

An efficient machine learning approach

Authors

Wei Xu¹, Yuehuan Chen², Conrad Coleman²,
and Thomas F. Coleman¹



**GLOBAL
RISK
INSTITUTE**

For Risk Management of large complex portfolios arising in insurance

Machine Learning (ML) is a rapidly developing technology with applications in a wide variety of areas. This technology has great potential in the efficient evaluation (and hedging) of large complex portfolios arising in the insurance industry. For example, we estimate a portfolio of 200,000 variable annuity contracts which would require 1.5 years to risk-analyze using standard techniques whereas the machine learning approach we describe here takes 1.5 hours!

Due to the scale and complexity of some insurance portfolios, computing the corresponding risk indicators can be an enormous computational challenge faced by financial institutions. In this white paper, to illustrate the potential of ML technology, we outline a new machine learning technique that efficiently computes deltas, value-at-risk (VaR), and conditional-value-at-risk (CVaR) measures of large variable annuity (VA) portfolios.

Variable annuity products are unit-linked investments with some form of guarantee, traditionally sold by insurers or banks into the retirement and investment market. They are very popular and represent huge portfolios. For example, in 2015 new VA sales in the US were 133 billion dollars^[3] while the sales in the UK were 433.3 million pounds.^[4] Similar popularity of VAs can be observed in several other countries including Canada, Japan, and South Korea. All

major insurance companies in these countries are managing VA portfolios of significant (and growing) sizes. However, determining how to hedge the risk of a large VA portfolio and determining the corresponding required capital poses a significant computational challenge to insurance companies. Existing valuation/hedging methods used in the risk management of an individual VA contract cannot feasibly be extended to a large portfolio of VAs due to the subsequent computational cost. For example, insurance companies typically follow a market-to-model approach and rely heavily on simulations. Nested simulations are used to determine the probability distribution of loss from mismatching in order to calculate required capital. However, the nested simulation approach has a significant computational cost. For example, if we run 1,000 scenarios on a 30 year contract and use 1,000 paths at each annual node, then we end up with a computational problem with $1,000 \times 30 \times 1,000 = 30$ million scenarios. This represents a lot of computing.

One effective way to save computational time is to reduce the number of scenarios over the contract, where the key question is how to achieve this reduction without decreasing the accuracy of the simulation. Here we first introduce a moment-matching scenario selection method^[1], whose selected scenarios match the first four moments of the stochastic scenario generation model. Let us consider an example to compute delta, 99 per cent VaR and 99 per cent CVaR for an individual VA contract. The contract has both a guaranteed minimum withdrawal benefit (GMWB) and a guaranteed minimum death benefit (GMDB) expiring in 19 years with a withdrawal rate of eight percent and an account value of \$272,934.25. The policy



**GLOBAL
RISK
INSTITUTE**



holder is a 41-year old male. Figure 1 shows the computed results for each year from our moment-matching method with 50 and 150 well selected scenarios and the nested simulations method with 10,000 scenarios. It shows that our moment-matching method can provide results comparable to nested simulation, but only takes three seconds while the nested simulation takes more than five minutes.

The moment-matching method is much faster than the nested simulation method, but it alone is still too expensive for portfolios with a large number of VA contracts. We observe that every VA contract is unique in terms of gender, age, time to maturity, guarantee type and fund type. Therefore, we combine a machine learning approach with the moment matching method. First, we select a relatively small number of VA contracts, and compute all of the risk indicators accurately by the moment-matching method. Second, the “machine is trained” with a standard machine learning method, such as a neural network or tree a regression. Finally, the risk indicators of the remaining contracts are estimate via the trained machine. For example, consider a portfolio with 10,000 VA contracts, whose attributes are randomly selected from Table 1. It takes a nested simulation more than 50 CPU hours to

obtain deltas, 99 per cent VaR and 99 per cent CVaR for each case, but it only requires 30 minutes for the machine learning approach with a 1,000-contract training set. Figure 2 shows that the computed results obtained from the two approaches are quite close.

Furthermore, the machine learning approach can easily handle huge portfolios (which cannot be handled via the nested simulation method due to cost). For example, a portfolio with 200,000 VA contracts, only requires a 2,000-contract training set in order to produce accurate risk indicator estimates. In conclusion, our proposed machine learning/moment matching approach appears to be a remarkably efficient alternative to the standard nested simulation methodology to hedge and manage the risk of large portfolios arising in the insurance industry. Further development and testing is needed on a real data set with a larger number of attributes with additional variation. Finally, we expect there will be application of similar ideas to other large complex portfolios of financial instruments.

For a complete technical paper on this subject, see ^[2].

Table 1
Description of Variable
Annuity attributes

ATTRIBUTE	VALUES	DISTRIBUTION
Guarantee Type	{GMDB, GMDB+GMWB}	50%, 50%
Gender	{Male, Female}	50%, 50%
Age	{20, 21, 22, ... , 60}	Uniform
Account Value	{10,000, 50,000}	Uniform
Withdrawal Rate	{0.04, 0.05, 0.06, 0.07, 0.08}	Uniform
Maturity	{10, 11, 12, ... , 25}	Uniform

Figure 1. Computed Deltas, 99% VaRs and 99% CVaRs for an individual VA contract from the moment matching method (MM) and the (traditional) nested simulations (NS), where MM50 stands for the moment matching method (our proposed ML approach) with 50 selected scenarios.

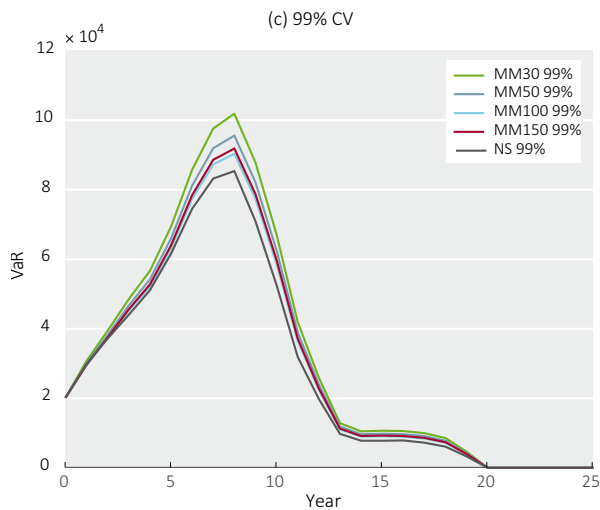
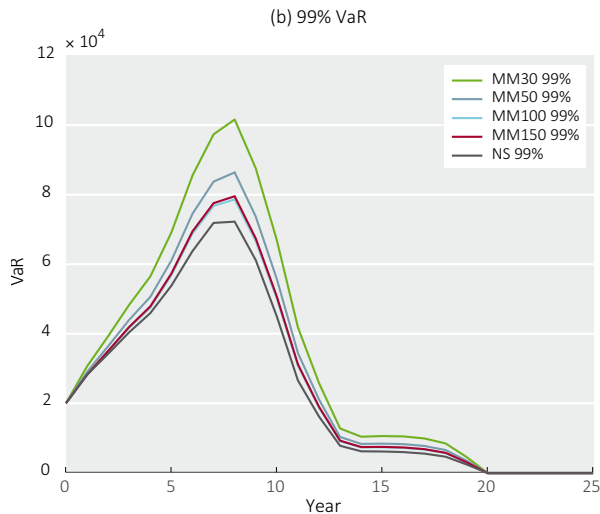
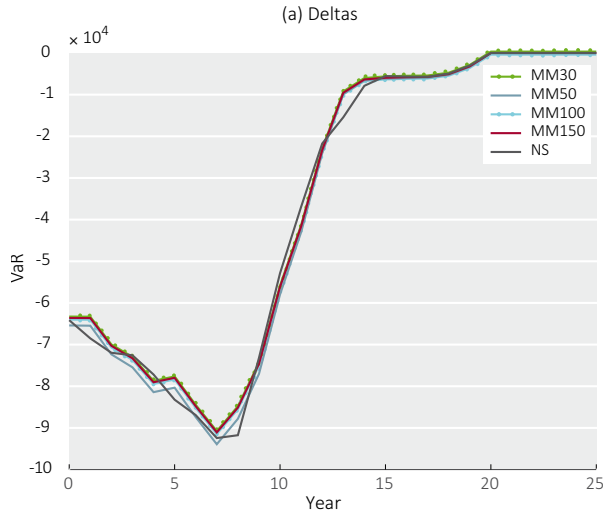
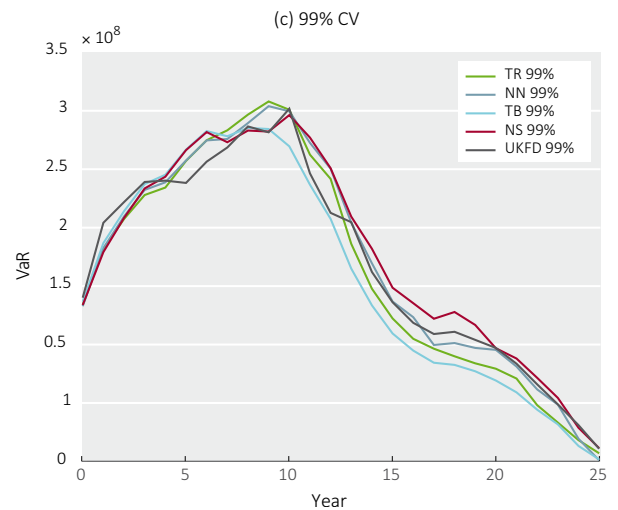
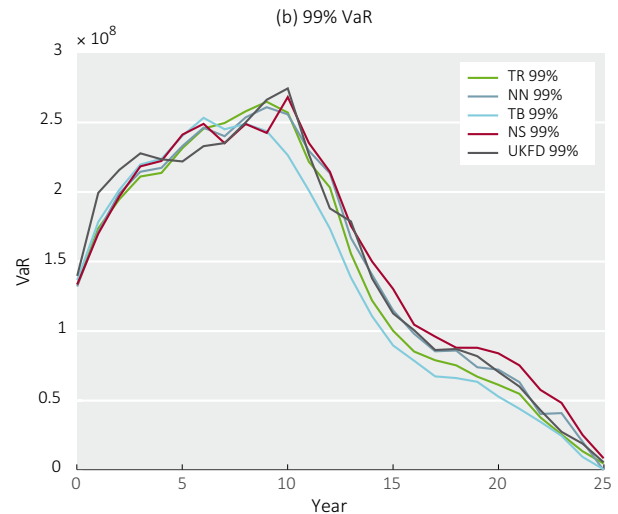
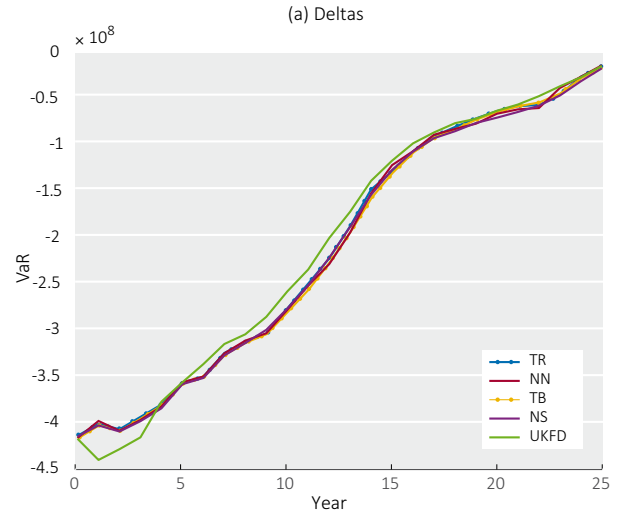


Figure 2. Computed Deltas, 99% VaRs and 99% CVaRs for a portfolio with 10,000 VA contract from the proposed machine learning approach based on Tree Regression (TR), Neural network (NN), the nest simulations (NS) and Unified Kriging for Function Data (UKFD)



FOOTNOTE

1. Waterloo Research Institute in Insurance, Securities, and Quantitative Finance, University of Waterloo (WatRISQ)
2. Cayuga Research Inc., Waterloo, On, Canada.

REFERENCE

1. Wei Xu and Yufang Yin, Pricing American Options by Willow Tree Method under Jump-Diffusion Process, Journal of Derivatives, Vol. 22, No. 1, 46-56, 2014
2. Wei Xu, Yuehuan Chen, Conrad Coleman, Thomas F. Coleman, Efficient Machine Learning Methods for Risk Management of Large Variable Annuity Portfolios, Technical Report, WatRISQ, 2015.
3. LIMRA, U.S. Individual Annuity Sales 2015
4. Willis Towers Watson, UK sales of enhanced annuities: Sales Survey

This article is supported by the Global Risk Institute, Toronto, Canada

.....