Variable annuity portfolio valuation with Shapley additive explanations

Gayani Thalagoda, Katja Hanewald, Andrés Villegas, Jonathan Ziveyi
School of Risk And Acuarial Studies and CEPAR, UNSW Sydney

6^{th} Insurance Data Science Conference
Stockholm, Sweden.
17-18 June 2024

Outline

Motivation

- Variable annuities are attractive retirement income products: equity participation & downside protection.
- Embedded financial guarantees expose VA insurer's liabilities to significant market risks.
- Regulators encourage dynamic hedging (NAIC, USA, APRA, Australia, IFRS17).
- Monte Carlo valuations alone aren't enough.
- Speeding up portfolio valuations has become a key concern for VA insurers.

Under certain conditions, the regulator permits group-level modelling.

- “Grouping of contracts shall be reflective of the quantity being measured.”

  AG 43/VM-21: Requirements for PBR of VA

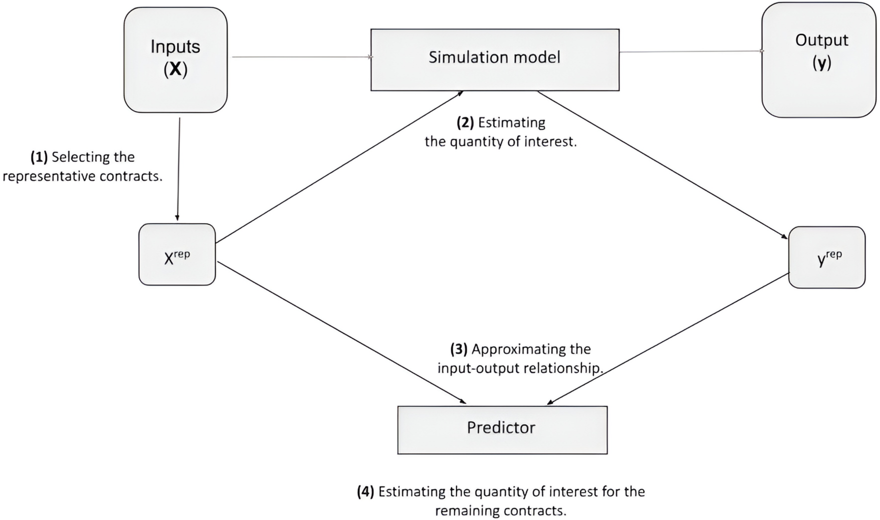
- “Are policies grouped correctly for the modelling purpose?”

  AAA: Principle-Based Reserves Checklist

- “Where an insurer does not have reasonable information to group contracts into sets of contracts, calculations must be performed on an individual contract-by-contract basis.”

  APS 117: Capital Adequacy Requirements

Existing metamodelling approach



Existing metamodelling approach: Strengths and limitations

Strengths

- Computationally efficient
- Scalable
- Accurate for homogeneous portfolios

Limitations

- Sampling is not reflective of the modelling task
- Poor explainability
- Inefficient sampling

Our work

“How to select the representative sample relative to the modelling task in an explainable manner?”

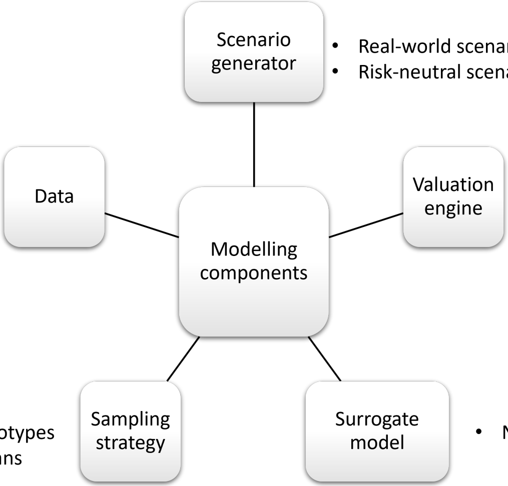
We study:

1. How to use existing knowledge to select the representative sample.
2. How to reformulate learned knowledge in an informative and explainable manner.

Modelling contribution:

1. New metamodelling framework that can select representative samples in an explainable manner.
2. Framework allows grouping contracts to suit the modelling task.
3. Framework de-noises the undue influence of categorical variables during sampling.

Model components



- Real-world scenarios (Hardy, 2001)
- Risk-neutral scenarios (Carmona, 2006)

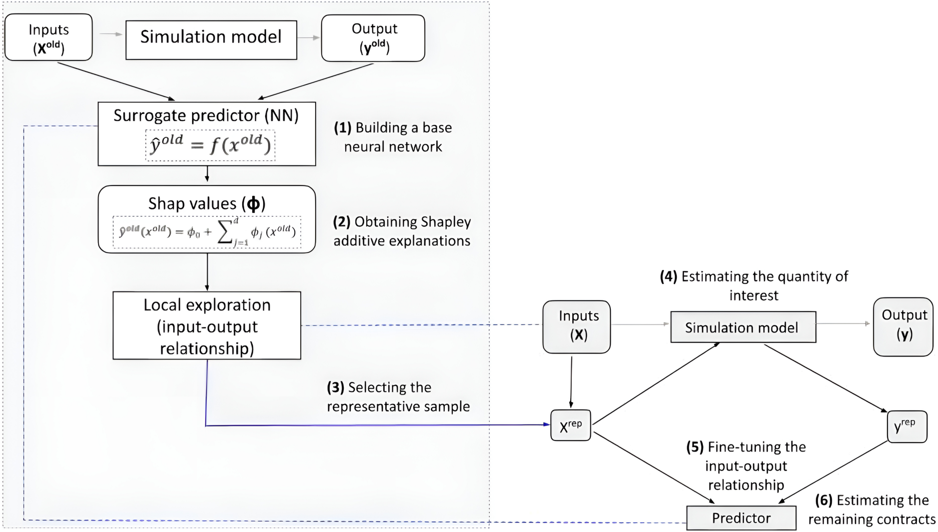
- Synthetic portfolio (Gan & Valdez, 2018)

- Portfolio of VA contracts
- Fund mapping
- Cash flow projection (benefits, fees)
- Mortality tables

- k-prototypes
- hkmeans

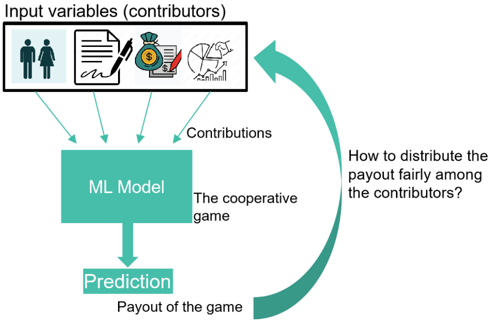
- Neural networks

Proposed metamodelling approach



Shapley framework

- Introduced by Lloyd Shapley in 1953.
- Assigning payouts to the players based on their contribution to the total payout.



Game theory	**Model explanation**
Game	Prediction task
Players	Input features
Payout	Actual prediction

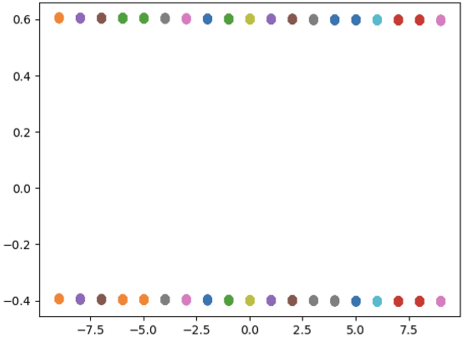
Dataset

- Synthetic dataset of variable annuities constructed by Gan and Valdez (2018)
- 38,000 synthetic VA contracts described by 34 features

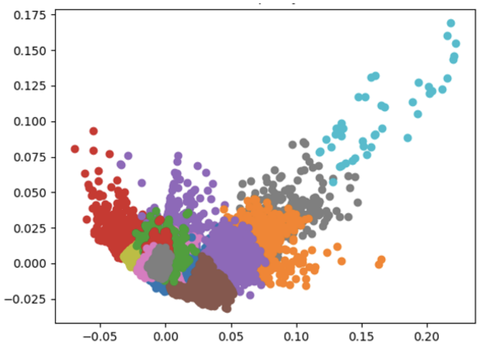
Variable	Description
Gender	Gender of the policyholder
Age	Age of the policyholder
Product Type	Product type of the VA policy
GMWB Balance	Guaranteed minimum withdrawal benefit (GMWB) balance
GB Amount	Guaranteed benefit amount
Fund Value i	Account value of the i^{th} fund, for $i = 1, 2, ..., n$
Time to Maturity	Time to maturity in years

Source: https://www2.math.uconn.edu/~gan/software.html

Cluster formation

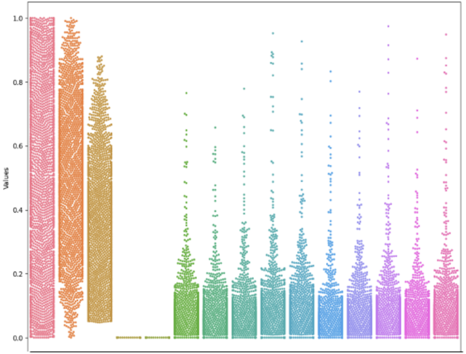


Existing method

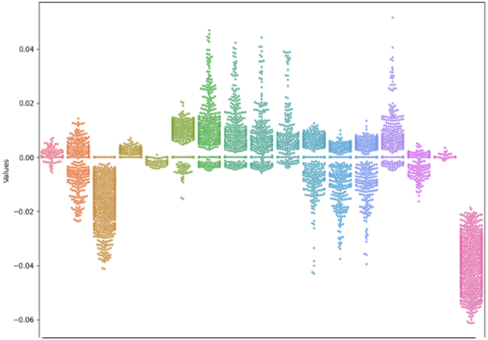


Proposed method

Cluster explainability



Existing method



Proposed method

Prediction accuracy

Portfolio level (PE)

PE	380	760	950	1500
Proposed	1.13%	0.78%	0.45%	0.35%
Existing	5.88%	4.65%	3.53%	1.74%

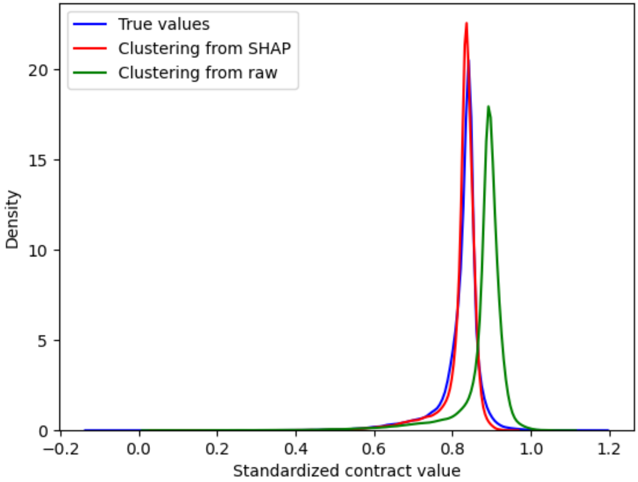
Contract level (RMSE)

RMSE	380	760	950	1500
Proposed	18504.120	17722.978	17596.437	17366.742
Existing	28450.483	22892.192	22031.004	18150.030

Goodness of fit (R^2)

R-Squared	380	760	950	1500
Proposed	0.765	0.768	0.773	0.773
Existing	0.497	0.529	0.590	0.606

Approximating the portfolio behaviour



Conclusion

- We study the practical challenges of the existing metamodelling framework.
- We propose a framework that selects contracts to better reflect the modelling task.
- We show how using existing information improves improves the explainability at the sampling step.
- We show how considering the feature importance can overcome the inefficient sampling.
- The framework can assist users in selecting the representative sample in an informative and explainable manner.

Thank you

Gayani Thalagoda
g.thalagoda@student.unsw.edu

References

[1] Godin, F., Hamel, E., Gaillardetz, P., & Ng, E. (2023).Risk allocation through shapley decompositions, with applications to variable annuities. *ASTIN Bulletin: The Journal of the IAA*, *53*(2), 311–331.

[2] Gan, G., & Valdez, E. A. (2018).Nested stochastic valuation of large variable annuity portfolios: Monte carlo simulation and synthetic datasets. *Data*, *3*(3), 31.

[3] Lundberg, S. M., & Lee, S. (2017).A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

[4] Lin, X., & Yang, S. (2020).Efficient dynamic hedging for large variable annuity portfolios with multiple underlying assets. *ASTIN Bulletin: The Journal of the IAA*, *50*(3), 913–957.

[5] Feng, R., Goujun, G., & Zhang, N. (2022).Variable annuity pricing, valuation, and risk management: A survey. *Scandinavian Actuarial Journal*, *2022*(10), 867–900.

KernelSHAP

Lundberg et al. (2017) showed that if we define an explanation model for an instance z' ;

$$g(z') = \psi_0 + \sum_{j=1}^{D} \psi_j z'_j \ ,$$

the estimated coefficients of the model ψ_j's are the Shapley values.

We obtain the values by training the linear model g with loss function L:

$$L(f, g, \pi_x) = \sum_{z' \in Z} [f(h_x(z')) - g(z')]^2 \pi_x(z')$$

where

$$\pi_x(z') = \frac{(D-1)}{\binom{D}{|z'|} |z'| (D - |z'|)}.$$