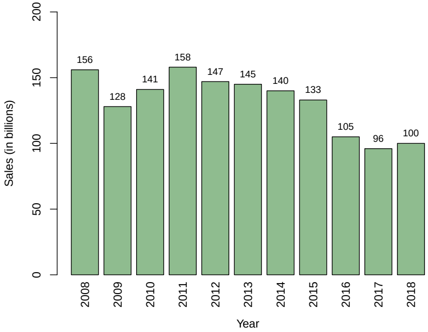
Metamodels and the valuation of large variable annuity portfolios

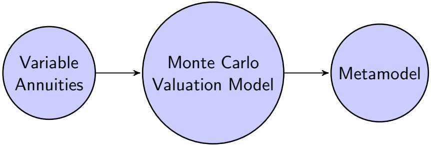
Emiliano A. Valdez, PhD, FSA
University of Connecticut

2nd Financial Mathematics Team Challenge (FMTC)
Fundaçcão Getúlio Vargas (FGV)
July 23 - August 3, 2019

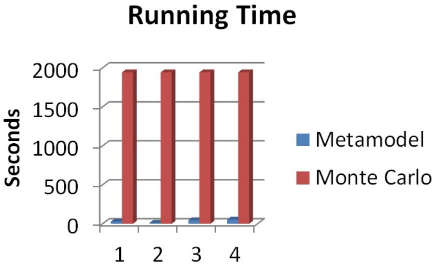
Efficient valuation of large variable annuity portfolios



1. A challenge



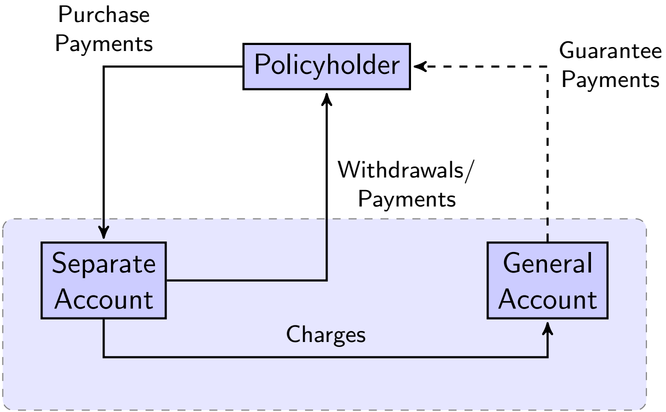
2. A metamodeling approach



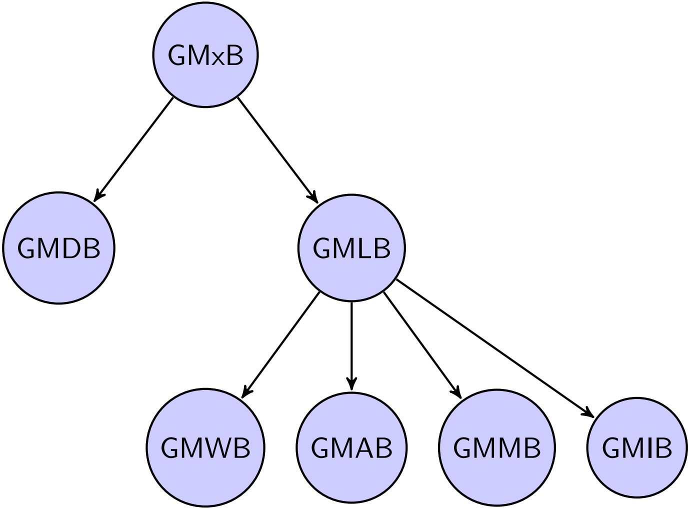
3. Numerical results

What is a variable annuity?

A variable annuity is a retirement product, offered by an insurance company, that gives you the option to select from a variety of investment funds and then pays you retirement income, the amount of which will depend on the investment performance of funds you choose.



Variable annuities come with guarantees



Insurance companies have to make guarantee payments under bad market conditions

Example (An immediate variable annuity with GMWB)

- Total investment and initial benefits base: $100,000
- Maximum annual withdrawal: $8,000

Policy Year	INV Return	Fund Before WD	Annual WD	Fund After WD	Remaining Benefit	Guarantee CF
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	r	0	8,000	0	36,000	8,000
⋮	⋮	⋮	⋮	⋮	⋮	⋮
12	r	0	8,000	0	4,000	8,000
13	r	0	4,000	0	0	4,000

Dynamic hedging

Dynamic hedging is a popular approach to mitigate the financial risk, but

- Dynamic hedging requires calculating the dollar Deltas of a portfolio of variable annuity policies within a short time interval.
- The value of the guarantees cannot be determined by closed-form formula.
- The Monte Carlo simulation model is time-consuming.

There is also the additional computational issue related to reflect the effect of dynamic hedging in (quarterly) financial reporting.

Use of Monte Carlo method

Using the Monte Carlo method to value large variable annuity portfolios is time-consuming:

Example (Valuing a portfolio of 100,000 policies)

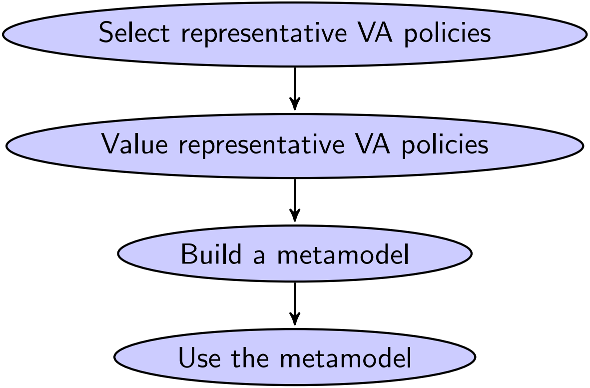
- 1,000 risk neutral scenarios
- 360 monthly time steps

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

$$\boxed{\frac{3.6 \times 10^{10} \text{ projections}}{200,000 \text{ projections/second}} = 50 \text{ hours!}}$$

Metamodeling

- A metamodel, also a surrogate model, is a model of another model.
- Metamodeling has been applied to address the computational problems arising from valuation of variable annuity portfolios: a number of work published by co-author G. Gan.
- It involves four steps:



Selecting representative policies

An important step in the metamodeling process is the selection of representative policies. Gan and Valdez (2016) compared five different experimental design methods for the GB2 regression model:

- Random sampling
- Low-discrepancy sequence
- Data clustering (hierarchical k-means)
- Latin hypercube sampling
- Conditional Latin hypercube sampling

Some metamodels proposed/examined

We have studied and proposed some metamodels for the valuation of large VA portfolios:

- Ordinary kriging
- Universal kriging
- GB2 regression model
- Rank-order kriging (quantile kriging)
- Linear models with interactions
- Tree-based models

Kriging has its origins in geostatistics or spatial analysis. It is in some sense an interpolation method that is closely related to the idea of regression.

Some publications on metamodeling approaches

Publication	Experimental Design	Metamodel
Gan (2013)	Clustering	Kriging
Gan and Lin (2015)	Clustering	Kriging
Gan (2015)	LHS	Kriging
Hejazi and Jackson (2016)	Uniform sampling	Neural network
Gan and Valdez (2016)	Clustering, LHS	GB2 regression
Gan and Valdez (2017)	Clustering	Gamma regression
Gan and Lin (2017)	LHS, conditional LHS	Kriging
Hejazi eta al. (2017)	Uniform sampling	Kriging, IDW, RBF
Gan and Huang (2017)	Clustering	Kriging
Xu et al. (2018)	Random sampling	Neural network, regression trees
Gan and Valdez (2018)	Clustering	GB2 regression
Quan, Gan and Valdez (2019)	Clustering	Regression trees

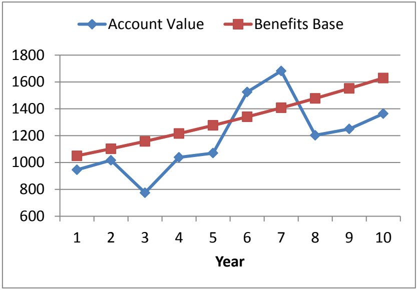
A portfolio of synthetic variable annuity policies

Feature	Value
Policyholder birth date	[1/1/1950, 1/1/1980]
Issue date	[1/1/2000, 1/1/2014]
Valuation date	1/1/2014
Maturity	[15, 30] years
Account value	[50000, 500000]
Female percent	40%
Product type	DBRP, DBRU, BBSU, etc.
Fund fee	30, 50, 60, 80, 10, 38, 45, 55, 57, 46bps for Funds 1 to 10, respectively
Base fee	200 bps
Rider fee	depends on product type
Number of funds invested	[1, 10]

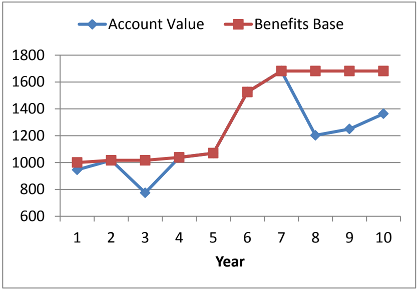
VA product types in the synthetic portfolio

Product	Description	Rider Fee
DBRP	GMDB with return of premium	0.25%
DBRU	GMDB with annual roll-up	0.35%
DBSU	GMDB with annual ratchet	0.35%
ABRP	GMAB with return of premium	0.50%
ABRU	GMAB with annual roll-up	0.60%
ABSU	GMAB with annual ratchet	0.60%
IBRP	GMIB with return of premium	0.60%
IBRU	GMIB with annual roll-up	0.70%
IBSU	GMIB with annual ratchet	0.70%
MBRP	GMMB with return of premium	0.50%
MBRU	GMMB with annual roll-up	0.60%
MBSU	GMMB with annual ratchet	0.60%
WBRP	GMWB with return of premium	0.65%
WBRU	GMWB with annual roll-up	0.75%
WBSU	GMWB with annual ratchet	0.75%
DBAB	GMDB + GMAB with annual ratchet	0.75%
DBIB	GMDB + GMIB with annual ratchet	0.85%
DBMB	GMDB + GMMB with annual ratchet	0.75%
DBWB	GMDB + GMWB with annual ratchet	0.90%

VA provides guaranteed appreciation of the benefits base

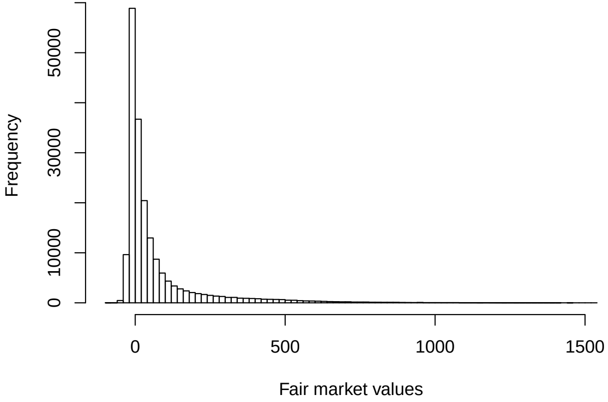


(Roll-up)



(Ratchet)

Fair market values of the guarantees



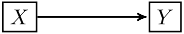
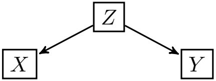
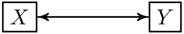
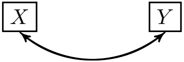
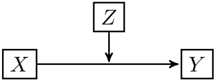
	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
fmv	-68.37	-5.55	64.63	11.7	64.84	1210.32

Training set - summary statistics - continuous variables

Response variables	Description	Min.	1st Q	Mean	Median	3rd Q	Max.
gmwbBalance	GMWB balance	0	0	27.8	0	0	422.26
gbAmt	Guaranteed benefit amount	51.88	183.98	323.29	306.89	437.36	920.62
FundValue1	Account value of the 1st fund	0	0	32.02	12.62	46.76	629.89
FundValue2	Account value of the 2nd fund	0	0	36.54	16.08	56.31	571.59
FundValue3	Account value of the 3rd fund	0	0	26.78	11.81	36.64	458.78
FundValue4	Account value of the 4th fund	0	0	25.8	10.48	38.29	539.36
FundValue5	Account value of the 5th fund	0	0	22.29	10.54	34.71	425.92
FundValue6	Account value of the 6th fund	0	0	37.15	19.64	53.96	654.64
FundValue7	Account value of the 7th fund	0	0	28.78	12.88	42.56	546.89
FundValue8	Account value of the 8th fund	0	0	31.27	15.59	46.24	529.57
FundValue9	Account value of the 9th fund	0	0	31.93	13.9	45.17	599.44
FundValue10	Account value of the 10th fund	0	0	32.6	13.86	45.09	510.43
age	Age of the policyholder	34.52	42.86	50.29	51.36	57.21	64.46
ttm	Time to maturity in years	0.75	10.09	14.61	14.6	19.12	27.52

Learning interactions

Six basic types of causal relationships:

Relationship	Example
Direct causal relationship	
Indirect causal relationship	
Spurious relationship	
Bidirectional causal relationship	
Unanalyzed relationship	
Moderated causal relationship	

Description of the model

Let Y be a continuous response variable. Let X_1, X_2, ..., X_p be p explanatory variables, which include continuous and categorical variables.

Then the first-order interaction model is given by:

$$E[Y|X_1, X_2, \ldots, X_p] = \beta_0 + \sum_{j=1}^{p} \beta_j X_j + \sum_{s<t} \beta_{s:t} X_{s:t},$$

where the term $X_{s:t} = X_s X_t$ denotes the interaction effect between X_s and X_t and terms X_1, X_2, ..., X_p denote the main effects.

- Main effects can be viewed as deviations from the global mean and interaction effects can be viewed as deviations from the main effects.
- As a result, it rarely makes sense to have interactions without main effects.
- This means that hierarchical interaction models are usually preferred.

Lasso

Let $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ denote the vector of responses and let $\mathbf{X}$ denote the design matrix.

Then the lasso is formulated as the following optimization problem:

$$\widehat{\boldsymbol{\beta}}^{LASSO} = \arg\min_{\boldsymbol{\beta}} \left(\frac{1}{2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \right),$$

where $\| \cdot \|_2$ denotes the ℓ^2-norm, $\| \cdot \|_1$ denotes the ℓ^1-norm, and λ is a tuning parameter that controls the amount of regularization.

The ℓ^1-norm has a nice property that it induces sparsity in the solution by forcing some coefficients to zero.

A larger value of λ implies more regularization, i.e., more coefficients will be zero.

The lasso is designed for selecting individual explanatory variables but not for factor variables.

Group-lasso

The group-lasso was proposed to select important factor variables.

Suppose that there are p groups of explanatory variables. For $j = 1, 2, \ldots, p$, let $\mathbf{X}_j$ denote the feature matrix for group j.

The group-lasso is formulated as follows:

$$\widehat{\boldsymbol{\beta}}^{GLASSO} = \arg\min_{\boldsymbol{\beta}} \left(\frac{1}{2} \|\mathbf{y} - \beta_0 \mathbf{1} - \sum_{j=1}^{p} \mathbf{X}_j \boldsymbol{\beta}_j \|_2^2 + \lambda \sum_{j=1}^{p} \gamma_j \|\boldsymbol{\beta}_j\|_2 \right),$$

where $\mathbf{1}$ is a vector of ones and $\lambda, \gamma_1, \ldots, \gamma_p$ are tuning parameters.

When each group contains one continuous variable, the group-lasso reduces to the lasso.

The group-lasso has an attractive property that if $\widehat{\boldsymbol{\beta}}_j$ is nonzero, then all its components are typically nonzero.

Overlapped group-lasso

The group-lasso described above may not lead to hierarchical interaction models. In order to obtain hierarchical interaction models, one way is to add an overlapped group-lasso penalty to the loss function.

The resulting method is called the overlapped group-lasso, which is formulated as the following constrained optimization problem:

$$
\begin{aligned}
&\widehat{\boldsymbol{\beta}}^{OGLASSO} \\
= \; & \arg\min_{\boldsymbol{\beta}} \left(\frac{1}{2} \left\| \mathbf{y} - \beta_0 \mathbf{1} - \sum_{j=1}^{p} \mathbf{X}_j \boldsymbol{\beta}_j - \sum_{s<t} (\mathbf{X}_s \tilde{\boldsymbol{\beta}}_s + \mathbf{X}_t \tilde{\boldsymbol{\beta}}_t + \mathbf{X}_{s:t} \boldsymbol{\beta}_{s:t}) \right\|_2^2 \right. \\
& \left. + \lambda \left(\sum_{j=1}^{p} \|\boldsymbol{\beta}_j\|_2 + \sum_{s<t} \sqrt{L_s \|\tilde{\boldsymbol{\beta}}_s\|_2^2 + L_t \|\tilde{\boldsymbol{\beta}}_t\|_2^2 + \|\boldsymbol{\beta}_{s:t}\|_2^2} \right) \right),
\end{aligned}
$$

- continued

subject to the following sets of constraints:

- if X_j is categorical,

$$\sum_{l=1}^{m_j} \beta_j^{(l)} = 0, \quad \sum_{l=1}^{m_j} \tilde{\beta}_j^{(l)} = 0,$$

- if X_j is categorical and X_t is continuous,

$$\sum_{l=1}^{m_j} \beta_{t:j}^{(l)} = 0,$$

- if X_j and X_t are both categorical,

$$\sum_{l=1}^{m_j} \beta_{t:j}^{(l,k)} = 0 \quad \forall k, \quad \sum_{k=1}^{m_t} \beta_{t:j}^{(l,k)} = 0 \quad \forall l.$$

- continued

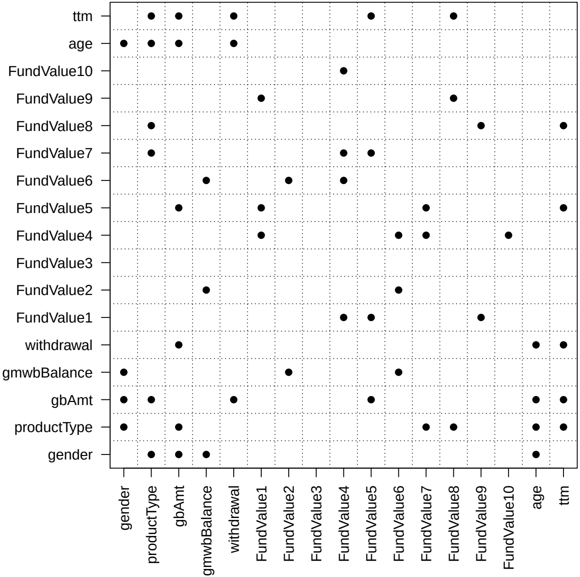
Here:

- m_j is the number of levels of X_j, m_t is the number of levels of X_t,
- $\beta_j^{(l)}$ is the lth entry of $\boldsymbol{\beta}_j$,
- $\tilde{\beta}_j^{(l)}$ is the lth entry of $\tilde{\boldsymbol{\beta}}_j$, and
- $\beta_{t:j}^{(l,k)}$ is the lkth entry of $\boldsymbol{\beta}_{t:j}$.

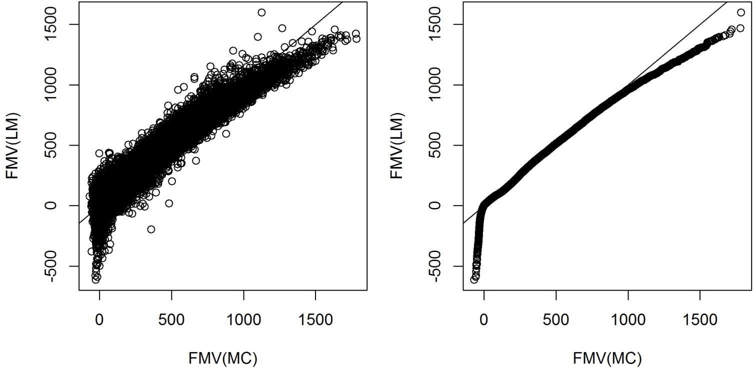
In addition, $\mathbf{X}_1$, $\mathbf{X}_2$, ..., $\mathbf{X}_p$ denote the feature matrices of the p group of variables.

- If X_j is continuous, then $\mathbf{X}_j$ is just a one-column matrix containing the values of X_j.
- If X_j is categorical, then $\mathbf{X}_j$ contains all the dummy variables associated with X_j.
- The matrix $\mathbf{X}_{s:t}$ denotes the feature matrix of the interaction term.

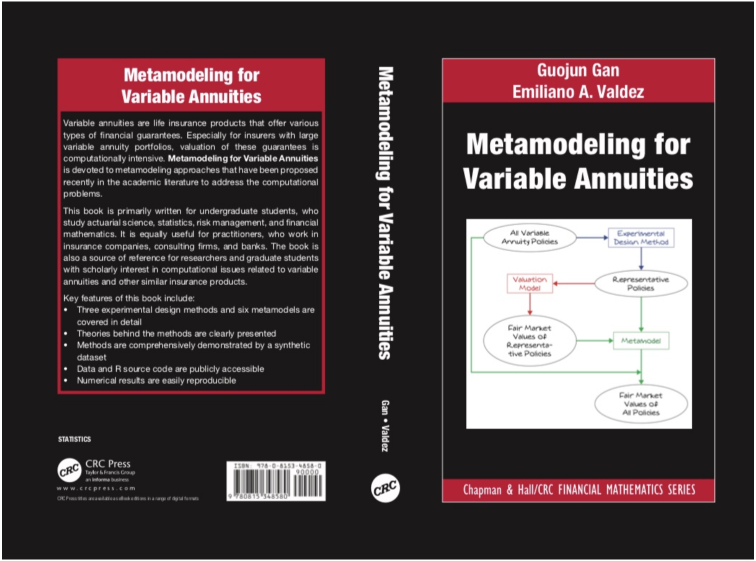
Pairwise interactions



Scatterplot and QQ plot



Metamodeling book



Synthetic data set can be found in:

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

There are two files containing the representative policies:

`hkmeans.csv`

`hkmeans2.csv`

Description of variables in the synthetic portfolio

The synthetic portfolio contains 190,000 VA policies, with 10,000 for each of the guarantee types. There are a total 45 variables, includig 10 fund values, 10 fund numbers, and 10 fund fees.

Field	Description
`recordID`	Unique identifier of the policy
`survivorShip`	Positive weighting number
`gender`	Gender of the policyholder
`productType`	Product type
`issueDate`	Issue date
`matDate`	Maturity date
`birthDate`	Birth date of the policyholder
`currentDate`	Current date
`baseFee`	M&E (Mortality & Expense) fee
`riderFee`	Rider fee
`rollUpRate`	Roll-up rate
`gbAmt`	Guaranteed benefit
`gmwbBalance`	GMWB balance
`wbWithdrawalRate`	Guaranteed withdrawal rate
`withdrawal`	Withdrawal so far
`FundValue`i	Fund value of the ith investment fund
`FundNum`i	Fund number of the ith investment fund
`FundFee`i	Fund management fee of the ith investment fund

Appendix: Validation measures

Validation measure	Description	Interpretation
Gini Index	$Gini = 1 - \frac{2}{N-1}\left(N - \frac{\sum_{i=1}^{N} i\tilde{y}_i}{\sum_{i=1}^{N}\tilde{y}_i}\right)$ where $\tilde{y}$ is the corresponding to y after ranking the corresponding predicted values $\widehat{y}$.	Higher Gini is better.
Coefficient of Determination	$R^2 = 1 - \frac{\sum_{i=1}^{N}(\widehat{y}_i - y_i)^2}{\sum_{i=1}^{N}\left(y_i - \frac{1}{n}\sum_{i=1}^{n} y_i\right)^2}$ where $\widehat{y}$ is predicted values.	Higher R^2 is better.
Concordance Correlation Coefficient	$CCC = \frac{2\rho\sigma_{\widehat{y}_i}\sigma_{y_i}}{\sigma^2_{\widehat{y}_i} + \sigma^2_{y_i} + (\mu_{\widehat{y}_i} - \mu_{y_i})^2}$ where $\mu_{\widehat{y}_i}$ and μ_{y_i} are the means $\sigma^2_{\widehat{y}_i}$ and $\sigma^2_{y_i}$ are the variances ρ is the correlation coefficient	Higher CCC is better.
Mean Error	$ME = \frac{1}{N}\sum_{i=1}^{N}(\widehat{y}_i - y_i)$	Lower $\lvert ME \rvert$ is better.
Percentage Error	$PE = \frac{\sum_{i=1}^{N}\widehat{y}_i - \sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} y_i}$	Lower $\lvert PE \rvert$ is better.
Mean Squared Error	$MSE = \frac{1}{N}\sum_{i=1}^{N}(\widehat{y}_i - y_i)^2$	Lower MSE is better
Mean Absolute Error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert\widehat{y}_i - y_i\rvert$	Lower MAE is better.

References

Gan, G. and Valdez, E.A. (2019). *Metamodeling for Variable Annuities*. CRC Press: Boca Raton, FL.

Gan, G. and Valdez, E.A. (2017). Valuation of large variable annuity portfolios: Monte Carlo simulation and synthetic datasets. *Dependence Modeling*. 5:354-374.

Gan, G. and Valdez, E.A. (2018). Regression modeling for the valuation of large variable annuity portfolios. *North American Actuarial Journal*. 22(1):40-54.

Quan, Z., Gan, G. and Valdez, E.A. (2019). Tree-based models for variable annuity valuation: Parameter tuning and empirical analysis. Submitted for publication.