Application of data clustering and machine learning in variable annuity valuation

Guojun Gan *

Global Variable Annuity Hedging Department, Manulife Financial, 200 Bloor Street East, Toronto, ON, M4W 1E5, Canada

H I G H L I G H T S

- We study the pricing of a large portfolio of VA policies.
- A clustering method is used to select representative policies.
- A machine learning method is used to estimate the guarantee value.
- The proposed method performs well in terms of accuracy and speed.

A R T I C L E   I N F O

A B S T R A C T

The valuation of variable annuity guarantees has been studied extensively in the past four decades. However, almost all the studies focus on the valuation of guarantees embedded in a single variable annuity contract. How to efficiently price the guarantees for a large portfolio of variable annuity contracts has not received enough attention. This paper fills the gap by introducing a novel method based on data clustering and machine learning to price the guarantees for a large portfolio of variable annuity contracts. Our test results show that this method performs very well in terms of accuracy and speed.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Variable annuity (VA), also known as segregated fund, guaranteed investment fund, unit-linked life insurance, or equity-linked life insurance (Armstrong, 2001), is a very popular insurance product. In a VA contract, the policyholder's premiums are invested in the assets of a fund. The benefit of the contract at maturity is equal to the market value of the accumulated premiums. The investments of the fund may consist of bonds and/or equities.

The VA contract comes with guarantees. For example, almost every VA contract contains the guaranteed minimum death benefit (GMDB). Other common guarantees include the guaranteed minimum withdrawal benefit (GMWB), the guaranteed minimum maturity benefit (GMAB), and the guaranteed minimum income benefit (GMIB). All these guarantees are financial guarantees, which cannot be adequately addressed by traditional actuarial approaches (Boyle and Hardy, 1997; Hardy, 2000). Stochastic simulation (i.e., the Monte Carlo method) and option pricing are the two

approaches that have been used to value VA contracts (Boyle and Hardy, 1997).

In the stochastic simulation approach, a stochastic model (i.e., Black–Scholes model) is used to simulate the future performance of the fund and generate a sample of the accumulated premiums of a VA contract at maturity. The mean of the discounted sample is the fair market value of the VA contract. In the option pricing approach, a dynamic investment strategy is used to adjust the assets of the portfolio on an ongoing basis so that the guarantee is met at maturity. In practice, the Monte Carlo method is used to value VA contracts since there are no closed-form formulas to value the VA contract.

The financial risk under the guarantees is non-diversifiable (Hardy, 2000). When the investment performance is poor, many maturing contracts may require additional funds to cover the guarantee at the same time. In fact, the claim can be quite significant in such case. Boyle and Hardy (1997) presented three methods to manage the financial risk associated with VA contracts.

The first method is actuarial reserving. In this method, an amount of money is held in risk free bonds until the policy matures. The amount of money is calculated to be sufficient to meet the guarantee with a given probability. The second method is dynamic

* Tel.: +1 6479383661.
E-mail address: Guojun.Gan@gmail.com.

hedging. In this method, the guarantee is treated like a put option. To hedge the put option, a hedge portfolio of financial instruments (e.g., futures and swaps) is purchased and actively managed in such a way that any shortfalls from the guarantee will be financed by the hedge portfolio under all possible financial market situations. The third method is static hedging. In this method, put options are bought to hedge the guarantee cost and the problem of dynamic hedging is passed to a third party.

The second method, dynamic hedging, is a popular risk management approach for VA contracts and is adopted by many insurance companies. However, this method requires determining the sensitivities of the option value to risk factors. These sensitivities are called "Greeks". Delta and Rho are two commonly used Greeks for dynamic hedging.

Since the Monte Carlo method is used to price the value of guarantees in practice, one major challenge of making dynamic variable annuity hedging work is to calculate the sensitivities of the option value (i.e., the value of the guarantee) for a large portfolio of VA contracts within a limited time interval. In order to complete the calculation in time, insurance companies employ many computing servers to conduct the calculation in parallel. In other words, insurance companies approach the computing problem from the perspective of hardware.

In this paper, we introduce a novel method based on data clustering and machine learning to calculate the sensitivities of the option value for a large portfolio of VA contracts. This method does not require using the Monte Carlo method to calculate the sensitivities for every VA contract in the large portfolio. As a result, this method is able to reduce the computing time significantly. Since the proposed method does not depend on any specific features of variable annuities, it can be used to value portfolios of other products such as exotic options.

The remaining of the paper is structured as follows. Section 2 gives a brief review of the academic papers relevant to variable annuity pricing and hedging. Section 3 gives a brief description of a clustering algorithm used to cluster mixed-type data. Section 4 presents a machine learning technique. Section 5 presents some test results of the clustering method and the machine learning method. Section 6 concludes the paper and gives a survey of future work.

2. Literature review

The academic literature on valuing and hedging guarantees in VA contracts is extensive. Since the seminal work on option pricing (Black and Scholes, 1973), many papers on variable annuity pricing and hedging have been published. In this section, we present a brief review of the relevant papers published in the past four decades.

Brennan and Schwartz (1976) introduced the Brennan–Schwartz model to study the equilibrium pricing of equity-linked life insurance policies with an asset value guarantee by decomposing the benefit into a sure amount and an immediately exercisable call option on the reference portfolio, where the call option is priced by the option pricing model of Black–Scholes (Black and Scholes, 1973). Boyle and Schwartz (1977) presented a theoretical framework for valuing death benefit guarantees and maturity benefit guarantees under equity-linked contracts. Bacinello and Ortu (1993) extended the Brennan–Schwartz model to the case when minimum guarantees that are functions of the premiums paid. Nielsen and Sandmann (1995) introduced a model with stochastic interest rates to price equity-linked life insurance contracts with periodic premiums. Ekern and Persson (1996) studied the valuation of complex unit-linked insurance contracts by decomposing the benefit pattern into simple components of increasing complexity and pricing these components separately.

Armstrong (2001) studied the guarantee reset feature in segregated funds using a discrete time Markov chain model and analyzed the optimal strategy of resetting the guarantee. Boyle et al.

(2001) presented a Monte Carlo simulation method to value the reset options embedded in some segregated funds. Windcliff et al. (2001) studied the valuation of VA contracts with reset features using a method based on the numerical solution of a set of linear complementarity problems. Windcliff et al. (2002) studied the effects of volatility, interest rates, and product design on the cost of providing a guarantee and the implication of the finding on the reset feature. Kijima and Wong (2007) studied the pricing of Ratchet equity-linked annuities in the presence of stochastic interest rates.

Milevsky and Promislow (2001) studied the options to annuities that are embedded in VA contracts under both the discrete and continuous-time pricing frameworks. Charupat and Milevsky (2002) studied the optimal asset allocation problem in variable annuities and derived the optimal utility-maximizing asset allocation between a risky and risk-free asset within a VA contract. Lee (2003) derived explicit pricing formulas for equity-linked annuities with path-dependent options using the method of Esscher transforms. Young (2003) used the principle of equivalent utility to study the valuation of equity-linked life insurance contracts.

Bacinello (2003a,b) studied variable annuities with surrender options using a binomial option pricing model. Bacinello (2005) presented an endogenous model to price equity-linked life insurance policies embedding surrender options. Cheung and Yang (2005) studied the optimal surrender strategy for holders of equity-linked investment products embedding surrender options. Gaillardetz and Lin (2006) used binomial models to value equity-linked insurance contracts. Costabile et al. (2008) also used binomial models to value equity-linked policies with surrender options. Bacinello et al. (2009) proposed a least squares Monte Carlo method to price life insurance contracts embedding surrender options. Costabile et al. (2009) proposed a bivariate model to price equity-linked policies with surrender options and maturity guarantees. Qian et al. (2010) studied the valuation of equity-indexed annuities under stochastic mortality and interest rate environment. Li and Szimayer (2011) proposed a method to determine the price bounds of unit-linked life insurance contracts by analyzing the upper and lower bounds of the mortality intensity.

Gerber and Shiu (2003) derived closed-form formulas for pricing complex guarantees embedded in some equity-linked annuities by studying European lookback options. Jacques (2003) studied the valuation of a single equity-linked contract without mortality diversification and analyzed a variety of situations depending on the guarantee, the volatility, and the age of the policyholder. Lin et al. (2009) studied the pricing problem of equity-linked annuities and variable annuities under a regime-switching model and used the Esscher transform to obtain an equivalent martingale measure for valuation in incomplete markets.

Milevsky and Posner (2001) used risk-neutral option pricing theory to value the GMDB in VA contracts. Bélanger et al. (2009) proposed a method for pricing the GMDB rider embedded in some VA contracts that allow for partial withdrawals, where the GMDB pricing problem was modeled as an impulse control problem. Wang (2009) studied quantile hedging for the GMDB rider presented in VA contracts. Gerber et al. (2012) proposed a method to price the GMDB rider by calculating the expected discounted value of a payment at time of death.

Milevsky and Salisbury (2006) developed a variety of methods to study the cost and value of the GMWB embedding in many VA contracts. Chen and Forsyth (2008) presented an impulse stochastic control formulation for pricing the GMWB. Dai et al. (2008) developed a singular stochastic control model to value VA contract with the GMWB rider. Xu and Wang (2009) proposed a model based on a two-dimensional partial differential equation to price the GMWB rider. Peng et al. (2010) studied the valuation of the GMWB rider under the Vasicek stochastic interest rate framework and derived analytic approximation solutions to the fair value of the GMWB rider. Gao and Ulm (2012) studied the valuation of the GMDB rider using a utility-based approach. Kolkiewicz and Liu

(2012) proposed a method to hedge the GMWB rider by finding the closest path-independent option to the guarantee and constructing a portfolio of traded European options that approximates the optimal option. Yang and Dai (2013) proposed a tree model to price the GMWB rider embedded in deferred life annuity contracts.

Jiang and Chang (2010) proposed an analytical solution of the cost of the GMAB rider by applying the Black–Scholes model. Ko et al. (2010) studied the valuation of maturity guarantee with dynamic withdrawal benefit under the Black–Scholes framework. Marshall et al. (2010) studied the valuation of the GMIB rider in a complete market by decomposing the GMIB value and analyzed the sensitivity of the GMIB value to the financial variables.

Bauer et al. (2008) proposed a general framework to price a variety of guarantees in a consistent manner. Consiglio and Giovanni (2008) developed a stochastic programming model to value the minimum guarantee in some life insurance products under the assumption that the markets are incomplete. Bacinello et al. (2011) proposed a unifying framework to price various types of guarantees. Ng and Li (2011) used a multivariate extension of the regime-switching conditional Esscher transform to study the valuation of VA guarantees written on multiple funds. Ng and Li (2013) proposed a multivariate framework for pricing VA guarantees written on multiple assets and demonstrated how dynamic hedging can be implemented in such a multivariate framework.

In all the papers mentioned above, emphasis is put on the valuation of a single VA contract. How to price efficiently a large portfolio that contains hundreds of thousands of VA contracts is not studied. This paper aims to fill this gap by proposing an efficient method for pricing a large portfolio of VA contracts.

3. A data clustering method

Data clustering (Gan et al., 2007) refers to the process of dividing a set of items into groups or clusters such that items in the same cluster are similar to each other and items in different clusters are distinct. The k-prototypes algorithm (Huang, 1998) is a clustering algorithm that is suitable to cluster mixed-type data. In this section, we give a brief description of the k-prototypes algorithm.

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ denote the portfolio of VA contracts, where n is the number of VA contracts and $\mathbf{x}_i$ represents the ith VA contract. Without loss of generality, we assume that a VA contract is characterized by d attributes (e.g., gender, age, account value, etc.) and that the first d_1 attributes are numeric and the last $d_2 = d - d_1$ attributes are categorical. Then the distance between two records $\mathbf{x}$ and $\mathbf{y}$ in X can be defined as (Huang, 1998)

$$D(\mathbf{x}, \mathbf{y}, \lambda) = \sqrt{\sum_{h=1}^{d_1} (x_h - y_h)^2 + \lambda \sum_{h=d_1+1}^{d} \delta(x_h, y_h)}, \tag{1}$$

where x_h and y_h are the hth component of $\mathbf{x}$ and $\mathbf{y}$, respectively, λ is a balance weight used to avoid favoring either type of attribute, and $\delta(\cdot, \cdot)$ is the simple matching distance defined as

$$\delta(x_h, y_h) = \begin{cases} 0, & \text{if } x_h = y_h, \\ 1, & \text{if } x_h \neq y_h. \end{cases}$$

Note that the numerical values in the above distance definition are normalized so that for each $h = 1, 2, \ldots, d_1$, the standard deviation of $x_{1h}, x_{2h}, \ldots, x_{nh}$ is one, where x_{ih} is the hth component of the ith VA contract $\mathbf{x}_i$.

The objective function that the k-prototypes algorithm tries to minimize is defined as

$$P_\lambda = \sum_{j=1}^{k} \sum_{\mathbf{x} \in C_j} D^2(\mathbf{x}, \boldsymbol{\mu}_j, \lambda), \tag{2}$$

where $D(\cdot, \cdot, \lambda)$ is defined in Eq. (1), k is the number of clusters, C_j is the jth cluster, and $\boldsymbol{\mu}_j$ is the center or prototype of cluster C_j.

The k-prototypes algorithm works iteratively in order to find a solution that minimizes the objective function defined in Eq. (2). In other words, the k-prototypes algorithm repeats updating the cluster memberships given the cluster centers and updating the cluster centers given the cluster memberships until some stop condition is satisfied.

Mathematically, the k-prototypes algorithm can be described as follows.

1. *Initialize cluster center.* At this step, the algorithm initializes the k cluster centers by selecting k distinct records from the dataset X randomly. Suppose $\boldsymbol{\mu}_1^{(0)}, \boldsymbol{\mu}_2^{(0)}, \ldots, \boldsymbol{\mu}_k^{(0)}$ are the k initial cluster centers.

2. *Update cluster memberships.* At this step, the algorithm updates the cluster memberships $\gamma_1, \gamma_2, \ldots, \gamma_n$ as follows:

$$\gamma_i^{(0)} = \underset{1 \leq j \leq k}{\operatorname{argmin}} D(\mathbf{x}_i, \boldsymbol{\mu}_j^{(0)}, \lambda), \tag{3}$$

where $D(\cdot, \cdot, \lambda)$ is defined in Eq. (1).

3. *Update cluster centers.* At this step, the algorithm updates the cluster centers as follows:

$$\mu_{jh}^{(1)} = \frac{1}{|C_j|} \sum_{\mathbf{x} \in C_j} x_h, \quad h = 1, 2, \ldots, d_1, \tag{4a}$$

$$\mu_{jh}^{(1)} = \operatorname{mode}_h(C_j), \quad h = d_1 + 1, \ldots, d, \tag{4b}$$

where $C_j = \left\{ \mathbf{x}_i \in X : \gamma_i^{(0)} = j \right\}$ for $j = 1, 2, \ldots, k$, and $\operatorname{mode}_h(C_j)$ is the most frequent categorical value of the hth attribute in cluster C_j. Let $A_{h1}, A_{h2}, \ldots, A_{h,m_h}$ be the distinct values the hth attribute can take, where m_h is the number of distinct values the hth attribute can take. Let $f_{ht}(C_j)$ be the number of records in cluster C_j, whose hth attribute takes value A_{ht} for $t = 1, 2, \ldots, m_h$. That is,

$$f_{ht}(C_j) = |\{\mathbf{x} \in C_j : x_h = A_{ht}\}|, \quad t = 1, 2, \ldots, m_h.$$

Then

$$\operatorname{mode}_h(C_j) = \underset{1 \leq t \leq m_h}{\operatorname{argmax}} f_{ht}(C_j), \quad h = d_1 + 1, \ldots, d.$$

4. Repeat Step 2 and Step 3 until the cluster memberships do not change between two iterations or the maximum number of iterations is reached.

Suppose that the cluster centers obtained from the k-prototypes algorithm are denoted by $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_k$. Then we select the representative VA contracts $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ as follows:

$$\mathbf{z}_j = \underset{\mathbf{x} \in X}{\operatorname{argmin}} D(\mathbf{x}, \boldsymbol{\mu}_j, \lambda).$$

That is, the representative VA contract $\mathbf{z}_j$ is the VA contract in X that is closest to the cluster center $\boldsymbol{\mu}_j$. We assume that these k representative VA contracts are mutually distinct, that is,

$$D(\mathbf{z}_r, \mathbf{z}_s, \lambda) > 0$$

for all $1 \leq r < s \leq k$.

If n and k are large (e.g., $n > 10000$ and $k > 20$), the k-prototypes algorithm will be very slow as it needs to perform many distance calculations. In such cases, we divide the portfolio of VA contracts into many subsets and find a few clusters from each subset. For example, if we want to cluster a portfolio of 200,000 VA contracts into 100 clusters, we first divide the portfolio into about 33 subsets and cluster each subset into 3 or 4 clusters. In this way, the clustering algorithm is very fast to find 100 representative contracts. If the portfolio is large and the policies are evenly distributed, this approach produces similar clustering results as the direct k-prototypes algorithm does. However, if the portfolio is small, this approach will produce different clustering results. For

example, if we want to cluster the 9 points given in Fig. 1 into 3 clusters, the direct k-prototypes algorithm gives us {1, 2, 3}, {4, 5, 6}, and {7, 8, 9}. If we divide the 9 points into {1, 2, 3, 4, 5} and {6, 7, 8, 9}, and find 2 clusters from {1, 2, 3, 4, 5} and 1 cluster from {6, 7, 8, 9}, then the approach gives us {1, 2, 3}, {4, 5}, and {6, 7, 8, 9}.

4. A machine learning method

Machine learning (Mitchell, 1997) refers to the process of constructing and studying systems that can learn from data. The Kriging method (Isaaks and Srivastava, 1990) is a machine learning method and is known as the Gaussian process predictor in the machine learning domain (Rasmussen and Williams, 2005). In this section, we present a brief description of the ordinary Kriging method (Isaaks and Srivastava, 1990).

Let $\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_k$ be the representative VA contracts obtained from the clustering algorithm. For every $j = 1, 2, \ldots, k$, let y_j be the fair value of $\mathbf{z}_j$ that is calculated by the Monte Carlo method. Then we use the Kriging method to estimate the fair value of the VA contract $\mathbf{x}_i$ as

$$\hat{y}_i = \sum_{j=1}^{k} w_{ij} \cdot y_j, \tag{5}$$

where $w_{i1}, w_{i2}, \ldots, w_{ik}$ are the Kriging weights.

The Kriging weights $w_{i1}, w_{i2}, \ldots, w_{ik}$ are obtained by solving the following linear equation system:

$$\begin{pmatrix} V_{11} & \cdots & V_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ V_{k1} & \cdots & V_{kk} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_{i1} \\ \vdots \\ w_{ik} \\ \theta_i \end{pmatrix} = \begin{pmatrix} D_{i1} \\ \vdots \\ D_{ik} \\ 1 \end{pmatrix}, \tag{6}$$

where θ_i is a control variable used to make sure the sum of the Kriging weights is equal to one,

$$V_{rs} = \alpha + \exp\left(-\frac{3}{\beta} D(\mathbf{z}_r, \mathbf{z}_s, \lambda)\right), \quad r, s = 1, 2, \ldots, k,$$

and

$$D_{ij} = \alpha + \exp\left(-\frac{3}{\beta} D(\mathbf{x}_i, \mathbf{z}_j, \lambda)\right), \quad j = 1, 2, \ldots, k.$$

Here the distance function $D(\cdot, \cdot, \lambda)$ is defined in Eq. (1), and $\alpha \geq 0$ and $\beta > 0$ are two parameters. Since $D(\mathbf{z}_r, \mathbf{z}_s, \lambda) > 0$ for all $1 \leq r < s \leq k$, the above linear equation system has a unique solution (Isaaks and Srivastava, 1990).

The fair value of the portfolio X is equal to the sum of the fair values of all VA contracts in X, i.e.,

$$\hat{Y} = \sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} \cdot y_j = \sum_{j=1}^{k} w_j \cdot y_j, \tag{7}$$

where

$$w_j = \sum_{i=1}^{n} w_{ij}.$$

If we are interested in only the fair value $\hat{Y}$ of the portfolio, we can calculate $\hat{Y}$ in an efficient way by obtaining $w_1, w_2, \ldots, w_k$ from the following linear equation system:

$$\begin{pmatrix} V_{11} & \cdots & V_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ V_{k1} & \cdots & V_{kk} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_k \\ \theta \end{pmatrix} = \begin{pmatrix} D_1 \\ \vdots \\ D_k \\ n \end{pmatrix}, \tag{8}$$
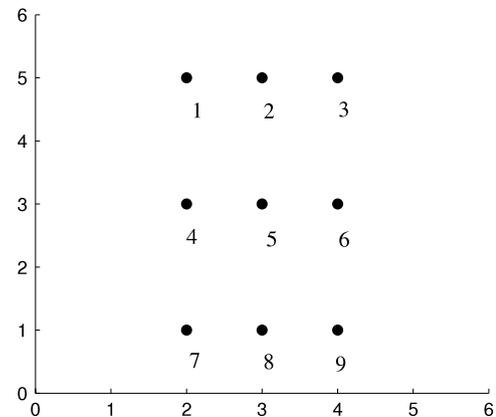
Fig. 1. An example of 9 points.

Table 1
Variable annuity contract specification. Here $\mathbf{N}$ and $\mathbf{R}$ denote the set of natural number and the set of real numbers, respectively.

Attribute	Values
Guarantee type	GMDB only, GMDB + GMWB
Gender	Male, Female
Age	$\mathbf{N} \cap [20, 60]$
Premium	$\mathbf{R} \cap [10000, 500000]$
GMWB withdrawal rate	0.04, 0.05, 0.06, 0.07, 0.08
Maturity	$\mathbf{N} \cap [10, 25]$

where

$$D_j = \sum_{i=1}^{n} D_{ij}, \quad j = 1, 2, \ldots, k.$$

In fact, Eq. (8) is obtained by summing both sides of Eq. (6) from $i = 1$ to n. In this way, we only need to solve one linear equation system.

5. Application in VA valuation

In this section, we present some test results of this method based on synthetic data. We first introduce how we generate a large portfolio of synthetic VA contracts. Then we introduce the Monte Carlo valuation of these VA contracts. Finally we present the performance of the clustering method and the machine learning method in terms of accuracy and speed.

5.1. Synthetic VA contracts

We generate a portfolio of 200,000 VA contracts, which are specified in Table 1. We consider only two types of guarantees: GMDB and GMWB. Since VA is a type of life insurance, every contract has the GMDB rider. The VA contracts are randomly generated by selecting values uniformly from the ranges specified in Table 1.

5.2. Monte Carlo valuation

To price the value of guarantees embedded in the VA contracts, we follow the Monte Carlo method proposed by Bauer et al. (2008). To describe the Monte Carlo method for pricing the GMDB rider and the GMWB rider, we used the mathematical symbols given in Table 2.

Since we consider only the GMDB rider and GMWB rider, there are two possible types of events (Bauer et al., 2008):

- the policyholder withdraws money as a guaranteed withdrawal of the GMWB rider;
- the policyholder dies.

Table 2
Mathematical symbols for the Monte Carlo method.

Symbol	Meaning
S_t	The underlying mutual fund at time t of the variable annuity
A_t	The account value at time t
W_t	The withdrawal benefit at time t
D_t	The death benefit at time t
G_t^W	The remaining total amount that can be withdrawn after time t
G_t^E	The maximum amount that can be withdrawn annually
G_t^D	The guaranteed minimum death benefit at time t
x_W	The proportion of the premium that can be withdrawn annually
T	The maturity of the contract

We use $(\cdot)_t^-$ and $(\cdot)_t^+$ to denote the value of a state variable (e.g., A_t) immediately before and after the occurrence of such events, respectively.

In our Monte Carlo valuation of the VA contracts, we assume the following:

- the underlying mutual fund is simulated as

$$S_0 = 1, \qquad S_t = S_{t-1} \exp\left(\left[r - \frac{1}{2}\sigma^2\right] + \sigma Z\right)$$

  for $t = 1, 2, \ldots, 40$, where r is the interest rate, σ is the volatility of the underlying mutual fund, and Z is a standard normal random variable. In our tests, we use $r = 3\%$ and $\sigma = 20\%$ and simulate 1000 paths;
- for a contract with the GMWB rider, the policyholder takes maximum annual withdrawals;
- all the events happen only at anniversary date;
- there are no fees;
- there are no lapses;
- the mortality follows the 1996 IAM mortality tables provided by the Society of Actuaries.

At time $t = 0$, we have

$$G_0^W = A_0, \qquad G_0^E = x_W A_0, \qquad G_0^D = A_0.$$

For $t = 0, 1, \ldots, T-1$, the evolution of the state variables between t^+ and $(t+1)^-$ is described as follows. The account value evolves as

$$A_{t+1}^- = A_t^+ \frac{S_{t+1}}{S_t}.$$

The guaranteed minimum death benefit, the maximum amount that can be withdrawn annually, and the remaining total amount that can be withdrawn do not change, i.e.,

$$G_{t+1}^{D-} = G_t^{D+}, \qquad G_{t+1}^{E-} = G_t^{E+}, \qquad G_{t+1}^{W-} = G_t^{W+}.$$

The evolution of the state variables between $(t+1)^-$ and $(t+1)^+$ is described as follows. The death benefit at time $t+1$ is calculated as

$$D_{t+1} = \max\left(0, G_{t+1}^{D-} - A_{t+1}^-\right).$$

Since we assume that the policyholder takes maximally available withdrawal annually, the withdrawal amount at year $t+1$ is given by

$$E = \min\left(G_{t+1}^{E-}, G_{t+1}^{W-}\right),$$

and the maximum amount that can be withdrawn annually does not change, i.e., $G_{t+1}^{E+} = G_{t+1}^{E-}$. The withdrawal benefit at time $t+1$ is given by

$$W_{t+1} = \max\left(0, E - A_{t+1}^-\right).$$

The account value becomes

$$A_{t+1}^+ = \max\left(0, A_{t+1}^- - E\right).$$

Table 3
Results of the Monte Carlo valuation. Numbers in brackets are negative numbers.

Measurement	Value
Market value	3,003,947,180
Dollar Delta	(8,150,275,955)
Dollar Rho	(9,736,358)
Time used (seconds)	1942.22

The remaining total amount that can be withdrawn after time $t+1$ becomes

$$G_{t+1}^{W+} = \max\left(0, G_{t+1}^{W-} - E\right).$$

The guaranteed minimum death benefit will be adjusted pro rata as follows:

$$G_{t+1}^{D+} = \frac{A_{t+1}^+}{A_{t+1}^-} G_{t+1}^{D-}.$$

Then the present value of the GMDB and the GMWB benefits is given by

$$V(S_1, S_2, \ldots, S_{40}) = \sum_{t=1}^{T+1} {}_{t-1}p_{x_0}(1 - q_{x_0+t-1})W_t e^{-rt}$$
$$+ \sum_{t=1}^{T+1} {}_{t-1}p_{x_0} q_{x_0+t-1} D_t e^{-rt}, \tag{9}$$

where x_0 is the age of the policyholder. The value of the GMDB rider and the GMWB rider is the average of $V(S_1, S_2, \ldots, S_{40})$ along all paths. If a contract does not have the GMWB rider, the above formula still applies by letting the withdrawal rate to be zero.

5.3. Test results

To measure the performance of the clustering method and the machine learning method, we first use the Monte Carlo method described above to calculate the market value, dollar Delta, and dollar Rho of the guarantees of 200,000 VA contracts and record the time elapsed. These numbers and the time elapsed will be our benchmarks used to measure the accuracy and speed of the new method, respectively.

The benchmark results from the Monte Carlo method are shown in Table 3. The market value, dollar Delta, and dollar Rho are the aggregate numbers for the whole portfolio of 200,000 VA contracts. That is, we used the Monte Carlo method to calculate the market value, dollar Delta, and dollar Rho of each individual contract and added them up to get the numbers in the portfolio level. From the table we see that it takes the Monte Carlo method 1942.22 s (or 32 min 22.22 s) to price all the contracts.

To test the clustering and machine learning methods, we first use the k-prototypes algorithm to select a small set of representative policies. Then we use the Monte Carlo method to calculate the market value, dollar Delta, and dollar Rho of each representative policy. Finally, we use the Kriging method to estimate the market value, dollar Delta, and dollar Rho of each policy in the portfolio. In the clustering step and the Kriging step, we use the z-score method to normalize the numerical attributes (e.g., age, premium, etc.) so that the distance is not dominated by a single attribute. In the Monte Carlo step, however, we still use the original data as normalized data do not make sense in this step.

We conducted four tests for the clustering and machine learning methods and used different numbers of clusters in different tests. In particular, we tested 100 clusters, 500 clusters, 1000 clusters, and 2000 clusters. The values of other parameters used in the tests are shown in Table 4. The results of these tests are shown in Table 5. From the table we see that the numbers calculated by the clustering and machine learning methods approximate very well

Table 4
Values of the parameters used in the tests. Since the numerical attributes are normalized by the z-score method, we set $\lambda = 1$ so that categorical attributes and numerical attributes are balanced in the distance calculation. The values of α and β are selected according to Isaaks and Srivastava (1990).

Parameter	Value
k	100, 500, 1000, 2000
λ	1
α	0
β	95th percentile of the distances $D(\mathbf{z}_r, \mathbf{z}_s, 1)$, $1 \le r < s \le k$

Table 6
Time used by the clustering method and the machine learning method. The numbers are in seconds.

	Number of clusters			
	100	500	1000	2000
k-prototypes	20.41	12.86	10.22	8.37
Monte Carlo	1.63	5.38	9.52	19.79
Kriging	5.83	26.8	52.69	106.86
Total	27.87	45.04	72.43	135.02

the benchmark numbers given in Table 3. In addition, the accuracy improves when the number of clusters increases. As we can see from Table 5, the percentage differences of the market value, dollar Delta, and dollar Rho are around 1% when 2000 clusters are used.

From Table 5, we see that the clustering and machine learning methods always overestimate the market value, dollar Delta, and dollar Rho. Such bias is caused by the fact that the set of representative polices does not contain boundary policies and the fact that the Kriging method does not perform well for extrapolation.

In Table 6, we present the time used by the clustering and machine learning methods. In this new method, we first cluster the portfolio of VA contracts to find a set of representative contracts. Then we use the Monte Carlo method to calculate the market value and Greeks of the representative contracts. Finally we use the machine learning method to calculate the market value and Greeks for all the contracts in the portfolio. Hence we present the time used by each step of the new method.

From Table 6 we see that the time used in the clustering step decreases as the number of clusters increases. This is consistent with the way of clustering the VA contracts (see Section 3). The time used by both the Monte Carlo method and the Kriging method increases as the number of clusters increases. The total time used by the method increases as the number of cluster increases. While it takes the Monte Carlo method 1942.22 s to calculate the market value, dollar Delta, and dollar Rho for the whole portfolio, it takes the new method (with 2000 clusters) about 135.02 s to calculate those quantities. The new method is 14 times faster than the original Monte Carlo method.

In our test, we used four different values of k (the number of clusters) to show the performance of the new method. Usually, if more numbers of clusters are used, the results will be closer to the benchmarks. In the extreme case, if all policies in the portfolio are used as representative policies, the results will be identical to the benchmarks (Isaaks and Srivastava, 1990). However, a large number of representative policies would make solving the linear equation system in Eq. (8) impractical because solving a large linear equation system requires lots of computer memory and time. For example, a 5000×5000 matrix of double precision floating-point numbers uses 191 megabytes of memory. In practice, the number

of clusters is selected in a way that accuracy and speed are balanced.

6. Concluding remarks

In this paper we proposed a novel method to calculate the market value and Greeks of the variable annuity guarantees for a large portfolio of variable annuity contracts. The method involves three steps:

1. cluster the large portfolio of VA contracts to get a small set of representative contracts;
2. use the Monte Carlo method to calculate the market value and Greeks of every representative contract;
3. use the machine learning method to calculate the market value and Greeks of every contract in the portfolio.

Since the number of representative contracts is small and the clustering method and machine learning method are fast, the new method can reduce the computing time significantly. Our tests on synthetic VA data show that the new method performs very well in terms of accuracy and speed.

The new method treats the valuation system (e.g., the Monte Carlo method) as a black box and learns the valuation system through inputs (e.g., the representative VA contracts) and outputs (e.g., the market value and Greeks). Since the synthetic VA contracts in our method are very simple and we used annual time steps, the Monte Carlo method is fast and can complete pricing 200,000 VA contracts in 1942.22 s. Even in this case, the new method (with 2000 clusters) is 14 times faster than the Monte Carlo method.

In cases when the valuation system is more time consuming, the new method will achieve even better performance in terms of speed because the clustering method and the machine learning method are independent of the performance of the valuation system. For example, if the Monte Carlo simulation model is based on a monthly time step basis, then the computing time would be approximately 12 times slow. In such case, it would take the Monte Carlo method $1942.22 \times 12 = 23306.64$ s or 6.47 h to price the 200,000 VA contracts. The time used by the clustering and machine

Table 5
Results of the clustering and machine learning methods. The dollar difference and percentage difference are calculated against the benchmark numbers obtained from the Monte Carlo method.

Cluster	Measurement	Value	Difference ($)	Difference (%)
100	Market value	3,304,156,326	300,209,146	9.99%
	Dollar Delta	(8,924,741,716)	(774,465,760)	9.50%
	Dollar Rho	(10,843,115)	(1,106,757)	11.37%
500	Market value	3,153,864,868	149,917,688	4.99%
	Dollar Delta	(8,584,797,784)	(434,521,829)	5.33%
	Dollar Rho	(10,304,245)	(567,887)	5.83%
1000	Market value	3,090,778,168	86,830,989	2.89%
	Dollar Delta	(8,438,870,353)	(288,594,398)	3.54%
	Dollar Rho	(10,043,320)	(306,961)	3.15%
2000	Market value	3,024,267,213	20,320,033	0.68%
	Dollar Delta	(8,233,152,864)	(82,876,909)	1.02%
	Dollar Rho	(9,805,786)	(69,427)	0.71%

Table 7
Time used by the clustering method and the machine learning method if monthly time steps are used in the Monte Carlo method. The numbers are in seconds.

	Number of clusters			
	100	500	1000	2000
k-prototypes	20.41	12.86	10.22	8.37
Monte Carlo	19.56	64.56	114.24	237.48
Kriging	5.83	26.8	52.69	106.86
Total	45.8	104.22	177.15	352.71

learning methods would be shown in Table 7. The new method (with 2000 clusters) is about 66 times faster than the Monte Carlo method.

In future we would like to improve and test the new method under more complex settings. We would like to consider more complex guarantees and multiple underlying mutual funds. In addition, we would like to test other machine learning methods such as neural networks.

Acknowledgments

References

Armstrong, M., 2001. The reset decision for segregated fund maturity guarantees. Insurance: Mathematics and Economics 29, 257–269.

Bacinello, A., 2003a. Fair valuation of a guaranteed life insurance participating contract embedding a surrender option. Journal of Risk and Insurance 70 (3), 461–487.

Bacinello, A., 2003b. Pricing guaranteed life insurance participating policies with annual premiums and surrender option. North American Actuarial Journal 7 (3), 1–17.

Bacinello, A., 2005. Endogenous model of surrender conditions in equity-linked life insurance. Insurance: Mathematics and Economics 37, 270–296.

Bacinello, A., Biffis, E., Millossovich, P., 2009. Pricing life insurance contracts with early exercise features. Journal of Computational and Applied Mathematics 233 (1), 27–35.

Bacinello, A., Millossovich, P., Olivieri, A., Pitacco, E., 2011. Variable annuities: a unifying valuation approach. Insurance: Mathematics and Economics 49 (3), 285–297.

Bacinello, A., Ortu, F., 1993. Pricing equity-linked life insurance with endogenous minimum guarantees. Insurance: Mathematics and Economics 12, 245–257.

Bauer, D., Kling, A., Russ, J., 2008. A universal pricing framework for guaranteed minimum benefits in variable annuities. ASTIN Bulletin 38 (2), 621–651.

Bélanger, A., Forsyth, P., Labahn, G., 2009. Valuing the guaranteed minimum death benefit clause with partial withdrawals. Applied Mathematical Finance 16 (6), 451–496.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81 (3), 637–654.

Boyle, P., Hardy, M., 1997. Two approaches to reserving for maturity guarantees. Journal of the Staple In Actuarial Society 21, 113–127.

Boyle, P., Kolkiewicz, A., Tan, K., 2001. Valuation of the reset options embedded in some equity-linked insurance products. North American Actuarial Journal 5 (3), 1–18.

Boyle, P., Schwartz, E., 1977. Equilibrium prices of guarantees under equity-linked contracts. Journal of Risk and Insurance 44 (4), 639–660.

Brennan, M., Schwartz, E., 1976. The pricing of equity-linked life insurance policies with an asset value guarantee. Journal of Financial Economics 3, 195–213.

Charupat, N., Milevsky, M., 2002. Optimal asset allocation in life annuities: a note. Insurance: Mathematics and Economics 30, 199–209.

Chen, Z., Forsyth, P., 2008. A numerical scheme for the impulse control formulation for pricing variable annuities with a guaranteed minimum withdrawal benefit (GMWB). Numerische Mathematik 109 (4), 535–569.

Cheung, K., Yang, H., 2005. Optimal stopping behavior of equity-linked investment products with regime switching. Insurance: Mathematics and Economics 37, 599–614.

Consiglio, A., Giovanni, D., 2008. Evaluation of insurance products with guarantee in incomplete markets. Insurance: Mathematics and Economics 42, 332–342.

Costabile, M., Gaudenzi, M., Massabó, I., Zanette, A., 2009. Evaluating fair premiums of equity-linked policies with surrender option in a bivariate model. Insurance: Mathematics and Economics 45, 286–295.

Costabile, M., Massabó, I., Russo, E., 2008. A binomial model for valuing equity-linked policies embedding surrender options. Insurance: Mathematics and Economics 42, 873–886.

Dai, M., Kwok, Y., Zong, J., 2008. Guaranteed minimum withdrawal benefit in variable annuities. Mathematical Finance 18 (4), 595–611.

Ekern, S., Persson, S.-A., 1996. Exotic unit-linked life insurance contracts. The Geneva Papers on Risk and Insurance Theory 21 (1), 35–63.

Gaillardetz, P., Lin, X., 2006. Valuation of equity-linked insurance and annuity products with binomial models. North American Actuarial Journal 10 (4), 117–144.

Gan, G., Ma, C., Wu, J., 2007. Data Clustering: Theory, Algorithms and Applications. In: ASA-SIAM Series on Statistics and Applied Probability, SIAM Press, Philadelphia, PA, USA.

Gao, J., Ulm, E., 2012. Optimal consumption and allocation in variable annuities with guaranteed minimum death benefits. Insurance: Mathematics and Economics 51 (3), 586–598.

Gerber, H., Shiu, E., 2003. Pricing lookback options and dynamic guarantees. North American Actuarial Journal 7 (1), 48–67.

Gerber, H., Shiu, E., Yang, H., 2012. Valuing equity-linked death benefits and other contingent options: a discounted density approach. Insurance: Mathematics and Economics 51 (1), 73–92.

Hardy, M., 2000. Hedging and reserving for single premium segregated fund contracts. North American Actuarial Journal 4 (2), 63–74.

Huang, Z., 1998. Extensions to the k-means algorithm for clustering large data sets with categorical values. Data Mining and Knowledge Discovery 2 (3), 283–304.

Isaaks, E., Srivastava, R., 1990. An Introduction to Applied Geostatistics. Oxford University Press, Oxford, UK.

Jacques, M., 2003. A risk management approach to the pricing of a single equity-linked contract. Journal of Pension Economics and Finance 2 (2), 159–179.

Jiang, S., Chang, M., 2010. Variable annuity with guarantees: valuation and simulation. Journal of Money, Investment and Banking 14, 74–83.

Kijima, M., Wong, T., 2007. Pricing of ratchet equity-indexed annuities under stochastic interest rates. Insurance: Mathematics and Economics 41, 317–338.

Ko, B., Shiu, E., Wei, L., 2010. Pricing maturity guarantee with dynamic withdrawal benefit. Insurance: Mathematics and Economics 47, 216–223.

Kolkiewicz, A., Liu, Y., 2012. Semi-static hedging for gmwb in variable annuities. North American Actuarial Journal 16 (1), 112–140.

Lee, H., 2003. Pricing equity-indexed annuities with path-dependent options. Insurance: Mathematics and Economics 33, 677–690.

Li, J., Szimayer, A., 2011. The uncertain mortality intensity framework: Pricing and hedging unit-linked life insurance contracts. Insurance: Mathematics and Economics 49 (3), 471–486.

Lin, X., Tan, K., Yang, H., 2009. Pricing annuity guarantees under a regime-switching model. North American Actuarial Journal 13, 316–338.

Marshall, C., Hardy, M., Saunders, D., 2010. Valuation of a guaranteed minimum income benefit. North American Actuarial Journal 14 (1), 38–59.

Milevsky, M., Posner, S., 2001. The titanic option: valuation of the guaranteed minimum death benefit in variable annuities and mutual funds. Journal of Risk and Insurance 68 (1), 93–128.

Milevsky, M., Promislow, S., 2001. Mortality derivatives and the option to annuitise. Insurance Mathematics and Economics 29, 299–318.

Milevsky, M., Salisbury, T., 2006. Financial valuation of guaranteed minimum withdrawal benefits. Insurance Mathematics and Economics 38, 21–38.

Mitchell, T., 1997. Machine Learning. McGraw-Hill, New York, NY.

Ng, A.C.-Y., Li, J.S.-H., 2011. Valuing variable annuity guarantees with the multivariate esscher transform. Insurance: Mathematics and Economics 49 (3), 393–400.

Ng, A.C.-Y., Li, J.S.-H., 2013. Pricing and hedging variable annuity guarantees with multiasset stochastic investment models. North American Actuarial Journal 17 (1), 41–62.

Nielsen, J., Sandmann, K., 1995. Equity-linked life insurance: a model with stochastic interest rates. Insurance Mathematics and Economics 16, 225–253.

Peng, J., Leung, K., Kwok, Y., 2010. Pricing guaranteed minimum withdrawal benefits under stochastic interest rates. Quantitative Finance 1–9.

Qian, L., Wang, W., Wang, R., Tang, Y., 2010. Valuation of equity-indexed annuity under stochastic mortality and interest rate. Insurance: Mathematics and Economics 47 (2), 123–129.

Rasmussen, C., Williams, C., 2005. Gaussian Processes for Machine Learning. In: Adaptive Computation and Machine Learning, The MIT Press, Cambridge, MA.

Wang, Y., 2009. Quantile hedging for guaranteed minimum death benefits. Insurance: Mathematics and Economics 45 (3), 449–458.

Windcliff, H., Forsyth, P., Vetzal, K., 2001. Valuation of segregated funds: shout options with maturity extensions. Insurance: Mathematics and Economics 29 (1), 1–21.

Windcliff, H., Le Roux, M., Forsyth, P., Vetzal, K., 2002. Understanding the behavior and hedging of segregated funds offering the reset feature. North American Actuarial Journal 6 (2), 107–124.

Xu, Z., Wang, Q., 2009. Pricing guaranteed minimum withdrawal benefits: a pde approach. In: International Conference on Management and Service Science, pp. 1–4.

Yang, S., Dai, T.-S., 2013. A flexible tree for evaluating guaranteed minimum withdrawal benefits under deferred life annuity contracts with various provisions. Insurance: Mathematics and Economics 52 (2), 231–242.

Young, V., 2003. Equity-indexed life insurance: pricing and reserving using the principle of equivalent utility. North American Actuarial Journal 7 (1), 68–86.