Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

APPLICATION OF METAMODELING TO THE VALUATION OF LARGE VARIABLE ANNUITY PORTFOLIOS

Guojun Gan

Department of Mathematics
University of Connecticut
196 Auditorium Rd, U-3009
Storrs, CT, 06269, USA

ABSTRACT

Variable annuities are long-term investment vehicles that have grown rapidly in popularity recently. One major feature of variable annuities is that they contain guarantees. The guarantees embedded in variable annuities are complex and the values of the guarantees cannot be obtained from closed-form formulas. Insurance companies rely heavily on Monte Carlo simulation to calculate the fair market values of the guarantees. Valuation and risk management of a large portfolio of variable annuities are a big challenge to insurance companies because the Monte Carlo simulation model is very time consuming. In this paper, we propose to use a metamodeling approach to speed up the valuation of large portfolios of variable annuities. Our numerical results show that the metamodeling approach can reduce the runtime significantly and produce accurate approximations.

1 INTRODUCTION

A variable annuity (VA) refers to an attractive life insurance product that provides upside participation and downside protection in both bull and bear markets. Once an investor enters into a variable annuity contract with an insurance company, the investor agrees to make one lump-sum or a series of purchase payments to the insurance company and the insurance company agrees to make benefit payments to the investor beginning immediately or at some future date. In a variable annuity contract, the investor's money is invested in a basket of mutual funds, which include bond funds and equity funds. When a variable annuity matures, the benefit of the contract is equal to the market value of the accumulated purchase payments. Variable annuity has other names such as segregated fund, guaranteed investment fund, unit-linked life insurance, equity-linked life insurance, or participating life insurance (Armstrong 2001)

A main feature of variable annuities is that they contain guarantees. For example, almost every VA contract contains the guaranteed minimum death benefit (GMDB) (Gerber, Shiu, and Yang 2013). VA contracts also include the guaranteed minimum withdrawal benefit (GMWB) (Yang and Dai 2013), the guaranteed minimum maturity benefit (GMAB) (Jiang and Chang 2010), and the guaranteed minimum income benefit (GMIB) (Bacinello, Millossovich, Olivieri, and Pitacco 2011). These guarantees are optional in that a policyholder can purchase these guarantees for additional fees. Due to the attractive guarantee features, variable annuities have grown rapidly in popularity recently. Figure 1 shows the annual sales of variable annuities from 2010 to 2013 in the US. From the figure we see that the annual sale of variable annuities was more than 140 billion dollars in the past few years.

The guarantees embedded in variable annuities are financial guarantees that cannot be adequately addressed by traditional actuarial approaches (Hardy 2000), which rely on diversification. Table 1 shows the cash flows of a variable annuity policy with a GMWB rider under a specific economic scenario. Because of the guarantee, the policyholder can withdrawal the guaranteed amount every year even when

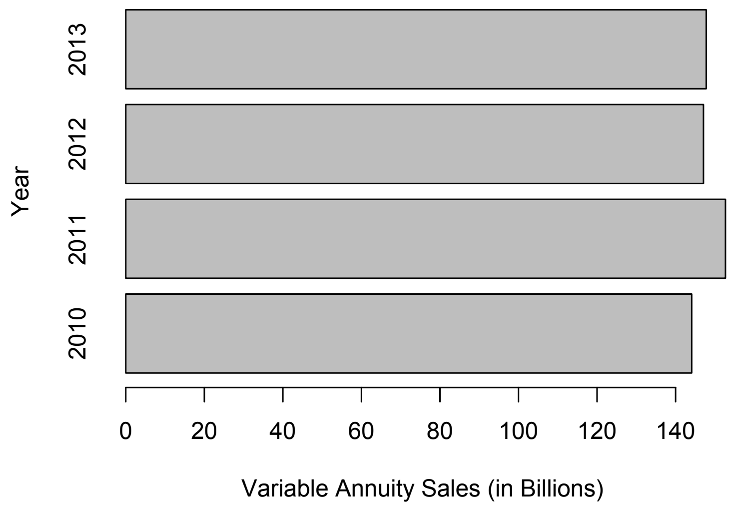


Figure 1: Variable annuity sales in the US. The numbers are in billions of dollars. (Source: LIMRA)

the investment fund goes to zero. The last column shows the guarantee cash flows, which are claims paid to the policyholder by the insurance company. From the example we see that the insurance company will loss money on all policies when market goes down. Dynamic hedging (Hardy 2003) is a popular risk management approach for variable annuities and is adopted by many insurance companies.

Since VA contracts embedding guarantees are relatively complex, the calculation of their fair market values cannot be done in closed form except for special cases (Gerber and Shiu 2003, Feng and Volkmer 2012). In practice, insurance companies rely on the Monte Carlo simulation method to determine the fair market values of VA contracts. However, using the Monte Carlo simulation method to value a large portfolio of VA contract is time consuming because every VA contract needs to be projected over many scenarios for a long time horizon. For example, using a Monte Carlo simulation method with 1,000 scenarios and 360 monthly time steps to calculate the fair market value of a portfolio consisting of 100,000 VA policies involves the following number of cash flow projections:

$$100,000 \times 1,000 \times 360 = 3.6 \times 10^{10}.$$

If a computer can process 200,000 projections per second, then it would take this computer 50 hours to finish the calculation. That is only the runtime for calculating the fair market value under a single market condition. To calculate the fair market values under 100 different market conditions, it would take this computer 5,000 hours to complete the calculation.

To make dynamic hedging work for a large portfolio of VA policies, an insurance company needs to calculate the Greeks (e.g., dollar Delta and dollar Rho) of the big portfolio on a daily basis in order to incorporate the changes in the portfolio and the market. In particular, the insurance company needs to complete the calculation of the Greeks over night between today's market close and tomorrow's market open. In order to complete the computationally intensive calculation, insurance companies employ many computers to do the calculation. For example, GPUs (Graphics Processing Unit) have been used to value VA contracts (Phillips 2012, NVIDIA 2012).

Although using many computers or GPUs can speed up the calculation, this approach is not scalable. In other words, if the number of VA policies in a portfolio doubles, then the insurance company needs to double the number of computers or GPUs in order to complete the calculation within the same time interval. In addition, buying or renting many computers or GPUs is expensive and can cost the insurance company a lot of money annually.

Table 1: Cash flows of a variable annuity policy with a GMWB rider under a specific economic scenario. In this sample variable annuity policy, the initial investment is 100,000 dollars, the GMWB amount is equal to the initial investment, and the policyholder is allowed to withdrawal 8% of the initial investment until the initial investment is recovered.

Policy Year	Investment Return	Fund Before Withdrawal	Annual Withdrawal	Fund After Withdrawal	Remaining Benefit	Guarantee Cash Flow
1	-10%	90,000	8,000	82,000	92,000	0
1	-10%	90,000	8,000	82,000	92,000	0
2	10%	90,200	8,000	82,200	84,000	0
3	-30%	57,540	8,000	49,540	76,000	0
4	-30%	34,678	8,000	26,678	68,000	0
5	-10%	24,010	8,000	16,010	60,000	0
6	-10%	14,409	8,000	6,409	52,000	0
7	10%	7,050	8,000	0	44,000	950
8	-	0	8,000	0	36,000	8,000
9	-	0	8,000	0	28,000	8,000
10	-	0	8,000	0	20,000	8,000
11	-	0	8,000	0	12,000	8,000
12	-	0	8,000	0	4,000	8,000
13	-	0	4,000	0	0	4,000

In this paper, we apply a metamodeling approach to address the computational problem mentioned above. In particular, we adopt a metamodel by using a Latin hypercube sampling method (McKay, Beckman, and Conover 1979, Pistone and Vicario 2010, Petelet, Iooss, Asserin, and Loredo 2010, Viana 2013) and the ordinary kriging model (Isaaks and Srivastava 1990).

The remaining of the paper is structured as follows. Section 2 gives a brief review of simulation metamodeling and its application in finance. Section 3 introduces a Latin hypercube sampling method used to select representative VA contracts and the ordinary kriging model. Section 3 also presents some numerical results of the proposed metamodel. Section 4 concludes the paper and gives a survey of future work.

2 METAMODELING

In simulation modeling, a metamodel refers to a model of a simulation model (Friedman 1996). One main reason for building a model of a simulation model is that the simulation model is complicate and computationally intensive. Metamodels of a simulation model are much simpler and more computationally efficient than the simulation model. Metamodels are sometimes called response surface models or surrogate models.

Building a metamodel of a simulation model involves three steps: first, we use an experimental design method to select a small set of sample points from the input domain; second, we run the simulation model to generate outputs at these selected sample points; third, we choose an appropriate metamodel form and estimate the parameters of the metamodel using the selected sample points and the outputs of the simulation model at the selected sample points. The experimental design method and the metamodel are two interrelated components of metamodeling.

During the past six decades, many papers on metamodelling and its applications have been published. Early works in this area include (Kleijnen 1975), (Franke 1982), (Hoerl 1985), (Barton 1992), (Laslett 1994), (Barton 1994), (Madu and Kuei 1994), (Barton 1998), to name just a few. Kleijnen (1975) introduced the concept of metamodels for simulation models. Barton (1994) presented a review of metamodels for studying the behavior of computer simulations during that time. In particular, Barton reviewed several modeling approaches such as spline models, radial basis functions, kernel methods, and spatial correlation models. Barton (1998) discussed other metamodel types such as neural network metamodels.

Recent works in this area include (Kleijnen and Deflandre 2006), (Wu, Chen, Hu, Zhang, and Liang 2008), (Kleijnen 2009), (Ankenman, Nelson, and Staum 2010), (Khuri and Mukhopadhyay 2010), (Yin, Ng, and Ng 2011), (Razavi, Tolson, and Burn 2012), (Wei, Wu, and Chen 2012), and (Zhao, Yue, Liu, Gao, and Zhang 2014), to name just a few. Kleijnen (2009) presented a review of the Kriging metamodel. Ankenman, Nelson, and Staum (2010) extended the basic theory of Kriging to the stochastic simulation setting. Khuri and Mukhopadhyay (2010) provided a survey of the development of response surface methodology since its introduction in the early 1950s. Razavi, Tolson, and Burn (2012) presented a wide variety of metamodeling methods with an emphasis on the water resources field.

A number of books have been devoted to metamodels, response surface methodologies, and surrogate models: (Box and Draper 1987), (Khuri and Cornell 1987), (Friedman 1996), (Shore 2005), (Khuri 2006), (Box and Draper 2007), (Forrester, Sobester, and Keane 2008), (Myers, Montgomery, and Anderson-Cook 2009), and (Das 2014). Friedman (1996) presented a diverse set of scholarly materials relevant to the study of simulation metamodels, including usage, applications, and methodology of metamodels. Box and Draper (2007) is a successor volume to Box and Draper (1987) and covers many topics on response surface models. Das (2014) is an introductory book devoted to robust response surface methodology and contains a review of the existing literature on response surface methodology.

The concept of metamodeling has been applied to financial engineering recently. In (Baysal, Nelson, and Staum 2008), the authors used Latin hypercube designs and kriging to simulate hedging and trading strategies under nested simulation. Liu and Staum (2009) and Liu and Staum (2010) used stochastic kriging to estimate expected shortfall of a portfolio. Gan (2013) used a data clustering method (Gan 2011) and the ordinary kriging method to estimate the fair market values of a portfolio of variable annuities. Salle and Yildizoglu (2014) applied the kriging model to approximate two well known economic models. Gan and Lin (2015) used a data clustering method and a universal kriging method developed for functional data to estimate the fair market values of a portfolio of variable annuities under nested simulation.

3 AN APPLICATION OF METAMODELING TO VARIABLE ANNUITY

In this section, we apply a metamodeling method to attack the computational problem arising from the variable annuity area. In this example, we use Latin hypercube sampling and kriging. The kriging method is a popular metamodeling method and the Latin hypercube sampling method works well with the kriging method (Baysal, Nelson, and Staum 2008).

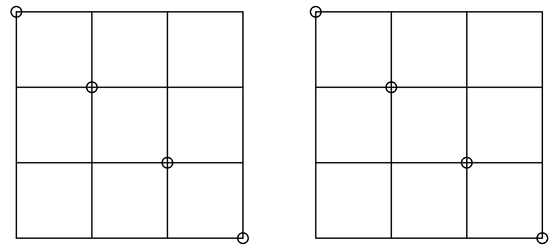
3.1 Latin Hypercube Sampling

Latin hypercube sampling (LHS) is a statistical method for generating plausible design points from multiple dimensional spaces that are used to conduct computer experiments. Figure 2 gives two examples of Latin hypercube designs with 4 points on a 2-dimensional area. From the figure we see that there is only one sample point in each row and each column. For more information about LHS, readers are referred to (McKay, Beckman, and Conover 1979), (Liefvendahl and Stocki 2006), (Minasny and McBratney 2006), (Pistone and Vicario 2010), (Petelet, Iooss, Asserin, and Loredo 2010), and (Viana 2013).

When the number of divisions and the number of variables increase, the number of Latin hypercubes increases exponentially (McKay and Wanless 2008). For example, there are

$$64 \times 4! \times (3!)^3 = 331,776$$

Latin hypercubes with 4 divisions and 3 variables. As a result, one way to find a good Latin hypercube design is to generate Latin hypercube samples randomly and select the best one from the samples.



(a) (b)

Figure 2: Two examples of Latin hypercube designs with 4 divisions and 2 variables.

Now let us introduce a LHS method for selecting representative VA policies, which are described by both categorical and numerical variables. The LHS method introduced here is able to handle both numerical and categorical variables. There are several ways to select an optimal Latin hypercube design (Liefvendahl and Stocki 2006). Here we select an optimal Latin hypercube design by maximizing the minimum distances.

To describe the LHS method, 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 numerical and the remaining $d_2 = d - d_1$ attributes are categorical. For $j = 1, 2, \ldots, d_1$, let L_j and H_j denote the minimum and maximum values that the jth numerical variable can take. That is,

$$L_j = \min\{x_j : \mathbf{x} \in X\}, \quad H_j = \max\{x_j : \mathbf{x} \in X\}, \tag{1}$$

where x_j denotes the jth component of $\mathbf{x}$ and $X = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ denote the portfolio of VA contracts. For $j = d_1 + 1,\ d_1 + 2,\ \ldots,\ d$, let N_j denote the number of distinct values that the jth categorical variable can take, i.e.,

$$N_j = \left|\{x_j : \mathbf{x} \in X\}\right|, \tag{2}$$

where $|\cdot|$ denote the number of elements in a set.

Suppose that we want to generate a Latin hypercube design with k design points, where $k \geq 2$. To do that, we first divide the range of each of the d_1 numerical variable into k divisions. For each $l = 1, 2, \ldots, k$, the lth division of the jth dimension is given by

$$I_l = \left(L_j + \left(l - \frac{3}{2}\right)\frac{H_j - L_j}{k-1}, L_j + \left(l - \frac{1}{2}\right)\frac{H_j - L_j}{k-1}\right].$$

Since

$$\bigcup_{l=1}^{k} I_l = \left(L_j - \frac{H_j - L_j}{2(k-1)}, H_j + \frac{H_j - L_j}{2(k-1)}\right] \subset [L_j, H_j],$$

the union of the k divisions covers the whole range of the jth variable. For each of the remaining categorical variables, we just treat each category as a division.

Let $\mathscr{H}$ be a set of d-dimensional points defined to be

$$\mathscr{H} = \{(a_1, a_2, \ldots, a_d)\} \tag{3}$$

such that for $j = 1, 2, \ldots, d_1$,

$$a_j \in \left\{L_j + (l-1)\frac{H_j - L_j}{k-1}, l = 1, 2, \ldots, k\right\},$$

and for $j = d_1+1, d_1+2, \ldots, d$,

$$a_j \in \left\{A_{jl}, l = 1,2,\ldots,N_j\right\},$$

where A_{j1}, A_{j2}, ..., A_{jN_j} are the distinct categories of the jth variable and L_j, H_j, and N_j are defined in Equations (1) and (2). There are many points in the set $\mathscr{H}$. In fact, we have

$$|\mathscr{H}| = k^{d_1} \prod_{j=d_1+1}^{d} N_j.$$

The first step of the LHS method is to select k points from the set $\mathscr{H}$ with the best score, which is to be defined. Let H be a subset of $\mathscr{H}$ with k elements. The score of the set H is defined to be the minimum distance between any pairs of distinct points in H. That is,

$$S(H) = \min\left\{M(\mathbf{a},\mathbf{b}) : \mathbf{a} \in H, \mathbf{b} \in H, \mathbf{a} \neq \mathbf{b}\right\}, \tag{4}$$

where $M(\mathbf{a},\mathbf{b})$ is the distance between $\mathbf{a}$ and $\mathbf{b}$ given by

$$M(\mathbf{a},\mathbf{b}) = \sum_{j=1}^{d_1} \frac{(k-1)|a_j - b_j|}{H_j - L_j} + \sum_{j=d_1+1}^{d} \delta(a_j, b_j), \tag{5}$$

where a_j and b_j are the jth components of $\mathbf{a}$ and $\mathbf{b}$, respectively, and $\delta(\cdot,\cdot)$ is defined in Equation (10). The larger the score, the better the Latin hypercube design. An optimal Latin hypercube design with k points is defined as

$$H^* = \operatorname*{argmax}_{H \subset \mathscr{H}, |H|=k} S(H). \tag{6}$$

Since the set $\mathscr{H}$ contains huge number of points, finding an optimal Latin hypercube design with k points from $\mathscr{H}$ is not easy. To find such an optimal Latin hypercube design, we randomly generate many (e.g., 500) Latin hypercube designs and select the one with the largest score. To generate a random Latin hypercube design $H = \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k\}$ with k points, we proceed as follows:

1. For each $j = 1,2,\ldots,d_1$, we randomly generate k uniform real numbers from the interval $[0,1]$. Suppose that these random numbers are r_{j1}, r_{j2}, ..., r_{jk}. Since these numbers are random real numbers, they are mutually distinct in general. We sort the k real numbers in an ascending order such that
$$r_{ji_1} < r_{jr_2} < \cdots < r_{jr_k},$$
where $(i_1, i_2, \ldots, i_k)$ is a permutation of $(1,2,\ldots,k)$. Then we define the first d_1 coordinates of the k design points as
$$a_{jl} = L_j + (i_l - 1)\frac{H_j - L_j}{k-1}, \quad j = 1,\ldots,d_1,\, l = 1,\ldots,k.$$
For each $j = 1,2,\ldots,d_1$, the coordinates of the k design points at the jth dimension are mutually distinct.
2. For each $j = d_1+1, d_1+2, \ldots, d$, we randomly generate k uniform integers from $\{1,2,\ldots,N_j\}$. For portfolios of variable annuity policies, we usually have $k > N_j$, that is, the number of design points is larger than the number of values that a categorical variable can take. Suppose that these random integers are $i_1, i_2, \ldots, i_k$. Then we define the remaining d_2 coordinates of the k design points as
$$a_{jl} = A_{ji_l}, \quad j = d_1+1,\ldots,d,\, l = 1,\ldots,k,$$
where A_{j1}, A_{j2}, ..., A_{jN_j} are the distinct categories of the jth variable.

Once we find a Latin hypercube design $H^* = \{\mathbf{a}_1^*, \mathbf{a}_2^*, \ldots, \mathbf{a}_k^*\}$ using the above procedure. The second step of the LHS method is to find k representative VA policies that are close to the k design points in H^*. In particular, the VA policy that is close to $\mathbf{a}_i^*$ is determined by

$$\mathbf{z}_i = \operatorname*{argmin}_{\mathbf{x}\in X} M(\mathbf{a}_i^*, \mathbf{x}), \quad i = 1, 2, \ldots, k,$$

where $M(\cdot,\cdot)$ is defined in Equation (5).

3.2 Ordinary Kriging

We use the ordinary kriging method (Isaaks and Srivastava 1990) to estimate the fair market value and the Greeks (e.g., sensitivities of the fair market values) of the whole portfolio from the representative VA policies.

Let $\mathbf{z}_1$, $\mathbf{z}_2$, ..., $\mathbf{z}_k$ 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

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

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

$$\begin{pmatrix} V_{11} & \cdots & V_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ V_{k1} & \cdots & V_{kk} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_{i1} \\ \vdots \\ w_{ik} \\ \theta_i \end{pmatrix} = \begin{pmatrix} D_{i1} \\ \vdots \\ D_{ik} \\ 1 \end{pmatrix}, \tag{8}$$

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

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

and

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

Here $\alpha \geq 0$ and $\beta > 0$ are two parameters, and the distance function $D(\cdot,\cdot)$ is defined as

$$D(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{h=1}^{d_1} (x_h - y_h)^2 + \sum_{h=d_1+1}^{d} \delta(x_h, y_h)}, \tag{9}$$

where x_h and y_h are the hth component of $\mathbf{x}$ and $\mathbf{y}$, respectively, and $\delta(\cdot,\cdot)$ is the simple matching distance defined as

$$\delta(x_h, y_h) = \begin{cases} 0, & \text{if } x_h = y_h, \\ 1, & \text{if } x_h \neq y_h. \end{cases} \tag{10}$$

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

$$\hat{Y} = \sum_{i=1}^{n} \hat{y}_i = \sum_{i=1}^{n} \sum_{j=1}^{k} w_{ij} \cdot y_j = \sum_{j=1}^{k} w_j \cdot y_j, \tag{11}$$

where

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

The fair value $\hat{Y}$ of the portfolio can be calculated efficiently by solving w_1, w_2, …, w_k from the following linear equation system

$$\begin{pmatrix} V_{11} & \cdots & V_{1k} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ V_{k1} & \cdots & V_{kk} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_k \\ \theta \end{pmatrix} = \begin{pmatrix} D_1 \\ \vdots \\ D_k \\ n \end{pmatrix}, \tag{12}$$

where

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

In fact, Equation (12) is obtained by summing both sides of Equation (8) from $i = 1$ to n.

3.3 Numerical Results

In this subsection, we present some numerical results of using the metamodel for VA portfolio valuation. To do the test, we follow the setup used in (Gan 2013). We generate a portfolio of 200,000 synthetic VA contracts. The attributes and their ranges of values are shown in Table 2. For each synthetic VA contract, the value of an attribute is generated from a uniform distribution with the corresponding range given in Table 2.

Table 2: Variable annuity attributes and their ranges of values. Here GMDB and GMWB refer to guaranteed minimum death benefit and guaranteed minimum withdrawal benefit, which are two major features of variable annuities.

Attribute	Values
Guarantee type	{GMDB only, GMDB + GMWB}
Gender	{Male, Female}
Age	{20, 21, 22, …, 60}
Account value	$[10000, 500000]$
GMWB withdrawal rate	{0.04, 0.05, 0.06, 0.07, 0.08}
Maturity	{10, 11, 12, …, 25}

We use the metamodel to estimate the fair market value, dollar Delta, and dollar Rho of the whole portfolio. In all the test cases, we used 500 iterations in the LHS method. In other words, 500 Latin hypercube designs are randomly generated and the one with the largest score is selected.

In our test, we used the LHS method to select a set of representative VA policies. Then we used the ordinary kriging method to estimate the fair market value, dollar Delta, and dollar Rho of the whole portfolio. In the ordinary kriging method, we set $\alpha = 0$ and β to be the 95th percentile of all distances between pairs of the representative VA policies as suggested in (Isaaks and Srivastava 1990). The accuracy of the metamodel is summarized in Table 3. The first row (MC) shows the fair market value, dollar Delta, and dollar Rho of the portfolio calculated by the Monte Carlo simulation model. The second row and the third row shows the numbers estimated by the metamodel. The last four rows show the dollar difference and the percentage difference. From the tables we see than most of the percentage differences are less than 0.5%.

Table 3: The fair market values, dollar Deltas, and dollar Rhos calculated by the Monte Carlo simulation model and those estimated by the metamodel with different number of representative VA policies. Numbers in the first five rows are in dollars. Numbers in brackets are negative numbers.

	Fair Market Value	Dollar Delta	Dollar Rho
MC	3,003,947,180	(8,150,275,955)	(9,736,358)
LHS100	3,016,679,402	(8,180,679,337)	(9,789,810)
LHS500	3,008,948,423	(8,181,891,156)	(9,751,202)
LHS100-MC	12,732,222	(30,403,381)	(53,451)
LHS500-MC	5,001,243	(31,615,200)	(14,844)
(LHS100-MC)/MC	0.42%	0.37%	0.55%
(LHS500-MC)/MC	0.17%	0.39%	0.15%

Table 4: Runtime used by the Monte Carlo simulation model and the metamodel with different number of representative VA policies. The numbers are in seconds. The LHS, MC, Kriging rows denote the runtime used by the Latin hypercube sampling method, the Monte Carlo simulation model, and the ordinary kriging method, respectively.

	Number of Rep. Policies		Entire Portfolio
	100	500	200,000
LHS	5.05	20.47	NA
MC	1.63	5.38	1942.22
Kriging	5.83	26.8	NA
Total	12.51	52.65	1942.22

Table 4 shows the runtime used by the Monte Carlo simulation model and that used by the metamodel. From the table we see that the metamodel is much faster than the Monte Carlo simulation model for valuing the portfolio. It took the Monte Carlo simulation model more than 30 minutes to calculate the fair market value, dollar Delta, and dollar Rho of the portfolio. In contrast, it took the metamodel less than one minute to produce accurate estimates of these numbers.

Although the VA policies considered in the numerical experiments are much simpler than the real VA policies, the numerical results show that metamodeling is a promising approach to address the computational problem arising from the VA area.

4 CONCLUSIONS

For an insurance company that has a big VA portfolio, a major challenge in risk management of the VA business is to calculate the fair market value and the Greeks of the VA portfolio in an efficient way. In this paper, we proposed a metamodeling approach to address the computational problem from the perspective of mathematical modeling instead of hardware. The idea of the metamodeling approach is to first select a small set of representative policies, then price the representative policies, and finally estimate the value of the whole portfolio. The method is efficient in that only a small set of representative policies is required to be priced by the time-consuming Monte Carlo simulation model.

To test the usefulness of the metamodeling method, we created a synthetic portfolio of VA policies and compared the accuracy and speed of the metamodel and the Monte Carlo simulation model using the synthetic portfolio. Our numerical results indicate that the metamodeling method is computationally efficient and is able to produce accurate approximations.

The metamodeling approach can be useful for insurance companies that have a VA business. Our simple numerical experiments show that the metamodeling approach has the potential to reduce the runtime significantly. In practice, the simulation model used by insurance companies is much more complex than the one used in this paper. For example, monthly time steps are usually used in practice and the cash flow projection is often complex for real variable annuity policies. The metamodeling approach can reduce the runtime even more in real applications than in the toy example presented in the paper. In addition, the metamodeling approach can also be used to address other computationally intensive issues such as calculating the economic capitals of a portfolio of variable annuities.

In future, we would like to test other experimental design methods and metamodeling techniques. In particular, we would like to conduct a comprehensive comparison of various experimental design methods, such as factorial design (Alam, McNaught, and Ringrose 2004), and metamodeling techniques, such as response surface methodology (Myers, Montgomery, and Anderson-Cook 2009).

REFERENCES

Alam, F. M., K. R. McNaught, and T. J. Ringrose. 2004. "A comparison of experimental designs in the development of a neural network simulation metamodel". *Simulation Modelling Practice and Theory* 12 (78): 559 – 578. Simulation in Operational Research.

Ankenman, B., B. L. Nelson, and J. Staum. 2010. "Stochastic Kriging for Simulation Metamodeling". *Operations Research* 58 (2): 371–382.

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

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

Barton, R. 1994, Dec. "Metamodeling: a state of the art review". In *Winter Simulation Conference Proceedings*, 237–244.

Barton, R. R. 1992. "Metamodels for Simulation Input-output Relations". In *Proceedings of the 24th Conference on Winter Simulation*, WSC '92, 289–299: ACM.

Barton, R. R. 1998. "Simulation Metamodels". In *Proceedings of the 30th Conference on Winter Simulation*, WSC '98, 167–176. Los Alamitos, CA, USA: IEEE Computer Society Press.

Baysal, R., B. Nelson, and J. Staum. 2008, Dec. "Response surface methodology for simulating hedging and trading strategies". In *Simulation Conference, 2008. WSC 2008. Winter*, 629–637.

Box, G. E. P., and N. R. Draper. 1987. *Empirical Model-Building and Response Surfaces*. Hoboken, NJ: Wiley.

Box, G. E. P., and N. R. Draper. 2007. *Response Surfaces, Mixtures, and Ridge Analyses*. 2nd ed. Hoboken, NJ: Wiley.

Das, R. N. 2014. *Robust Response Surfaces, Regression, and Positive Data Analyses*. Boca Raton, FL: CRC Press.

Feng, R., and H. Volkmer. 2012. "Analytical calculation of risk measures for variable annuity guaranteed benefits". *Insurance: Mathematics and Economics* 51 (3): 636–648.

Forrester, A., A. Sobester, and A. Keane. 2008. *Engineering Design via Surrogate Modelling: A Practical Guide*. West Sussex, UK: Wiley.

Franke, R. 1982. "Scattered Data Interpolation: Tests of Some Method". *Mathematics of Computation* 38 (157): 181–200.

Friedman, L. W. 1996. *The Simulation Metamodel*. Norwell, MA, USA: Kluwer Academic Publishers.

Gan, G. 2011. *Data Clustering in C++: An Object-Oriented Approach*. Data Mining and Knowledge Discovery Series. Boca Raton, FL, USA: Chapman & Hall/CRC Press.

Gan, G. 2013. "Application of data clustering and machine learning in variable annuity valuation". *Insurance: Mathematics and Economics* 53 (3): 795–801.

Gan, G., and X. S. Lin. 2015. "Valuation of large variable annuity portfolios under nested simulation: A functional data approach". *Insurance: Mathematics and Economics* 62:138–150.

Gerber, H., and E. Shiu. 2003. "Pricing Lookback Options and Dynamic Guarantees". *North American Actuarial Journal* 7 (1): 48–67.

Gerber, H., E. Shiu, and H. Yang. 2013. "Valuing equity-linked death benefits in jump diffusion models". *Insurance: Mathematics and Economics* 53 (3): 615–623.

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

Hardy, M. 2003. *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. Hoboken, New Jersey: John Wiley & Sons, Inc.

Hoerl, R. W. 1985. "Ridge Analysis 25 Years Later". *The American Statistician* 39 (3): 186–192.

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

Jiang, S., and M. Chang. 2010. "Variable Annuity with Guarantees: Valuation and Simulation". *Journal of Money, Investment and Banking* 14:74–83.

Khuri, A. I. (Ed.) 2006. *Response Surface Methodology And Related Topics*. Singapore: World Scientific.

Khuri, A. I., and J. A. Cornell. 1987. *Response Surfaces: Designs and Analyses*. New York, NY, USA: Marcel Dekker, Inc.

Khuri, A. I., and S. Mukhopadhyay. 2010. "Response surface methodology". *Wiley Interdisciplinary Reviews: Computational Statistics* 2 (2): 128–149.

Kleijnen, J. P. 2009. "Kriging metamodeling in simulation: A review". *European Journal of Operational Research* 192 (3): 707 – 716.

Kleijnen, J. P., and D. Deflandre. 2006. "Validation of regression metamodels in simulation: Bootstrap approach". *European Journal of Operational Research* 170 (1): 120 – 131.

Kleijnen, J. P. C. 1975. "A Comment on Blanning's "Metamodel for Sensitivity Analysis: The Regression Metamodel in Simulation"". *Interfaces* 5 (3): 21–23.

Laslett, G. M. 1994. "Kriging and Splines: An Empirical Comparison of Their Predictive Performance in Some Applications". *Journal of the American Statistical Association* 89 (426): 391–400.

Liefvendahl, M., and R. Stocki. 2006. "A study on algorithms for optimization of Latin hypercubes". *Journal of Statistical Planning and Inference* 136 (9): 3231 – 3247.

Liu, M., and J. Staum. 2009, Dec. "Estimating expected shortfall with stochastic kriging". In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, 1249–1260.

Liu, M., and J. Staum. 2010. "Stochastic kriging for efficient nested simulation of expected shortfall". *The Journal of Risk* 12 (3): 3–27.

Madu, C. N., and C.-H. Kuei. 1994. "Regression metamodeling in computer simulation - the state of the art". *Simulation Practice and Theory* 2 (1): 27 – 41.

McKay, B., and I. Wanless. 2008. "A Census of Small Latin Hypercubes". *SIAM Journal on Discrete Mathematics* 22 (2): 719–736.

McKay, M., R. Beckman, and W. J. Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code". *Technometrics* 21 (2): 239–245.

Minasny, B., and A. B. McBratney. 2006. "A conditioned Latin hypercube method for sampling in the presence of ancillary information". *Computers & Geosciences* 32 (9): 1378 – 1388.

Myers, R. H., D. C. Montgomery, and C. M. Anderson-Cook. 2009. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. 3rd ed. Hoboken, NJ: Wiley.

NVIDIA 2012. "People Like VAs Like GPUs". *Wilmott magazine* 2012 (60): 10–13.

Petelet, M., B. Iooss, O. Asserin, and A. Loredo. 2010. "Latin hypercube sampling with inequality constraints". *AStA Advances in Statistical Analysis* 94 (4): 325–339.

Phillips, P. 2012. "Lessons Learned About Leveraging High Performance Computing for Variable Annuities". In *Equity-Based Insurance Guarantees Conference*. Chicago, IL.

Pistone, G., and G. Vicario. 2010. "Comparing and generating Latin Hypercube designs in Kriging models". *AStA Advances in Statistical Analysis* 94 (4): 353–366.

Razavi, S., B. A. Tolson, and D. H. Burn. 2012. "Review of surrogate modeling in water resources". *Water Resources Research* 48 (7): 1–32.

Salle, I., and M. Yildizoglu. 2014. "Efficient Sampling and Meta-Modeling for Computational Economic Models". *Computational Economics* 44 (4): 507–536.

Shore, H. 2005. *Response Modeling Methodology: Empirical Modeling for Engineering and Science*. Singapore: World Scientific.

Viana, F. 2013. "Things you wanted to know about the Latin hypercube design and were afraid to ask". In *10th World Congress on Structural and Multidisciplinary Optimization*. Orlando, FL.

Wei, X., Y.-Z. Wu, and L.-P. Chen. 2012. "A new sequential optimal sampling method for radial basis functions". *Applied Mathematics and Computation* 218 (19): 9635 – 9646.

Wu, B., L. Chen, Z. Hu, W. Zhang, and J. Liang. 2008, Oct. "A relevance vector regression based metamodeling approach for complex system analysis". In *7th International Conference on System Simulation and Scientific Computing*, 612–619.

Yang, S., and T.-S. Dai. 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.

Yin, J., S. Ng, and K. Ng. 2011. "Kriging metamodel with modified nugget-effect: The heteroscedastic variance case". *Computers & Industrial Engineering* 61 (3): 760 – 777.

Zhao, H., Z. Yue, Y. Liu, Z. Gao, and Y. Zhang. 2014. "An efficient reliability method combining adaptive importance sampling and Kriging metamodel". *Applied Mathematical Modelling*.

AUTHOR BIOGRAPHIES

GUOJUN GAN is an assistant professor in the Department of Mathematics at the University of Connecticut, where he has been since August 2014. Prior to that, he worked at a large life insurance companies in Toronto, Canada for six years. He received a BS from Jilin University, China, in 2001 and MS and PhD degrees from York University, Canada, in 2003 and 2007, respectively.