DE LA RECHERCHE À L'INDUSTRIE

Inverse uncertainty quantification of input model parameters in thermal-hydraulic simulation

Guillaume Damblin

CEA Saclay, Energies division

Workshop on calibration of numerical code

May 31, 2023

Prelude

Thank you to my colleagues and co-authors for the great collaboration!

Summary

1. Context and motivations
2. The CIRCE method
3. Multi-group CIRCE
4. Related works

Context and motivations

Thermal-hydraulic simulation in the nuclear field

- Needed to simulate power plants (innovative or in operation) as well as for safety analyses of hypothetical accidental scenarios.
- Simulations run by **Best Estimate** computer codes with a great effort to V&V (ex: the CATHARE code in CEA).
- Based on balance equations (mass, momentum and energy) which require **closure models** due to time and space averaging.
- Example: the energy equation applied to a control volume:

$$\rho \overbrace{\frac{\partial I}{\partial t}}^{(1)} = - \overbrace{\text{grad}(\dot{\mathbf{q}}'')}^{(2)} + \overbrace{\dot{q}'''}^{(3)} + \overbrace{\frac{\partial P}{\partial t}}^{(4)} + \overbrace{\phi}^{(5)}$$

(1) variation in time of enthalpy, **(2) and (3) are heat fluxes which should be modeled by (semi)-empirical models**, (4) variation in time of the pressure, (5) dissipation function.

From modeling to simulation

"Nominal" closure models:

- Established by means of both expertise and well-chosen experimental data, and denoted by

$$M_{nom}(\mathbf{x}),$$

with $\mathbf{x}$ being thermal-hydraulic and design variables,

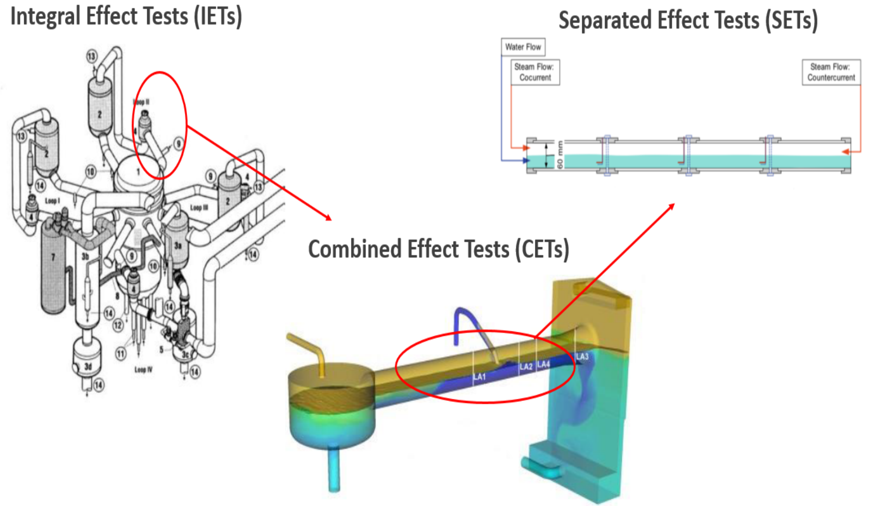
- Experimental uncertainty on $\mathbf{x}$ may occur but is often neglected in practice.

Simulations:

i) Implemented from appropriate closure models $\Longrightarrow$ numerical uncertainties to check (verification task).

ii) Comparisons between the simulations and corresponding experimental data (validation task).

iii) Assessment of model uncertainty from the discrepancy between the two $\Longrightarrow$ **IUQ (inverse uncertainty quantification).**


Integral Effect Tests (IETs)
Separated Effect Tests (SETs)
Combined Effect Tests (CETs)
Water Flow
Steam Flow: Cocurrent
Steam Flow: Countercurrent
60 mm
Loop I
Loop II
Loop III
Loop IV
LA1
LA2
LA4
LA3

The CIRCE method

CIRCE $=$ *Calcul des Incertitudes Relatives aux Corrélations Élémentaires* (De Crécy and Bazin, 2001).

Main assumptions:

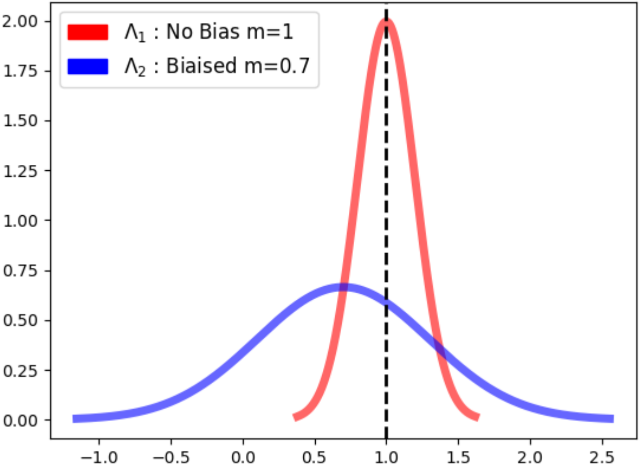
- Model uncertainty is multiplicative:

$$M_\lambda(\mathbf{x}) = \lambda \times M_{nom}(\mathbf{x})$$

- λ is modeled as a probability distribution $\Longrightarrow M_\lambda$ is aleatory,
- λ is Gaussian $\mathcal{N}(m, \sigma^2)$,
- Model uncertainty is known as "unbiased" if $m = 1$.

We wish that the bias, equal to $1 - \hat{m}$, is as small as possible.

Biased versus unbiased distributions



CIRCE is an inverse problem

Probabilistic link between the experimental data and simulations:

- $y_i \in \mathbb{R}$ the experimental QoI at $\mathbf{x}_i$ (for $i \in [\![1; n]\!]$),
- G the CATHARE code (used as a black-box),
- We assume that

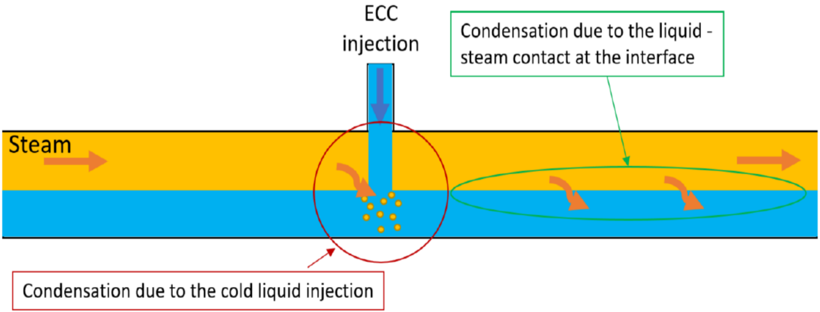
$$\begin{aligned} y_i &= G(\overbrace{M_{\lambda_{1,i}}(\mathbf{x}_i), \cdots, M_{\lambda_{p,i}}(\mathbf{x}_i)}^{Aleatory\ closure\ models}) + \epsilon_i \\ &= G_{\lambda_i}(\mathbf{x}_i) + \epsilon_i \end{aligned}$$

where

- $\lambda_i = (\lambda_{1,i}, \cdots, \lambda_{p,i})^T \in \mathbb{R}^p$ with $\lambda_{j,i} \sim \mathcal{N}(m_j, \sigma_j^2)$, $j \in [\![1; p]\!]$.
- $\epsilon_i \sim \mathcal{N}(0, \sigma_{\epsilon_i}^2)$.

The CIRCE method jointly estimates all the m_j and σ_j^2 (SETs: $p = 1$, CETs: $p \geq 2$).

Ex : Emergency Core Cooling injection (Cocci, 2022)



This is a CET: $p = 2$ condensation phenomena, modeled respectively by $M_{\lambda_1}(.)$ and $M_{\lambda_2}(.)$, take place at the same time.

CIRCE consists of three main steps:

1) Linearization of the simulations in λ at the nominal model, i.e. $\lambda^\star = \mathbf{1}_p$

$$y_i - G_{\lambda^\star}(\mathbf{x}_i) = h_i^T(\lambda_i - \lambda^\star) + \epsilon_i, \quad i \in [\![1; n]\!].$$

2) Joint computation of **Maximum Likelihood estimates** $(\hat{m}_j, \hat{\sigma}_j^2)$ if $H = [h_1, \cdots, h_n]^T \in \mathcal{M}_{n,p}(\mathbb{R})$ is full rank.

3) Confirmation of the results through a posterior inspection of both linearity and normality assumptions.

In this presentation, we are focusing on the second step!

Notations used in the sequel

- Mean parameters $m := (m_1, \cdots, m_p)^T \in \mathbb{R}^p$,
- Variance parameters $\sigma^2 := (\sigma_1^2, \cdots, \sigma_p^2)^T \in \mathbb{R}^p$,
- Shifted observed data $Y := (y_1 - G_{\lambda^\star}(\mathbf{x}_1), \cdots, y_n - G_{\lambda^\star}(\mathbf{x}_n))^T \in \mathbb{R}^n$,
- Shifted latent data $\lambda := (\lambda_1 - \lambda^\star, \cdots, \lambda_n - \lambda^\star)^T \in \mathcal{M}_{n,p}(\mathbb{R})$,
- Shifted complete data $Z := \{Y, \lambda\}$.

Complete and marginal likelihood

- Complete likelihood: $L(Z|m,\sigma^2) = \overbrace{L(Y|\lambda,m,\sigma^2)}^{Gaussian}\overbrace{L(\lambda|m,\sigma^2)}^{Gaussian}$ with

$$L(Y|\lambda,m,\sigma^2) \propto \prod_{i=1}^{n} \exp\Big(-\frac{1}{2}\frac{\big(Y_i - h_i^T\lambda_i\big)^2}{\sigma_{\epsilon_i}^2}\Big)$$

and

$$L(\lambda|m,\sigma^2) \propto \prod_{i=1}^{n} |\text{diag}(\sigma^2)|^{-1/2} \exp\Big(-\frac{1}{2}(\lambda_i - m)^T\text{diag}(\sigma^2)^{-1}(\lambda_i - m)\Big).$$

- Marginal likelihood: integrating over λ leads to the likelihood of the observed data only, still Gaussian:

$$L(Y|m,\sigma^2) \propto \prod_{i=1}^{n}(h_i^T\text{diag}(\sigma^2)h_i+\sigma_{\epsilon_i}^2)^{-1/2} \exp\Big(-\frac{1}{2}\frac{(Y_i - h_i^T m)^2}{h_i^T\text{diag}(\sigma^2)h_i + \sigma_{\epsilon_i}^2}\Big).$$

The ECME algorithm for MLE

ECME $=$ *Expectation-Conditional Maximization Either* (Celeux et al., 2010)

1. Step of Expectation (E): calculation of

$$Q((m,\sigma^2),(m^k,\sigma^{2,k})) = \mathbb{E}_\lambda[l(Z|m,\sigma^2)|Y,m^k,\sigma^{2,k}].$$

   The expectation is taken with respect to the distribution of λ conditional on $(Y, m^k, \sigma^{2,k})$.

2. Steps of Conditional Maximization (CM):

   - CM1: $\sigma^{2,k+1} = \underset{\sigma^2}{\operatorname{argmax}}\, Q((m,\sigma^2),(m^k,\sigma^{2,k}))$,
   - CM2: $m^{k+1} = \underset{m}{\operatorname{argmax}}\, l(Y|m^k,\sigma^{2,k+1})$.

- CM1 and CM2 have **analytic expressions** as functions of $(m^k, \sigma^{2,k})$.
- Starting from a first sample (m_0, σ_0^2), the convergence of the ECME algorithm is **faster than that of EM.**

CIRCE using CETs in practice

CIRCE on CETs

- $p \geq 2$ factors estimated jointly from Y, often in a small data context $(50 \leq n \leq 200)$,
- If $H_{\star 1} >> H_{\star 2}$ (case $p = 2$), then the estimators $(\hat{m_2}, \hat{\sigma}_2^2)$ and $(\hat{m_1}, \hat{\sigma}_1^2)$ may be respectively inaccurate and degraded.

Multi-stage CIRCE: (Cocci et al., 2022)

- CETs can still be used, but to estimate only the dominant factor, say λ_1, while neglecting the other ones,
- If being known, the uncertainty of the other factors λ_j $(2 \leq j \leq p)$ adds up to the experimental uncertainty (case $p = 2$ below):

$$Y_i = h_{i1}\lambda_{1,i} + \epsilon_i \quad \text{with} \quad \epsilon_i \sim \mathcal{N}\left(h_{i2}m_2, h_{i2}^2\sigma_2^2 + \sigma_{\epsilon_i}^2\right).$$
$$\Longleftrightarrow$$
$$Y_i - h_{12}m_2 = h_{i1}\lambda_{1,i} + \epsilon_i \quad \text{with} \quad \epsilon_i \sim \mathcal{N}\left(0, h_{i2}^2\sigma_2^2 + \sigma_{\epsilon_i}^2\right).$$

Multi-group CIRCE

Multi-group CIRCE (Damblin et al., 2023)

Motivation:

The model uncertainty may not be the same across the whole set of experimental tests Y. How to statistically check on it ?

- Y is now made up of s groups of different experimental setups:

$$Y := (Y_1, \cdots, Y_s, \cdots, Y_l)^T \in \mathbb{R}^n \quad ; \quad 1 \leq s \leq l$$

- A variance parameter σ_s^2 is estimated for each group jointly to a mean parameter m common to every group.
- For example, Y_s may have a specific geometry or thermal-hydraulic input range.
- Let i_s denote the last index of the s-th group. Then,

$$i_{s-1} + 1 \leq i \leq i_s \implies \lambda_i \sim \mathcal{N}(m, \sigma_s^2)$$

Multi-group complete likelihood:

$$L(Z|m,\sigma_1^2,\cdots,\sigma_l^2) = L(Y|\lambda,m,\sigma_1^2,\cdots,\sigma_l^2)L(\lambda|m,\sigma_1^2,\cdots,\sigma_l^2)$$

with

$$L(\lambda|m,\sigma_1^2,\cdots,\sigma_l^2) \propto \prod_{s=1}^{l}\prod_{i=i_{s-1}+1}^{i_s}\Big[|\text{diag}(\sigma_s^2)|^{-1/2} \exp\Big(-\frac{1}{2}(\lambda_i-m)^T\text{diag}(\sigma_s^2)^{-1}(\lambda_i-m)\Big)\Big].$$

Both E. and CM. steps of the multi-group ECME are **still analytic**, and thus the MLE is readily computable.

The Wald test in the multi-group CIRCE

We aim to evaluate the degree of statistical evidence that the variances of the groups are different to one another.

- Null hypothesis: $\mathcal{H}_0 \ : \ \sigma_s^2 - \sigma_{s'}^2 = 0, \quad 1 \leq s \neq s' \leq l.$
- The **Wald's statistic** is written as:

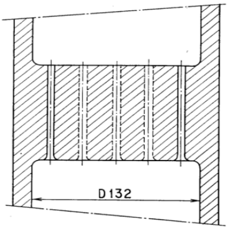
$$W = \frac{(\hat{\sigma}_s^2 - \hat{\sigma}_{s'}^2)^2}{\mathbb{V}[\hat{\sigma}_s^2] + \mathbb{V}[\hat{\sigma}_{s'}^2] - 2\text{Cov}(\hat{\sigma}_s^2, \hat{\sigma}_{s'}^2)} \sim \chi^2(1) \ \text{ under } \ \mathcal{H}_0,$$

  with $\chi^2(1)$ denoting the chi-square distribution with one degree of freedom.

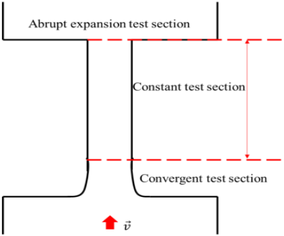
- The test can be applied to each pair of indexes $1 \leq s \neq s' \leq l$.

Application to the critical mass flow

- Discharge of coolant flow due to pressure drop at the break.
- The mass flow rate reaches a maximum value called **critical mass flow** (or chocked flow).
- Several types of SETs for this phenomenon, including BETHSY Nozzle 2 (B2) and BETHSY Nozzle 6 (B6).



a) BETHSY 2"



b) BETHSY 6"

Application to the critical mass flow

Is the uncertainty influenced by the geometry?

- $p = 1 : M_{nom}$ is the so-called flashing model,
- $Y = Y_{B2} \cup Y_{B6}$ with $n_{B2} = 25$ and $n_{B6} = 24$,
- Simulations run with the CATHARE code,
- Log-Linearization was more accurate $\Longrightarrow \lambda \sim \mathcal{LN}(m, \sigma^2)$,
- The multi-group ECME gives:

$$\hat{m} = 0.57 \text{ and } (\hat{\sigma}^2_{B2}, \hat{\sigma}^2_{B6}) = (0.31, 0.13).$$

- $W = 3.62$ and $\mathbb{P}[\chi^2(1) \leq 3.84] = 0.95$. The equality of variances is thus not rejected at the 5% level.

Related works

Related works

- Non-linear CIRCE (Barbillon et al., 2011), Bayesian CIRCE (Damblin and Gaillard, 2020).
- Assessment of the **adequacy** of experimental databases through the criteria of representativeness and completeness (Baccou et al., 2019).
- On going-OCDE project, named **ATRIUM**, on the realization of IUQ methods by numerous worldwide participants.
- Scaling issue: do the uncertainties remain valid on IETs or ultimately on an actual power plant?

Bibliography

Baccou, J., Zhang, J., Fillion, P., Damblin, G., Petruzzi, A., Mendizábal, R., Reventós, F., Skorek, T., Couplet, M., Iooss, B., Oh, D.-Y., and Takeda, T. (2019). Development of good practice guidance for quantification of thermal-hydraulic code model input uncertainty. *Nuclear Engineering and Design*, 354 :110173.

Barbillon, P., Celeux, G., Grimaud, A., Lefebvre, Y., and De Rocquigny, E. (2011). Nonlinear methods for inverse statistical problems. *Computational Statistics and Data Analysis*, 55(1) :132–142.

Celeux, G., Grimaud, A., Lefèbvre, Y., and de Rocquigny, E. (2010). Identifying intrinsic variability in multivariate systems through linearized inverse methods. *Inverse Problems in Science and Engineering*, 18.

Cocci, R. (2022). *Statistical Learning and inverse uncertainty quantification in nuclear thermal-hydraulic simulation : application to the condensation modelling at the safety injection*. PhD thesis, Université Paris Saclay.

Cocci, R., Damblin, G., Ghione, A., Sargentini, L., and Lucor, D. (2022). Extension of the circe methodology to improve the inverse uncertainty quantification of several combined thermal-hydraulic models. *Nuclear Engineering and Design*, 398.

Damblin, G., Bachoc, F., Gazzo, S., Sargentini, L., and Ghione, A. (2023). A generalization of the circe method for quantifying input model uncertainty in presence of several groups of experiments. *Submitted to Nuclear Engineering and Design*.

Damblin, G. and Gaillard, P. (2020). Bayesian inference and non-linear extensions of the circe method for quantifying the uncertainty of closure relationships integrated into thermal-hydraulic codes. *Nuclear Engineering and Design*, 369(6).

De Crécy, A. and Bazin, P. (2001). Determination of the uncertainties of the constitutive relationship of the CATHARE 2 code. *M&C Salt Lake City, Utah, USA*.