Aggregation criteria in risks' modeling service

Aurélien Couloumy – ADDACTIS Worldwide – Oslo Colloquium 2015

Aurelien.couloumy@addactis.com

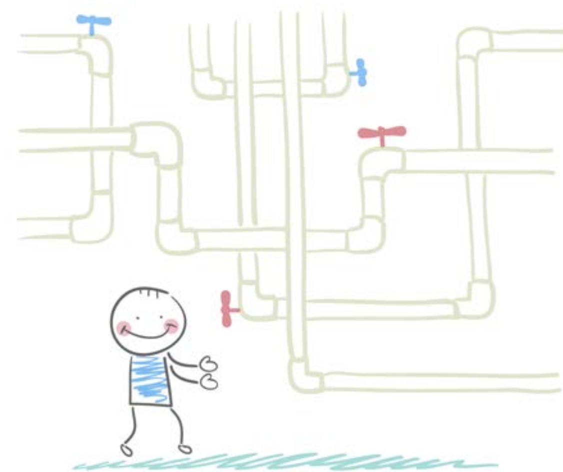
CONTENTS

Part 1 – Issues and classifications methods

Context and goals (1/2)

- A huge volume of information to feed as well as possible risk management model

- But the impossibility to use it optimally



Context and goals (2/2)

- Consequently , technical teams reccur to simplified assumptions

- Current methods are easy to use but not really under control

Context and goals (2/2)

- The aim of the study will be twofold

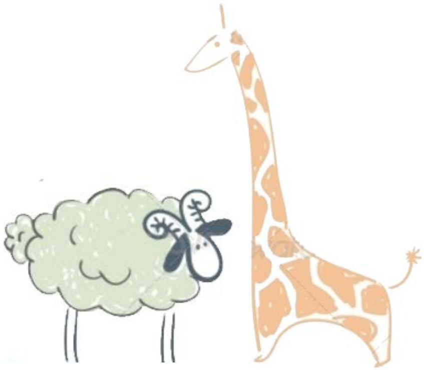
1. Suggest solutions to optimize and to improve input creation process for risk modeling

2. Give some information and analysis on aggregation criteria influence compared to classical methods

Classification methods (1/3)

- Classification is about segmenting a sample of individuals in homogeneous groups according to specifications

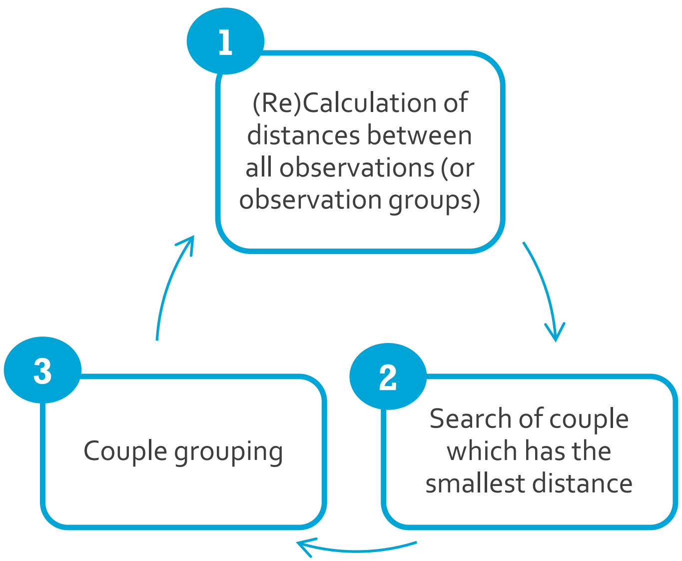
- The concept is very useful in all scientific fields

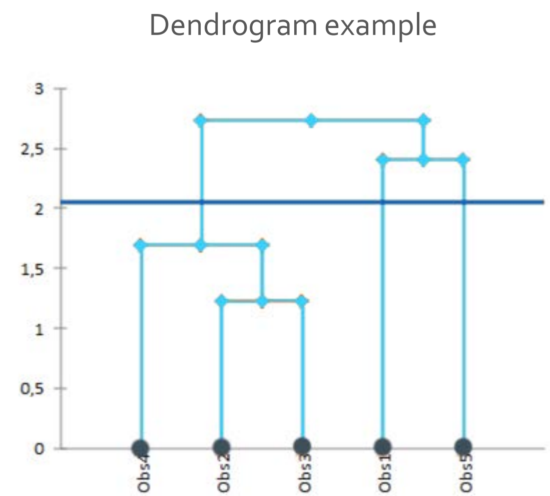


- A lot of classification families are availale

Classification methods (2/3)

- Hierarchical ascending classification (HAC)



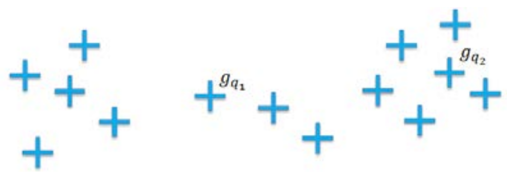


- With variations regarding distance calculation method: simple link, complete link , average link or Ward criteria

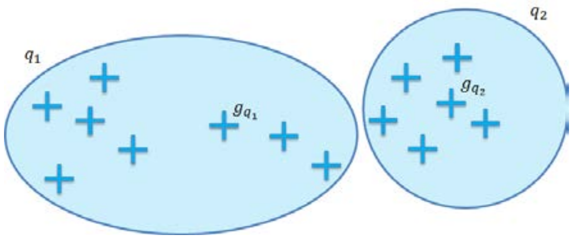
Classification methods (3/3)

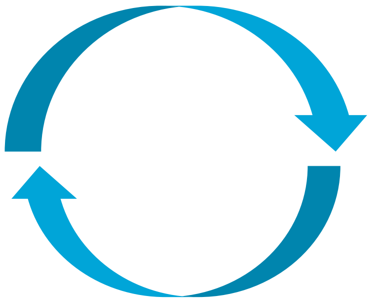
- K-means method

Step 0: Centers choices

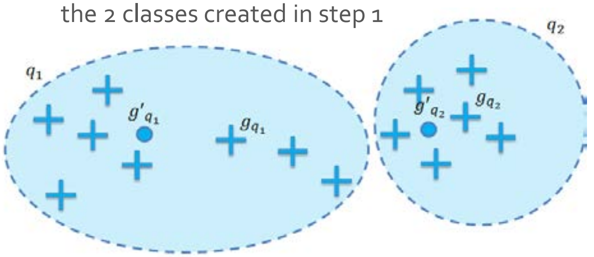


Step 1: Creation of classes around centers
q_1: points closer to g_{q_1} than g_{q_2}
q_2: points closer to g_{q_2} than g_{q_1}

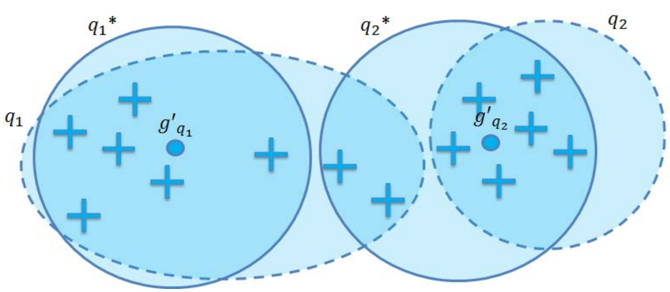




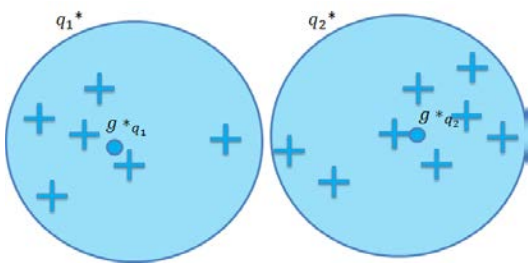
Step 2: Barycentric centers calculation ($g_{q_1}{}'$ and $g_{q_2}{}'$) of the 2 classes created in step 1



Step 3: New classes creation around barycentric points $g_{q_1}{}'$ and $g_{q_2}{}'$ (same as the step 1)



Step 4: Barycentric points calculation for classes created in step 2. If there aren't modifications of classes, algorithm is finished



Some comparisons

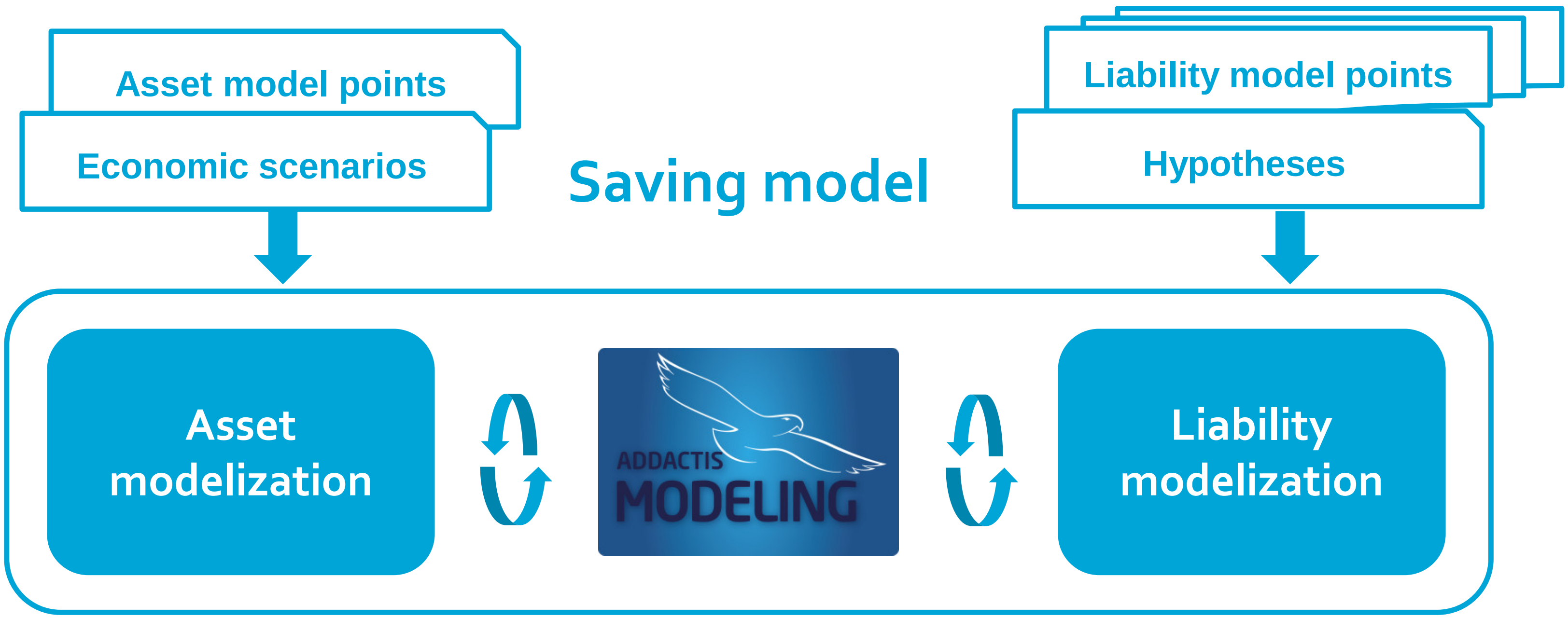
- Pro and cons for each family

Classical method	- Very easy to understand	- Necessity to define discriminant variables - Impossibility to choose a targeted class number a priori - Time calculation
HAC	- Possibility to define a targeted class number	- Aggregation quality depends on link type - Low dimensional acceptance - Time calculation
K-means	- Very fast algorithm - Acceptable quality - Unlimited dimensional acceptance	- Initialization - Not sure to find the optimal solution

- In order to have the best of both worlds, **cumulative strategies** are commonly used

Part 2 – Setting up of the application

Modeling and assumptions



- Simulated risks are **the life undertaking risk** and the market risk
- Annual frequency, projection on **40 years**, 1000 simulations
- Portfolio of **35 000 policyholders** described through **26 variables**

Process proposal

- A thorough adaptable process for any model assumptions aggregation case

1. Preliminary study: Elementary statistics to understand observation characteristics

2. PCA pretreatment: Avoid correlation, scale effect, eventually to reduce dimensions , justify us of euclidian distance

3. Aggregation: Parametrization and choice of the aggregation method

4. Completion of tables: Translation of the aggregated tables in the classical model language : use of barycentric objects

5. Automation and calculations: Automation of parametrization of model scenarios with ADDACTIS Workflow

6. Results: Comparison between scenario results

Process precision for the study

2. PCA pretreatment

- In the light of the PCA results, 2 options will be proposed
  - O1: 4 factors are used after the PCA, which means 45% of the initial information
  - O2: 15 factors are used after the PCA, which means 100% of the initial information
- Contributions of initial variables in PCA factors allow to « weight » a variable rather than another during the clustering process

3. Aggregation

- Cumulative strategies: K-mean to generate **3000 model points** then HAC with different link nuances
- Simple K-means aggregation: random initial centers, **10 repeats** of the global aggregation and **500 iterations**
- Classical average reference method: discriminant variables are age, sex and fiscal maturity

Select cases

- At the end, several liability scenarios are built

2 options
depending on PCA choices

- O1 : 4 factors
- O2 : 15 factors

1 option
no special option

4 cumulative strategies
K-mean then HAC depending on the link type

- Simple link
- Complete link
- Average link
- Ward criteria

1 non cumulative method
K-mean only

1 classical method with weighted average

7 truncation levels
depending on different aggregation levels

- 2000 model points
- 1500 model points
- 1000 model points
- 500 model points
- 100 model points
- 10 model points
- 2 model points

= **77 model points tables are inserted in the saving model**

Part 3 – Application and results

Application results

- Results are divided into 2 parts, the explanations of which are indissociable

Aggregation results

- Aggregation time for each strategies
- Aggregation quality (R^2)

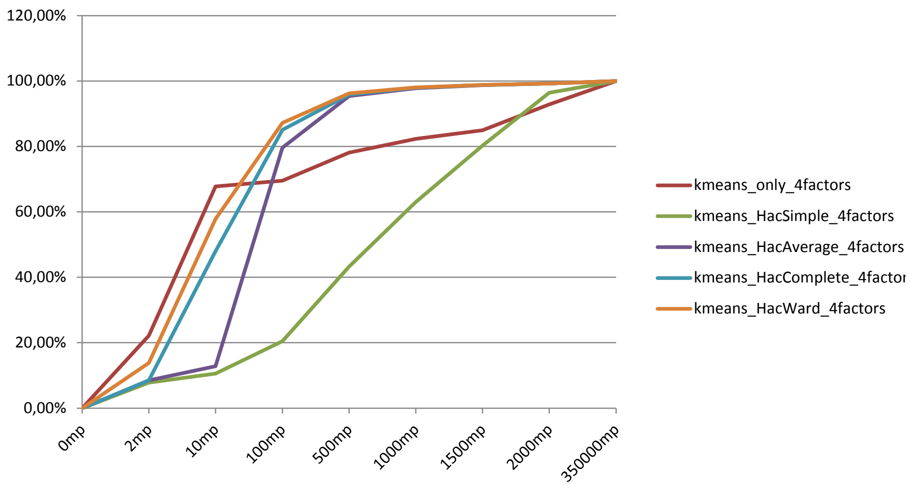
Model results

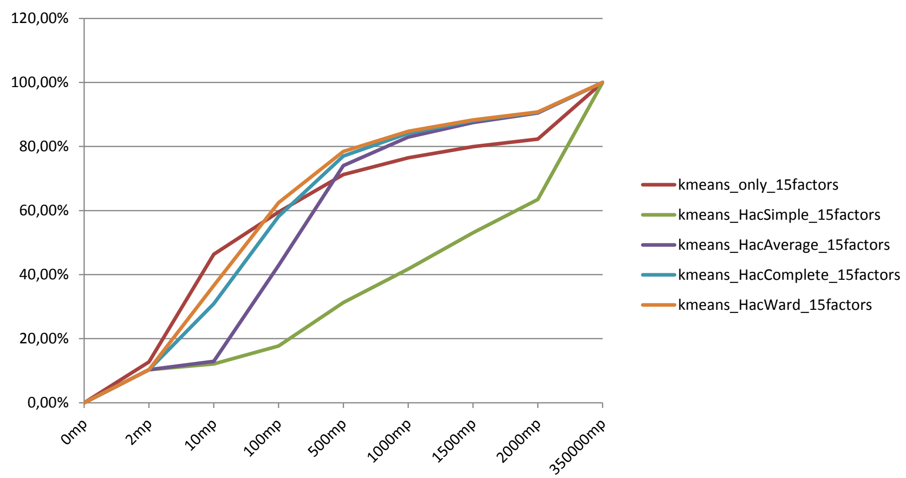
- Calculation time of the model
- Impacts on model results: BE, BOF

Conclusions

Aggregation results

- **Aggregation time**: PCA, types of link do not have influence
- About **macro studies** on R^2 curves

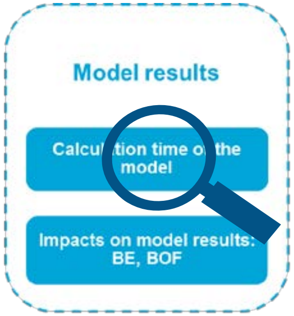




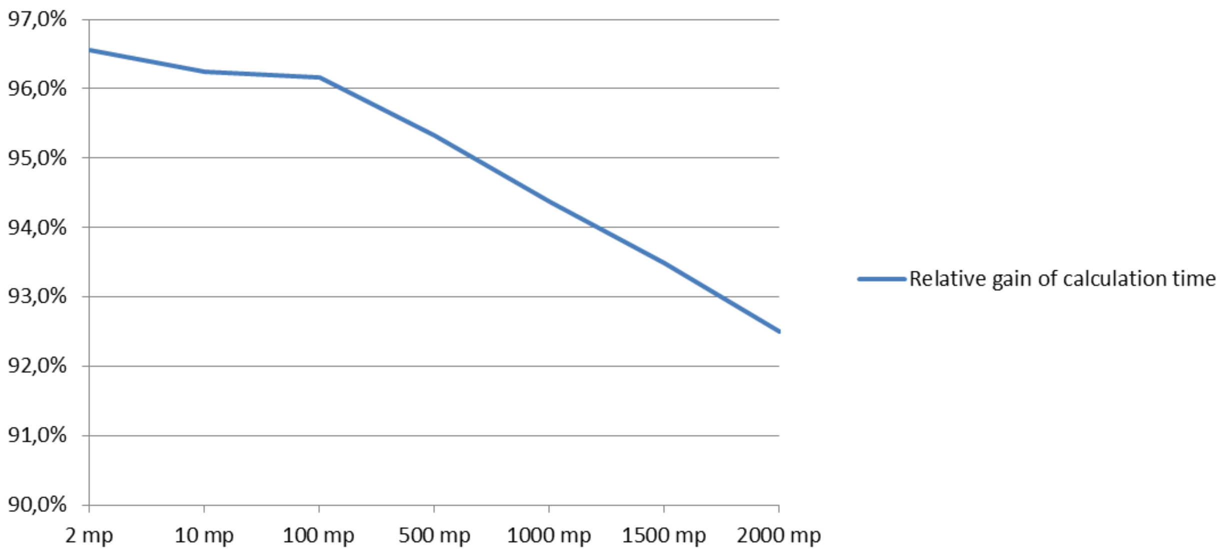
- About **micro studies** on descriptive variables
  - High volatility of variables for low truncation levels
  - Dispersion percentage which has been kept has an impact

Model results : calculation time

- **Calculation time is high** when the per-head table is used

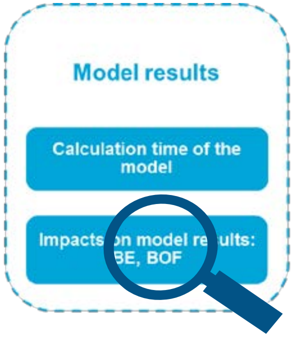


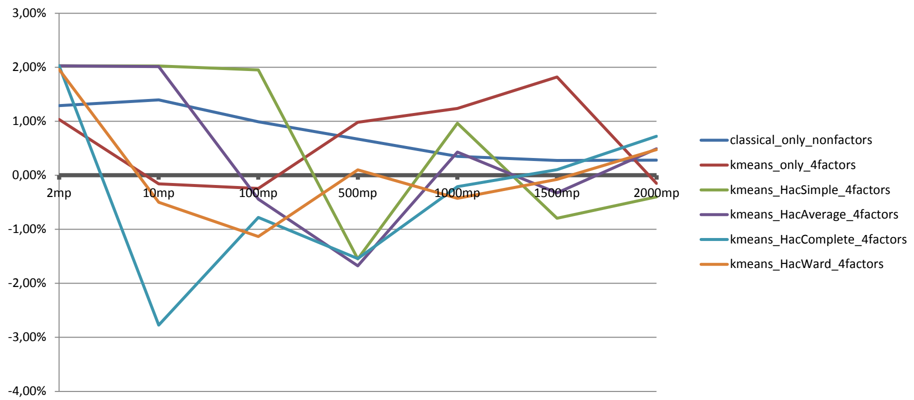
- **Calculation time is reduced** when liability tables are aggregated : no effect regarding clustering method choices

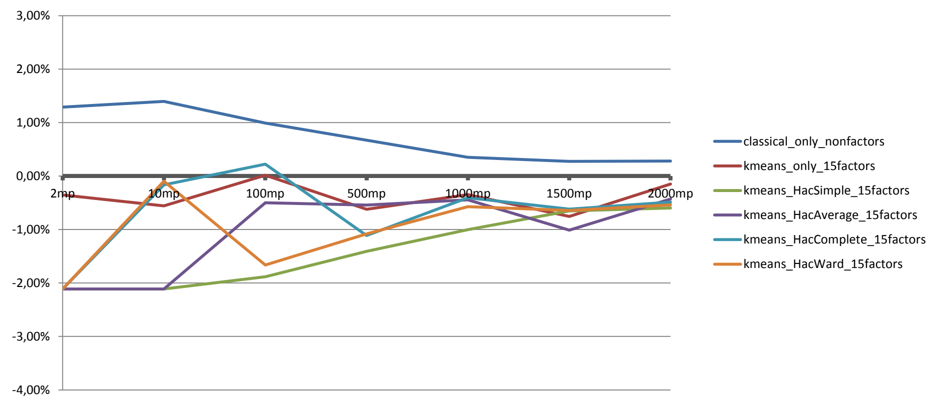


Model results : central BE

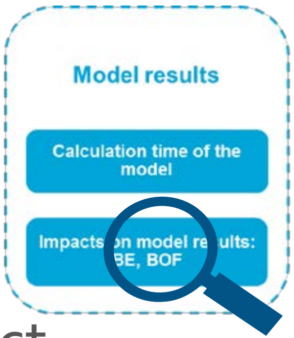
- Good results for aggregated cases
- O2 is globally better than O1, and volatility of O1 results remains high
- A hierarchy is visible between cumulative strategies
- Clustering porfolios are more benificial for the insurer, so BE differences are negative



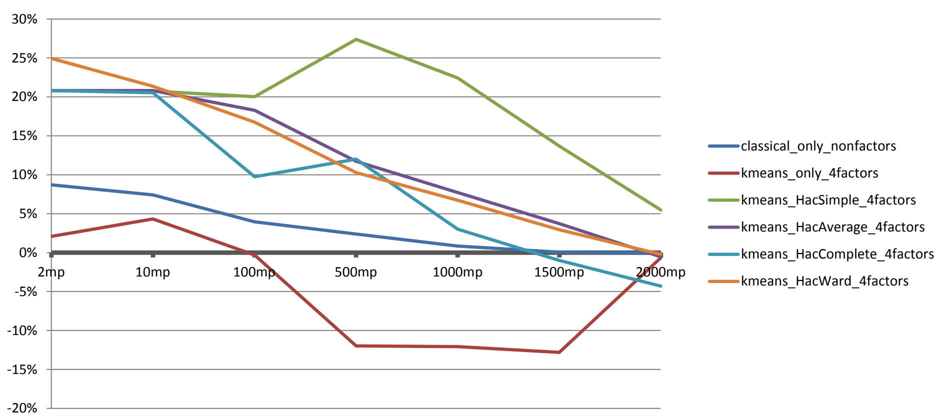


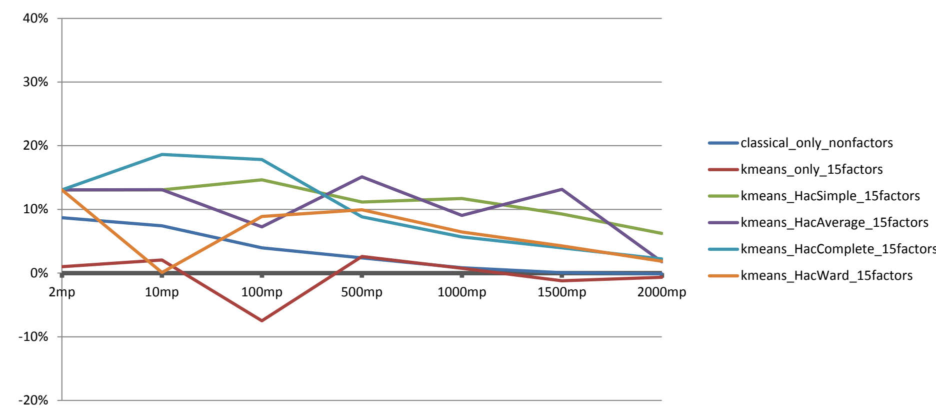


Model results : Lapse shock

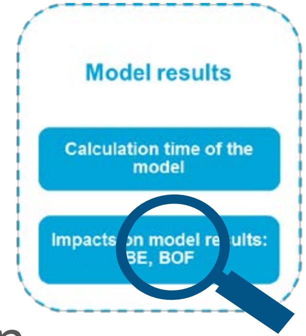


- Same observations as previous charts
- Classical method has good results ! But be careful
- O1 converges as fast as O2 : the PCA weighting effect
- As aggregated portefolio are more benefit, in case of redemption, impact is more disadvantageous

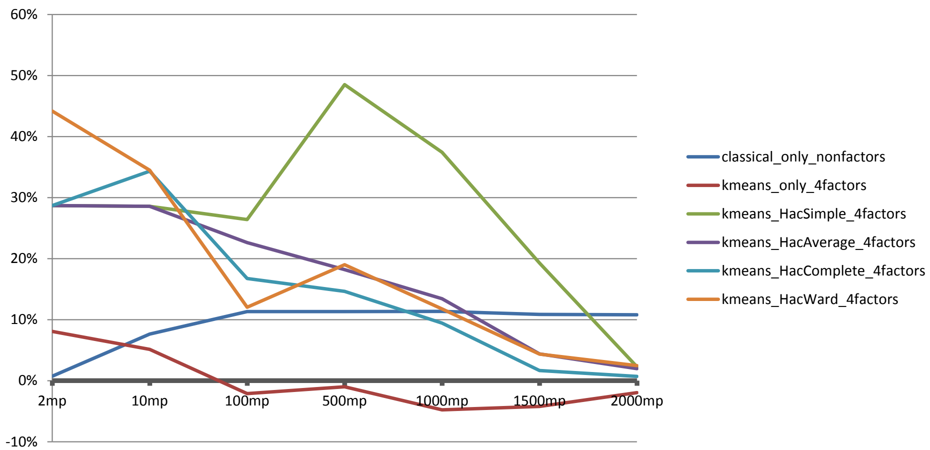


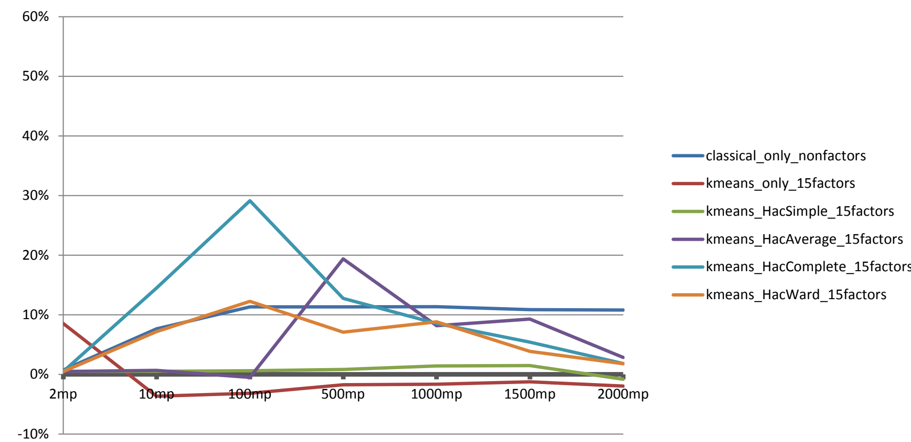


Model results : Expense shock

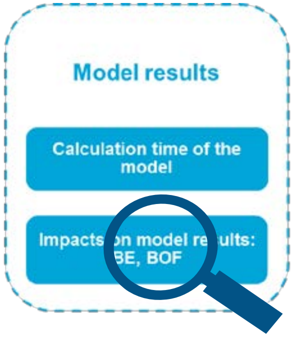


- Classical method hasn't good results this time
- O2 clearly gives better results than O1
- Some difficulties to understand correlation between aggregated characteristics and convergence results





Model results : Other shocks



- Additional informations about global indicator results are available in the article (SCR)

- Concerning other life risks **convergence conclusions are quite similar**

- And for the others shocks (market):
  - Δ BOF are negative and do not take into account in SCR calculations (for instance: longevity risk)
  - Δ BOF are positive and perfectly converge: aggregations on liability dataset don’t have an effect on capital charge (for instance: equity risk)
  - Δ BOF depend on too many interactions (for instance: interest up risk)

Limits

- Some information must be put in perspective

Limits on dataset

- Tables dimensions
- Type of data

Limits on the model

- Level of sophistication of the model
- Number of outputs

Limits on methodology

- Aggregation methods
- Choices during the PCA

Conclusions

Conclusions

Efficient process to measure information loss and to aggregate precisely liability tables

- Gain time around 94%
- Dimension reduction around 97%

Hierarchies between clustering strategies' quality

Classical method sometines good but often uncertain

Some approximations to study in a stricter context

Thank you !

ADDACTIS Worldwide
13/15 boulevard de la Madeleine
75001 PARIS
France
Tel. : +33 (0)4 81 92 13 00
Fax. : +33 (0)4 81 92 13 01

SAS au capital de 100 000 Euros - RCS: PARIS 529 256 695 - TVA intracommunautaire FR 86529256695

Reception

+33(0) 4 81 92 13 00

Worldwide actuarial software. European expertise. Local solutions.