Department:
Group Risk Management

Country:
Head Office

Internal Model Approach for Risk Management

8th European Congress of Mathematics
Portorož, June 21st, 2021

Johannes Schoenenwald

Vincenzo Russo

Department:
Group Risk Management

Country:
Head Office

Internal Model Approach for Risk Management - part I

8th European Congress of Mathematics
Portorož, June 21st, 2021

Johannes Schoenenwald

Agenda

Internal Model Approach for Risk Management

- Introduction and background
- An overview of the Internal Model framework
- Deep dive: life liability loss modelling and the LSMC method

Introduction

In the context of risk management needs in the insurance industry, **mathematical and statistical methods are widely used to evaluate, measure and manage risks** that may arise from the insurance business

In particular, since the Solvency II directive entered in force, insurance companies have been entitled to **develop** and implement proprietary **Internal Models for the evaluation and monitoring of their risk profile**

Key objectives are to:

- **Provide an overview of the framework around the development and implementation of an Internal Model**
- **Present some of the specific methodological and technical aspects that characterize the design of the Generali Internal Model**

Background

Solvency II

Solvency II is a harmonized risk-based supervisory regime applied across the European Union since 2016. It covers various aspects of Risk Management, requires insurers to measure, report and monitor risks and is primarily focused on the **amount of capital that EU insurance companies must hold to reduce the risk of insolvency (Solvency Capital Requirement)**

Solvency Capital Requirement

The Solvency Capital Requirement (SCR) is described as the **amount of own funds that ensures the absorption of losses in a one-year time horizon with a confidence level of 99.5%**, i.e.

$$SCR = VaR_{0.995}(L) = \underset{x}{\operatorname{argmin}}\{P(L > x) \leq 0.005\}$$

with $L = C_0 - C_1 (1+i)^{-1}$, where C_0 and C_1 represent the available capital at times 0 and 1 and i is the risk-free rate on [0,1].

As C_0 and i can be considered as known, the stochastic component of the loss L is C_1

Agenda

Internal Model Approach for Risk Management

- Introduction and background
- An overview of the Internal Model framework
- Deep dive: life liability loss modelling and the LSMC method

Internal Model Framework
What is an Internal Model?

An **Internal Model** is developed in order to evaluate and monitor the risk profile of an insurance company through a **detailed statistical description of the possible future movements in its own funds**

The requirements

- Be theoretically sound, appropriately **justified** and use **reliable data**
- Enable the calculation of the **full profit and loss distribution**
- Be **widely used** in the company's system of governance
- Be **fully documented** such that a third party can replicate all the results

The challenges

- Obtain a **precise yet not overly complex representation** of the quantifiable risks to which the company is exposed
- Achieve a **forward-looking model** when using historical data
- Manage complex calculations in a **computationally efficient** manner

The objectives

- Be fully **integrated** into company **processes** and business **decisions**
- Improve the **understanding** of the business and its risk profile
- Allow for **fast and transparent communication** of results internally and externally
- Assess the **impact of** different risks or **strategic decisions**

The outcome

- Produce an estimation of the full multivariate statistical **distribution of one-year losses**
- Allow for **optimization of risk profile** within the limits of the risk appetite, as well as for optimization of strategic asset allocation, asset liability management and risk mitigation strategies

GENERALI

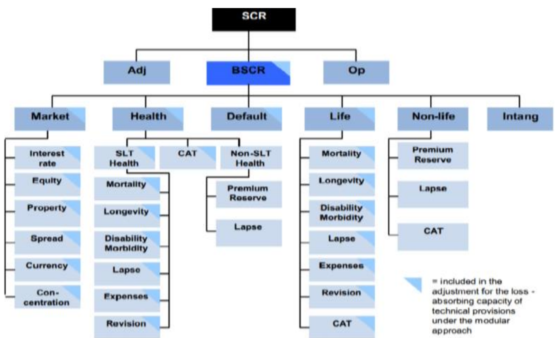
Internal Model Framework

Why is an Internal Model needed?

The development of an Internal Model is not mandatory, in fact the regulation provides a **Standard Formula** which all companies can apply to evaluate the SCR. This formula enables a simplified calculation of capital requirements but does not allow a company to have a full understanding of its risk profile

Standard Formula

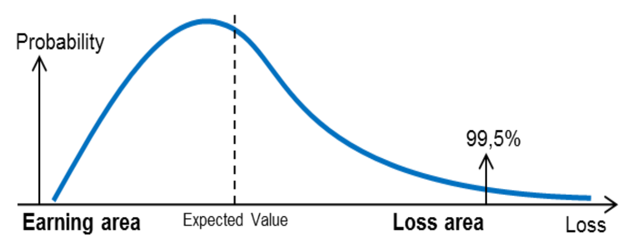
$$SCR = \sqrt{\sum_{i,j} SCR_i SCR_j \rho_{i,j}}$$



Internal Model

$$SCR = \underset{x}{\operatorname{argmin}}\{P(L > x) \leq 0.005\}$$

Probability Distribution Forecast of SII ΔOF

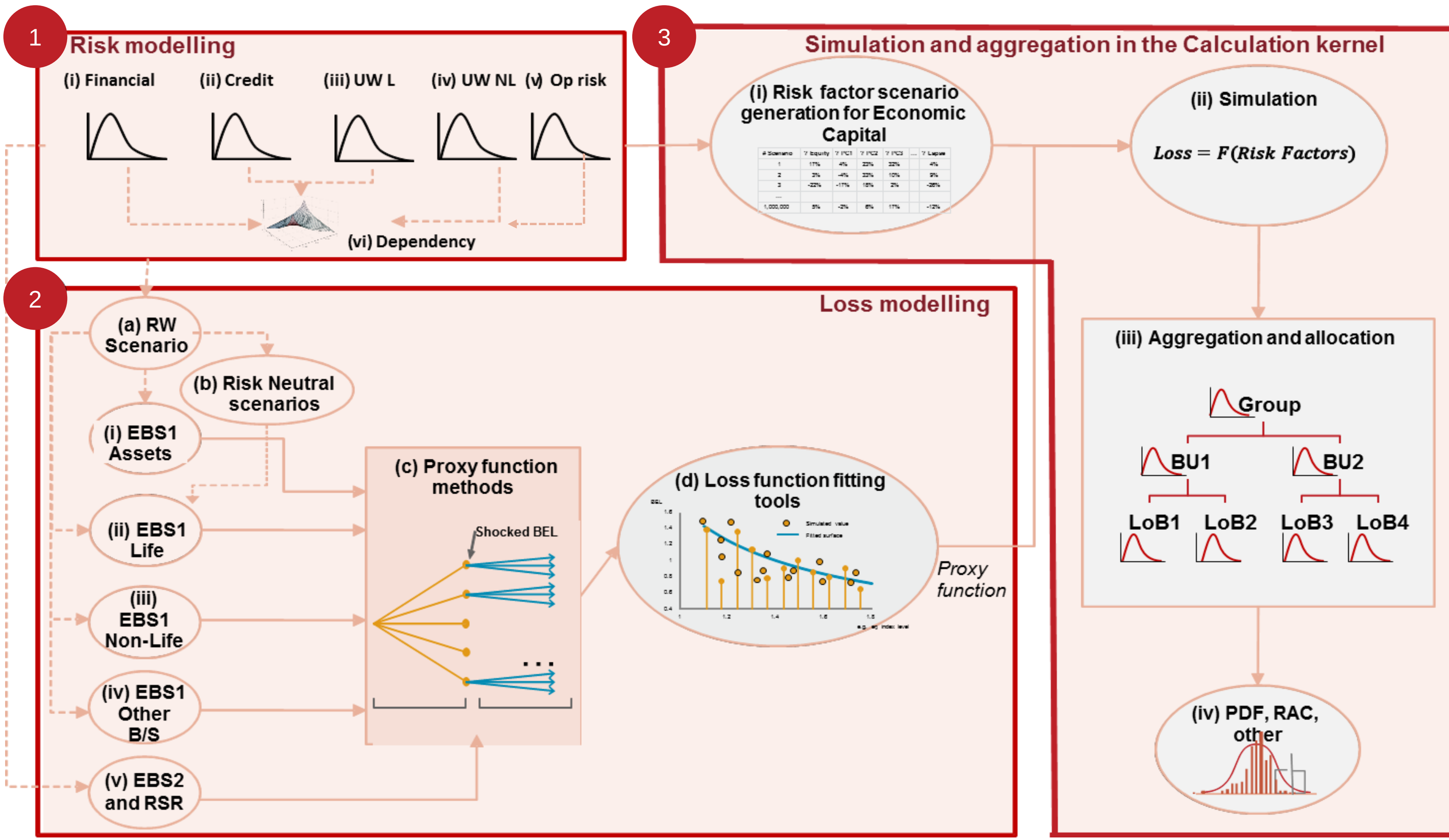


- Based on **deterministic stress scenarios**
- **Linearity** of loss function
- Assumes **joint ellyptical distribution** of risk factors
- **Not able to retrieve a full profit and loss distribution**
- **Easy implementation**

- Based on **stochastic stress scenarios**
- **No constraints** on loss function form
- **Any** analytical or empirical **distribution** is possible for risk factors
- **Outputs a full multivariate distribution** for profit and loss
- **More challenging to implement** in terms of costs and time

Internal Model Framework

How does an Internal Model work?



Risk Modelling

Deriving standalone probability distributions and dependencies

1 Risk factor identification

- Identify the risks that might impact the value of assets or liabilities held by the insurance company (e.g. equity, interest rate, mortality risks)
- Identify the granularity at which each risk should be modelled, i.e. the number of distinct risk factors for which a probability distribution should be derived
- In case it is not feasible to calibrate the required number of risk factors, consider dimensionality reduction techniques

Focus in next slides

2 Data collection and preparation

- Identify internal/external data sources to retrieve historical evolutions of the underlying phenomena
- Perform quality checks on data (e.g. presence of outliers, missing values)
- Enrich or transform data if needed (e.g. smoothing, inter-extrapolation, projection)
- Choose length of data series to use in the model (point-in-time vs through-the-cycle calibrations)
- Choose quantity to be modelled (e.g. absolute/relative variations)

3 Model and dependency calibration

- For each risk factor, calibrate a statistical model which – using the available data – provides as an output either an analytical or an empirical probability distribution representing possible one-year standalone movements in the modelled quantity

Focus in next slides

- Calibrate the correlations between different risk factors
- Model dependency structure between risk factors

Focus in next slides

Risk Modelling

Focus on dimensionality reduction

Many phenomena require lots of distinct risk factors in order to be described adequately. However, introducing lots of risk factors into the model may result in additional complexity in subsequent steps of the process: correlations need to be set between each new risk factor and all others, numerical approaches used in loss modelling and aggregation are less stable, etc.

Example: how to model the risk-free spot interest rate curve with maturity up to 100 years? Theoretically, 100 distinct risk factors would be needed to describe every possible movement in the curve.

Possible solution	Methodological approach	Advantages & Limitations
Volatility anchoring	▪ A distribution is derived only for a single data series ("anchor"), best representing the underlying phenomenon ▪ For each other series, the volatility of the distribution is re-scaled on the basis of the ratio between the volatility of the series and that of the anchor ▪ Correlations are set to 100%, volatility may be reduced ex-post to account for diversification	▪ Simple, easily interpretable and computationally straightforward ▪ Only uses 1 risk factor to model the phenomenon ▪ Assumes same distribution for all series ▪ Assumptions on correlation are very restrictive, reducing volatility ex-post lessens interpretability
Aggregated calibration and allocation of stresses	▪ A distribution is derived only for an aggregated quantity (e.g. the series obtained as the average of all available data series representing the underlying phenomenon) ▪ For each granular series, the volatility of the distribution is re-scaled on the basis of regression analyses, through which the aggregated stress is distributed across granular series ▪ Correlations are set to 100%, but diversification is implicitly embedded into the computation of the aggregated quantity	▪ Computationally straightforward ▪ Only uses 1 risk factor to model the phenomenon ▪ Assumes same distribution for all granular series ▪ Assumptions on correlation are very restrictive, implicitly embedded diversification lessens interpretability ▪ The model does not envisage "standalone" distributions at granular level
Principal Components Analysis	▪ A Principal Components Analysis is performed on all available data series ▪ The first *n* principal components (PCs) are taken as risk factors, where *n* is chosen in order to cover a significant part of the phenomenon's variability ▪ A distribution is derived for each selected PC ▪ By generating simulated realizations of the PCs and by using the eigenvectors, empirical distributions and quantiles can be derived also for the original series	▪ Robust theoretical framework ▪ Slightly more challenging computationally ▪ Might need several risk factors ▪ Correlations need to be set between PCs ▪ More difficult interpretation of risk factors

Risk Modelling

Focus on modelling annual risk factor distributions

The objective of the risk modelling is deriving a probability distribution describing annual variations in each risk factor, with a focus on the extreme tails of the distribution. Using simple annual historical variations might not be feasible due to the unavailability of enough data to perform a reliable fit and/or presence of structural breaks.

Possible solution	Methodological approach	Advantages & Limitations
Modelling annual overlapping variations	▪ Annual overlapping variations are computed on a rolling basis (e.g. if data has a monthly granularity, variations are computed between the 13th observation and the 1st, between the 14th and the 2nd and so on) ▪ The annual overlapping variations are subject to the distribution fitting, which will directly represent one-year movements	▪ Simple and computationally straightforward, no need for additional modelling steps ▪ Uses same amount of data as in the original set minus the length of the rolling window (e.g. 12 observations in case of monthly data) ▪ Introduces very strong synthetic autocorrelation into the modelled data, which might lead to bias in parameter estimates
Modelling granular data and annualising parameters	▪ A distribution fitting is performed on granular data (e.g. monthly) and the autocorrelation function is computed ▪ The parameters of the distribution are annualized, keeping in mind that: $E(X_i) = \mu \longrightarrow E(X_1 + \ldots + X_{12}) = 12\mu$ $V(X_i) = \sigma^2 \longrightarrow V(X_1 + \ldots + X_{12}) = \sigma^2 \sum\sum Cor(X_i, X_j)$	▪ Simple and computationally straightforward, no need for additional modelling steps ▪ Uses same amount of data as in the original set ▪ Distributional form selected for granular data is assumed to work also for annual data, except for the parameters ▪ Only works with elliptical distributions with finite variance
Modelling granular data and performing stochastic annualisation	▪ A distribution fitting is performed on granular data (e.g. monthly) and the autocorrelation function is computed ▪ A great number of joint one-year sequences (e.g. 12 consecutive monthly observations) are simulated from the monthly distribution by means of a copula approach, taking the autocorrelation between the data into account ▪ The corresponding annual variation is derived in each simulation (e.g. by summing up the 12 monthly variations in case of simple or log changes) ▪ A distribution fitting is performed again on annual simulated data	▪ Computationally challenging, requires choice of copula and potentially increases model risk ▪ Uses same amount of data as in the original set ▪ Allows to change distributional form between granular and annual data ▪ Works with every analytical or empirical distribution

GENERALI

Risk Modelling

Focus on modelling the dependency structure

The **joint distribution of risk factors is obtained by means of a copula**, i.e. a function $C: [0,1]^n \rightarrow [0,1]$ such that $P(X_1 \leq x_1, \dots, X_n \leq x_n) = C\big(P(X_1 \leq x_1), \dots, P(X_n \leq x_n)\big)$.

To **simulate from the joint distribution**, a sample $(U_1, \dots, U_n)$ from C is drawn and then $(X_1, \dots, X_n)$ is set equal to $(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$, with $F_i = P(X_i \leq x)$

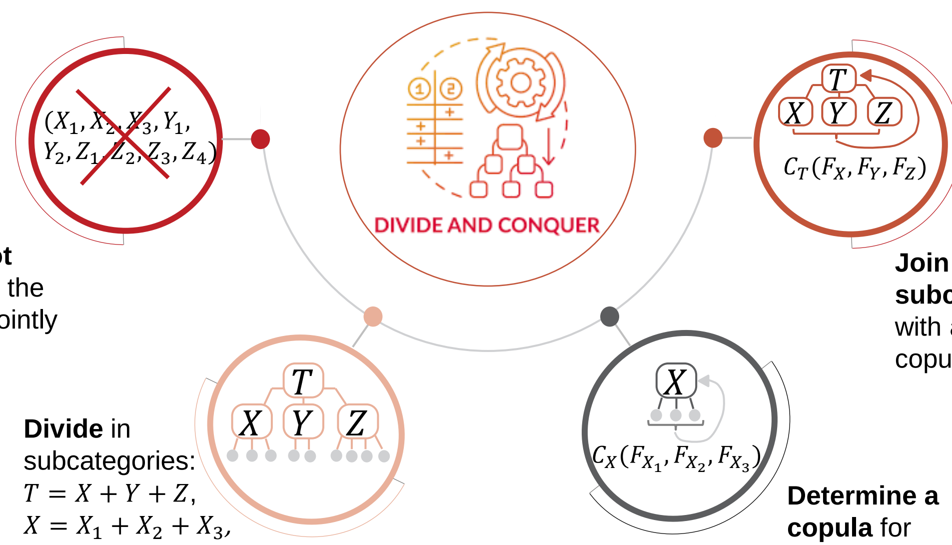
Problem

High dimensional copulae present many downsides:

- Too symmetric dependence structure
- Difficult to calibrate
- Numerically slow simulation
- Hard to justify (to managers, regulators, etc…)

Possible solution: **hierarchical risk aggregation** approach

Suppose we have risk from three categories: X_i, Y_j and Z_h, and let for example the total risk be $T = X_1 + X_2 + X_3 + Y_1 + Y_2 + Z_1 + Z_2 + Z_3 + Z_4$



Do not model the risks jointly

Divide in subcategories:
$T = X + Y + Z$,
$X = X_1 + X_2 + X_3$,
$Y = Y_1 + Y_2$,
$Z = Z_1 + Z_2 + Z_3 + Z_4$

Determine a copula for each subcategory

Join the subcategories with another copula

In the following way **a high dimensional copula is substituted** with many low dimensional ones

Loss Modelling

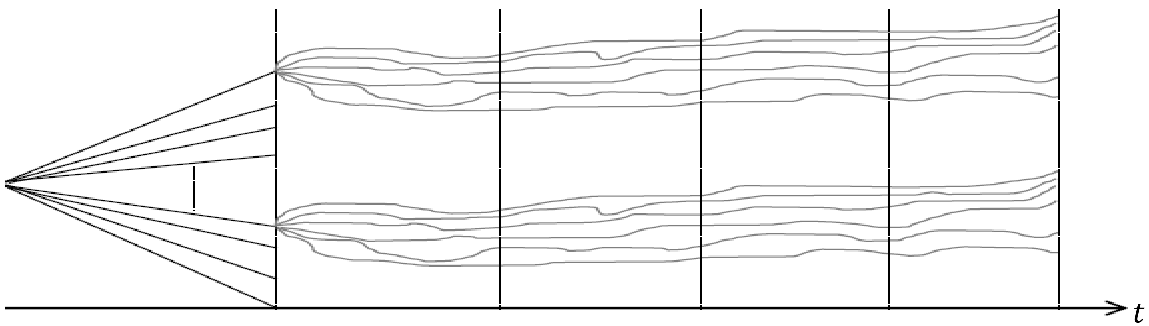
The **impact of the risk factors on the assets and liabilities held by the company** now needs to be determined. In mathematical terms, we need to find a function f such that:

$$L = f(X_1, \dots, X_n)$$

where L denotes the loss the company is facing and $X_1, \dots, X_n$ all risk factors towards which the company is exposed.

This presents many challenges:

- Since $X_1, \dots, X_n$ are random variables, given the complexity of the model, the distribution of the loss L can only be derived via a **Monte Carlo approach**. Hence we need to find f such that $l = f(x_1, \dots, x_n)$ in each scenario $x_1, \dots, x_n$. Given the vastly different behaviour of balance sheet items, the loss also needs to be derived at the most granular level
- Given a generic scenario $x_1, \dots, x_n$, deriving the loss is very complex especially for life liabilities and some derivatives, for which **no closed valuation formulae exist** and asset pricing engines and cash flow projection models often need to perform a Monte Carlo approach as well to derive the loss ("Nested stochastic simulations")

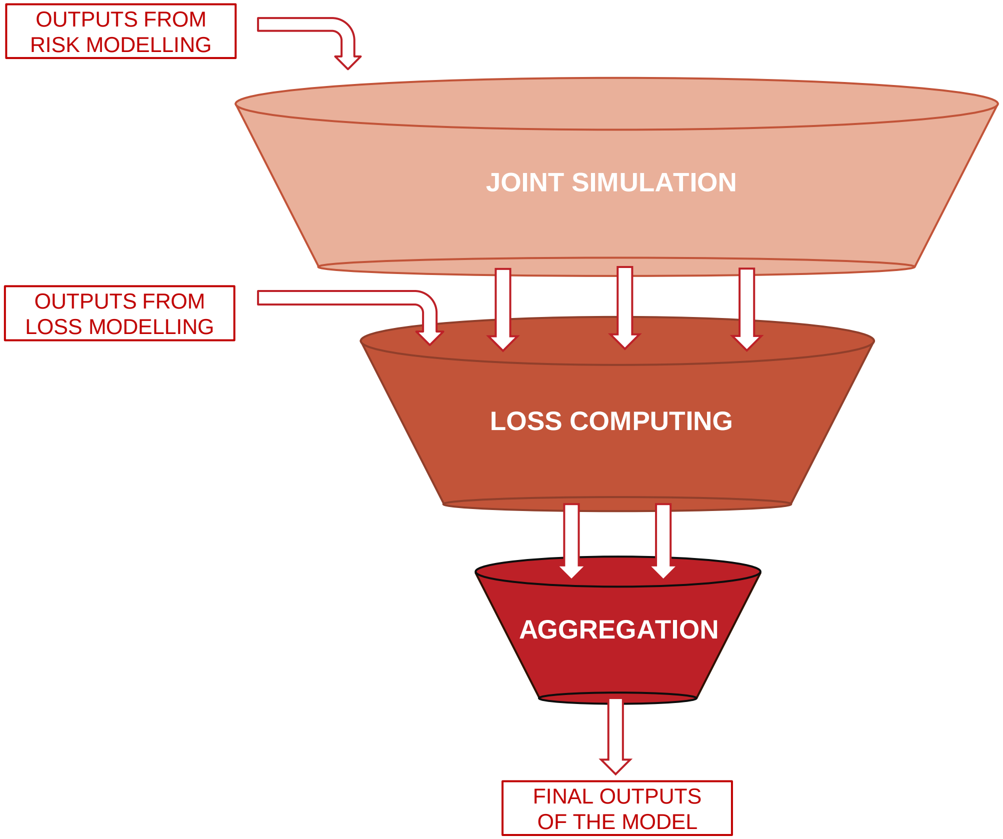


- Nested simulations are computationally not feasible, therefore proxy approaches are needed. These usually are built upon two different approximations of the loss function: first, the value of the loss in each scenario is approximated by a linear combination of basis functions of the risk factors, then the related coefficients are estimated.

$$L \approx \sum_{j=1}^{m} \beta_j f_j(x_1, \dots, x_n) \approx \sum_{j=1}^{m} \hat{\beta}_j f_j(x_1, \dots, x_n)$$

Simulation and Aggregation

The outputs of risk and loss modelling are then used as inputs in the context of **simulation and aggregation**, which is the final step of the process that **leads to the outputs of the Internal Model**



A joint simulation of the movements of the risk factors in a 1-year view is performed, taking advantage of the outputs of risk modelling, i.e. standalone distributions, correlations and copulae

In every simulated scenario, the functions derived through the loss modelling process are used to **obtain a simulation of the loss at the lowest granularity possible**.

The losses are finally **aggregated** at a higher and higher level, in order to obtain a detailed **loss distribution for every line of business, company as well as for the Group as a whole**.

Department:
Group Risk Management

Country:
Head Office

Internal Model Approach for Risk Management - part II

8th European Congress of Mathematics
Portorož, June 21st, 2021

Vincenzo Russo

Agenda

Internal Model Approach for Risk Management

- Introduction and background
- An overview of the Internal Model framework
- Deep dive: life liability loss modelling and the LSMC method

Deep dive: life liability loss modelling and the LSMC method

How do we evaluate the balance sheet items ?

Non-life liabilities and majority of assets	▪ Closed form valuation formulas are available: the balance sheet item value in each real-world (RW) scenario can be easily computed …

Life liabilities

- For life liabilities, a stochastic valuation approach based on cash-flow projection models is needed. This requires a high number of risk neutral (RN) scenarios of risk factor realizations for the whole life span of the run-off portfolio, for each single RW scenario

- Under Solvency II, liabilities in life insurance are valued on the base of the market-consistency principle taking into account:
  - ✓ financial options and guarantees
  - ✓ future management actions, e.g. profit sharing rules
  - ✓ the policyholder's behavior
  - ✓ underwriting risks (i.e. mortality, longevity, lapse, …)
  - ✓ financial risks (i.e. interest rates, equities, credit spreads, …)

- A stochastic Asset Liability Management (ALM) model based on Monte-Carlo balance sheet projection is generally implemented to compute the best estimate of liabilities (BEL) in relation to a specific RW scenario:

$$BEL_t = \mathbb{E}^Q\left[\sum_{u>t} \delta_u CF_u\right]$$

  - ➢ δ_u , the stochastic discount factor (deflator) at time u
  - ➢ CF_u , the net payment cash-flows at time u
  - ➢ $\mathbb{E}^Q$, the expectation is calculated under the “risk-neutral” probability measure Q

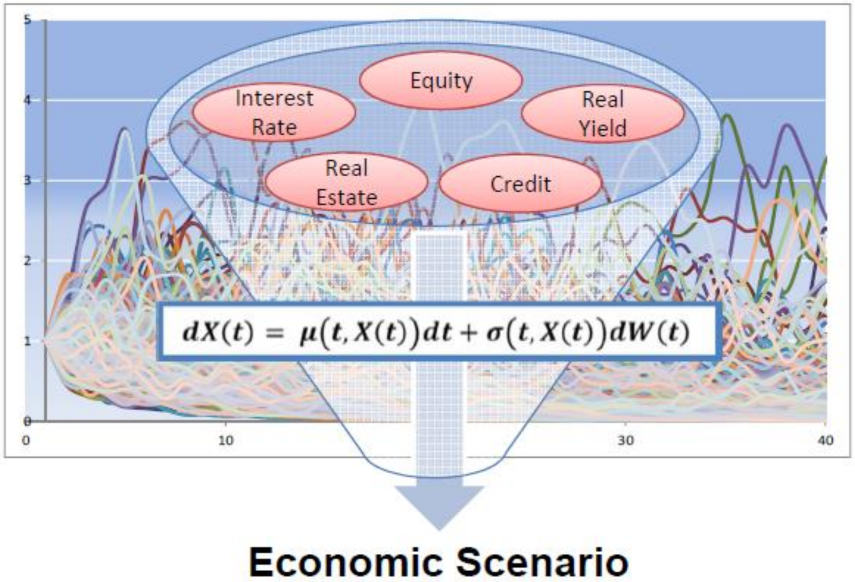
Deep dive: life liability loss modelling and the LSMC method

Market consistent / Risk-neutral scenarios

- Liabilities are evaluated using **market consistent / risk-neutral valuation methods** under the following assumptions:

Market consistency	• Ability to reproduce market prices • Calibration of volatilities is usually based on implied market data (i.e. swaption)
Risk-neutrality	• All assets earn the risk-free rate of interest: pricing of payoff not dependent on expected returns in the real world • All assets, when discounted at the cash rollup, are martingales.

- Brownian motions typically form the basis of stochastic differential equations that describe economic variables (*): interest rates (nominal, real, inflation) , credit transitions, equity returns, property returns:
  - **Libor Market Model plus (LMM+)** for nominal risk-free yield curves, which allows for negative interest rates;
  - **Two-Factor Vasicek model** for real yield curves;
  - **G2 model** (an extension of JLT model) for Corporate bonds' credit spreads;
  - **Time Varying Deterministic Volatility model** for equity indexes;
  - Constant Volatility model for real estate indexes;
  - Mean reverting process for equity dividend yields and real estate income return.



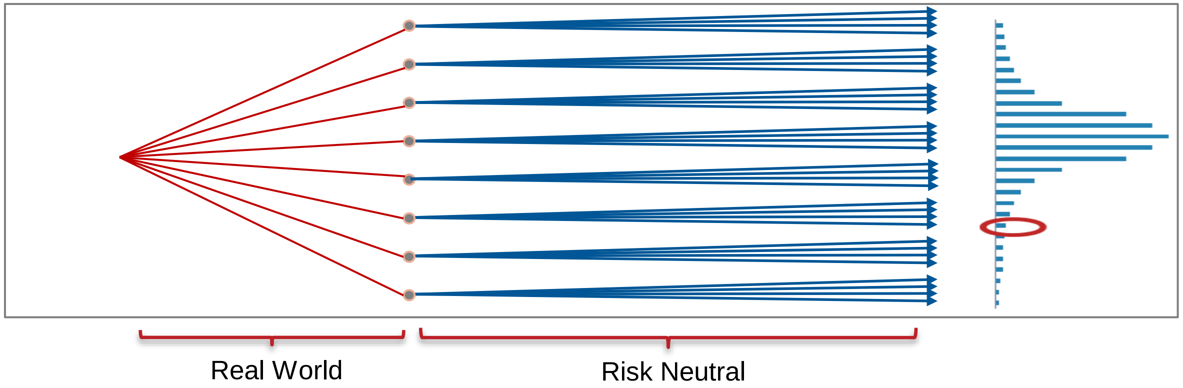
(*) A Brownian motion is a process for describing the evolution of a normally distributed random variable.

Deep dive: life liability loss modelling and the LSMC method

Nested stochastic approach

- Under a traditional Monte Carlo approach, the calculation of economic capital for complex life insurance liabilities requires a nested stochastic approach. This involves running a large number of real-world (outer) simulations, with each outer simulation being the basis for a large number of risk-neutral (inner) simulations:

NESTED STOCHASTIC SIMULATIONS



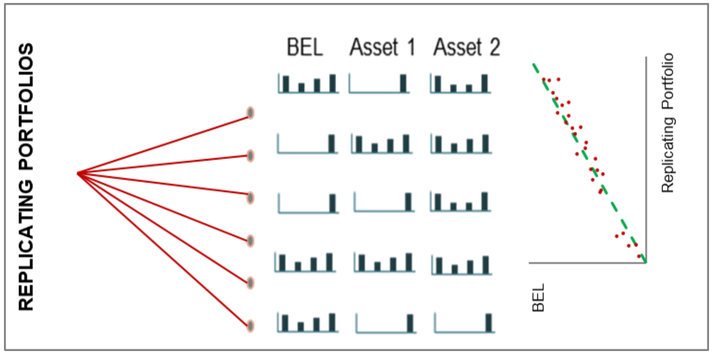
- However, the nested stochastic approach can rapidly become very time consuming and require a large amount of technical resource: e.g. if 1000 real-world simulations are used, each with 1000 risk-neutral simulations, then the total number of simulations run will be 1,000,000, which in most models will take a long time to calculate and create a huge computational burden.

Due to time constraints, however, this approach is computationally not feasible even with state of the art machines. To overcome this limitation, different techniques are proposed in recent literature, as we will see in the following slides.

Deep dive: life liability loss modelling and the LSMC method

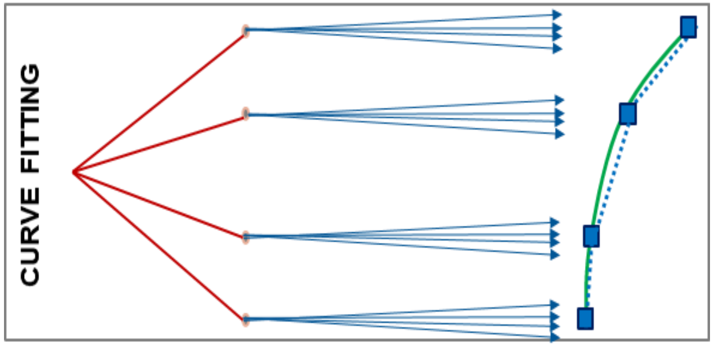
Different techniques available as proxy approaches (1/2)

Replicating Portfolio



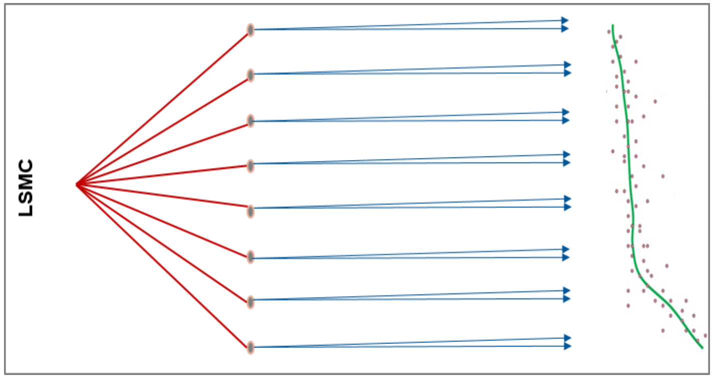
- The idea behind this method is to use a portfolio of simple-to-price financial instruments whose cash-flows are able to replicate those of a liability portfolio with sufficient precision in each stochastic scenario.
- Replicating Portfolio is a robust theoretical solution widely used in the fields of mathematical finance and insurance (e.g. hedging) but it is not able for modelling non-financial risks.

Curve Fitting



- The calibration is performed evaluating the liability portfolio value precisely (with a huge number of risk neutral scenarios) in a few real world / deterministic scenarios, and then determining a curve through regression or interpolation methods.
- Curve fitting is a very practical solution but suffers of the high arbitrariness in the choice of the real-world scenarios with extremely significant impact on the result.

Least-Squares Monte Carlo (LSMC)



- A large number of real world scenarios is generated, and in each scenario the liability portfolio is valued approximately (with just a few risk neutral scenarios). Then, using least-squares estimation, a polynomial regression curve is derived.
- Least-Squares Monte Carlo (LSMC) represents an evolution of the curve fitting method based on the homonymous technique developed in quantitative finance to price American options.

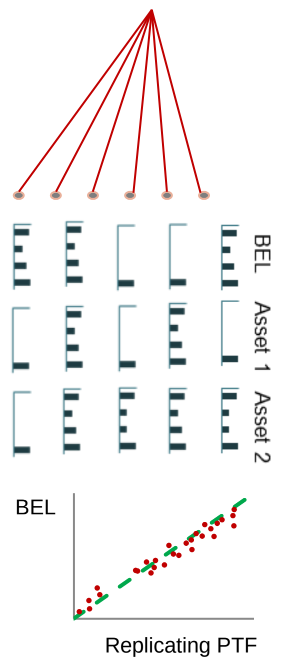
Deep dive: life liability loss modelling and the LSMC method

Different techniques available as proxy approaches (2/2)

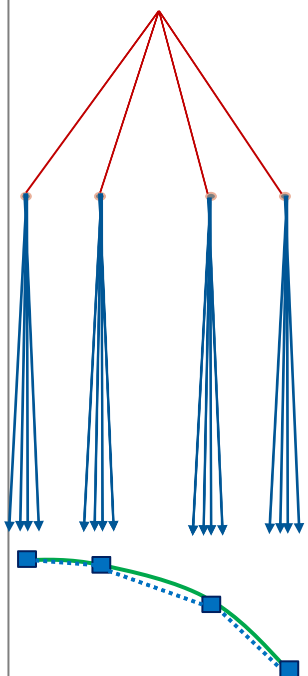
Focus on Loss Modelling

Least-Squares Monte Carlo resulted as the best choice to be adopted in terms of risks coverage, robustness, and feasibility

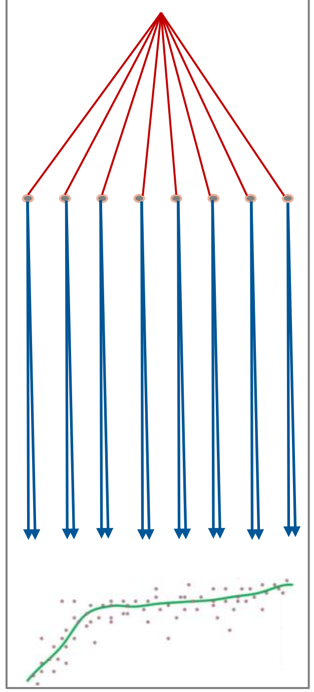
REPLICATING PTF



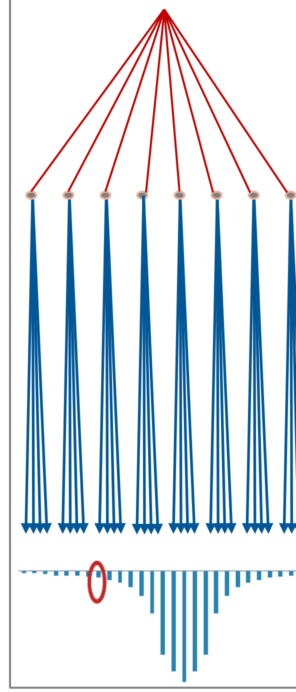
CURVE FITTING



LSMC



NESTED STOC.



	REPLICATING PTF	CURVE FITTING	LSMC	NESTED STOC.
covers all the risks	x	✔	✔	✔
approach's robustness	✔	x	✔	✔
feasibility (system / time)	x	✔	✔	x

GENERALI

Deep dive: life liability loss modelling and the LSMC method

Least-Squares Monte Carlo (LSMC) methodology

The LSMC methodology consists in carrying out a Least Squares regression on the outputs of a number of Monte Carlo simulations.

LS | MC

LEAST SQUARES REGRESSION

Linear regression is an approach used to model the relationship between a scalar dependent variable Y and one or more explanatory variables denoted as X:

$$\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Ordinary least squares is one of the methods for estimating the unknown parameters $\boldsymbol{\beta}$:

$$\widehat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T \boldsymbol{y}$$

In order to model non-linear relationships and interaction effects between regressors, a polynomial approach is adopted:

$$\boldsymbol{Y_i} = \boldsymbol{\beta_0} + \boldsymbol{\beta_1 x_1} + \boldsymbol{\beta_2 x_2} + \boldsymbol{\beta_3 x_1^2} + \boldsymbol{\beta_4 x_2^2} + \boldsymbol{\beta_5 x_1 x_2} + \boldsymbol{\varepsilon_i}$$

CENTRAL LIMIT THEOREM

$$\frac{\sum_{i=1}^{n} x_i}{n}, X_i\ iid \xrightarrow{n\to\infty} N\left(\mu, \frac{\sigma^2}{n}\right)$$

The arithmetic mean of a large number of independent random variables identically distributed (with a finite variance) is approximately normally distributed, regardless of the underlying distribution.

MONTE CARLO SIMULATION

Monte Carlo algorithms make use of repeated random sampling to carry out a numerical computation.

They can be used to solve any problem with a probabilistic interpretation: integrals described by the expected value of some random variable can be approximated by taking the simple average of independent samples of the variable (law of large numbers).

Thanks to the Central limit theorem, the method shows a speed of convergence equal to $1/\sqrt{N}$, where N is the number of simulations.

As powers of the same risk factors are considered as different regressors in the linear model, strong **multicollinearity** will be present in the model. Therefore, the design matrix of the model will potentially be ill-conditioned as the determinant of $X^T X$ could be close to zero.
An approach used to increase the robustness of the estimation process is the **orthogonalization** of the design matrix. This can be achieved, for example, using the $\boldsymbol{QR}$ **decomposition**, which states that a full rank matrix can be seen as matrix product between an orthogonal matrix Q (i.e. $Q^T Q = I$) and an upper triangular matrix R.

Deep dive: life liability loss modelling and the LSMC method

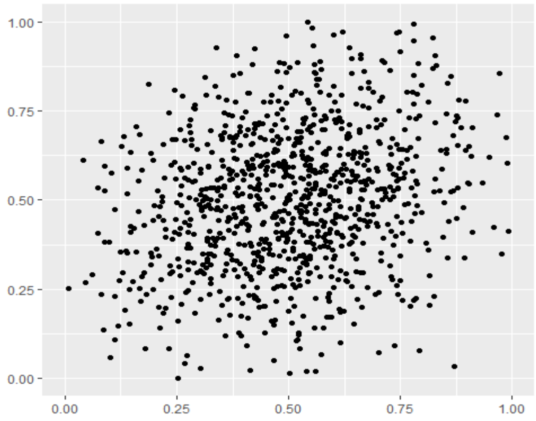
Real-world scenarios

The outer fitting scenarios for the calibration of the LSMC model could be taken from the real-world distribution of risk factors, as for the traditional nested stochastic method. However, as we want the function to be precise expecially in the tails of the distribution, other techniques are more effective.

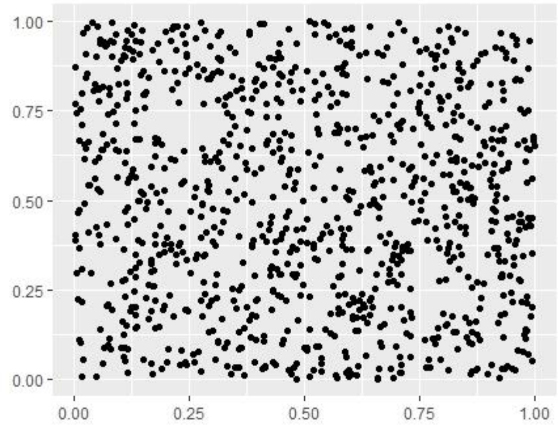
An idea is to simulate from the uniform distribution (pseudo-random numbers) in order to have lots of fitting points also on the tails, but this solution can still be improved.

We define **discrepancy** as a measure of how much a sequence of numbers differs from uniformity. A lower discrepancy generally leads to more accurate fits. Therefore instead of trying to mimic casuality with pseudo-random numbers, we can try to minimize the discrepancy of a sequence, which leads to quasi-random numbers (low discrepancy sequences) like **Sobol sequences**.

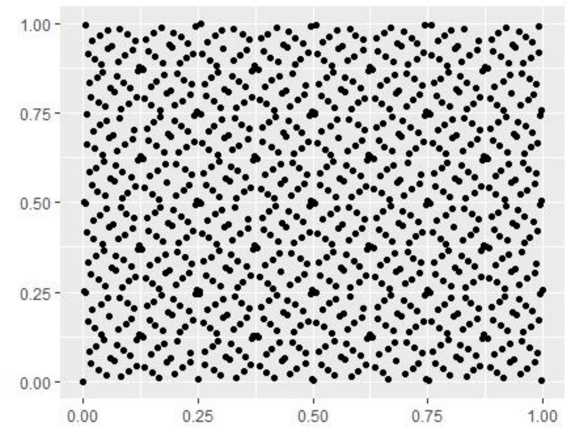
REAL-WORLD SIMULATIONS



PSEUDO-RANDOM NUMBERS



QUASI-RANDOM NUMBERS



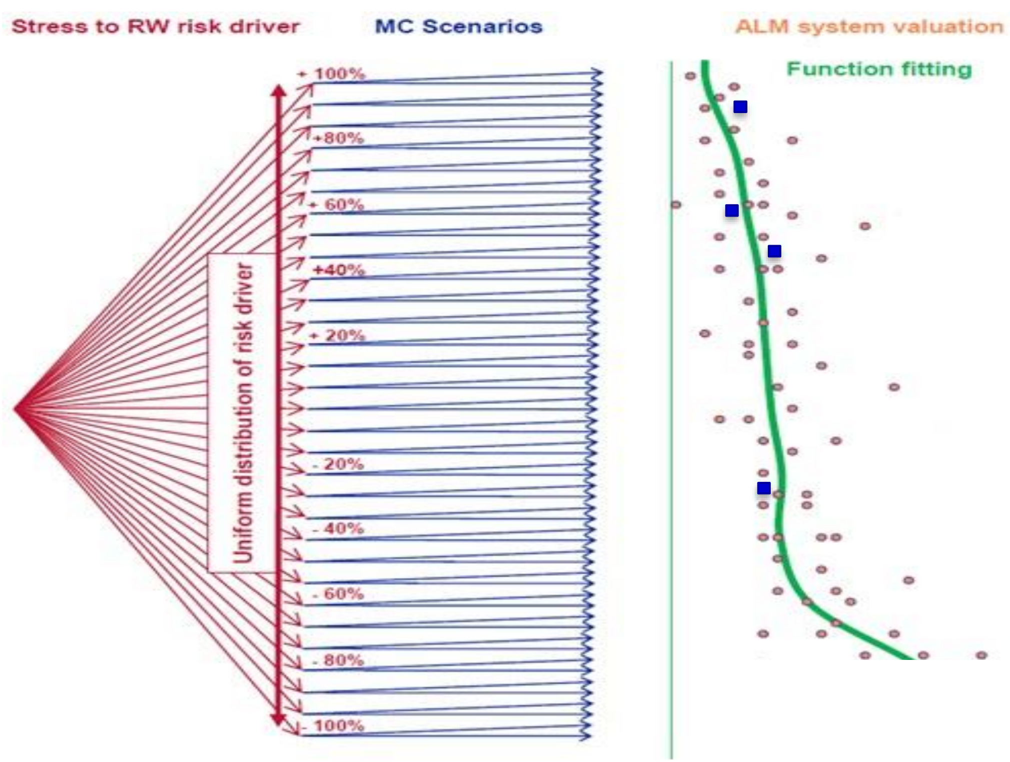
Focus on Loss Modelling

Deep dive: life liability loss modelling and the LSMC method

Fitting process to provide Liability Proxy Functions (LPF)

Fitting process

In order to fit the proxy, an imprecise value of liability is evaluated according to each of the numerous real word scenarios via MC simulation over a small number of risk neutral scenarios (**fitting scenarios**). Then, a LS regression algorithm is applied to get the proxy.

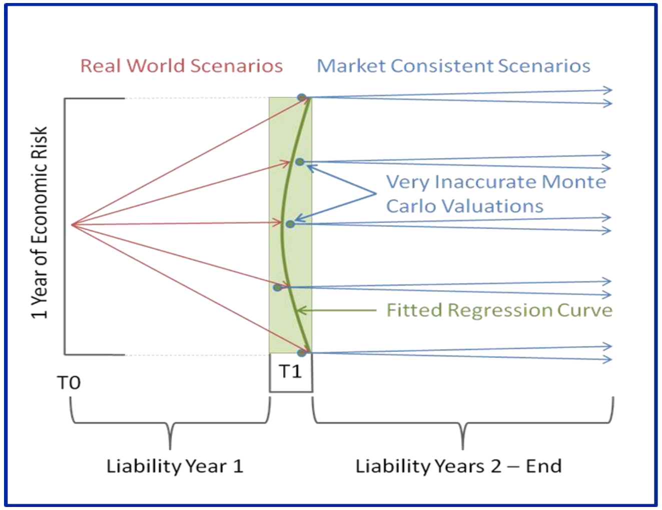


Approximate value for every fitting point

Real-world projection: Sobol random numbers used to simulate real-world scenarios

Risk-neutral projection: Antithetic Variates are used. The Antithetic Variates method is a variance reduction technique which aims to reduce the variability of a Monte Carlo estimator without increasing the number of simulations.

Function fitting: Inaccurate value for every RW scenario; however least squares regression captures the overall shape of the curve



Focus on Loss Modelling

Deep dive: life liability loss modelling and the LSMC method

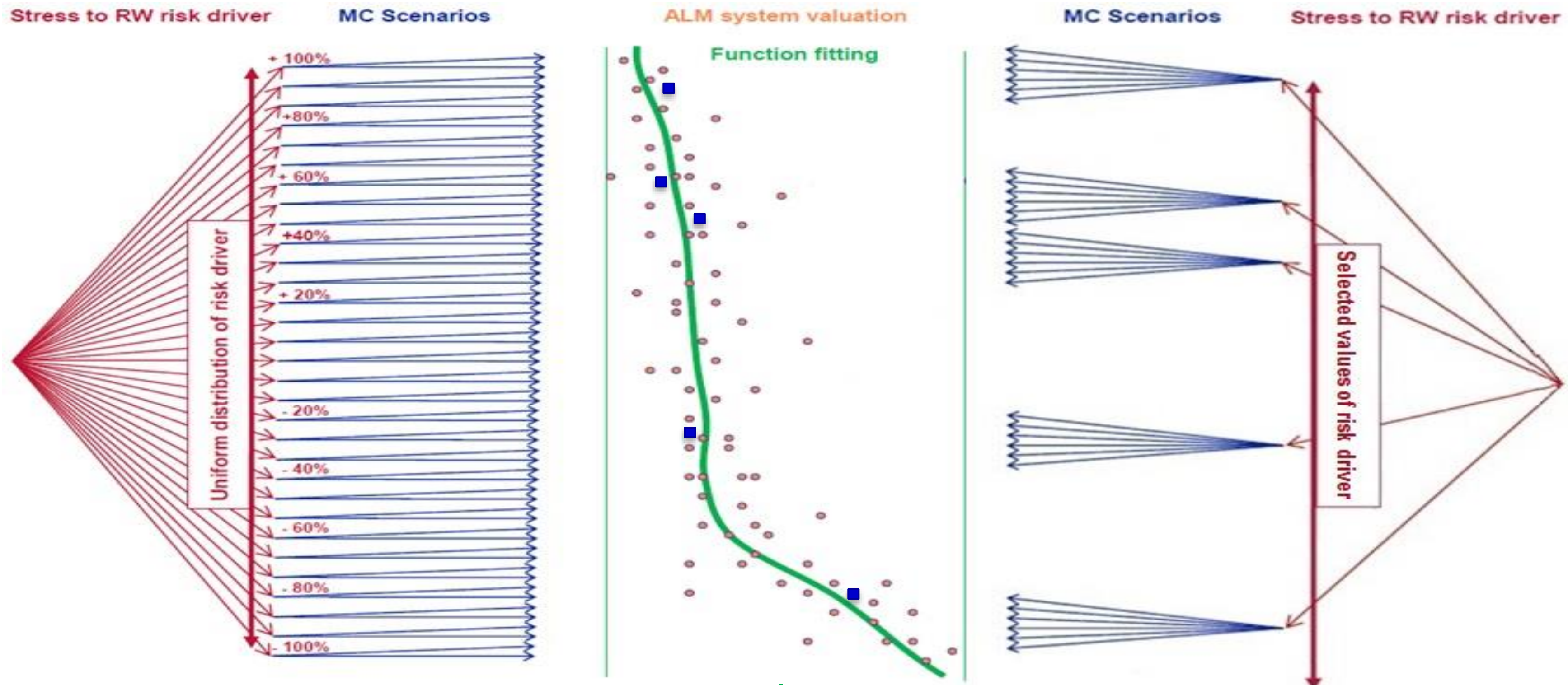
Fitting and validation processes

Fitting process

In order to fit the proxy, an imprecise value of liability is evaluated according to each of the numerous real word scenarios via MC simulation over a small number of risk neutral scenarios (**fitting scenarios**). Then, a LS regression algorithm is applied to get the proxy.

Validation process

An out-of-sample validation is performed. Specific scenarios (**validation scenarios**) are deterministically selected and an accurate valuation is carried out with many risk neutral scenarios. The goodness of the model is assessed looking at the differences between actual and predicted value.



Approximate value for every fitting point

LS regression captures the overall shape of the function

Accurate value for every validation point

GENERALI

Focus on Loss Modelling

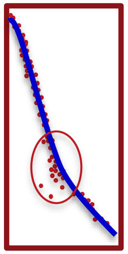
Deep dive: life liability loss modelling and the LSMC method

Under-fitting and over-fitting issues: a room for further researches …

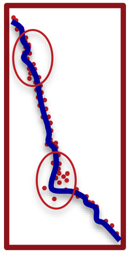
To avoid under- and over-fitting, a **variable selection** technique is needed in order to include only the relevant powers in the model.

Automatic selection procedures are proposed in the statistical literature (e.g. forward and stepwise algorithms) based on information criteria (AIC, BIC). These however need to be adapted for the polynomial approach used in the LSMC model.

Under-fitting
insufficiently high risk polynomial power



Over-fitting
too many terms selected (it also fits the random noise)



The **adaptive forward procedure** extends the forward algorithm to the polynomial regression building upon the statistical principle of marginality, for which higher order and cross terms shouldn't be included in the model if the main effects aren't present first.

ITERATION 0	ITERATION 1		ITERATION 2		ITERATION 3	
$y=\beta_0$	x, w, z	$y=\beta_0+\beta_1 x$	x^2, w, z	$y=\beta_0+\beta_1 x+\beta_2 z$	x^2, w, z^2, xz	$y=\beta_0+\beta_1 x+\beta_2 z+\beta_3 x^2$

ITERATION 4		ITERATION 5		ITERATION 6	
x^3, w, z^2, xz	$y=\beta_0+\beta_1 x+\beta_2 z+\beta_3 x^2+\beta_4 xz$	$x^3, w, z^2, x^2 z$	$y=\beta_0+\beta_1 x+\beta_2 z+\beta_3 x^2+\beta_4 xz+\beta_5 w$	$x^3, w^2, z^2, x^2 z, xwz$	…