100% ACTUAIRES & 100% DATA SCIENCE

INSTITUT DES ACTUAIRES

The projection of the solvency ratio : a machine learning approach to by-pass the operational constraints of nested simulations

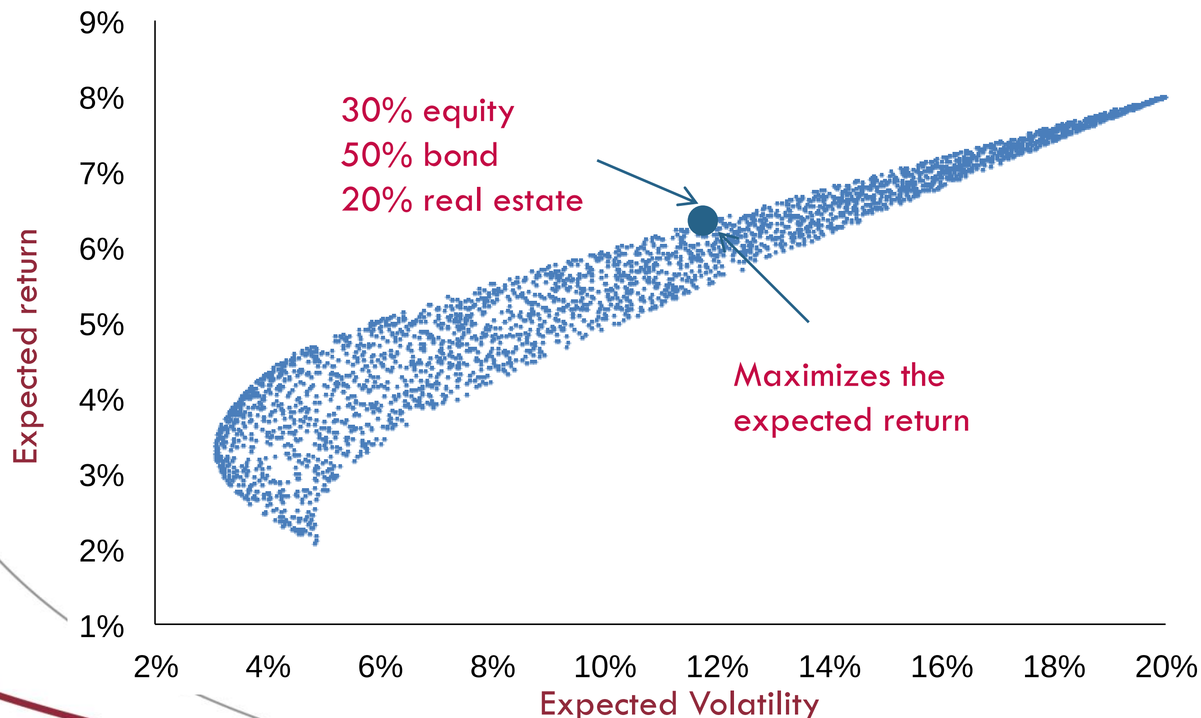
4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

Éclairer les risques, tracer l'avenir

Introduction

Traditional financial management *(Markowitz Theory)*

The insurer sets a **volatility budget** and selects the portfolio which maximizes the expected return.
Illustration for a volatility budget of 12%:



Assumptions for the portfolio

- Equity (σ = 20% ; μ = 8%)
- Bonds (σ = 5% ; μ = 2%)
- Real Estate (σ = 4% ; μ = 4%)

Assumption for the insurer commitments

- Assets and liabilities are independent

Financial management under Solvency 2

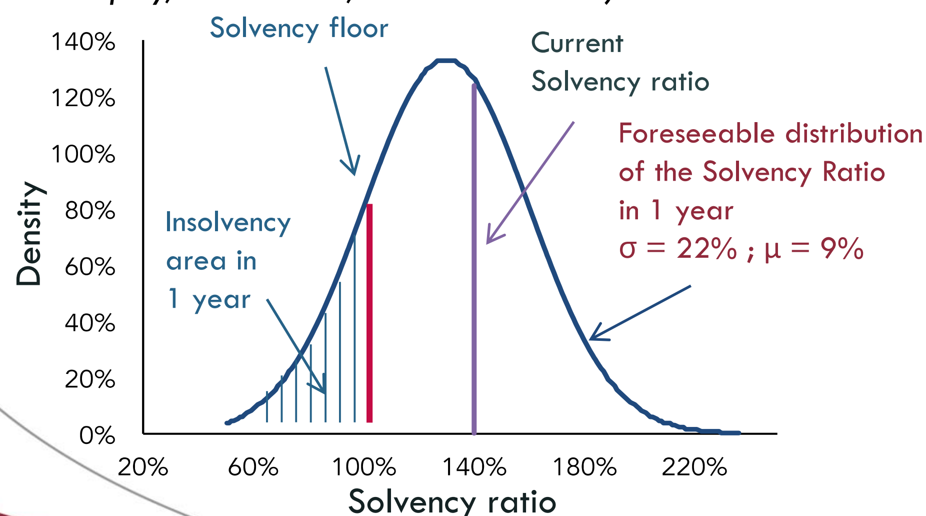
Must take into account the asset/liability interactions (Best Estimate, Own funds, SCR, Solvency ratio)

→ The solvency ratio relies on the asset allocation

→ Projecting the solvency ratio allows us to set an accepted risk level for the asset allocation

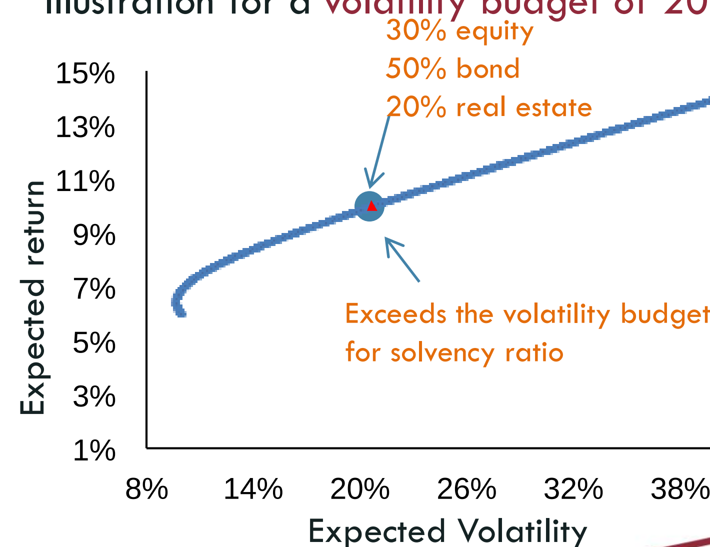
For 1 allocation

(30% equity, 50% bond, 20% real estate)



For all allocations

Illustration for a volatility budget of 20%:



Introduction

Under Solvency 2, life insurers must record on their prudential balance sheet all **contractual options** and **financial guarantees** included in their products.

The financial **guarantees** of a retirement contract only concern the **contractual profit sharing**.

From a mathematical point of view, it can be expressed as an expectation based on the evolution of financial variables (example):

E[max{ book yield – technical rate ; inflation }]

The assessment of such option calls for the Monte Carlo method (no closed formula)

Implementation complexity

Date	Nested simulations	Target assessment
Today	1	Prudential balance sheet
1 year	2	SCR internal models
2 years	3	Efficient border of the solvency ratio

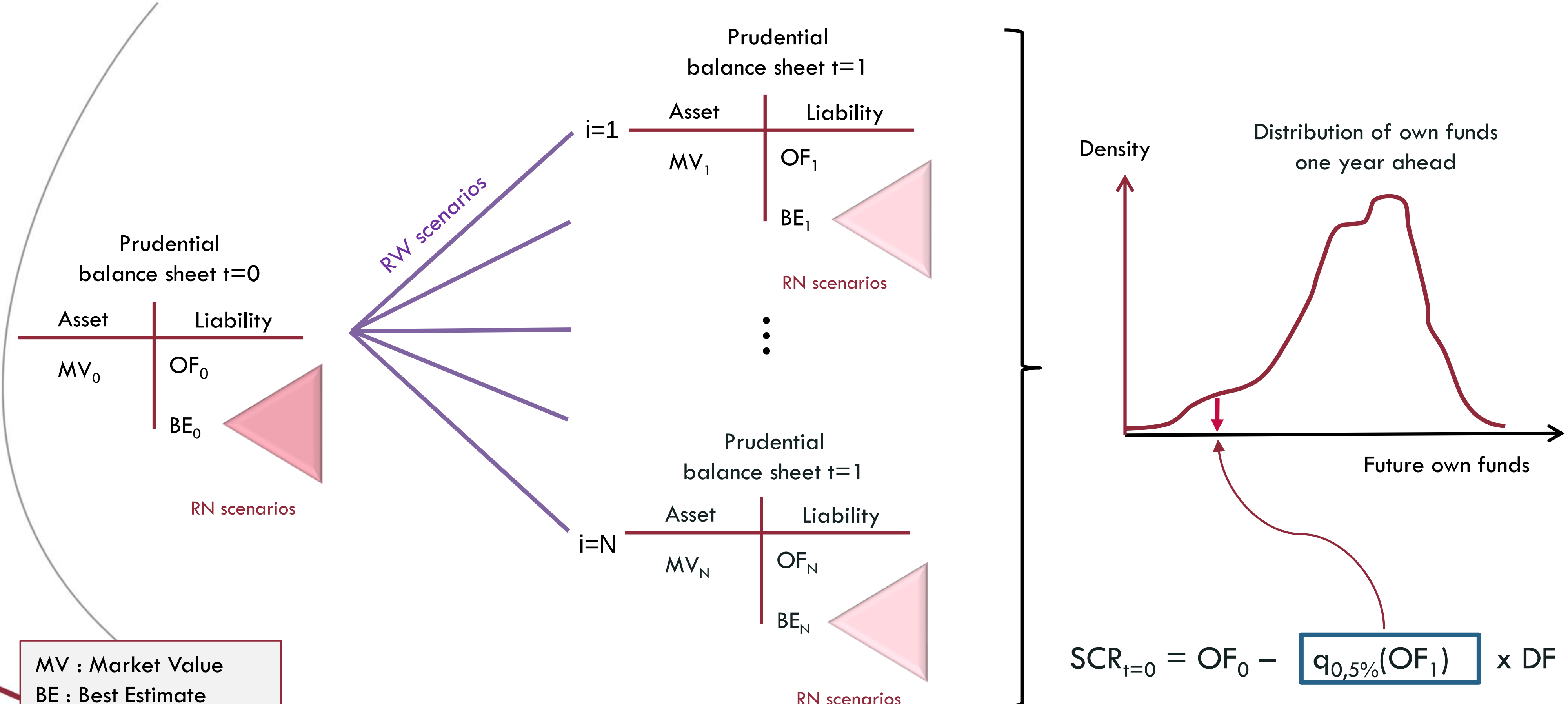
Machine learning can help to delete one level of simulation and to apply a closed formula

Projection of the solvency ratio

- Nested simulation approach

The Nested Simulations method

First and second levels of simulations



$$SCR_{t=0} = OF_0 - \boxed{q_{0,5\%}(OF_1)} \times DF$$

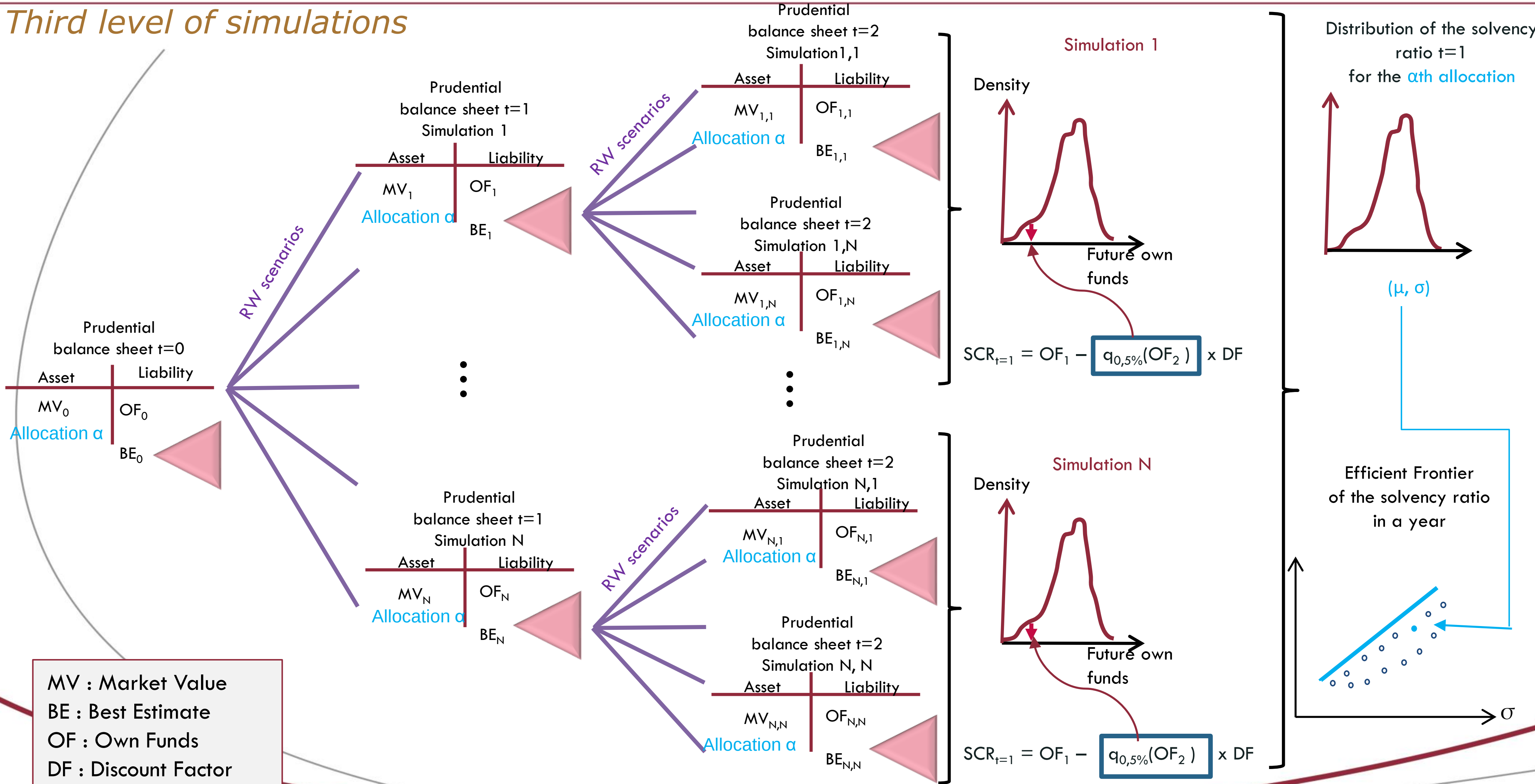
MV : Market Value
BE : Best Estimate
OF : Own Funds
DF : Discount Factor
RW : Real World
RN : Risk Neutral

100% ACTUAIRES & 100% DATA SCIENCE

INSTITUT DES ACTUAIRES

The Nested Simulations method

Third level of simulations



MV : Market Value
BE : Best Estimate
OF : Own Funds
DF : Discount Factor
RW : Real World
RN : Risk Neutral

Projection of the solvency ratio

- Data generation

Our goal

Simulate a large variety of economic and financial situations to generalize the *Best Estimate* learning at the starting date but also over time.

<u>Several options:</u>

- **Option n° 1**: learning the *Best Estimate* in 2 years from the current economic and financial situation
- **Option n° 2**: learning the *Best Estimate* in 1 year from the current and the stressed economic and financial situations
- …

→ We choose the option n° 1

In any case, we have to generate the full nested simulations once to assess the prediction capacity of our future model.

Data generation

Base content

	Explained variable	Explanatory variables						
Generated Observation	Best Estimate	Inflation	BE initial	Put option	…	Equity share	10 y gvt rates	Curve Slope
1	95,5€	2,2%	94,9€	2,3€	…	22,6%	1,6%	-1,5%
2	100,3€	2,3%	100,6€	3,6€	…	34,5%	2,9%	-1,3%
3	98,4€	3,5%	99,6€	2,5€	…	28,6%	1,5%	-1,6%
…	…	…	…	…	…	…	…	…
1099	109,6€	2,9%	108,6€	3,8€	…	21,9%	2,5%	-1,8%
1100	99,3€	3,1%	100,1€	2,1€	…	27,8%	2,8	-1,1%

Operational implementation

The estimation of the capital requirement requires crossing observed data (insurance portfolio) as well as generated data (financial variables).

The generation of financial data requires making choices concerning:

The code architecture with object-oriented programming for:

- Modularity
- The legacy
- Clarity / safety
- Reuse

The computer language that must satisfy constraints:

- Object creation
- Large matrices handling
- Execution speed on loops
- Calculations parallelism
- Implementation ease

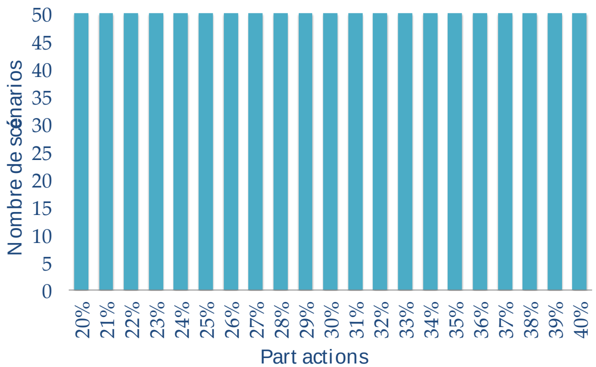
$$\hat{f}(x)$$

Modeling financial variables:

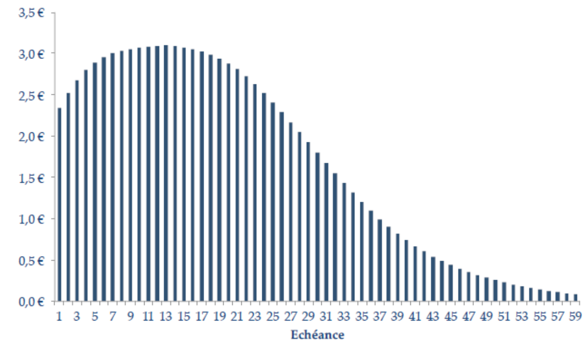
- For which we have an analytical formula to project, at a given time, financial variables in the future
- That are usable in several projection universes (real word and risk neutral)
- That can be correlated
- That satisfy regulatory constraints (martingale, market consistency)

Our approach

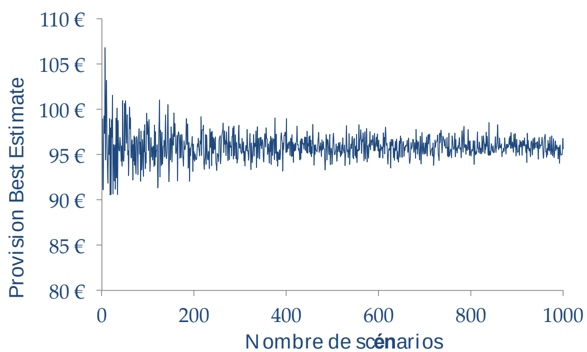
Simulated allocation: 50 simulations for each allocation between 20% and 40%



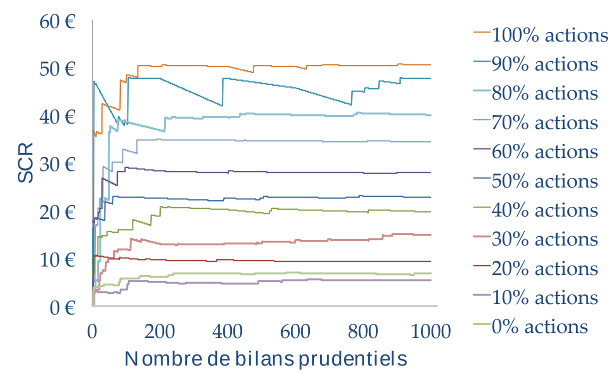
Simplified modeling: 2 asset classes, a revaluation policy dependent on book yield and inflation, etc.



Convergence study n° 1: <u>Number of scenarios</u> **for** ***Best Estimate***



Convergence study n° 2: <u>Number of prudential balance sheets</u> for SCR



Projection of the solvency ratio

- Machine learning principles

Differences with common statistics

In ***statistics***, we generally formulate a hypothesis on the data distribution (ex: Gaussian law), then we calibrate the model parameters on the observed data set:

- The model parameters are estimated so that the average residuals are null (so-called *« unbiased » estimator*).
- The model is evaluated on its ability to verify statistical tests.

In ***machine learning***, a model is evaluated on its ability to capture the structuring components of the data and to be generalizable:

- No hypothesis on the data distribution: the approach consists of selecting several models / algorithms that are candidates for the best generalization of data;
- Each model is calibrated on part of the observed data and then tested via a *« cost function »* (eg: mean squared error) on the other part of the data to assess its generalization;
- The model chosen is the one offering the best generalization on the test data (i.e. the lower cost).

4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

Attributes of a model

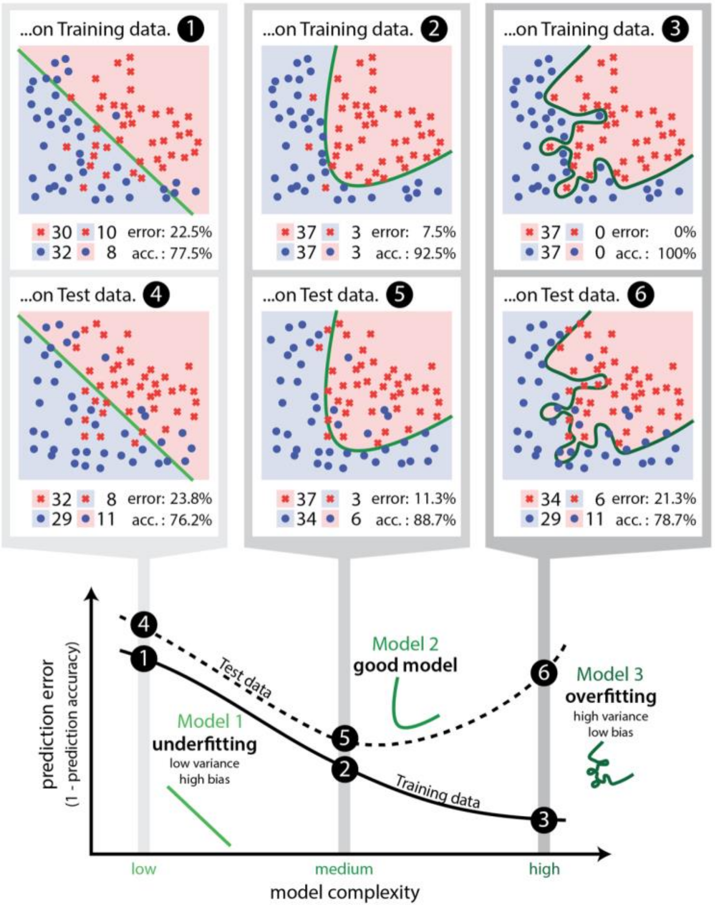
In ***machine learning***, a model is valued according to 5 attributes:

	Bias	Variance	Complexity	Flexibility	Generalization
Under learning: the model is too simplistic	High	Low	Low	Low	High
Overfitting: the model models noise	Low	High	High	High	Low

All the machine learning approach challenge is to manage the « ***bias/variance tradeoff*** »:

Estimation error = Bias + Variance + incompressible error

Bias – variance tradeoff



In machine learning, a model contains 2 types of parameters:

- The **hyper-parameters** whose set of discrete definition is established a priori by the *data scientist* for the adjustment of the bias variance tradeoff and the reduction of the number of dimensions. These parameters can be calibrated by a *« cross validation »* approach.
- **Other parameters** whose calibration results from an algorithm (gradient descent, etc.)

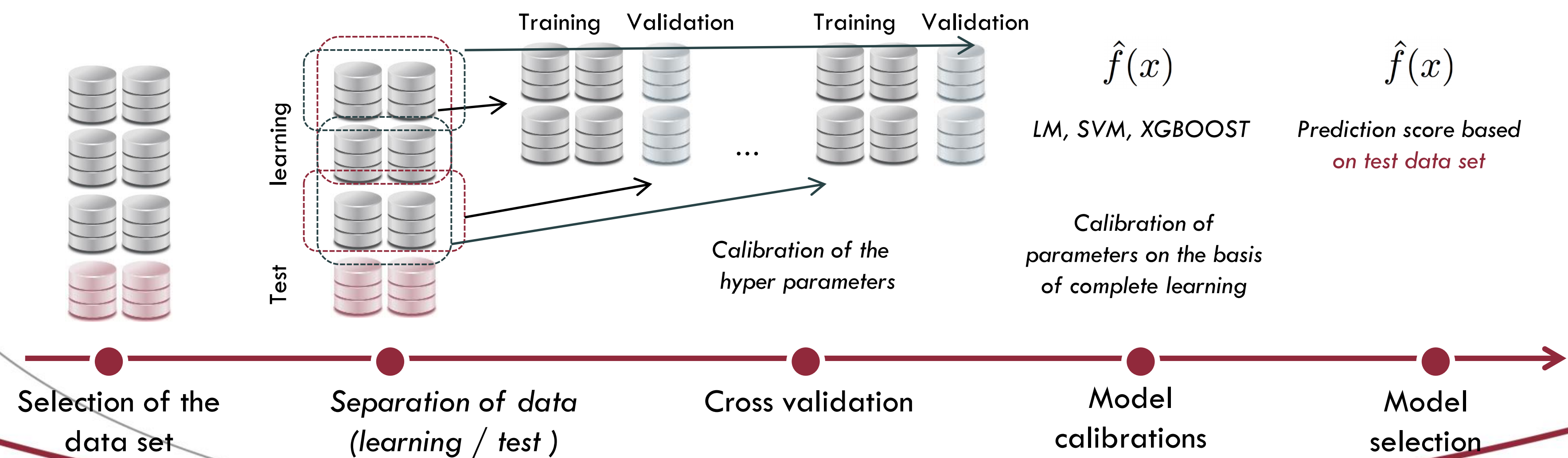
Model selection

Principle

No assumptions about the probability distribution of the data with:

- The search for a bias/variance tradeoff
- The minimization of the empirical risk via an objective function: cost function + regulation parameter

Implementation



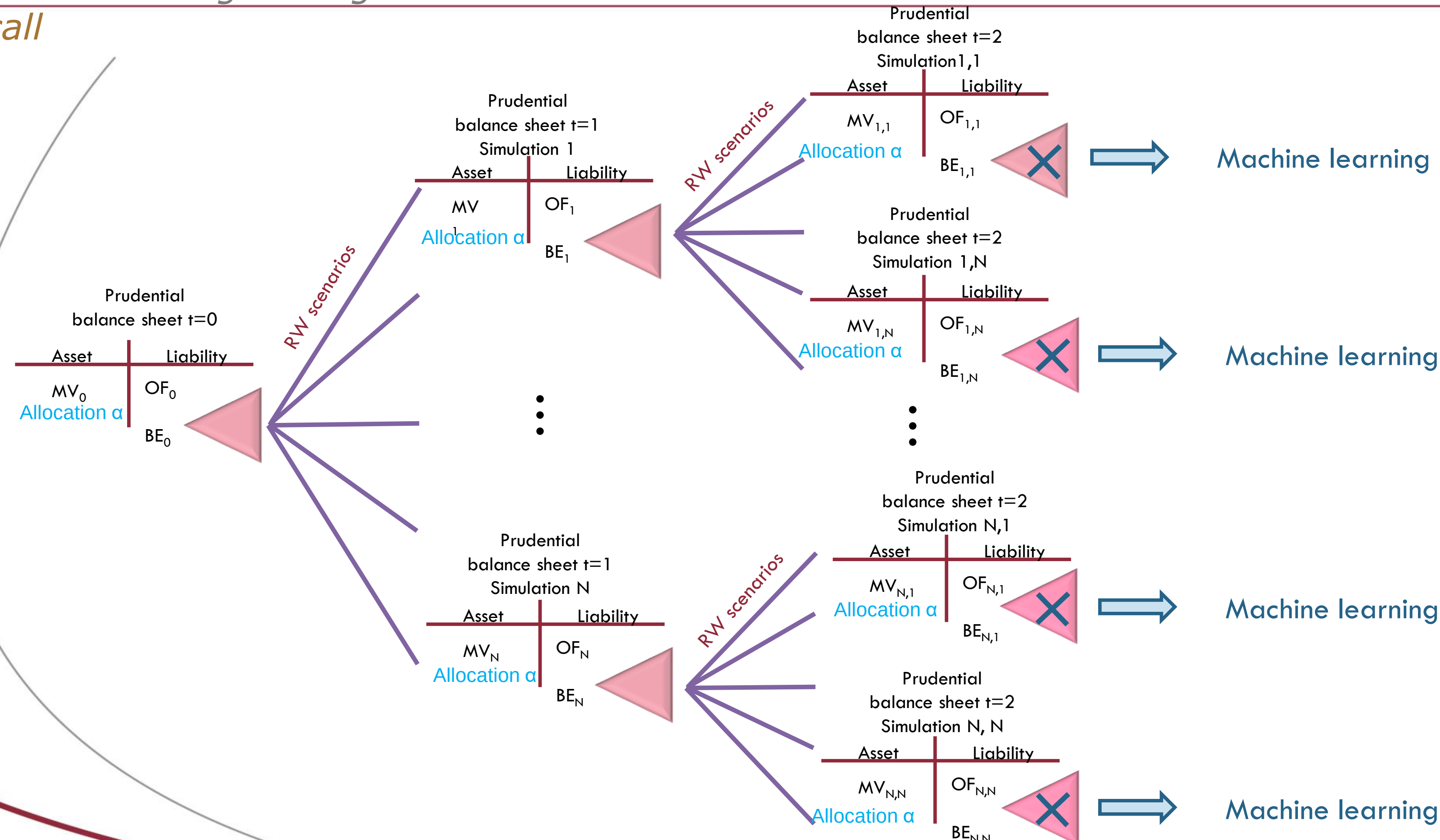
4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

Projection of the solvency ratio

- Machine learning testing

INSTITUT DES ACTUAIRES

Recall



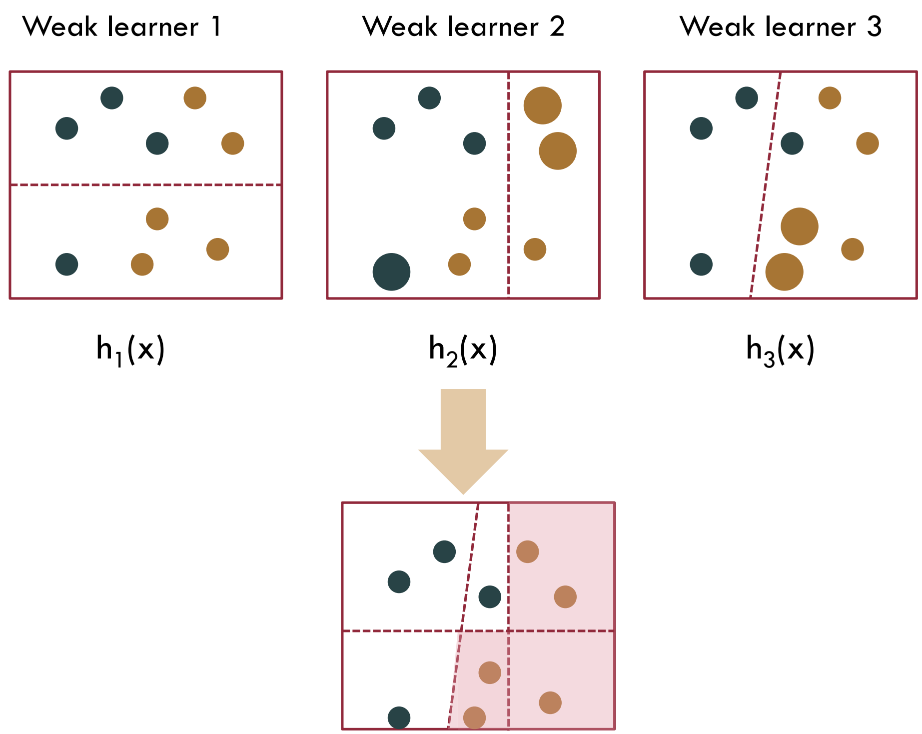
Boosting algorithm

Brief description

Boosting algorithms rely on the principle that many weak learners can be more efficient than the strongest of the predictors, by simultaneously mobilizing all their assets. These algorithms consist of repeatedly training a basic classification/regression model to minimize its loss function over the entire learning data base. The algorithm focuses on each new iteration on the misclassified data at the previous iteration, thanks to a mechanism of adjustment of the importance of all data of the learning data base.

These boosting algorithms are available:

- *In classifier or regression version*
- *To minimize cost function: with or without gradient-based descent*
- *To avoid over fitted: with or without regularization term*



$$H(x) = \text{sign}(\ \alpha_1 \times h_1(x) + \alpha_2 \times h_2(x) + \alpha_3 \times h_3(x)\ )$$

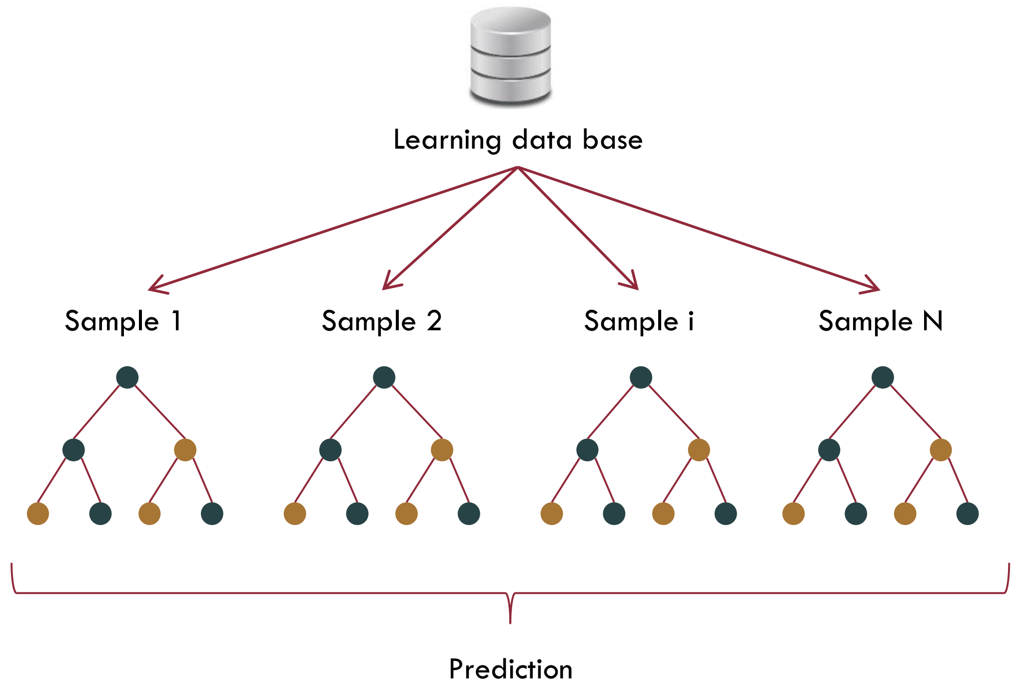
4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

INSTITUT DES ACTUAIRES

Random forest algorithm

Brief description

The random forest algorithm aggregates several decision trees built with a set of randomly drawn learning data base. For each tree, the construction of a node is performed on a subset of variables randomly drawn. Different segmentation criteria exist for three construction. The Gini index (CART algorithm) is one of the most used. The result of the model is a majority vote (classification) or an average (prediction). This algorithm can be particularly useful when using a large number of explanatory variables.



Classification: majority class predicted by N trees.
Regression: mean predicted by N trees.

→ **It is possible to visualize the variables with the strongest explanatory power on the variable to be explained.**

4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

100% ACTUAIRES & 100% DATA SCIENCE

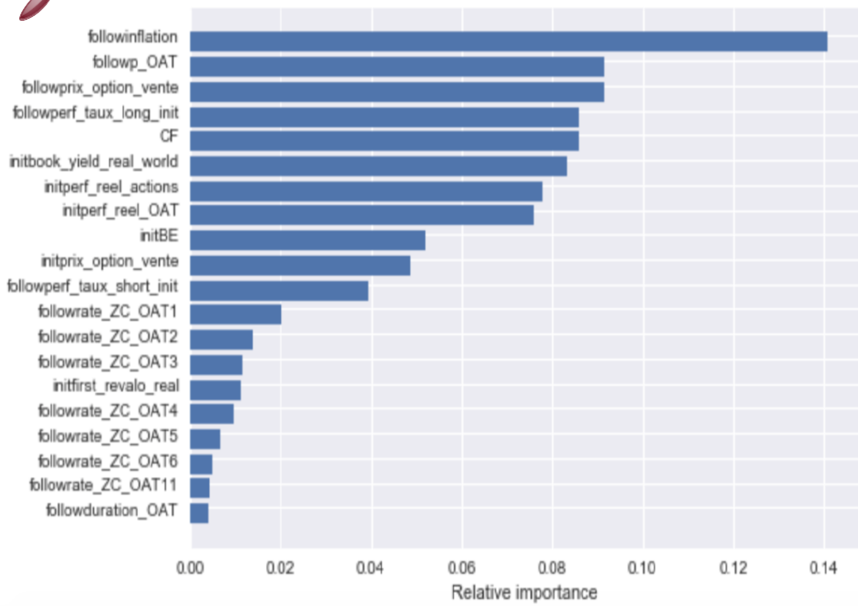
INSTITUT DES ACTUAIRES

1/2 Relative importance of explicative variables

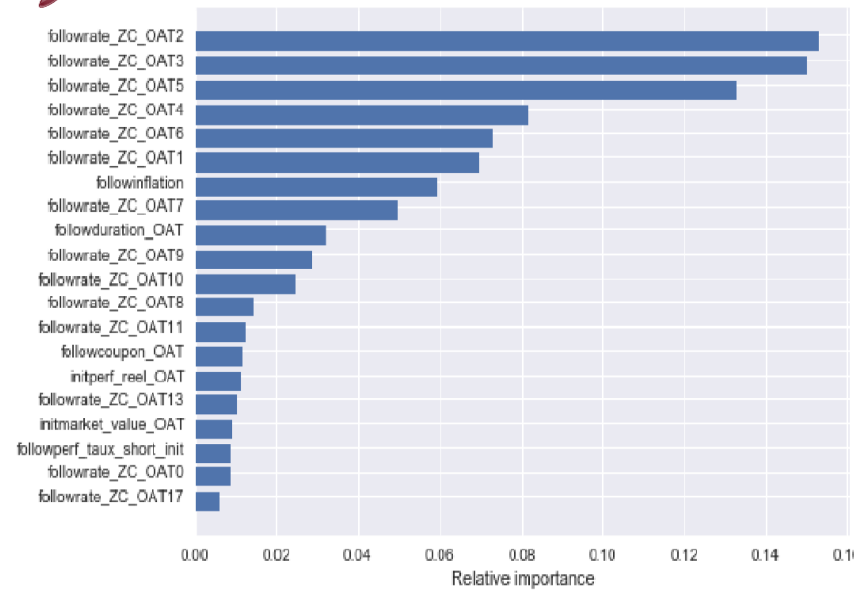
For Best Estimate

Boosting algorithm focus more on inflation whereas random forest is more interested in risk-free curve:

XGBoost



Random Forest



We can assume that XGBoost predicts the Best Estimate with a first tree based on future inflation and then corrects it given the equity share and other financial variables.

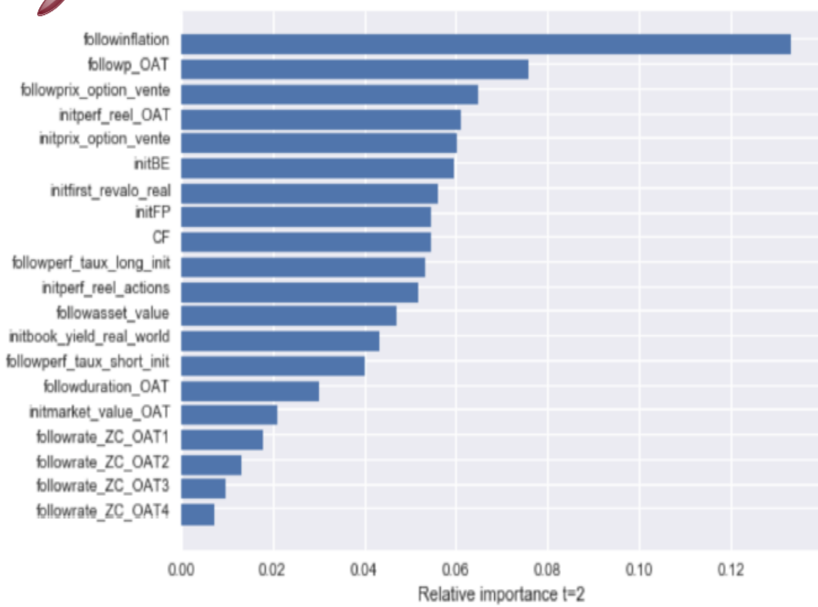
4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

2/2 Relative importance of explicative variables

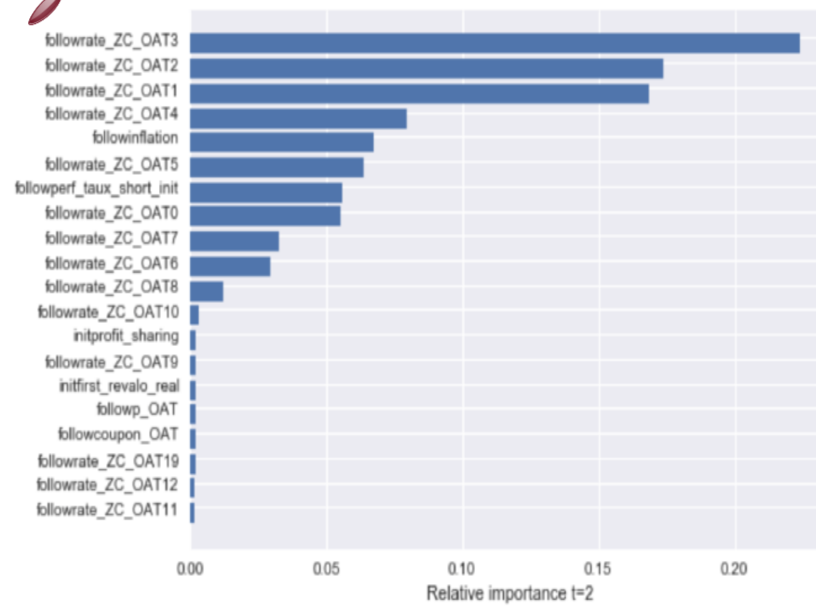
<u>For Solvency ratio</u>

The Xgboost model seems to use more explanatory variables to estimate the Best Estimate by two years as compared to the Random Forest model:

XGBoost



Random Forest



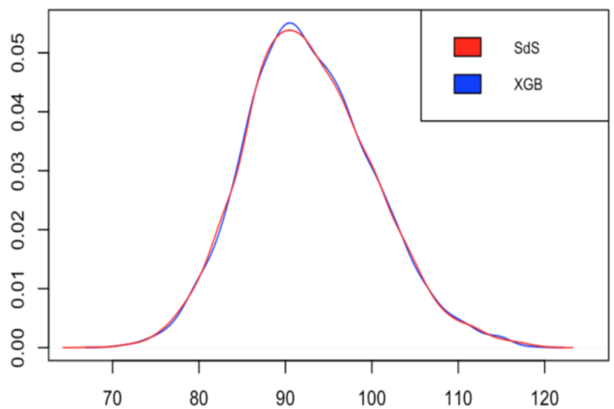
The Random Forest algorithm focuses almost exclusively on the early maturities of future rates and future inflation. We can dread a drop in the performance of this model.

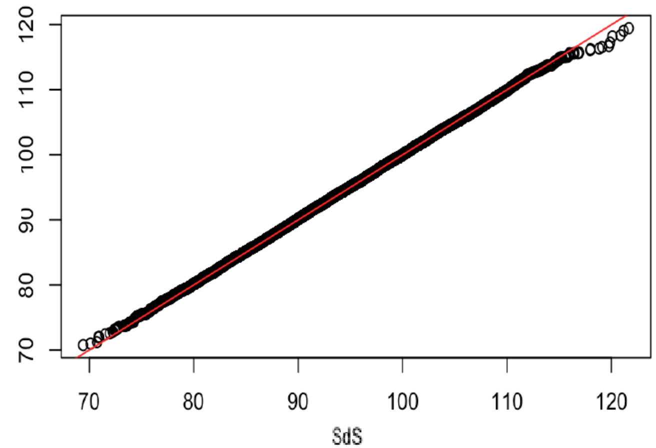
1/2 Prediction accuracy

For Best Estimate

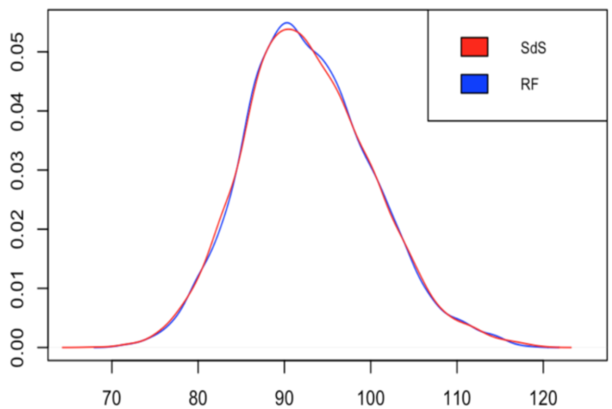
Predictions for XGBoost and Random Forest are very close:

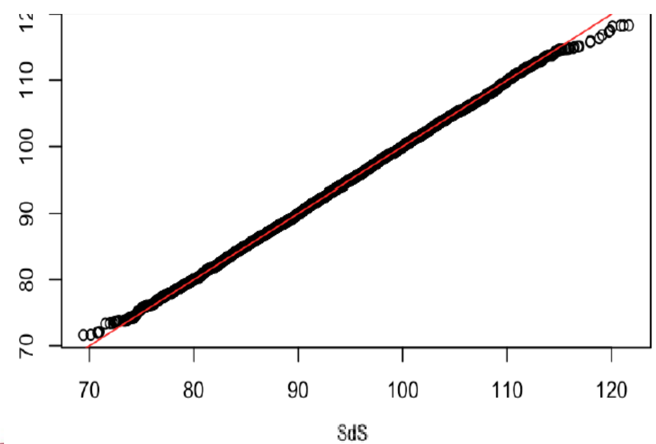
XGBoost





Random Forest



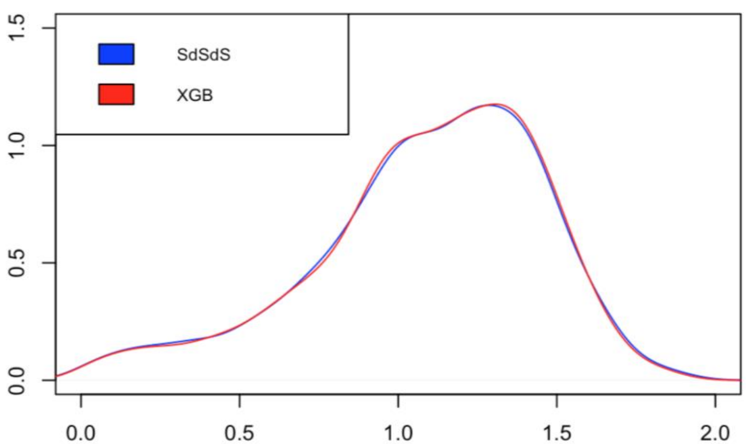


4, rue Chauveau-Lagarde - 75008 Paris - tél. 01 44 51 72 72 - fax 01 44 51 72 73
www.institutdesactuaires.com

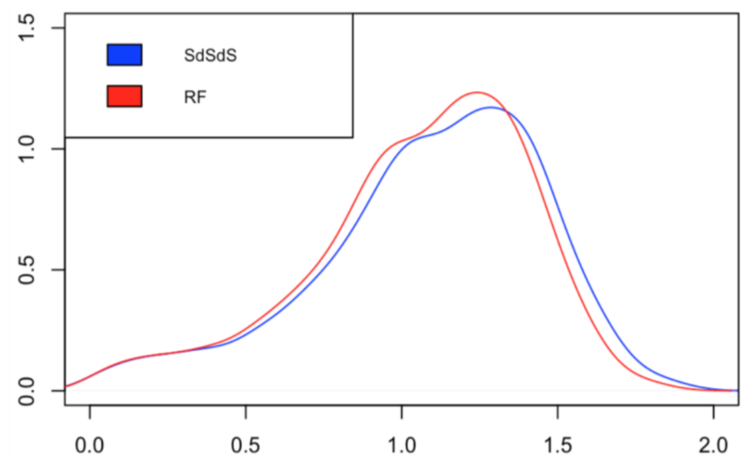
For Solvency ratio

XGBoost is far more accurate than Random Forest for the prediction:

XGBoost



Random Forest



Projection of the solvency ratio

- Building of efficient border

Calibration of the XGBoost model from a learning approach

To Do:

- Selection of an equity share interval: 20% - 40%
- Generation of N economic scenarios up to t = 1
- For each asset allocation:
  - Application of the XGBoost model in 2 years to by-pass nested simulations
  - Construction of the probability distribution of the solvency ratio in 1 year
  - Computation of the expectation and volatility from this distribution
- Construction of the efficient border

2/2 Prediction accuracy