

Société de Calcul Mathématique SA

*Outils d'aide à la décision*

*depuis 1995*



## **Le projet BEMUSE**

### **Analyse critique**

par Bernard Beauzamy

septembre 2015

# Résumé Opérationnel

## 1. Caractéristiques générales

Le projet "BEMUSE" (Best Estimate Methods – Uncertainty and Sensitivity Evaluation) regroupe 11 institutions participantes (pour la France : le CEA et l'IRSN). Il vise à comparer un certain nombre de codes de calcul, analogues à CATHARE pour dix d'entre eux, du point de vue de leur capacité à répondre à une évaluation de sûreté. Une exception est la méthode développée par l'Université de Pise, dont nous parlerons plus loin.

Notre commentaire porte sur le document "BEMUSE Phase III", daté février 2007.

Notre conclusion est simple : ce projet est entièrement dépourvu de pertinence scientifique ; il n'aurait jamais dû être lancé. Il consiste simplement à jeter des runs au hasard dans d'énormes codes de calcul et à se demander quelle est la valeur de l'exploration ainsi réalisée.

Pour répondre à cette question, prenons une comparaison très claire. La surface de la Sibérie est d'environ 13 millions de km<sup>2</sup>, soit à peu près 10<sup>19</sup> mm<sup>2</sup>. Imaginons que dix opérateurs envoient chacun une fléchette, qui explore 1 mm<sup>2</sup> : c'est une toute petite fléchette, qui n'explore pas grand'chose. La question posée par BEMUSE est : ces explorateurs auront-ils tous la même vision de la Sibérie et ces visions sont-elles comparables entre elles ?

En procédant de la sorte, nous avons une vision optimiste de BEMUSE. En effet, avec les fléchettes, la fraction explorée pour chacun est de l'ordre de 10<sup>-19</sup> (quotient surface explorée/surface à explorer). Avec un code de calcul, s'il y a 40 paramètres, chacun discrétisé en dix valeurs, la configuration à explorer représente 10<sup>40</sup> cas possibles ; si on fait 1000 runs, la fraction explorée est de l'ordre de 10<sup>-37</sup> ; le quotient  $\frac{10^{-19}}{10^{-37}} = 10^{18} = 1\,000\,000\,000\,000\,000\,000$  soit un milliard de milliards de fois plus faible.

On peut raisonnablement penser que dix opérateurs qui explorent la Sibérie en envoyant chacun une fléchette qui voit 1 mm<sup>2</sup> ont peu de chance de réaliser une bonne exploration et encore moins de chances de voir la même chose. La situation avec Cathare et les autres codes est un milliard de milliards de fois pire.

Un programme de recherche qui dirait : nous allons explorer la Sibérie avec dix fléchettes voyant chacune 1 mm<sup>2</sup> aurait peu de chances d'être financé, peu de chances d'être pris au sérieux.

## 2. Un défaut de gouvernance

Ce que montre avant tout le projet BEMUSE, c'est un défaut de gouvernance au plus haut niveau. Les différents acteurs se contentent de jeter des runs au hasard, sans la moindre réflexion quant à la stratégie globale. Personne, par exemple, ne prend l'initiative d'une décision (pour tant essentielle) : définissons correctement la situation d'usage (les modalités de fonctionnement du réacteur au moment de l'accident). Un ensemble de réunions préliminaires visant à se mettre

d'accord sur la situation à considérer, puis sur les valeurs à prendre pour les différents paramètres, était pourtant essentiel.

### **3. Les méthodes mathématiques utilisées**

Il n'y a absolument aucune réflexion quant aux méthodes mathématiques utilisées pour le traitement. Tous les traitements statistiques utilisés reposent sur la méthode de Wilks, dont nous avons montré qu'elle était incorrecte. Voir :

[http://www.scmsa.eu/archives/BB\\_Wilks\\_2016\\_01\\_11.pdf](http://www.scmsa.eu/archives/BB_Wilks_2016_01_11.pdf)

Le projet BEMUSE ne comporte aucune réflexion, aucune interrogation, quant à cette méthode. Dans d'autres cas, on envisage la mise en œuvre de tests statistiques spécifiques, sans aucune réflexion quant à leur validité.

Il y a là encore un défaut méthodologique majeur. Les méthodes probabilistes, comme la méthode de Wilks, sont du niveau Terminale ; elles n'ont aucun caractère de confidentialité, bien au contraire ! Les participants auraient pu se demander si, après avoir obtenu leurs résultats, ils les traitaient avec les outils appropriés. Il n'existe aucune amorce de réflexion sur ces questions dans le projet BEMUSE. On a le sentiment que les connaissances probabilistes des participants sont extrêmement faibles, mais, en même temps, que c'est une question qu'ils traitent par le mépris.

### **4. Les résultats**

Comme l'exemple de la Sibérie le laissait attendre, les résultats montrent une très grande hétérogénéité. Les valeurs de température obtenues sont très différentes, mais aussi les intervalles de confiance associés. Comme dit le rapport, page 75 "il vaut mieux que les Autorités de Sûreté ne voient pas cela". L'ensemble du rapport semble discréditer complètement l'utilisation des gros codes de calcul, mais nous considérons que, en réalité, il montre que l'utilisation en aveugle est absurde. Nous maintenons donc les termes de notre précédent rapport : de tels codes peuvent être utiles, à condition de bien spécifier le scénario dans lequel on les emploie.

### **5. Le cas de l'Université de Pise**

Elle a développé une méthode appelée CIAU (code with capability of internal assessment of uncertainty) qui utilise de manière permanente une base de données expérimentale. Les différences entre valeurs observées et valeurs calculées sont propagées pour corriger les valeurs prévues par le code.

Nous ne connaissons pas le détail de cette méthode, mais il est certain qu'utiliser des valeurs expérimentales et s'en servir pour corriger les prévisions est de bon sens. Les autres méthodes ne font aucune référence aux valeurs expérimentales et n'ont aucun moyen de les utiliser.

## I. Description du projet

Le présent rapport concerne le document "BEMUSE PHASE III REPORT ; Uncertainty and Sensitivity Analysis of the LOFT L2-5 Test", daté février 2007. Les institutions participantes sont :

CEA, France ..... P. Bazin, A. de Crécy  
GRS, Germany..... H. Glaeser, T. Skorek  
IRSN, France ..... J. Joucla, P. Probst  
KAERI, South Korea..... B. Chung  
KINS, South Korea..... D.-Y. Oh  
NRI-1, Czech Republic..... M. Kyncl, R. Pernica  
NRI-2, Czech Republic..... J. Macek, R. Meca  
PSI, Switzerland ..... R. Macian  
UNIPI, Italy ..... F. D'Auria, A. Petruzzi  
UPC, Spain ..... M. Perez, F. Reventos  
JNES, Japan..... K. Fujioka

Le projet BEMUSE comporte six phases au total :

Phase 1: Presentation a priori of the uncertainty methodology to be used by participants

Phase 2: Best-estimate calculation of the LOFT L2-5 test, re-analysis of the ISP-13 exercise

Phase 3: Uncertainty and sensitivity analyses of LOFT L2-5, first conclusions on the methods and suggestions for improvement

Phase 4: Best-estimate calculation of the LB-LOCA transient applied to the NPP

Phase 5: Uncertainty and sensitivity analyses for this transient, taking into account the improvements resulting from Phase 3

Phase 6: Status report on the area, classification of the methods, conclusions and recommendations

Le but recherché est de comparer les codes de calcul permettant les évaluations de sûreté dans le cas de grosses brèches dans le système de refroidissement primaire d'un réacteur nucléaire. Plus précisément, les participants voudraient :

- To evaluate the practicability, quality and reliability of Best-Estimate (BE) methods including uncertainty and sensitivity evaluation in applications relevant to nuclear reactor safety
- To develop common understanding
- To promote and facilitate their use by the regulatory bodies and the industry.

La comparaison porte sur 5 codes différents (et certains ont plusieurs versions) :

ATHLET: 2 participants

CATHARE: 2 participants

MARS: 1 participant

RELAP5: 4 participants

TRACE: 2 participants

L'Université de Pise participait avec une méthode complètement différente, dont nous avons parlé brièvement dans le résumé.

Les résultats donnés par les codes de calcul peuvent être comparés avec les mesures faites sur un "modèle réduit" de réacteur, appelé LOFT. La variable de sortie principale est une température, notée Max\_TC, qui représente la température la plus élevée atteinte dans le réacteur (indépendamment de son niveau de puissance au moment de l'accident).

La description du projet est la suivante :

Part 1. The input parameters and their uncertainties.

Step 1. List the general sources of uncertainties considered for BEMUSE Phase 3.

Step 2. How is the list of uncertain input parameters established?

Step 3. How are the uncertainties of the uncertain input parameters quantified?

Step 4. List the uncertain input parameters and quantify their uncertainties: the synthesis.

Part 2. Uncertainty analysis.

Step 5. Sampling for the probabilistic approach.

Step 6. Running the code.

Step 7. First uncertainty analysis results.

Part 3. Sensitivity analysis.

Step 8. Sensitivity analysis.

Part 4. Complementary methods, assessment methodology.

Step 9. Complementary methods.

Step 10. Assessment methodology.

Pour chacun des paramètres d'entrée, les participants définissent une "incertitude", à dire d'expert.

## II. Conduite méthodologique

On constate, et ceci est vraiment essentiel, que les participants ne se mettent pas d'accord au préalable sur une situation de fonctionnement (par exemple : le réacteur est à son niveau de puissance maximal). Ceci est réellement absurde, et montre bien l'absence de coordination scientifique du projet. Le niveau de puissance est considéré par chacun des participants comme un paramètre à sa discrétion ; il ne sera pas choisi de la même façon par tous, selon les runs qui sont effectués. Mais le niveau de puissance au moment de l'accident conditionne évidemment la température qui sera atteinte.

Le nombre de paramètres d'entrée varie selon les participants : de 13 à 64 selon les cas. Ceci rend évidemment les comparaisons difficiles, parce que les paramètres d'entrée sont différents (en nombre et en nature). Il aurait fallu spécifier (en termes grossiers) une situation de fonctionnement, que chacun des participants aurait ensuite convertie en valeurs pour ses propres paramètres d'entrée. Cette condition, évidente et essentielle, n'a pas été remplie. Aucune allusion n'y est faite.

### III. Différences selon les codes

Un même code de calcul ne donne pas la même chose, même s'il est alimenté avec les mêmes valeurs d'entrée : cela dépend de la version et de l'utilisateur. On note d'importantes différences pour le code CATHARE : "The most important user's effect seems to be related to both users of CATHARE, but there is an explanation to this result: CEA has corrected a mistake in the friction form loss in the accumulator line, found after completion of phase 2 whereas IRSN has kept the results of phase 2 unchanged, with the aim of performing a blind study for phases 2 and 3. Before this correction, both time trends were very close." En effet, selon la figure, les valeurs données par le code peuvent différer de plus de 100°K selon les étapes. Il y a même des situations où les codes ne donnent rien, et on se demande alors quoi faire du résultat. Ceci s'appelle "code failures" ; voir table 9 dans le rapport.

La détermination des incertitudes, selon le rapport, se fait selon trois méthodes : "literature review, fitting of experimental data and expert judgement." Il y a très peu de données expérimentales, si bien que les deux autres méthodes sont les plus utilisées.

Conformément à l'approche retenue, les incertitudes sont globales. Elles ne prennent pas en compte les conditions de fonctionnement au moment de l'accident, ce qui constitue à l'évidence une faute méthodologique majeure. En effet, les incertitudes retenues "à dire d'expert" correspondent à des conditions générales de fonctionnement, qui peuvent être tout à fait différentes de celles rencontrées lors de l'accident.

Pour faire comprendre ceci en langage imagé, on peut s'intéresser à la capacité d'un conducteur à maintenir son véhicule en ligne droite, et conclure que, en temps normal, il ne dévie pas de plus de 20 cm à gauche ou à droite, par rapport à la trajectoire idéale. Ceci peut être vrai sur autoroute à 130 km/h, mais risque d'être complètement inexact sur route de montagne avec des lacets, surtout si la vitesse est excessive.

L'idée fondamentale, sous-jacente au projet BEMUSE, selon laquelle on peut faire une étude globale, c'est-à-dire faire l'économie de "scénarios" venant caractériser telle ou telle situation d'exploitation, est fondamentalement incorrecte.

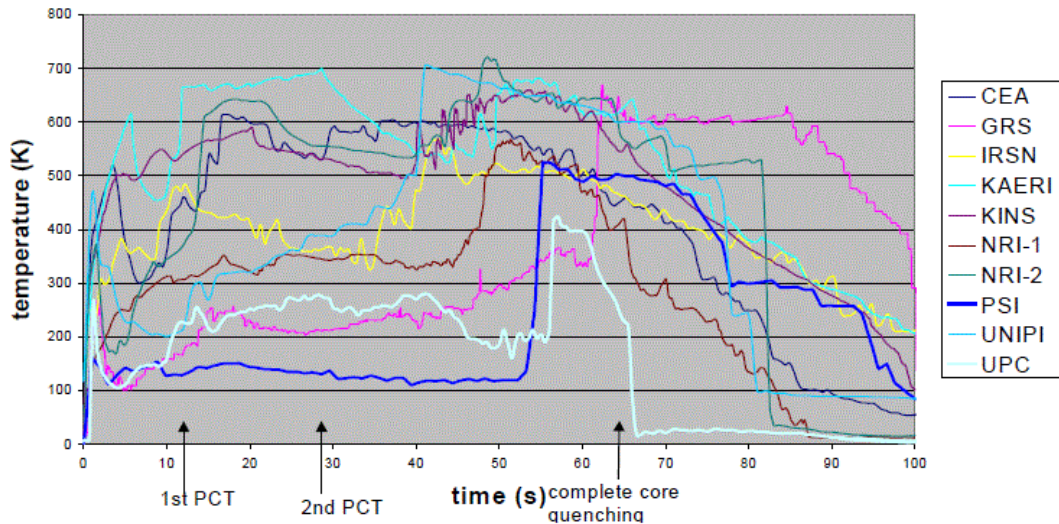
Les lois de probabilité relatives aux diverses incertitudes sont fixées "à dire d'expert" ; ce sont des lois uniformes, des lois normales ou log-normales tronquées. Ces lois sont données sans aucune justification.

### IV. Traitement mathématique

Tous les participants procèdent à des tirages aléatoires (chacun le sien, bien sûr), sans aucune homogénéisation des méthodes. Après quoi, les résultats découlent de l'utilisation de la méthode de Wilks (que nous avons critiquée, voir ref. plus haut) pour déterminer un intervalle de confiance pour chacun des paramètres. Bien entendu, en sortie, les résultats sont différents "No general trend can be observed for the form of the distributions", comme on pouvait s'y attendre depuis le début !

Pour la température maximale observée (la variable de sortie la plus importante), les résultats diffèrent considérablement d'un participant à l'autre, en valeur comme en forme de la courbe au cours du temps. Voici la largeur de l'incertitude, variant au cours du temps, selon les participants :

Figure 7. **Maximum cladding temperature: uncertainty band widths**



De tels résultats sont aberrants, et discréditent complètement l'approche retenue. Le rapport conclut néanmoins que "The majority of the participants find an uncertainty band which envelops the experimental data: for example in 8 cases out of 10 for the time trend "maximum cladding temperature" (Max\_TC), in roughly 6 cases for the time trend upper plenum pressure." Mais c'est au prix d'une valeur retenue pour l'incertitude qui est considérable, et totalement inexploitable en pratique.

## V. Analyse de sensibilité

Il s'agit en principe de voir quelle influence peut avoir une variation d'un paramètre sur le résultat final (une température). Si le code était une fonction explicite de plusieurs variables, la sensibilité correspondrait à l'évaluation d'une dérivée partielle. Ici, bien entendu, la fonction n'est pas explicite, et on ne dispose que d'un petit nombre de runs. Les participants font usage de divers tests statistiques qu'ils maîtrisent mal : Spearman's Rank Correlation Coefficients, Standardised Rank Regression Coefficients, Partial Rank Correlation Coefficients, etc., avec toujours cette erreur fondamentale : ils veulent faire une analyse de sensibilité globale, alors qu'elle dépend étroitement des conditions de fonctionnement. Pour reprendre notre exemple précédent, la sensibilité de la voiture aux écarts du conducteur sera différente selon que l'on est sur autoroute ou dans des lacets de montagne.

## VI. Hiérarchisation des paramètres

Cette question est considérée dans le rapport comme d'intérêt limité, dans la mesure où les paramètres ne sont pas les mêmes pour tous les participants, ni les lois de probabilité associées.

La SCM avait été consultée par l'IRSN sur la hiérarchisation des paramètres intervenant dans CATHARE : nous ne connaissions pas l'existence du projet BEMUSE et nous avons uniquement à notre disposition un tableau EXCEL représentant les différents runs. Nous avons mis en œuvre notre méthode de hiérarchisation, décrite dans le livre [NMP]. Nos résultats étaient différents de ceux obtenus par l'IRSN. Nous ne savons pas si nos résultats ont été utilisés par l'IRSN dans le cadre du projet BEMUSE.

Là encore, la remarque fondamentale (qui n'est pas faite) est que cette hiérarchisation dépend nécessairement de la situation. Une hiérarchisation globale n'a pas de sens. Les résultats obtenus diffèrent considérablement d'un participant à l'autre "Only some very crude trends can be observed". Le rapport conclut : "As a conclusion, the initial purpose of this study, separate in the influence the effect of the sensitivity from that of the range of variation, has not been successful."

## VII. Deux défauts majeurs

En résumé, les défauts majeurs de cette approche sont les suivants :

- Vouloir faire une exploration par runs choisis aléatoirement, sans se préoccuper d'aucun scénario. Le bon sens montre (avec l'exemple de l'exploration de la Sibérie par dix fléchettes) que cette approche ne peut produire aucun résultat ;
- Ne pas se préoccuper un instant de la validité des méthodes mathématiques utilisées pour traiter les résultats, alors que ces méthodes, bien connues et très élémentaires, peuvent faire l'objet d'une analyse critique distincte.

## VIII. Référence

[NMP] Bernard Beauzamy : Nouvelles Méthodes Probabilistes pour l'évaluation des risques. SCM SA. ISBN 978-2-9521458-4-8. ISSN 1767-1175, avril 2010.