Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points?

Pierre-Olivier GOFFARD*, Xavier GUERRAULT**

* AXA France and Aix-Marseille University, pierreolivier.goffard@axa.fr
** AXA France, xavier.guerrault@axa.fr

Abstract

An agregation method adapted to life insurance portfolios is presented. The aim of this work is to reduce the running time induced by the increasing complexity of actuarial models. The method is a two steps procedure. The first step consists in using statistical partitioning methods in order to gather insurance policies. The second step is the construction of a representative policy for each aforementioned groups. Running times reduction techniques are motivated by the production in a timely manner of more and more sensitivity analysis, required for both decision making and regulatory reporting. The efficiency of the agregation method is illustrated on a real saving contracts portfolio within the frame of a cash flows projection model used for best estimate liabilities and solvency capital requirements computations. The procedure is already part of AXA France valuation process.

Keywords: Solvency II, claim reserving, statistical partitioning and classification, functional data analysis, longitudinal data classification.

1 Introduction

Cash flows projection models are getting more and more sophisticated so as to reflect expected realistic future demographic, legal, medical, technological, social or economic developments. These models are designed to provide realistic evaluations of the financial reserves of an insurance company. Best Estimate Liabilities (BEL) are the probability weighted average of future cash flows taking into account the time value of money. The use of Monte-Carlo simulations techniques induces large running time for a policy-by-policy approach. In addition to baseline runs, several sensitivities analysis are needed to understand the dependency of resulting financial reserves on different input and parametrizations. Running time is of prime importance because these studies usually need to be performed within tight deadlines. Actuaries have therefore to develop efficient modeling techniques to tackle this issue. Grouping methods, which group policies into model cells and replace all policies in each groups with a representative policy, is the oldest form of modelling techniques. It has been advised in [6] and is commonly used in practice due to its availability in commercial actuarial softwares such as MG-ALFA and GGY-Axis, see [8]. Recently an alternative technique, based on statistical sampling and extrapolation, has been proposed. This method has been put into practice by Towers and Watson and is the subject of a recent PhD thesis [18]. In this paper, we present a new agregation method that is applied to the agregation of a life insurance portfolio. This work has been motivated by the recast of the valuation process in AXA France to prepare the enforcement of Solvency II. The agregation method needed to meet many practical requirements such as easy implementation and understanding, appliacability to the whole life insurance portfolio and good empirical and theoritical justification. We believe that many insurance companies face running time problems, and this paper may help practionners to solve it. Section 1 is an introduction that presents the context and the motivations of this work. In Section 2, we give a description of a cash flows projection model. In Section 3, we describe a statistical based way to group policies, followed by procedures to construct the representative policy. Section 4 illustrates the efficiency of the method on a portfolio of saving contracts on the french market.

2 Cash flows projection models and best estimate liabilities evaluations

We consider a classical asset-liability model to project the statutory balance sheet and compute mathematical reserves, described for instance in [1]. This model allows computations of reserves linked to any type of guarantee. For the sake of simplicity, we describe a cash flow projection model for life insurance of saving type. Consider a saving contract with surrender value $SV(0)$ at time $t = 0$, the surrender value at time t is defined as

$$SV(t) = SV(0) \times exp\left(\int_0^t r_a(s)ds\right), \tag{2.1}$$

where r_a denotes the instantaneous accumulation rate (including any guaranteed rate) modeled by a stochastic process. We are interested in the present surrender value at time t, we need therefore to add a discount factor in order to define the Present Surrender Value

$$\begin{aligned} PSV(t) &= SV(t) \times exp\left(-\int_0^t r_\delta(s)ds\right), \\ &= SV(0) \times exp\left(\int_0^t (r_a(s) - r_\delta(s))ds\right), \end{aligned} \tag{2.2}$$

where r_δ denotes the instantaneous discount rate, also modeled by a stochastic process. The spread between the accumulation and the discount rates in (2.2) is then a stochastic process. Let τ be the early surrender time, modeled by a random variable having probability density function f_τ on the positive half line, absolutely continuous with respect to the Lebesgue measure. The probability density function can be interpreted as an instantaneous surrender rate between times t and $t + dt$, which depends on the characteristics of the policy holder like for instance his age, his gender or the seniority of his contract. More specifically, in the case of a saving contract, the payment of the surrender value occurs in case of early withdrawal due to lapse or death, or expiration of the contract. In the cash flows projection model definition, an horizon of projection is usually specified. There are also life insurance contracts with a fixed term. Both of these instants are deterministic. We denote by T the minimum of the expiration date of the contract and the horizon of projection. The real surrender time $\tau \wedge T = min(\tau, T)$ is a random variable associated with a probability measure divided into the sum of a singular part and a continuous part

$$dP_{\tau\wedge T}(t) = f_\tau(t)d\lambda(t) + \overline{F_\tau}(T)\delta_T(t), \tag{2.3}$$

where λ is the Lebesgue measure on $[0, T]$, δ_T is the Dirac measure at T and $\overline{F_\tau}(t)$ denotes the survival function associated to the random instant τ. The BEL at time $t = 0$ is defined as

$$BEL(0,T) = E^{P_{\tau\wedge T}\otimes Q^f}(PSV(\tau \wedge T)), \tag{2.4}$$

where Q^f is a probability measure that governs the evolution of instantaneous rates. We refer to [10] for this definition of best estimate liabilities. Let $\mathbf{F}$ be a financial scenario, which correponds to a trajectory of the spread between the accumulation and discount rates. We write the BEL given a financial scenario, assuming no dependency between surrender probabilities and financial scenarios, as follows

$$\begin{aligned} BEL^{\mathbf{F}}(0,T) &= E^{P_{\tau\wedge T}}(PSV(\tau \wedge T)|\mathbf{F}) \\ &= \int_0^{+\infty} SV(0) \times exp\left(\int_0^t (r_a(s) - r_\delta(s))ds\right) dP_{\tau\wedge T}(t) \\ &= \int_0^{T} SV(0) \times exp\left(\int_0^t (r_a(s) - r_\delta(s))ds\right) f_\tau(t)dt \\ &+ \overline{F_\tau}(T) \times SV(0) \times exp\left(\int_0^T (r_a(s) - r_\delta(s))ds\right) \end{aligned} \tag{2.5}$$

In order to avoid tedious calculations of the integral in (2.5), time is often discretized. The BEL is therefore writen as

$$BEL^{\mathbf{F}}(0,T) \approx \left[\sum_{t=0}^{T-1} p(t,t+1) \prod_{k=0}^{t} \frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)} + p(T) \prod_{k=0}^{T-1} \frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)}\right] SV(0), \tag{2.6}$$

where $p(t,t+1)$ is the probability that surrender occurs between time t and $t+1$, and $r_a(t,t+1)$ and $r_\delta(t,t+1)$ are the accumulation and discount rates between time t and $t+1$. The probabilities $p(t,t+1)$ necessitate the evaluation of a mortality and a lapse rate from one period to another using classical actuarial tools, see [17]. Monte-Carlo methods for BEL evaluation consists in generating a set of financial scenarios under Q^f and compute the BEL for each one of them. The final estimation is the mean over the set of all scenarios. This procedure is fast enough for one policy, it becomes time consuming for a large portfolio. It is worth noting that we use this Cash flows projection model to illustrate the agregation procedure. We just describe here a model that is used by practitioners. The purpose of this work is not to comment the validity of the model, we agree on the fact that a lot of theoretical questions might arise from its definition.

3 Presentation of the agregation procedure

The goal of agregation procedures is to reduce the size of the input portfolio of the cash flows projection model. The first step consists in creating groups of policies sharing similar features. The second one is the definition of an "average" policy, called Model Point (MP), that represents each group and forms the agregated portfolio. The initial surrender value of the MP is the sum of the initial surrender values over the represented group. The method must be flexible so as to generate the best agregated portfolio under the constraint of a given number of MP. Let us consider two contracts having identical characteristics (same age, same seniority,...) and therefore having identical surrender probabilities. We build a contract having these exact characteristics and whose initial surrender value is the sum of the initial surrender values of the two aforementionned contracts. The BEL of this contract, given a financial scenario, is

$$
\begin{aligned}
BEL^F_{MP}(0,T) &= \left[\sum_{t=0}^{T-1} p(t,t+1)\prod_{k=0}^{t}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)} + p(T)\prod_{k=0}^{T-1}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)}\right] \\
&\times\ SV_{MP}(0) \\
&= \left[\sum_{t=0}^{T-1} p(t,t+1)\prod_{k=0}^{t}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)} + p(T)\prod_{k=0}^{T-1}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)}\right] \\
&\times\ \sum_{i=1}^{2} SV_{C_i}(0) \\
&= \sum_{i=1}^{2}\left[\sum_{t=0}^{T-1} p(t,t+1)\prod_{k=0}^{t}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)} + p(T)\prod_{k=0}^{T-1}\frac{1+r_a(k,k+1)}{1+r_\delta(k,k+1)}\right] \\
&\times\ SV_{C_i}(0) \\
&= \sum_{i=1}^{2} BEL^F_{C_i}(0,T),
\end{aligned}
$$

where $BEL^F_{C_i}(0,T)$ and $SV_{C_i}(0)$ are the best estimate liability and initial surrender value of the contract $i \in \{1,2\}$. The idea behind the grouping strategy lies in this additivity property of the BEL. The agregation of contracts having the same surrender probabilities leads to an exact evaluation of the BEL of the portfolio. The creation of an agregated portfolio by grouping the policies having identical characteristics leads to a portfolio that is usually still too big to perform an valuation. Nevertheless, as it is smaller than the input portfolio the use of partitioning algorithms will be faster and one valuation might be doable in order to get a benchmark value for the BEL and assess the accuracy of the agregation procedure in the validation phase. We describe in the first subsection how to gather contracts having close surrender probabilities and in the second subsection how to build a representative contract for each group.

3.1 The partitioning step

A portfolio is a set of contracts $\mathcal{P} = \{\mathbf{x}_i\}_{i\in 1,\dots,n}$, where n is the size of the portfolio and each observation $\mathbf{x}_i$ is a vector. We aim to partition the n contracts into k sets $\mathcal{C} = \{C_1,\dots,C_k\}$. The idea is to use clustering algorithms widely used in datamining to identify sub-populations. A choice has to be made concerning the d variables that caracterize each observation and the metric that permits to measure the dissimilarity between two observations. In order to get closer to the additivity of the BEL, every individuals in the portfolio is represented by its sequence of surrender probabilities

$$\mathbf{x}_i = \left(p_i(0,1), p_i(1,2), \dots, p_i(T-1,T), p_i(T)\right). \tag{3.2}$$

Policies are therefore located in a vector space of size $T+1$, the variables are quantitative and fall between 0 and 1. The natural dissimilarity measure between two observations is the euclidean distance defined as

$$||\mathbf{x}_i - \mathbf{x}_j||_2 = \sqrt{(p_i(0,1)-p_j(0,1))^2 + \dots + (p_i(T)-p_j(T))^2}. \tag{3.3}$$

Partitioning algorithms are designed to find the k-sets partition that minimizes the Within Cluster Sum of Square (WCSS), which characterises the homogeneity in a group,

$$\tilde{\mathcal{C}} = \arg\min_{\mathcal{C}} \sum_{j=1}^{k} \sum_{\mathbf{x} \in C_j} ||\mathbf{x} - \mu_j||_2^2, \tag{3.4}$$

where μ_j is the mean of the observations that belongs to the set C_j. The two foreseen methods are the so called KMEANS procedure and the agglomerative hierarchical clustering procedure with Ward criterion. These two standard methods are described in [11, 7], and Ward criterion has been introduced in [19].
The KMEANS algorithm starts with a random selection of k initial means or centers, and proceeds by alternating between two steps. The Assignement step permits to assign each observation to the nearest mean and the update step consists in calculating the new means resulting from the previous step. The algorithm stops when the assignements no longer change. The algorithm that performs an agglomerative hierarchical clustering uses a bottom up strategy in the sense that each observation forms its own cluster and pairs of cluster are merged sequentially according to Ward criterion. At each step, the number of clusters decreases of one unit and the WCSS increases. This is due to the Huygens theorem that divides the Total Sum of Square (TSS), into Between Cluster Sum of Square (BCSS) and WCSS,

$$\begin{array}{ccccc} \sum_{\mathbf{x} \in \mathcal{P}} ||\mathbf{x} - \mu||_2^2 & = & \sum_{j=1}^{k} \sum_{\mathbf{x} \in C_j} ||\mathbf{x} - \mu_j||_2^2 & + & \sum_{j=1}^{k} ||\mu_j - \mu||_2^2 \\ TSS & = & WCSS(k) & + & BCSS(k) \end{array}$$

Note that TSS is constant and that BCSS and WCSS evolve in opposite directions. The application of Ward's criterion leads to the agregation of the two individuals that goes along with the smallest increase of WCSS at each step of the algorithm. The best way to visualize the data is to plot "surrender trajectories" associated with each policy as in Figure 4. Our problem is analogous to the problem of clustering longitudinal data that arises in biostatistics and social sciences. Longitudinal data are obtained by doing repeated measurements on a same individual over time. We choose a non parametric approach, also chosen in [9, 12]. A parametric approach is also possible by assuming that the dataset comes from a mixture distribution with a finite number of components, see [16] for instance. We believe that the non parametric approach is easier to implement and clearer from a practionner point of view.
The KMEANS method, that takes the number of clusters as a parameter, seems to be more suited to the problem than the agglomerative hierarchical clustering method. In the Agglomerative Hierarchical Clustering (AHC) algorithm the grouping in k sets depends on the previous grouping. Furthermore, the KMEANS algorithm is less greedy in the sense that fewer distances need to be computed. However the KMEANS algorithm is an optimization algorithm and the common problem is the convergence to a local optimum due to bad initialization. To cope with this problem, we initialize the centers as the means of clusters builded by the AHC, thus we ensure that the centers are geometrically far from each other. The value of the BEL is highly correlated to the initial surrender value, thus a bad representation of some policy having a significant initial surrender value gives rise to a significant negative impact on the estimation error after agregation. We decide to define a weight according to the initial surrender value

$$w_{\mathbf{x}} = \frac{SV_{\mathbf{x}}(0)}{\sum_{\mathbf{x} \in \mathcal{P}}^{n} SV_{\mathbf{x}}(0)}, \tag{3.5}$$

and define the Weighted Within Cluster Inertia - WWCI as

$$WWCI(k) = \sum_{j=1}^{k} \sum_{\mathbf{x} \in C_j} w_{\mathbf{x}} ||\mathbf{x} - \mu_j||_2^2, \tag{3.6}$$

where μ_j becomes the weighted mean over the set C_j. Each surrender path is then stored in a group.

Remark 1. *A time continuous approach within the cash flows projection model would leave us with probability density function to group. The problem would be analogous to the clustering of functional data that have been widely studied in the literature. A recent review of the different techniques has been done in [13].*

3.2 The agregation step

The agregation step leads to the definition of a representative policy for each group resulting from the partitioning step. Probabilities of surrender depend on characteristics of the contracts and of the policyholders. The best choice as a representative under the least square criterion is the barycenter. Its surrender probabilities are defined through a mixture model

$$f_{\tau_C}(t) = \sum_{\mathbf{x} \in C} w_i f_{\tau_{\mathbf{x}}}(t), \tag{3.7}$$

where C is a group of policies and τ_C is the random early surrender time for every member of C. The equivalent within a discrete vision of time is a weighted average of the surrender probabililities with respect to each projection year. The probability density function of surrender of a given contract is associated to its age and seniority. The PDF defined in (3.7) is not associated to an age and a seniority. This fact might give rise to an operational problem if every MP in the agregated portfolio need to have an age and a seniority. The number of suitable combinations of age and seniority fall into a finite set given the possible features of policies. It is then possible to generate every "possible" surrender probability density functions in order to choose the closest to the barycenter. This optimal density function might be associated with a policy (or equivalently a combination of an age and a seniority) that does not exist in the initial portfolio.

4 Illustration on a real life insurance portfolio

The procedure is illustrated within the frame of a saving contracts portfolio extracted from AXA France portfolio. The mechanism of the product is quite simple. The policyholder makes a single payment when subscribing the contract. This initial capital is the initial surrender value that will evolve during the projection, depending on the investment strategy and the financial scenario. The policy owner is free to withdraw at any time. In case of death, the surrender value is payed to the designated beneficiaries. The contractual agreement does not specify a fixed term. The projection horizon is equal to 30 years. Best estimate liabilities are obtained under a discrete vision of time. Mortality rates are computed using an unisex historical life table. Mortality depends therefore only on the age of the insured. Lapse probabilities are computed with respect to the observed withdrawal and is assumed to be dependent on the seniority of the contracts. The main driver that explains the withdrawal behavior is the specific tax rules applied on french life insurance contracts. Financial scenarios are generated through stochastic modeling. The instantaneous interest rates are stochastic processes and simulations are completed under a risk neutral probability. The number of policies and the amount of the initial reserves in the portfolio are given in Table 1. Surrender probabilities depend only on the age and

Number of policies	Mathematical provision (euros)
140 790	2 632 880 918

Table 1: Number of policies and amount of the initial surrender value of the portfolio

the seniority. The heterogeneity of the trajectories depends on the distribution of ages and seniorities in the portfolio. Statistical descriptions are given in tables 2 and 3. In

Variable: AGE			
Mean	Standard deviation	Minimum	Maximum
49.09	18.57	1	102

Table 2: Statistical description of the variable AGE in the portfolio

Variable: SENIORITY			
Mean	Standard deviation	Minimum	Maximum
4.10	1.63	1	7

Table 3: Statistical description of the variable SENIORITY in the portfolio

Section 3, it has been pointed out that an exact evaluation of the BEL is obtained with a portfolio that agregates policies having the same key characteristics. In our modeling, policies that have identical age and seniority are grouped together in order to have a first agregation of the portfolio that provides an exact value of BEL, the number of MP in this agregated portfolio and the resulting value of BEL are reported in Table 4. We define the

Number of policies	BEL (euros)
664	2 608 515 602

Table 4: Number of MP and best estimate liability of the agregated portfolio

error by the difference between the exact BEL and the BEL obtained with an agregate portfolio. We also want to compare the two agregation ways discussed in Section 3.2. One corresponds to the exact barycenter of each group (METHOD=BARYCENTER), the other being the closest-to-barycenter policy associated with an age and a seniority (METHOD=PROXYBARYCENTER). These procedures are also compared to a more "naive" grouping method (METHOD=NAIVE) that consists in grouping policies having the same seniority and belonging to a given class of age. The classes are simply defined by the quartiles of the distribution of ages in the portfolio. The age and seniorities of the MP associated with each group are obtained by a weighted mean. The weights are defined as in (3.5). The "naive" method leads to an agregated portfolio with 28 MP. The errors for the three methods with 28 MP are given in Table 5. The two proposed methods outperform greatly the naive one. On Figure 1, one can notice that the proposed methods permit a better accuracy even for a smaller number of MP. The use of the barycenter is the best choice. The question of the optimal choice of the number of clusters arises naturally. The

BEL error (euros) METHOD=BARYCENTER	BEL error (euros) METHOD=PROXYBARYCENTER	BEL error (euros) METHOD=Naive
-10 880	-199 734	1 074 983

Table 5: Best estimate liabilities error with 28 model points depending on the agregation method

optimal number of clusters has been widely discussed in the literature, and there exists many indicators. The main idea is to spot the number of clusters for which the WWCI reaches a sort of plateau when the number of clusters is increasing, see Figure 2.

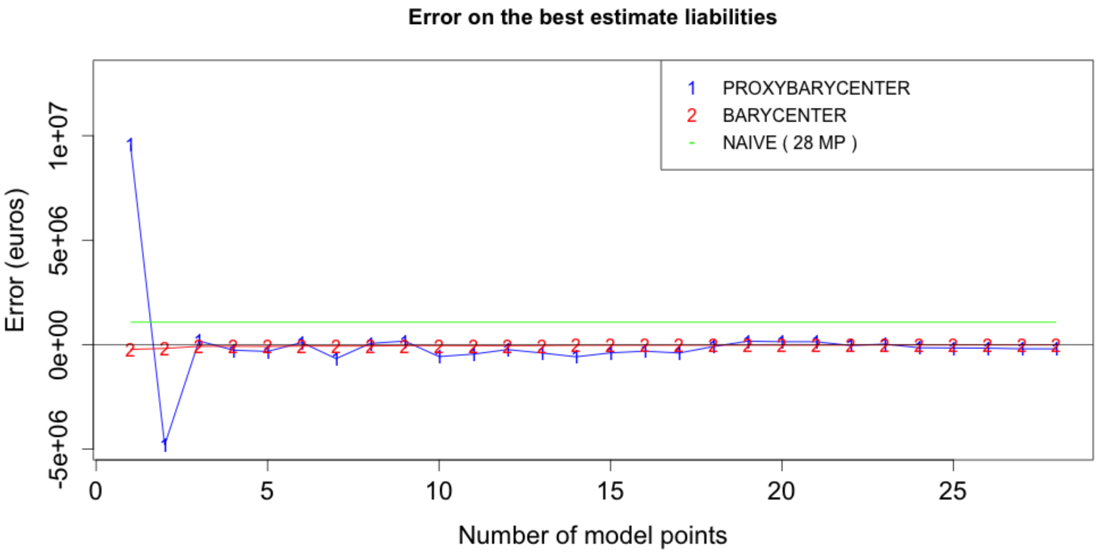

Figure 1: Error on the BEL evaluation depending on the number of model points and the agregation method

Optimal number of clusters might be 3 or 6. In order to automatize the choice, we can use indicators. The relevance of such indicators often depends on the partitionning problem. We need to choose the best suited to our problem. Among the indicators recommended in the literature, there is the index due to Calinsky and Harabasz, defined in [2] as

$$CH(k) = \frac{WBCI(k)/(k-1)}{WWCI(k)/(n-k)}, \tag{4.1}$$

where n is the number of observations, WBCI and WWCI denote the Weighted Between and Weighted Within Cluster Inertia. The idea is to find the number of clusters k that maximises CH. Note that $CH(1)$ is not defined. This indicator is quite simple to understand, as a good partition is characterized by a large WBCI and a small WWCI. Another indicator has been proposed in [15] as follows: First define the quantity

$$DIFF(k) = (k-1)^{2/p} \times WWCI(k-1) - k^{2/p} \times WWCI(k), \tag{4.2}$$

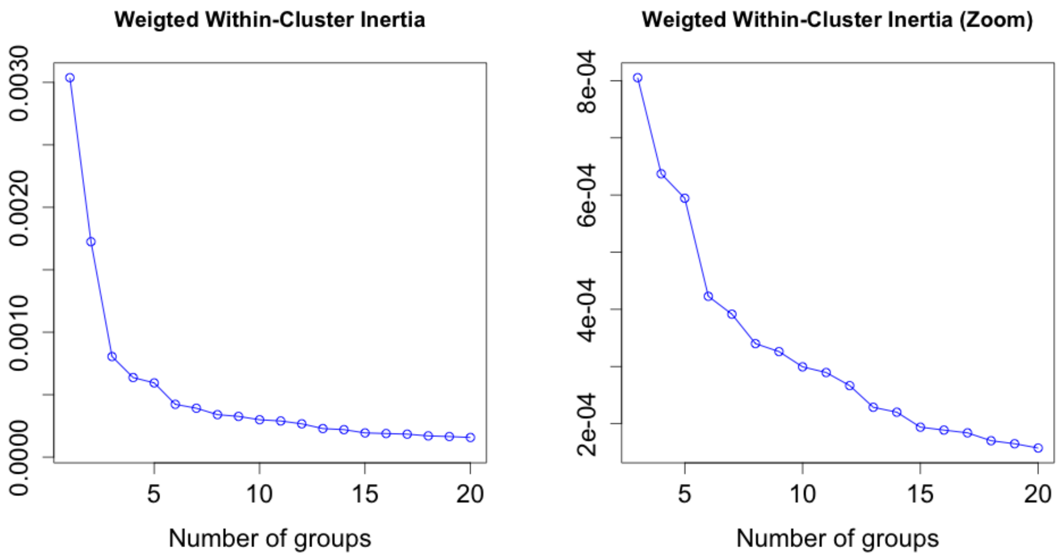

Figure 2: WWCI evolution depending on the number of clusters

and choose k which maximises

$$KL(k) = \left| \frac{DIFF(k)}{DIFF(k+1)} \right|. \tag{4.3}$$

This permits to compare the decreasing of WWCI in the data with the decreasing of WWCI within data uniformly distributed through space. The silhouette statistic has been introduced in [14] and is defined, for each observation i, by

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}, \tag{4.4}$$

where $a(i)$ is the average distance between i and the others points in its cluster, and $b(i)$ is the average distance from i to the data points in the nearest cluster besides its own. A point is well clustered when $s(i)$ is large. The optimal number of clusters maximises the average of $s(i)$ over the data set. The different indicators have been computed for every partition ranging from one to twenty groups, see Figure 3. The different indicators seems to retain 3 clusters (except for KL, that is maximized for 6 clusters but still have a large value for 3 clusters). The 3-groups partition of the portfolio portfolio can be visualized using the surrender trajectories, that are displayed in Figure 4. The errors on the BEL, normalized by its exact value and expressed in percentage are reported in Table 6. One can note again that the presented methods outperform greatly the naive one with less model points. Our

BEL error % BARYCENTER	BEL error % PROXYBARYCENTER	BEL error % NAIVE (28 MP)
-0.003 %	0.007 %	0.0412 %

Table 6: Best estimate liabilities error with 3 model points depending on the agregation method

agregation procedure does not only performed an accurate BEL evaluation but manages also to replicate the cash flows dynamics throughout the entire projection. Figure 5 shows the accuracy on the expected present value of the exiting cash flow associated with each projection year for the $3-MP$ agregated portfolio.

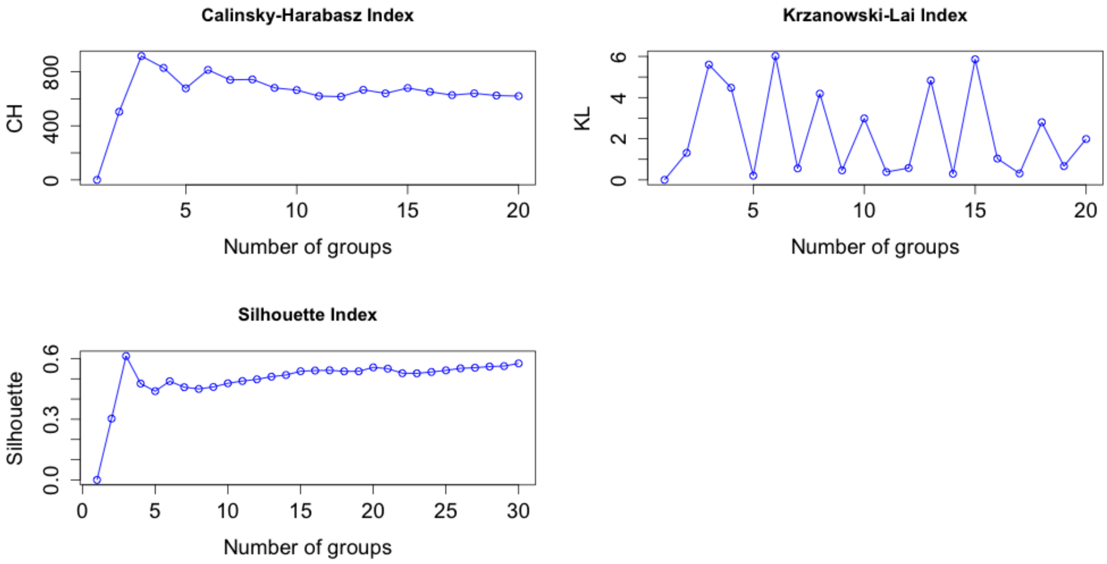

Figure 3: Partitionning quality indicators variations depending on the number of clusters

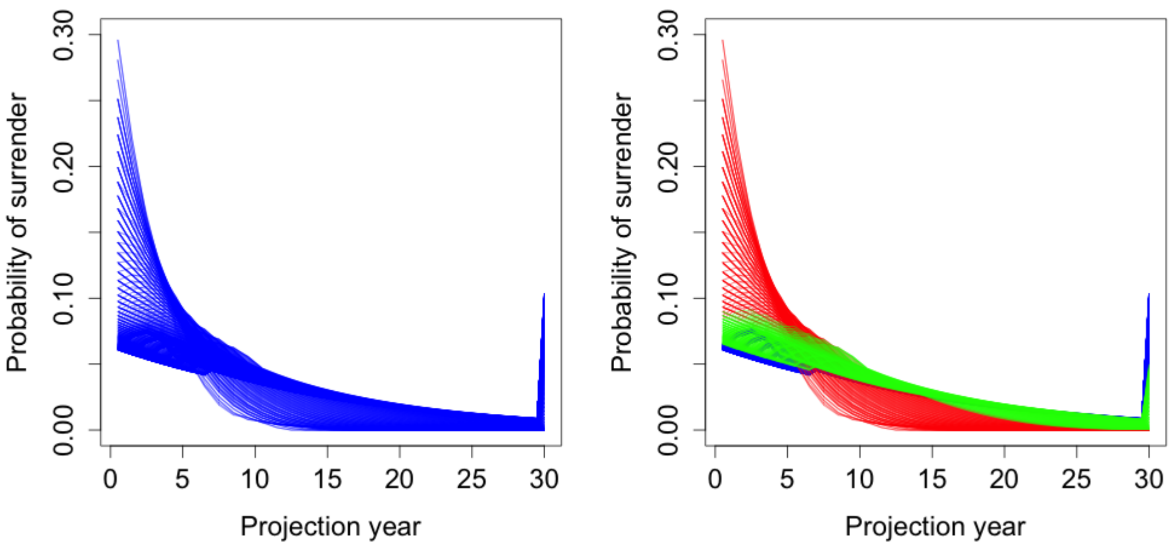

Figure 4: Portfolio visualization through its trajectories of surrender

From a practical point of view, the optimal number of clusters should be associated with a level of error chosen by the user. We did not manage to establish a clear link between

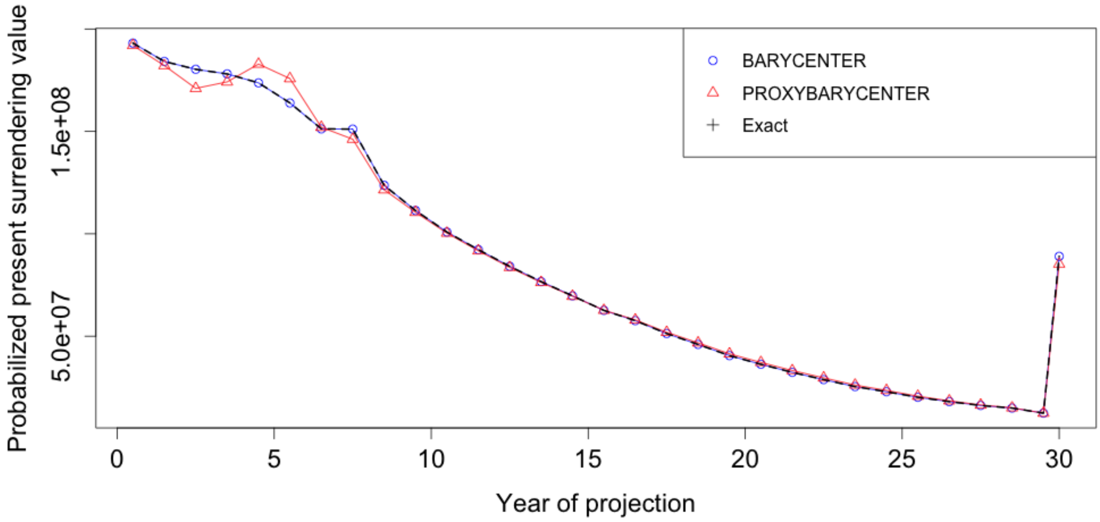

Figure 5: Expected present value of surrender during the projection

WCSS and the error on the BEL evaluations. Maybe it does not exist and we need to define another indicator, instead of WCSS, that we can compute from the probabilities of surrender and that is more linked to the evaluation error. This a topic of current research. We may also add that the optimal number of clusters is not necessarily a problem that needs much thoughts from a practitionner point of view. The number of model points in the agregated portfolio is often constrained by the capability of computers to deal with the valuation of large portfolios. The number of model points is then a parameter of the agregation process. Another technical requirement is to take into account the different lines of business. Two life insurance products cannot be grouped together because they may have characteristics, besides their probabilities of surrender, that impact the BEL computations. A solution is to allocate a number of model points for each line of business in proportion to their mathematical provision.

5 Conclusion

The agregation procedure permits a great computation time reduction that goes along with a limited loss of accuracy. The method is easy to understand and to implement as the statistical tools are available in most datamining softwares. The application extends to the entire scope of life insurance business, and can be tailored to many cash flows projection model based on the general definition of best estimate liabilities given in the introduction. This work represents the successful outcome of a Research and Development project in the industry. It is already implemented in AXA France, from a portfolio that contains millions of policies, the output is a portfolio of only a few thousands model points. It remains many rooms for improvement, especially at the partitioning level where distance other than euclidean might be better suited to quantify the distance between trajectories. The definition of an indicator that gives insight on the error resulting from the agregation would be as well a great improvement. For instance, the computation of this indicator may provide an optimal number of model points allocated to each line of business and therefore the whole portfolio.

6 Acknowledgments

The authors would like to thank the AXA research found for founding and allowed the diffusion of this work. Mohamed Baccouche, chief of the AXA France actuarial department for life business for his support during this project, Victor Intwali for his fine contribution to the implementation, Denys Pommeret and Stephane Loisel for their useful suggestions and comments.

References

[1] F. Bonnin, F. Planchet, and M. Julliard. Best estimate calculations of saving contracts by closed formulas: Application to the orsa. *Les cahiers de recherche de l'ISFA*, (2012.5), 2012.

[2] R. B. Calinsky and J. Harabasz. A dendrite method for cluster analysis. *Communications in Statistics*, 3:1–27, 1974.

[3] S. L. Christiansen. Representative interest rate scenarios. *North American Acturial Journal*, 2(3):29–44, 1998.

[4] Y. C. Chueh. Efficient stochastic modeling for large and consolidated insurance business: Interest rate algorithms. *North American Acturial Journal*, 6(3):88–103, 2002.

[5] L. Devineau and S. Loisel. Construction d'un algorithme d'accélération de la méthode des "simulations dans les simulations " pour le calcul du capital économique solvabilité ii. *Bulletin Français d'Actuariat*, 10(17):188–221, 2009.

[6] EIOPA. Technical specification for qis v. *European comission*, 2010.

[7] B. S. Everitt, S. Landau, and M. Leese. *Cluster Analysis.* A Hodder Arnld Publication, 4 edition, 1995.

[8] A. Fredman and C. Reynolds. Cluster analysis: A spatial approach to actuarial modeling. Research report, Milliman, August 2008.

[9] C. Genolini and B. Falissard. "kml: K-means for longitudinal data". *Computational Statistics*, 25(2):317–328, 2010.

[10] H.U. Gerber. *Life insurance mathematics.* Springer-Verlag, Berlin, 1990.

[11] J. Hartigan. *Clustering Algorithms.* Wiley, New York, 1975.

[12] B. Hejblum, J. Skinner, and R. Thiebaut. Application of gene set analysis of time-course gene expression in a hiv vaccine trial. *33rd annual conference of international society for clinical biostatistics*, 2012.

[13] J. Jacques and C. Preda. Functional data clustering: a survey. *Adavances in data analysis and classification*, pages 1–25, 2013.

[14] L. Kaufman and P. J. Rousseeuw. *Finding groups in Data: An introduction to cluster analysis.* Wiley, New York, 1990.

[15] W.J. Krzanowski and Y. T. Lai. A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics*, 44:23–34, Mars 1988.

[16] D. S. Nagin. Analysing developmental trajectories: a semiparametric group based approach. *Psychological Methods*, pages 139–157, 1999.

[17] P. Petauton, D. Kessler, and J.L. Bellando. *Théorie et pratique de l'assurance vie: manuel et exercices corrigés.* Dunod, 2002.

[18] M. Sarukkali. *Replicated Stratified Sampling for Sensitivity Analysis.* PhD thesis, University of Connecticut, 2013.

[19] J. H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 236-544:236–544, 1963.