




## L'Utilisation l'Extrapolateur Dynamique



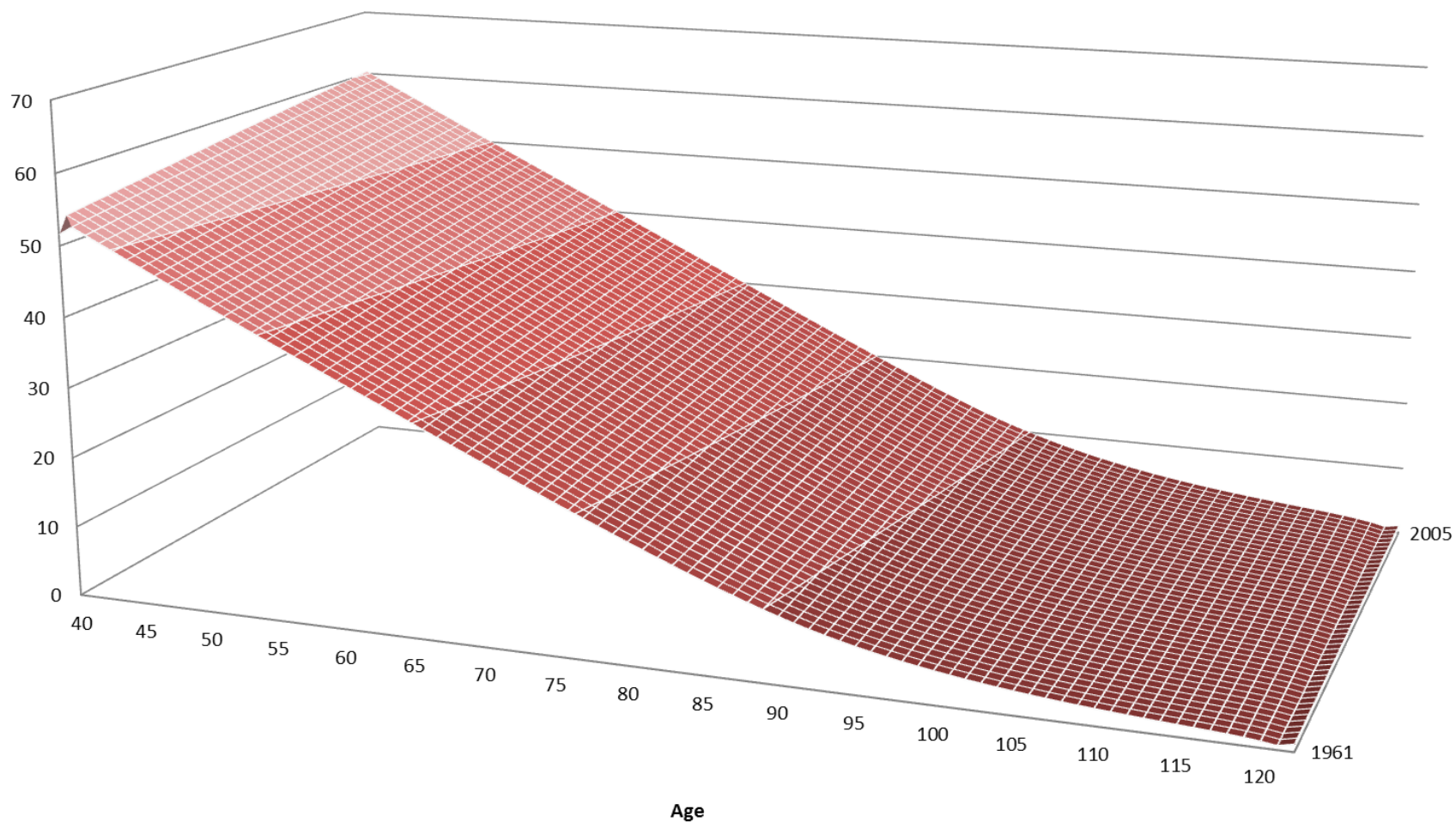
# L'Utilisation d'Extrapolateur Dynamique



# L'Utilisation d'Extrapolateur Dynamique

## Esperance de vie TGF 05 extrapolé

■ 0-10 ■ 10-20 ■ 20-30 ■ 30-40 ■ 40-50 ■ 50-60 ■ 60-70



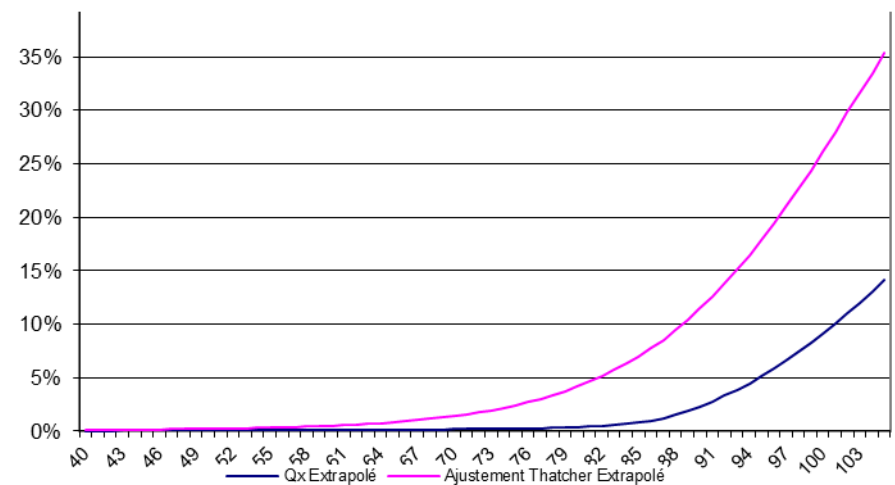
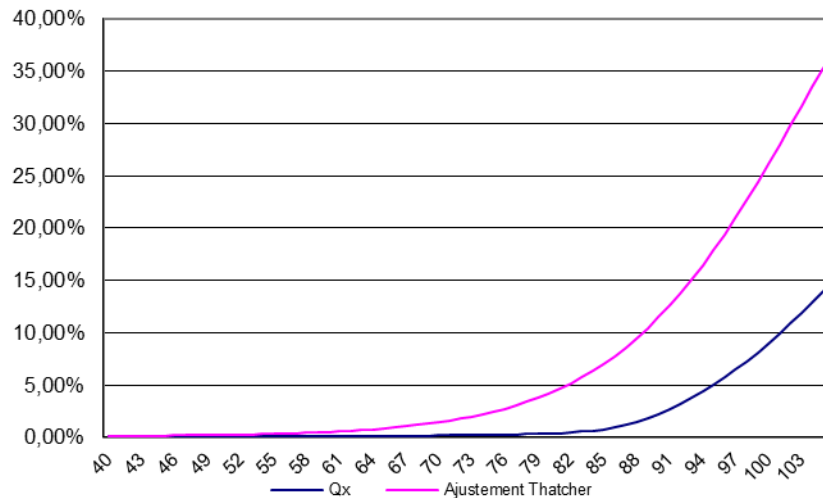
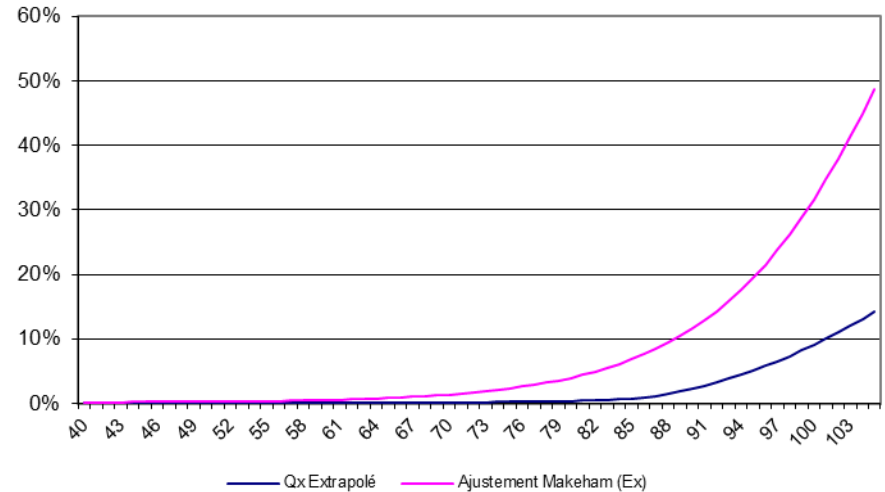
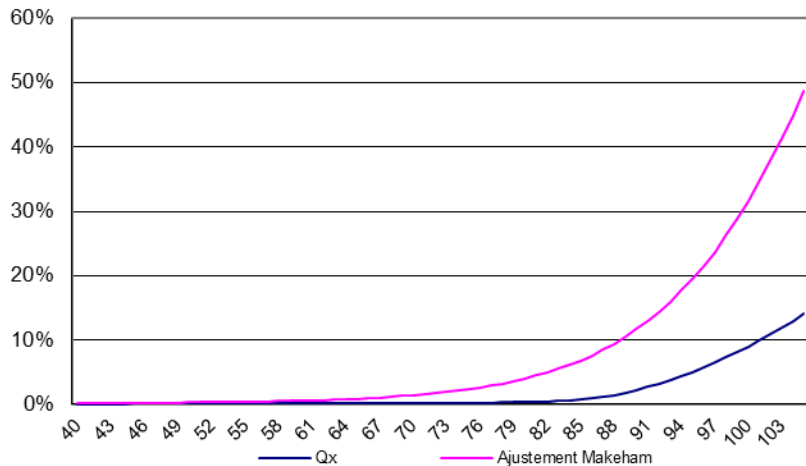
# Les méthodes tychastique: L'extrapolateur dynamique

Table de mortalité extrapolé TGF 05

Age	Lx TGF 05	Qx	Ex	Age	Lx TGF 05	Qx	Ex	Age	Lx TGF 05	Qx	Ex
0	100 000	0,26%	39	40	98 850	0,05%	38	80	93 391	0,37%	20
1	99 744	0,08%	38	41	98 804	0,05%	37	81	93 041	0,42%	19
2	99 661	0,04%	37	42	98 756	0,05%	36	82	92 648	0,49%	19
3	99 625	0,02%	36	43	98 701	0,06%	35	83	92 198	0,56%	18
4	99 605	0,01%	35	44	98 639	0,06%	34	84	91 678	0,67%	17
5	99 590	0,01%	34	45	98 576	0,07%	33	85	91 067	0,80%	16
6	99 579	0,01%	33	46	98 507	0,08%	32	86	90 338	0,98%	15
7	99 570	0,01%	32	47	98 432	0,08%	31	87	89 456	1,21%	14
8	99 561	0,01%	31	48	98 354	0,09%	30	88	88 377	1,50%	13
9	99 552	0,01%	30	49	98 270	0,10%	29	89	87 053	1,86%	12
10	99 544	0,01%	29	50	98 175	0,10%	28	90	85 435	2,29%	12
11	99 537	0,01%	28	51	98 074	0,11%	27	91	83 480	2,78%	11
12	99 528	0,01%	27	52	97 971	0,11%	26	92	81 161	3,31%	10
13	99 519	0,01%	26	53	97 864	0,11%	26	93	78 476	3,87%	10
14	99 510	0,01%	25	54	97 754	0,11%	25	94	75 440	4,48%	9
15	99 499	0,01%	24	55	97 643	0,12%	24	95	72 064	5,13%	8
16	99 486	0,01%	23	56	97 530	0,12%	23	96	68 366	5,83%	8
17	99 471	0,02%	22	57	97 413	0,12%	22	97	64 378	6,58%	7
18	99 454	0,02%	21	58	97 293	0,13%	21	98	60 141	7,38%	7
19	99 434	0,02%	20	59	97 170	0,13%	20	99	55 705	8,21%	6
20	99 412	0,02%	19	60	97 045	0,13%	19	100	51 129	9,10%	6
21	99 389	0,02%	18	61	96 918	0,13%	18	101	46 475	10,03%	6
22	99 365	0,02%	17	62	96 791	0,13%	17	102	41 813	11,01%	5
23	99 342	0,02%	16	63	96 663	0,13%	16	103	37 210	12,03%	5
24	99 319	0,02%	15	64	96 532	0,14%	15	104	32 735	13,09%	5
25	99 295	0,02%	14	65	96 398	0,14%	14	105	28 450	14,20%	4
26	99 271	0,02%	13	66	96 259	0,15%	13	106	24 409	15,35%	4
27	99 247	0,02%	12	67	96 115	0,16%	12	107	20 661	16,55%	4
28	99 222	0,03%	11	68	95 966	0,16%	11	108	17 242	17,79%	4
29	99 197	0,03%	10	69	95 810	0,17%	10	109	14 175	19,08%	3
30	99 172	0,03%	9	70	95 649	0,18%	9	110	11 470	20,40%	3
31	99 145	0,03%	8	71	95 480	0,19%	8	111	9 131	21,77%	3
32	99 118	0,03%	7	72	95 302	0,19%	7	112	7 143	23,19%	3
33	99 090	0,03%	6	73	95 117	0,21%	6	113	5 487	24,64%	2
34	99 061	0,03%	5	74	94 919	0,22%	5	114	4 135	26,14%	2
35	99 031	0,03%	4	75	94 711	0,24%	4	115	3 054	27,68%	2
36	98 999	0,03%	3	76	94 488	0,25%	3	116	2 209	29,27%	2
37	98 965	0,04%	2	77	94 248	0,28%	2	117	1 562	30,89%	1
38	98 929	0,04%	1	78	93 987	0,30%	1	118	1 080	32,54%	1
39	98 891	0,04%	39	79	93 704	0,33%	21	119	728	34,28%	1

# L'Utilisation d'Extrapolateur Dynamique

on peut aussi appliquer l'extrapolateur dynamique dans les méthodes d'ajustements du taux, voir ici Ajustement Makeham et Ajustement Thatcher:



# **Application de l'extrapolateur dans un modèle d'ajustement**



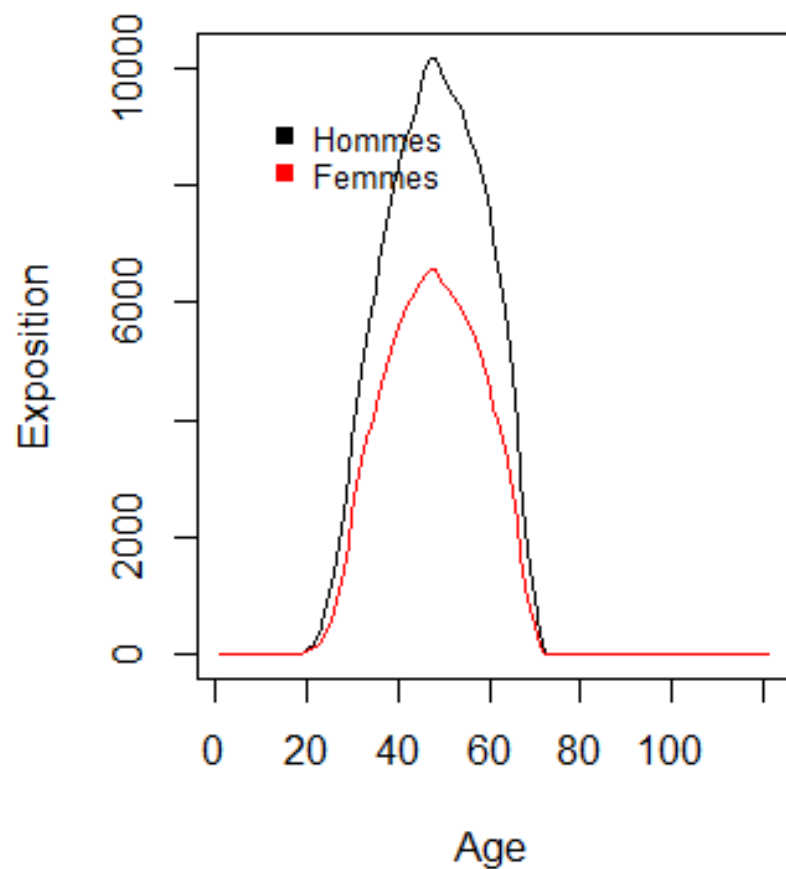
## Données Disponibles

- Les données proviennent des portefeuilles avec 148451 ligne de données, avec DateNaissance compris entre 01/01/1934 et 31/12/2015 , DateEntree compris entre 01/01/2008 et 31/12/2015 , DateSortie compris entre 01/01/2005 et 31/12/2015 , résumant les statistiques observées des données masculines et féminines respectivement. La période d'observation totale retenue s'étend du 01/01/2011 à 31/12/2015;
- Dans cet exemple, les tables TGF05 et TGH05 fournissent toujours les données brutes  $L_{2005,xF}$  et  $L_{2005,xH}$  comme données externes. Les générations étudiées sont compris entre  $[40, 80]$  .

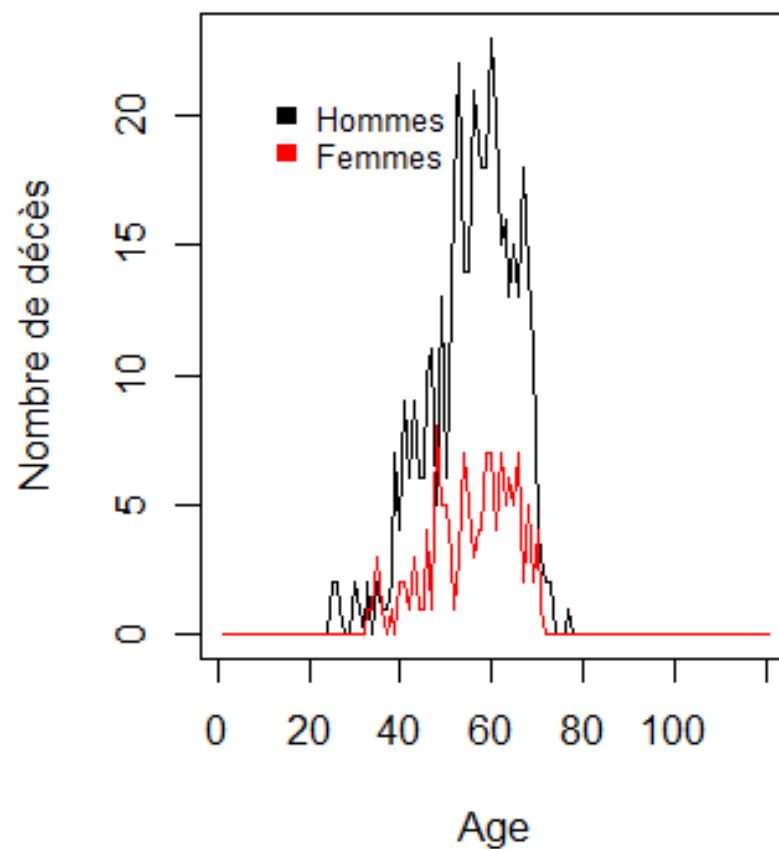


## Données Disponibles

### Exposition par âges



### Décès par âges

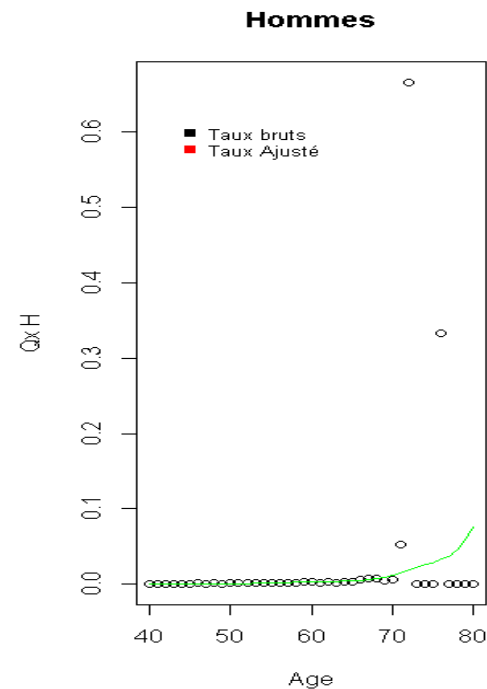
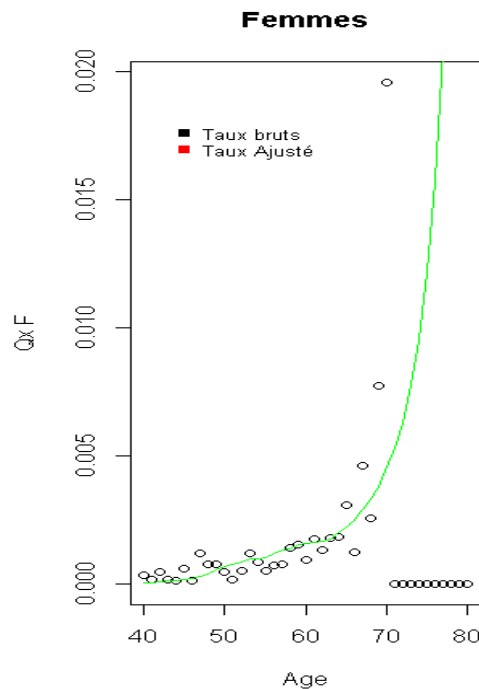




# Ajustement de taux

Dans cet exemple, prenons nous la méthode des modèles logit de **Brass** estime les paramètres  $\alpha$  et  $\beta$  dans un système relationnel associé à un standard judicieusement choisi, qui reproduit exactement les valeurs observées des indices de mortalité:

$$l(x) = \frac{1}{1 + \exp(2(\alpha + \beta Y^s(x)))}$$

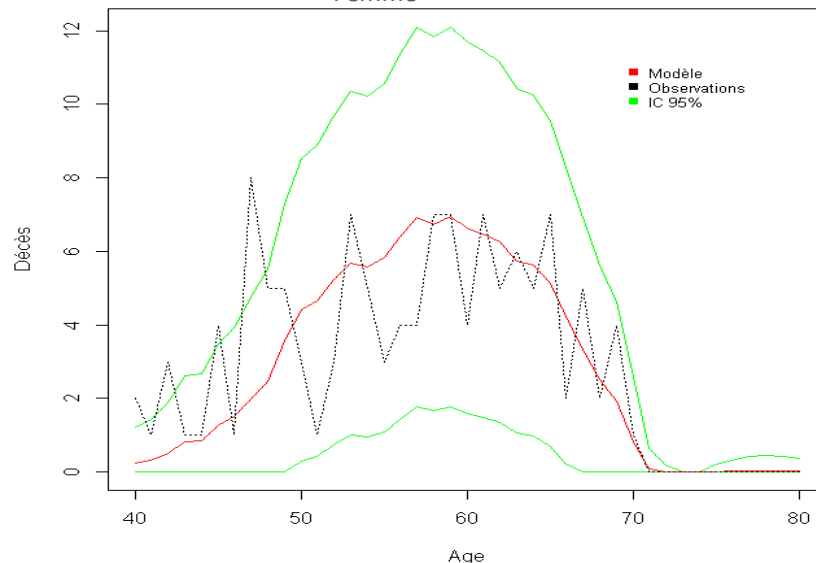


# **Comparaison des résultats**



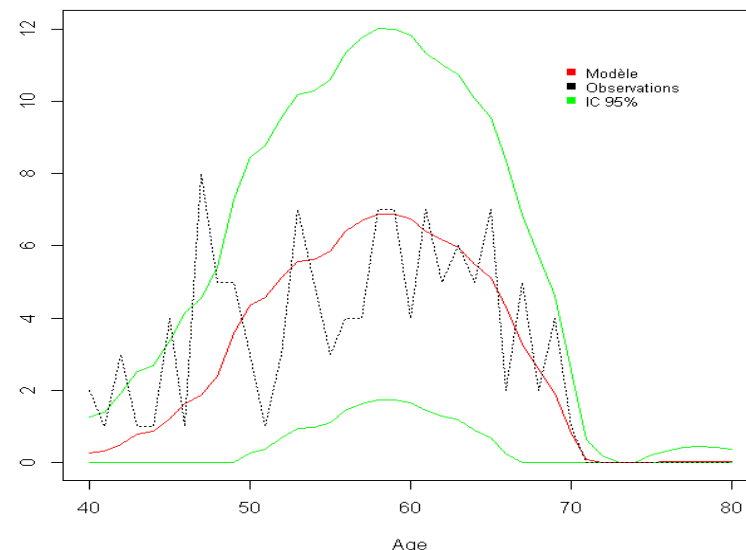
# Comparaison des résultats

Femme



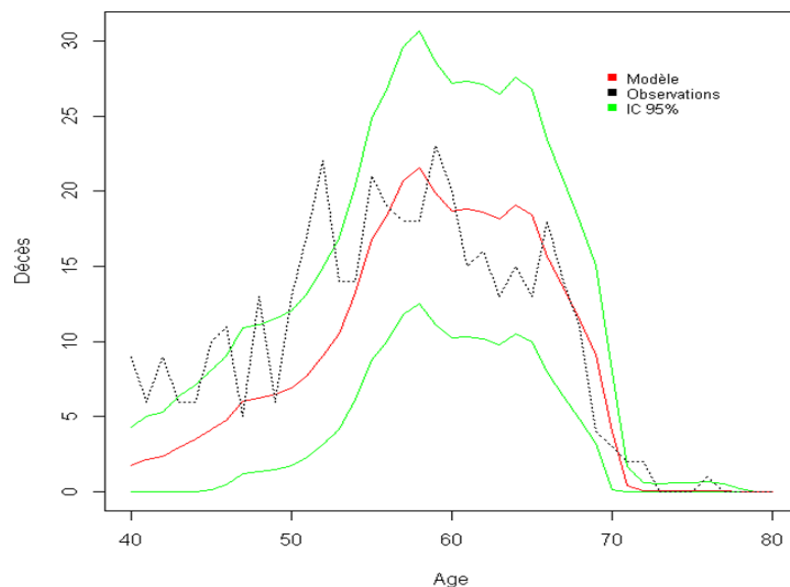
Ratio observations / modèle : 102 % - Incertitude d'échantillonnage pour les décès : 17.82 %

Femme extrapolé



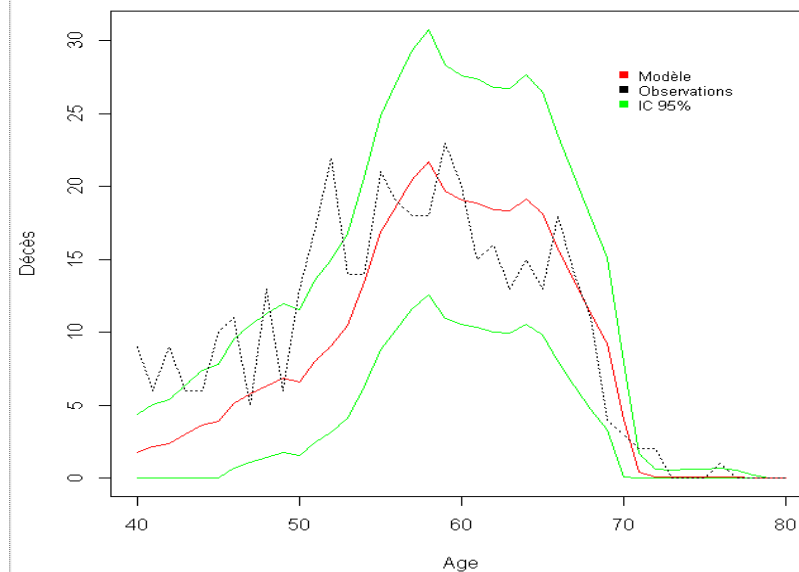
Ratio observations / modèle : 102 % - Incertitude d'échantillonnage pour les décès : 17.85 %

Homme



Ratio observations / modèle : 116 % - Incertitude d'échantillonnage pour les décès : 10.44 %

Homme extrapolé



Ratio observations / modèle : 115 % - Incertitude d'échantillonnage pour les décès : 10.43 %

# Comparaison de Ratio Observation/ Prédiction

Nous pouvons faire la comparaison des valeurs modélisées et réalisées par l'identification d' un ratio global d'observation par rapport à la prédiction calculée. Les ratios obtenus avec le modèle de Brass et le modèle de Brass extrapolé de calculs différents sont très proches.

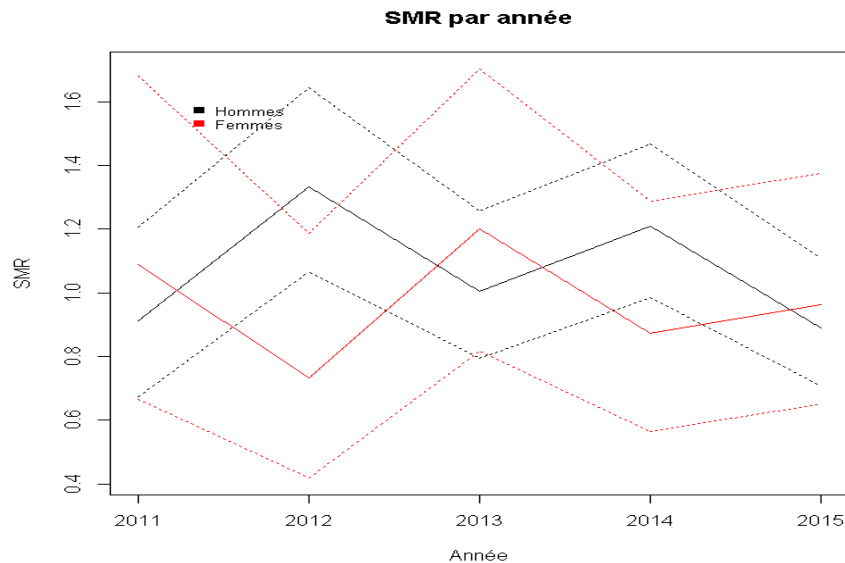
Ratio global observation/ prediction	Brass		Extrapolateur	
	Femme	Homme	Femme	Homme
2011	109,0%	91,2%	109,0%	91,0%
2012	73,1%	133,0%	73,4%	133,0%
2013	120,0%	101,0%	120,0%	100,0%
2014	87,1%	121,0%	87,4%	121,0%
2015	96,3%	89,1%	96,7%	88,9%



# Standardised Mortality Ratio - SMR

Le SMR est un ratio spécifique à la mortalité, il représente le rapport entre le nombre de décès observés, et le nombre de décès prévus, nous allons l'adapter à notre cas de figure :

$$SMR = \frac{\sum d_{x,t}}{\sum ER_{x,t} \hat{q}_{x,t}} \longrightarrow SMR^{extr} = \frac{\sum d_{x,t}}{\sum ER_{x,t} \hat{q}_{x,t}^{extr}}$$



Année	Brass		Extrapolé	
	SMR Femme	SMR Homme	SMR Femme	SMR Homme
2011	1,0885	0,9123	1,0931	0,9097
2012	0,7314	1,3317	0,7341	1,3284
2013	1,2003	1,0053	1,2047	1,0030
2014	0,8713	1,2084	0,8745	1,2059
2015	0,9634	0,8908	0,9666	0,8891



# **Autres Comparaisons des résultats**



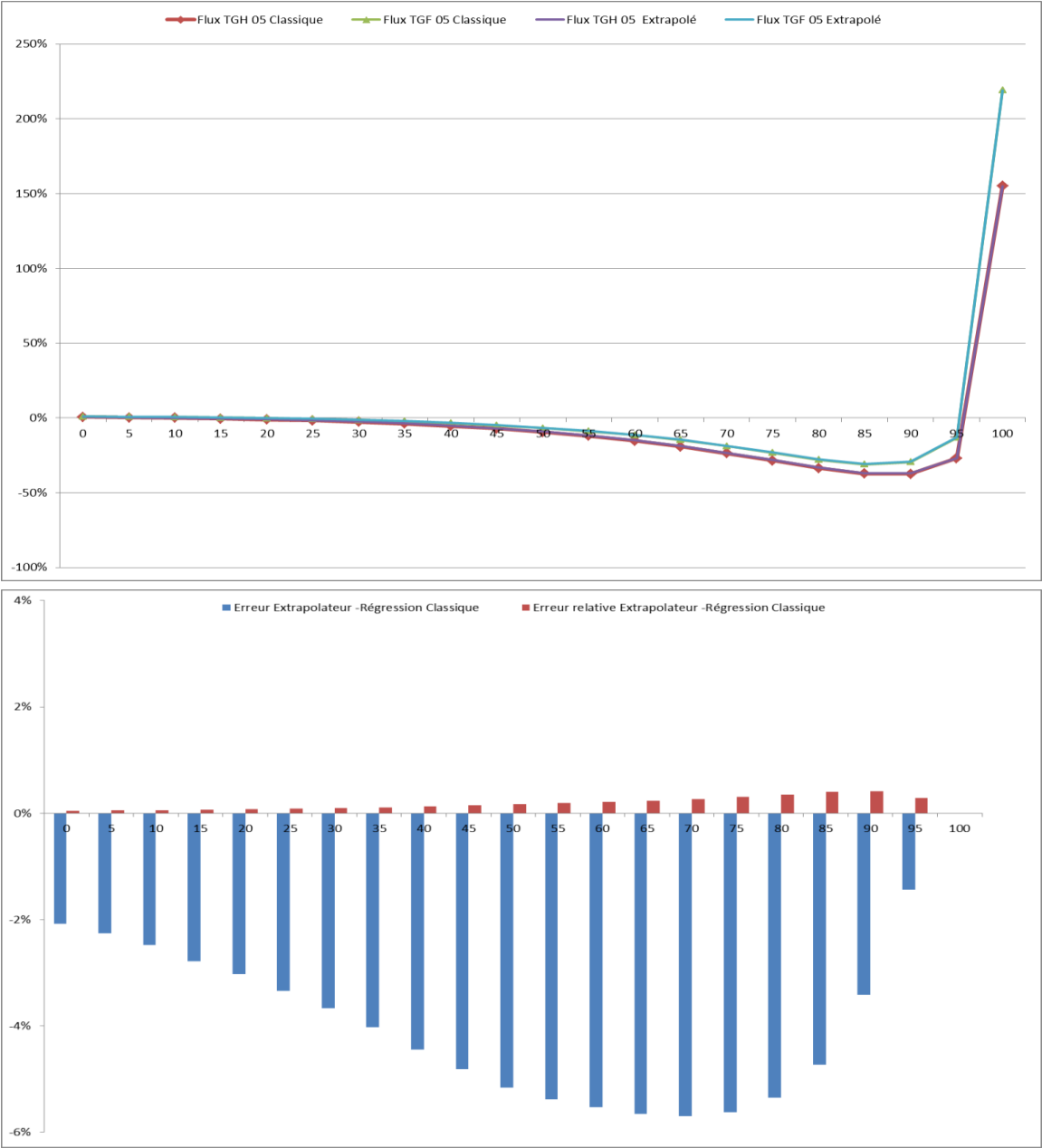
### Comparaison des coefficients de provisionnement calculés avec la table du moment et la table prospective

Considerons un taux d'actualisation de 2%, pour tous génération d'âge, nous estimons leur Capital constitutif prospectif et Capital constitutif instantané, et nous comparons l'écart entre ces deux capitaux.

Age 0:	Capital constitutif prospectif	42,88	43,35	42,88	43,35
	Capital constitutif instantané	42,65	42,95	42,61	42,93
	Ecart	0,55%	0,92%	0,63%	0,97%
Age 10:	Capital constitutif prospectif	41,06	41,64	41,06	41,64
	Capital constitutif instantané	41,09	41,42	41,06	41,39
	Ecart	-0,08%	0,54%	0,01%	0,60%
Age 20:	Capital constitutif prospectif	38,54	39,31	38,54	39,31
	Capital constitutif instantané	38,97	39,38	38,93	39,35
	Ecart	-1,11%	-0,17%	-1,00%	-0,10%
Age 30:	Capital constitutif prospectif	35,42	36,41	35,42	36,41
	Capital constitutif instantané	36,41	36,93	36,36	36,89
	Ecart	-2,74%	-1,42%	-2,61%	-1,32%
Age 40:	Capital constitutif prospectif	31,5	32,78	31,5	32,78
	Capital constitutif instantané	33,29	33,97	33,23	33,93
	Ecart	-5,37%	-3,51%	-5,21%	-3,38%
Age 50:	Capital constitutif prospectif	26,76	28,41	26,76	28,41
	Capital constitutif instantané	29,54	30,48	29,48	30,43
	Ecart	-9,43%	-6,80%	-9,23%	-6,65%
Age 60:	Capital constitutif prospectif	21,26	23,3	21,26	23,3
	Capital constitutif instantané	25,09	26,33	25,02	26,28
	Ecart	-15,26%	-11,51%	-15,02%	-11,33%
Age 80:		TGH 05	TGF 05	TGH 05	TGF 05
Age 80:	Capital constitutif prospectif	9,13	10,96	9,13	10,96
	Capital constitutif instantané	13,74	15,23	13,67	15,18
	Ecart	-33,57%	-28,01%	-33,23%	-27,76%
Age 100:	Capital constitutif prospectif	2,55	3,19	2,55	3,19
	Capital constitutif instantané	1	1	1	1
	Ecart	154,95%	219,20%	154,95%	219,20%



# Comparaison des résultats





# Conclusion

*Le tendomètre*, est un système d'alarmes qui incite à étudier les raisons des retournements de tendance.

*L'extrapoleur dynamique*, n'est ni meilleur, ni pire que les autres méthodes de prédiction, car il n'y a pas de critère de comparaison universel. Sa seule caractéristique qui le distingue des autres et qu'il tient compte non seulement des valeurs passées, mais de leurs dérivées, et donc, de l'aspect dynamique absent des multiples variantes des techniques de régression qui consistent à minimiser des erreurs.

