# 28th INTERNATIONAL CONGRESS OF ACTUARIES

**The international meeting of the actuarial profession**

# 28e CONGRÈS INTERNATIONAL DES ACTUAIRES

**Le rendez-vous international de la profession actuarielle**

# Two dimensionnal recovery curve estimation by using spline surfaces

*Frédéric PLANCHET / Pascal WINTER*

*with the kind help of Michael HILLARY*

WINTER & ASSOCIÉS

I.S.F.A.

# Context

The estimation of temporary disability recovery curves/surfaces is one of the main problems in the evaluation of mathematical reserves and in the quotation of insurance products.
The insurer's experience shows that the exit rate for temporary disability depends on both the age of the insurer and his actual length of disability (ancientness). This leads to consider two dimensional tables, with one dimension for the age at occurrence, and another for the actual length of disability.

This modeling has been chosen by the BCAC in its table defined by the French regulation.
On the raw recovery rates estimation, researches are being lead to generalise the Kaplan Meier estimator in a dimension greater than one (*cf.* Xie et Liu [2000]).

Nonetheless, the classical Kaplan Meier recovery rate estimation (age by age) is still the most robust method to produce raw recovery rates. Furthermore, the loss of information generated by not considering the inter dependence betweenn the two dimensions is small and in practice, negligible.

The aim of this presentation is to point out the use of bi dimensional splines. The bi dimensional splines were first used in the industrial field, and they have rapidly become a well known method (Risler [1991] and De Boor [1978]).
For an overview of statistical smoothing (graduation) methods, one can read Besse and Cardot [2001], an article based on the smoothing in functional spaces.

Nonetheless, the use of this method in the construction of recovery curves is less usual, even if in a one dimensional space it is a well known tool.

Although, smoothing splines in a one dimensional space are widely used for the contrsuction of recovery curves, but it remains more unusual in a two dimensions space.

# Data set description

- Characteristics :

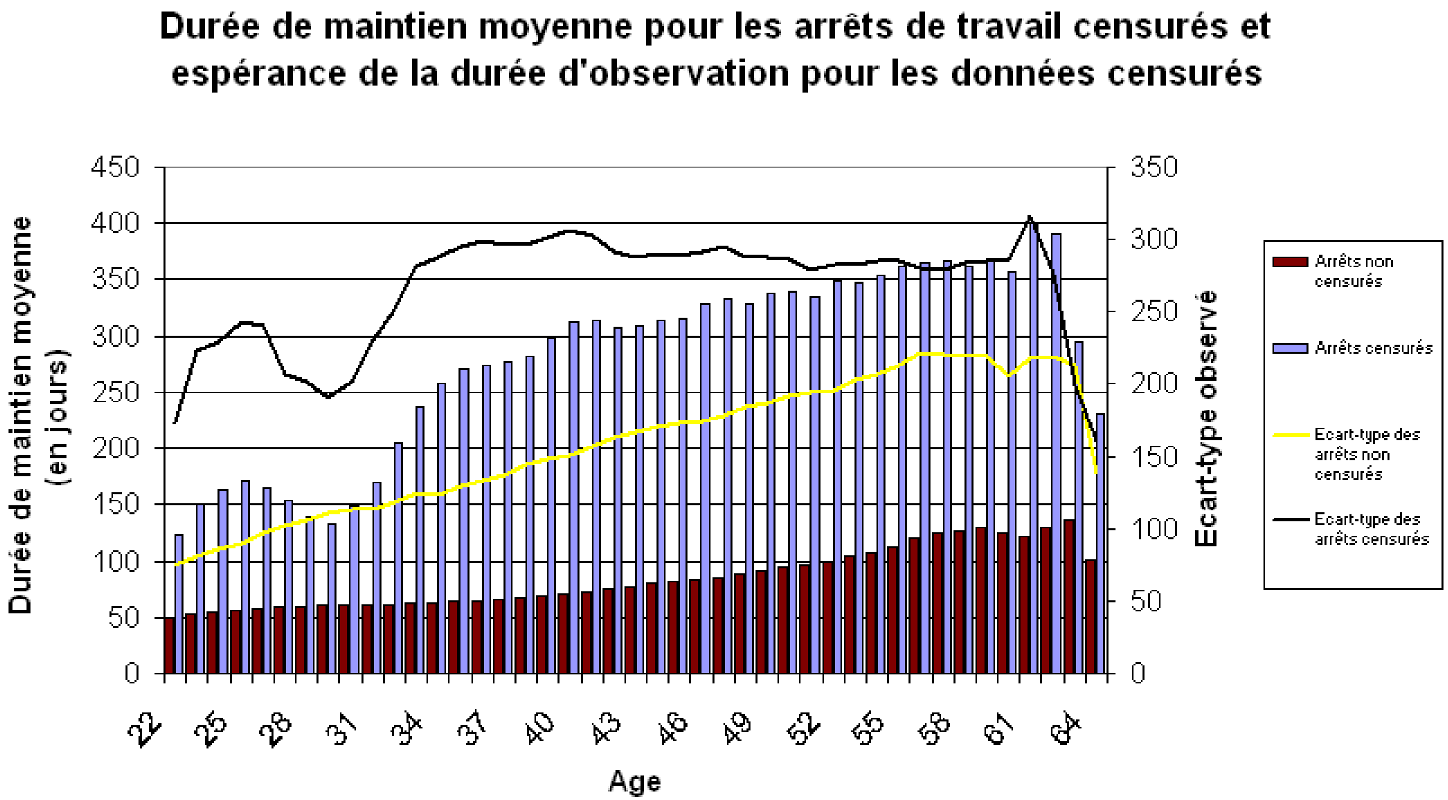
| | |
|---|---|
| ➢Volume | 160 000 |
| ➢Observation period length | 5 years and a ½ |
| ➢Observed variable | Claim length |
| ➢Other variables | Closed / ICOP |

- Data set check
- Basic statistics :

| | |
|---|---|
| ➢Average « closed » claim | 88 days (+10) |
| ➢Standard deviation « closed » claim | 176 days |
| ➢Average « ICOP » claim | 328 days (+10) |
| ➢Standard deviation « ICOP » claim | 284 days |

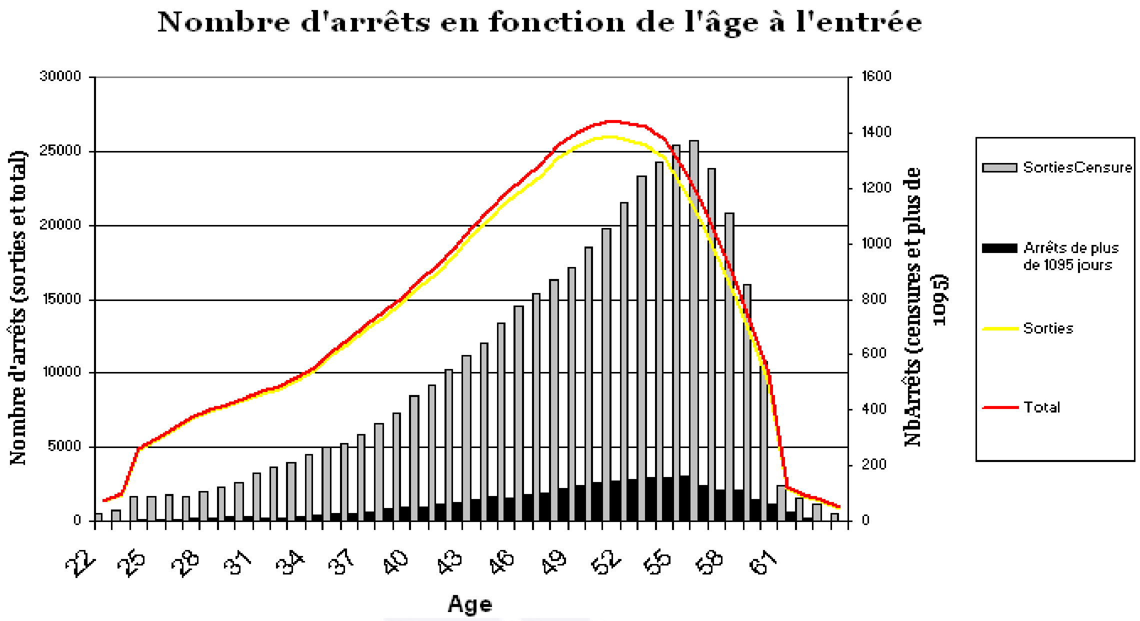
# Data set description

The average claim length has the following shape:



# Data set description

Here is a description of the number of claims by the age of the insured:



# Raw recovery rates computation

Kaplan-Meier estimator is computed age by age :

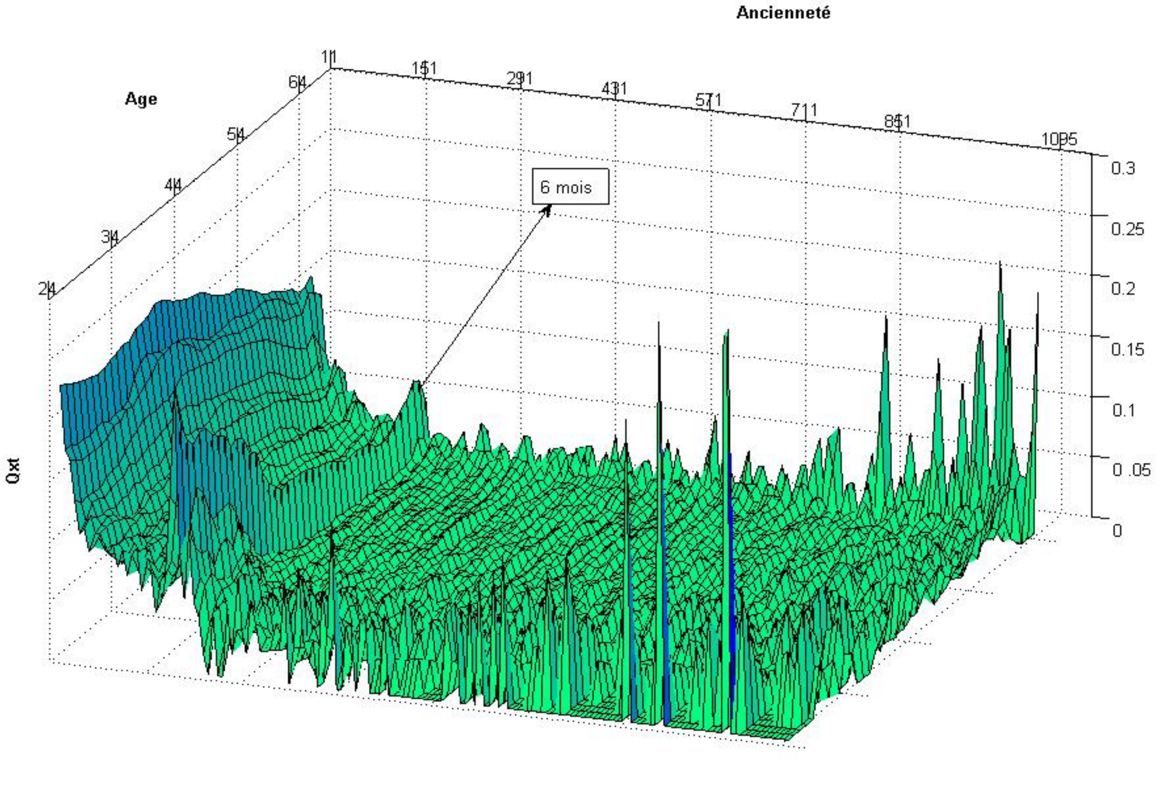
$$\forall t \in \{AncMin, \ldots\ldots, AncMax\} \quad \hat{S}_x(t) = \prod_{T_i \le t} \left(1 - \frac{d_x(T_i)}{n_x(T_i)}\right)$$

Since there are a lot of observed claims, and that the lengths are expressed in days, there are equal values. We have therefore used the following formula:

$$\forall t \in \{AncMin \ldots\ldots AncMax\} \quad \hat{S}_x(t) = \prod_{i=AncMin}^{t} \left(1 - \frac{d_x(i)}{n_x(i)}\right)$$

# Raw recovery rates computation

The following surface is obtained :

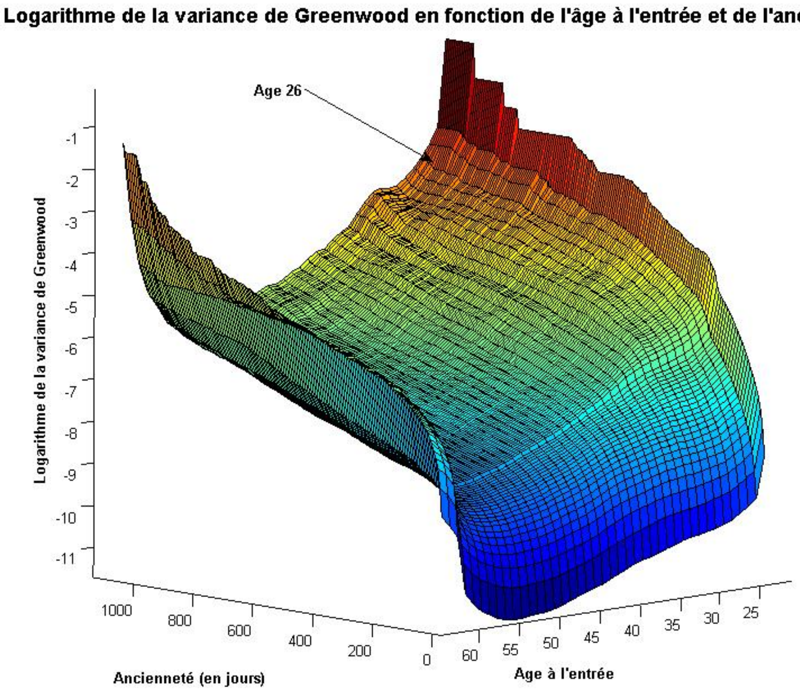



Qxt pour le maintien en Incapacité : taux KM bruts

# Raw recovery rates computation

Greenwood estimator enables us to quantify the raw exit rate volatility:

$$Var\hat{S}_x(t) = \hat{S}_x^2(t) \sum_{i=AncMin}^{t} \frac{d_x(i)}{n_x(i)(n_x(i) - d_x(i))}$$



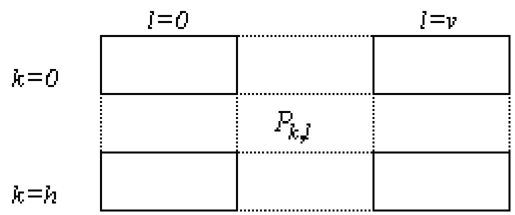

Conclusion:

• The highest and lowest ages have a  high volatilty, due to the lack of data for those class of policy holder

• The volatility tends to increase with the actual length of disability, mainly due to the deacrease of the observed population within the length of the claims

# Spline graduation

The principle of a spline graduation is to make a “patchwork” of subsets on the function area, and then to adjust a simple function (with the least square method) on each subset.
The functions created have to adhere to continuity hypothesis.
The graduation result is highly dependant on the choice of the subsets.
This method allows us to obtain a simpler function than one obtained with a single adjustment on the whole area.




Polynomial functions are simple, and can therefore be used for a spline graduation.
In practice, polynomial functions of degree 3 are used to build “cubic splines”.
The junction of those functions can be obtained by imposing some constraints on the border points.

$$P_{kl}(t,x) = \sum_{(i,j)\in\{0,\ldots,n\}^2} {}_{kl}p_{ij}\,t^i x^j \qquad\qquad P_{kl}(t,x) = {}^{\tau}T_n P_{k,l} X_n$$

# Spline graduation

The regularity criterion (expressed in the constraint equation below) is based on the minimisation of the second differentiate of the spline surface.
To achieve this, the surface has to be C2*. This condition can be expressed under a linear system involving the polynomial functions parameters constituting the spline surface.
By using matrix notations:

$$T_n = \begin{pmatrix} 1 \\ \vdots \\ t^n \end{pmatrix} \qquad X_n = \begin{pmatrix} 1 \\ \vdots \\ x^n \end{pmatrix} \qquad X'_n = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ nx^{n-1} \end{pmatrix} \qquad P_{k,l} = \left({}_{kl}p_{ij}\right)_{(i,j)\in\{0,\dots,n\}^2}$$

The system to solve is :

$$\forall (k,l) \in [\![1,h]\!] \times [\![1,v]\!] \quad \begin{cases} {}^{\tau}T_n(t^*_{k,l}).(P_{k,l} - P_{k-1,l}) = (0,\dots,0) \\ {}^{\tau}T'_n(t^*_{k,l}).(P_{k,l} - P_{k-1,l}) = (0,\dots,0) \\ {}^{\tau}T''_n(t^*_{k,l}).(P_{k,l} - P_{k-1,l}) = (0,\dots,0) \\ (P_{k,l} - P_{k,l-1}).X_n(x^*_{k,l}).= {}^{\tau}(0,\dots,0) \\ (P_{k,l} - P_{k,l-1}).X'_n(x^*_{k,l}).= {}^{\tau}(0,\dots,0) \\ (P_{k,l} - P_{k,l-1}).X''_n(x^*_{k,l}).= {}^{\tau}(0,\dots,0) \end{cases}$$

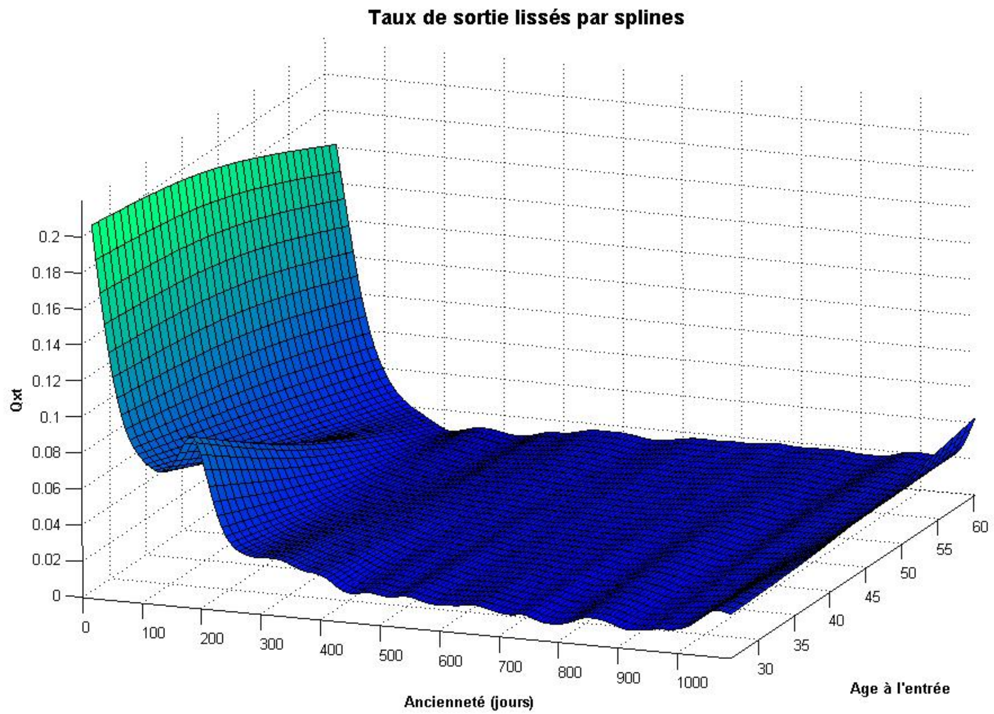
Under the constraint :

$$\alpha \sum_x \sum_t \omega_{x,t} \left(SP(t,x) - q_{x,t}\right)^2 + (1-\alpha) \iint \lambda(s,y) \left| D^2 P(s,y) \right|^2 dsdy$$

* : A C2 function is a function wich is two times differentiable and whose the second differentiate is continuous.

# Spline graduation

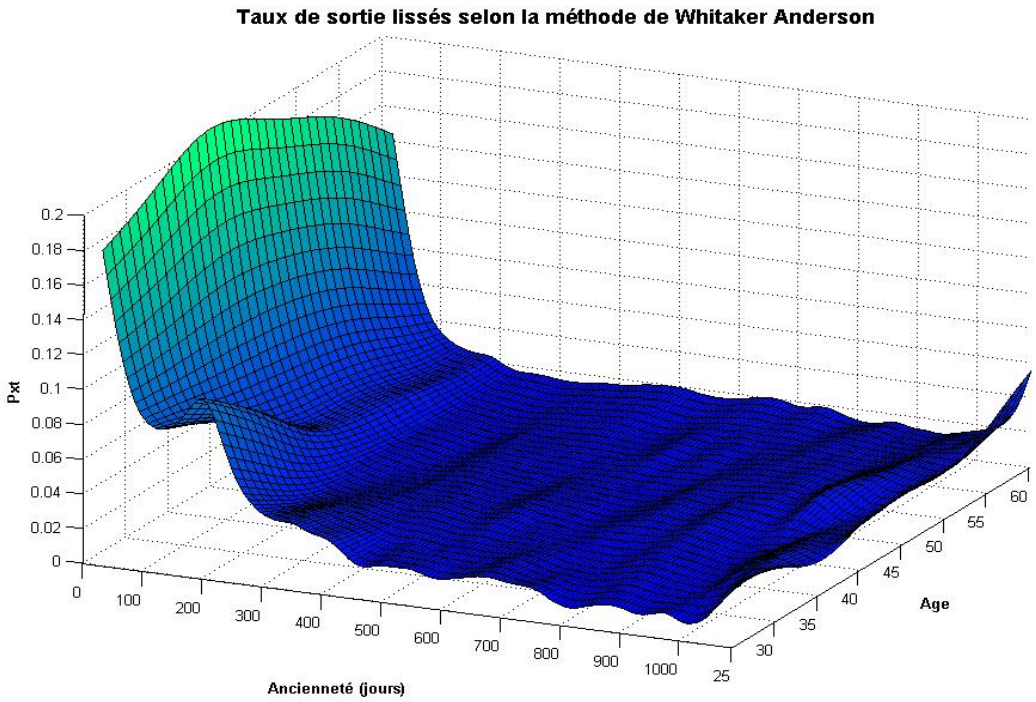
The following results are obtained :




We can a see a « peak » at 6 months, and also many little waves at certain durations. These are the consequences of the issue condition of temporary disability: the lengths of the claims do not take all the possible values, since the doctors tend to choose a standard unit (the month for example).

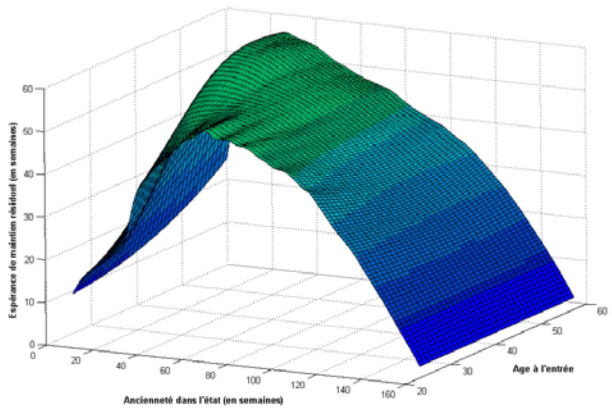
# Whittaker Henderson in two dimensions

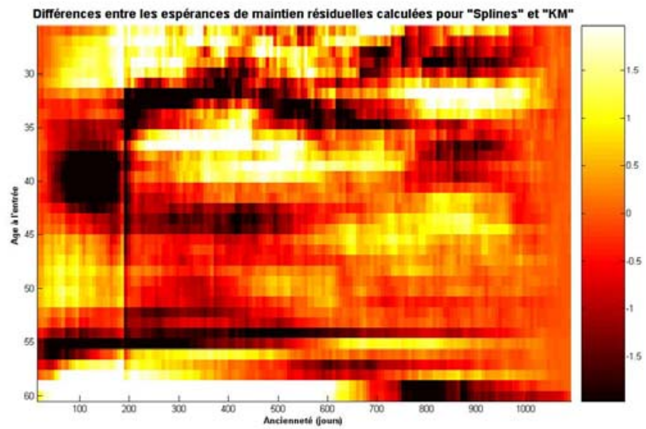
To compare, here are the results obtained with the Whittaker Henderson method on two dimensions:

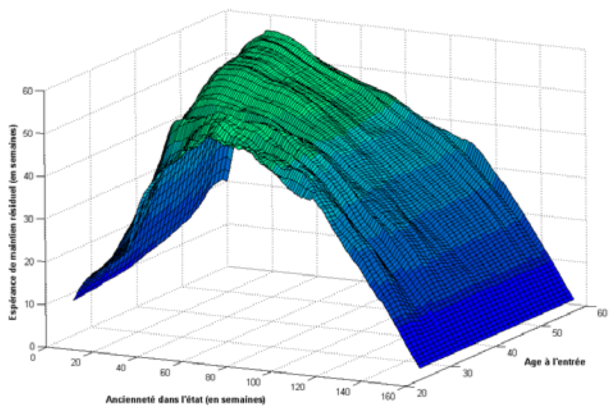


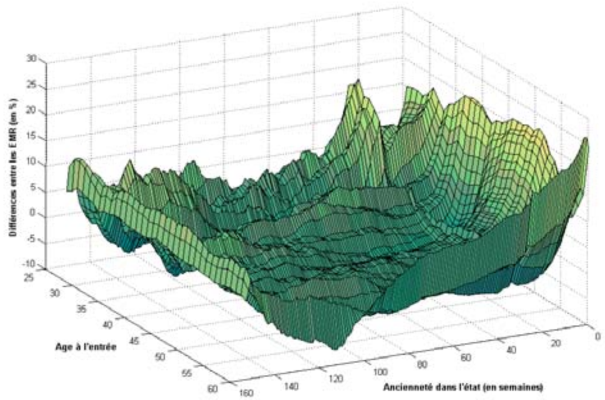
# Spline graduation

Comparison beetween the raw remaining expected durations and the smoothed ones.



# Adjustment quality

A Chi square statistic is used :

$$W_{k,l} = \sum_{x^*=1}^{l} \sum_{t^*=1}^{k} \frac{(D_{x^*,t^*} - \tilde{D}_{x^*,t^*})^2}{\tilde{D}_{x^*,t^*}}$$

The results of other smoothing methods are summarized in the following tab:

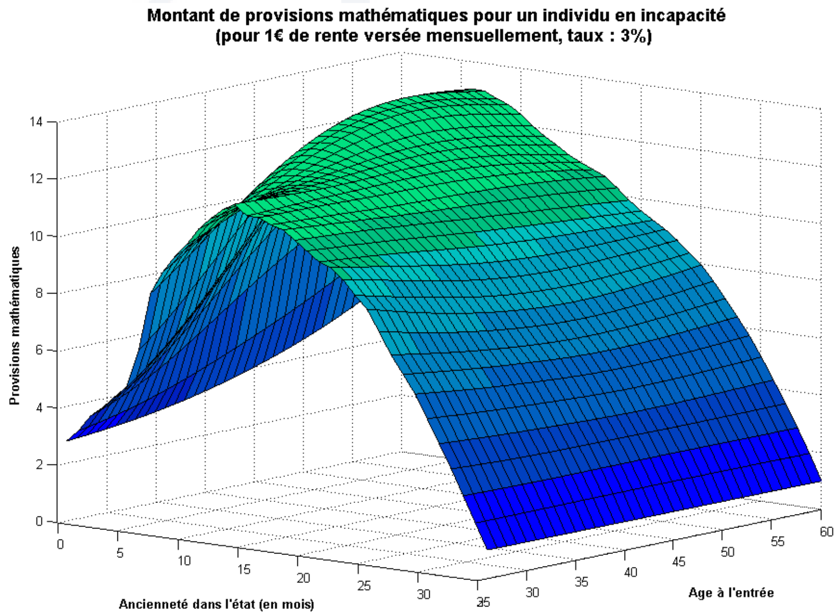
| Table de référence: | Kaplan Meier brut |
|---|---|
| Tables | Chi2 |
| Lee Carter | 151.39 |
| Lee Carter lissé | 130.07 |
| Log Poisson | 42.36 |
| Log Poisson Lissé | 52.07 |
| Spline1 | 7.46 |
| Spline2 | 10.27 |
| Whittaker-Henderson1 | 6.62 |
| Whittaker-Henderson2 | 5.92 |
| Inversion1 | 7.89 |
| Inversion2 | 6.97 |

# Reserving impact

The resulting table can be used for the computation of “In Course Of Payments” reserves and “Notified But Not Admitted” reserves, for temporary disability claims. For a given actuarial rate *i*, the reserve coefficient can be expressed with the following formula:

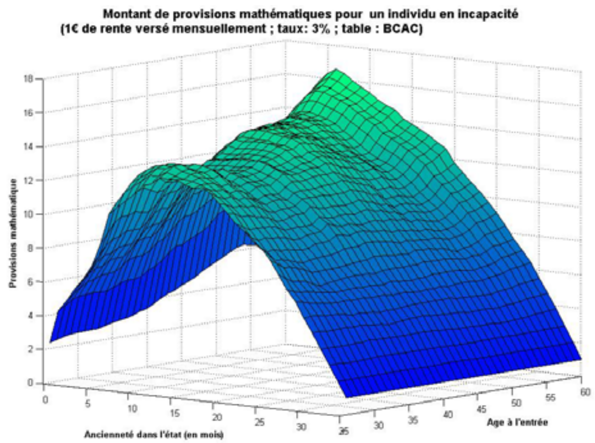
$$ {}_{INC}PM_y^x = \frac{1}{{}_{INC}L_y^x} \sum_{k=0}^{36-y} \frac{{}_{INC}L_{k+y}^x}{(1+i)^{\frac{k}{12}}} $$

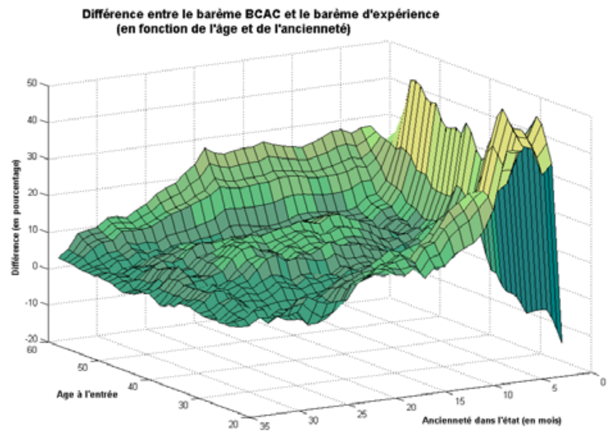
With a rate equal to 3%, we have the following surface :

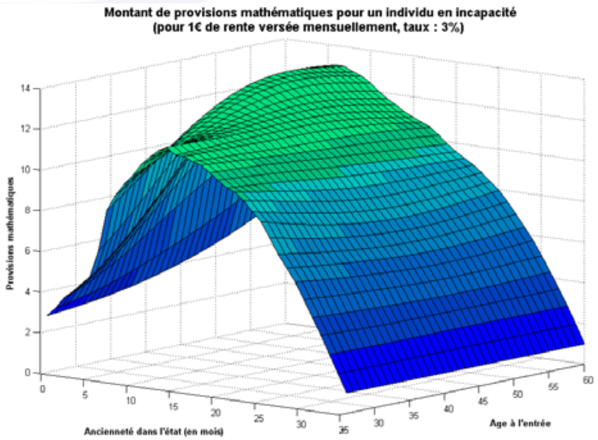


# Reserving impact

Comparison with the BCAC table




- Conclusions :

Structural differences between the 2 tables.
On average, the experience table gives reserves 5% to 10% lower than the ones computed with the BCAC table.

# Conclusion

A two dimensional adjustment on the raw exit rate with two dimensional splines allows us to build a continuous parametric surface for the recovery rates. Its numerical calculation is not that difficult with an adapted tool.

The advantage of this method, compared to others, is that it takes into account the dependencies between the two dimensions (age and disability actual duration) of the recovery curve.

We therefore obtain a direct representation of the interactions within the two directions.
Furthermore, this two dimensional approach allows us to estimate a recovery curve for ages for which there is a lack of data: polynomial functions are easy to extrapolate.

This method can be used for different applications than temporary disability, where the recovery curves (or the mortality tables) involve two dimensional dependencies

- Total disability mortality table
- Prospective mortality tables (estimating future mortality)

# Conclusion

The continuous parametric nature of this model also allows interpolation, which enables us to obtain tables for different time units. This can simplify the computation of the expected annuity payments, whatever the lag between the payments is.

The practical use of the technique presented here on temporary disability coverage is immediate. Work is in process to go further and build a table without waiting period (“on the first day”). The aim is to obtain a tool which would enables us to price individual contracts which would cover the first days of temporary disability. To achieve this, we are using the work by Bonche et al. [2005].