A Comprehensive Validation Approach Using The RAVEN Code[1]

Ivan Rinaldi*, Andrea Alfonsi†, Joshua Cogliatti†, Cristian Rabiti†, Fabio Giannetti*, Gianfranco Caruso*

*Sapienza, University of Rome, Rome, Italy
†Idaho National Laboratory, Idaho Falls, Idaho
rinaldi.1325668@studenti.uniroma1.it, andrea.alfonsi@inl.gov, joshua.cogliati@inl.gov, cristian.rabiti@inl.gov, fabio.giannetti@uniroma1.it, gianfranco.caruso@uniroma1.it

INTRODUCTION

The RAVEN computer code [1, 2, 3], developed at the Idaho National Laboratory, is a generic software framework to perform parametric and probabilistic analysis based on the response of complex system codes. RAVEN is a multi-purpose probabilistic and uncertainty quantification platform, capable to communicate with any system code.
A natural extension of the RAVEN capabilities is the implementation of an integrated validation methodology, involving several different metrics, that represent an evolution of the methods currently used in the field [4]. The state-of-art validation approaches use neither exploration of the input space through sampling strategies, nor a comprehensive variety of metrics needed to interpret the code responses, with respect experimental data. The RAVEN code allows to address both these lacks [5].
In the following sections, the employed methodology, and its application to the newer developed thermal-hydraulic code RELAP-7 [6], is reported.The validation approach has been applied on an integral effect experiment, representing natural circulation, based on the activities performed by EG&G Idaho [7]. Four different experiment configurations have been considered and nodalized.

PROPOSED VALIDATION METRICS

The validation activities of system codes are always fundamental processes in the development and assessment of the accuracy of the employed physical models. The state-of-art methodology is well described by Oberkampf et al.,2010 [4]. Such approach treats uncertainties individually (i.e. each uncertain parameter is considered distinctly from one another), while the proposed methodology path performs the exploration of the input space considering the associated uncertainties altogether and analyzes the responses through the use of several validation metrics. Uncertainties in the input space are taken into account separately from ones in the output space. As already mentioned, Such distinction is performed employing sampling of the input space. Such capability, available in the RAVEN code, permits to compare a larger sample of data. This approach allows multiple and more precise comparisons as an alternative of using comparison on single code runs data.

Parameter Identification

The first step in any code validation is the identification of the parameters that actively influence the response of the system. These parameters are going to be dived in two sets: dependent and independent. The independent parameters are the ones which are set up in the experiment, and are therefore going to be the initial conditions and boundary conditions of the model described in the input of the code under investigation. The dependent parameters instead, are the Figures Of Merit (FOM) relevant to measure. When the experiment is carried out, many field variables are monitored, but, generally, only few of them are considered to be representative of the physics of the experiment, and therefore objects of the comparison with the system code under validation.

Input Uncertainties

In order to assess the accuracy of the system code under consideration, it is, obviously, fundamental the inclusion of the experimental data uncertainties. As already mentioned, those uncertainties are going to be directly mirrored in the input space of the experiment, being modeled with the system code under investigation. The uncertainties associated to the input space can be represented by Probability Density Functions (PDFs); when dealing with experiment uncertainties, it is common practice to use the following PDFs, sorted in ascending order of "knowledge" regarding the uncertainty sources:

- Uniform PDF: used when no knowledge of the dispersion and mean of the data are available, but only the variation boundaries (i.e. lower and upper boundaries);
- Triangular PDF: utilized when the variation boundaries and the most probable mean are known, but no information on the dispersion of the data are available;
- Normal PDF: used when information about the mean and dispersion is available.

Output Uncertainties

As already mentioned, when performing validation activities, the FOMs that are considered in the "comparison" between the experimental and system code responses are affected by different sources of uncertainties.
As it is well known, the experimental data are affected by uncertanties coming from different sources:

- Determinate errors, also called "systematic" uncertainties, related to the measurement devices (e.g mis-calibration, hysteresis, full-scale, etc.);
- Human errors, introduced by the experimentalist;
- Indeterminate (Random) errors, related to variation in experimental conditions, operator bias, or other factors not easily accounted for.

The system code responses are affected by uncertainties that, obviously, come from the modeling approximation [4], such as:

[1] Risk Analysis and Virtual control ENvironment.

- Computer Round-off Error (introduced when exact solution can not be obtained from discrete equations);
- Iterative Convergence Error (introduced when truncation close to the solution occurs);
- Spatial and Temporal Discretization Error;
- Statistical Sampling Error;
- Response Surface Error.

IMPLEMENTATION OF VALIDATION METRICS

The final objective of a validation is to compare the results from the system code under investigation to the results of the experiments. Before such comparison can be achieved, the data has first to be sorted and then evaluated.

Binning

RAVEN uses an automatic binning algorithm in order to subdivide the output data and minimize the noise caused from statistical sampling. Given a certain response surface, as the one shown in Figure 1, RAVEN subdivides the domain of the output parameters in equally spaced intervals. In Figure 2 there is the subdivision of the data after RAVEN computed the optimal number of bins. The optimal number of bins is calculated using Eq. (1), but in RAVEN there is a wide range of options, depending on the distribution of the input.

$$k = \lceil log_2 n + 1 \rceil \tag{1}$$

Sensitivity Analysis

RAVEN, during the output analysis, has the capability to calculate all the sensitivity coefficients between chosen parameters [8]. The sensitivity matrix generated can be used to study the dependence of the system parameter from a certain input sampling. This method is used to check that the parameters sampled in the input space and the response of the system are dependent from one another, and if so, if they are correlated as expected. The sensitivity coefficients give the linear correlation between input and output parameters; by calculating the relative error between this linear regression and the actual output of the system, the interdependence among the parameter can be reconstructed. The equation used to linearize the problem is Eq. (2).

$$A_{lin} = \frac{\partial A}{\partial B}(B - \langle B \rangle) + \langle A \rangle \tag{2}$$

CDF and PDF Reconstruction

With all the output data divided in the optimized number of bins, the probability curves are constructed. The first function built is the Cumulative Distribution Function. The number of output counts in each bin, or intervals, as shown in Figure 2, is normalized to one and summed up bin by bin. The normalization occurs by summing up the total number of counts, and using this sum as a quotient for each channel's number. This normalization is needed in order to achieve the constraint of the final sum of the bin count to one, constraint needed for the PDF. A quadratic interpolation is then used to fit the data. A data sample is shown in Table I. The example shown represents the mass flow rate in the loop. The resulting function is referred as $R(x)$.

From the CDF the derivative is calculated, using the tree point derivation method, in order to compute the probability distribution function, referred as $R'(x)$. RAVEN then proceeds to build a function and plot of a distribution with the input given parameters, in our case a normal distribution given the mean value and the standard deviation. The CDF is referred as $E(x)$ and the PDF as $E'(x)$.

Functions Comparison

Having the continuous functions for the CDFs and PDFs for both the code output and the experiment RAVEN uses differerent comparison metrics together.

CDF Area Metric

The first metric, also called Minkowsky L_1 metric, calculates the area difference between the code output CDF and the experimental CDF using Eq. (3). An example is shown in Figure 4

$$d(R, E) = \int_{-\infty}^{\infty} |R(x) - E(x)| dx \tag{3}$$

Since the CDF is obviously normalized to one, whatever results from this metric gives an estimation in terms of the unit of measured used.

PDF Area Metric

This metric instead calculates the area underneath both PDFs. It gives out a degree of agreement between the two distributions. This metric is not unit sensitive, the result is given in percentace of agreement.

$$I = \int_{-\infty}^{\infty} R'(x) dx \cap \int_{-\infty}^{\infty} E'(x) dx \tag{4}$$

Difference of Continuous Functions

Being z a continuous random variable equal to the difference of two random variables, and being the two random variables statistically independent it can be calculated Eq. (5)

$$f_z(z) = \int_{-\infty}^{\infty} R'(x) E'(x - z) dx \tag{5}$$

Eq. (5) is a PDF representing the difference of the two PDFs. The domain is over the variable z, and it is unit measure sensitive. Calculating the mean of this distribution gives the value for which the two PDFs best overlap. Calculating the standard deviation of this distribution returns how much the two PDFs differ once they are at the best overlapping position.

TEST CASE

The study case presented is a natural circulation experiment [7]. In such kind of experiment the independent variables, the ones set up by the experimentalists, are the power in the core and the pressure in the primary and secondary loops. As described, being this case study a natural circulation experiment, the most relevant figures of merit to analyze are: mass

TABLE I. CDF and PDF recontruction for primary loop mass flow rate.

Bin Midpoint	Bin Count	Normalized Count	CDF	PDF
2.4313E-01	1	1.2000E-04	1.2000E-04	8.7400E-03
2.4451E-01	3	3.6000E-04	5.8000E-04	7.4970E-02
2.4589E-01	14	1.6900E-03	2.1700E-03	4.3748E-01
2.4727E-01	62	7.4900E-03	9.6600E-03	3.8936E+00
2.4865E-01	196	2.3680E-02	3.3340E-02	2.4499E+01
2.5003E-01	499	6.0280E-02	9.3620E-02	6.2035E+01
2.5142E-01	937	1.1319E-01	2.0681E-01	1.1099E+02
2.5280E-01	1393	1.6828E-01	3.7509E-01	1.4415E+02
2.5418E-01	1642	1.9836E-01	5.7345E-01	1.6808E+02
2.5556E-01	1631	1.9702E-01	7.7048E-01	1.6808E+02
2.5694E-01	1051	1.2696E-01	8.9744E-01	6.2998E+01
2.5832E-01	553	6.6800E-02	9.6424E-01	2.5943E+01
2.5970E-01	219	2.6450E-02	9.9070E-01	7.8310E+00
2.6108E-01	64	7.7300E-03	9.9843E-01	1.7062E+00
2.6246E-01	13	1.5700E-03	1.0000E+00	3.7480E-02

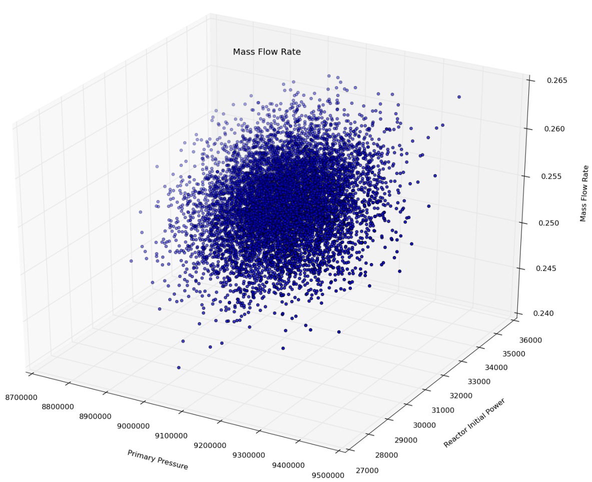

Fig. 1. RAVEN/RELAP-7 output: Mass flow rate response surface.

flow rate, hot leg and cold leg temperatures. Meanwhile the input variables are primary pressure and core power, with the following uncertainties.

Pressure	Power
±0.1 MPa	±1 kW

The optimized number of Monte Carlo samples was calculated to be 8000 runs for each input of the input deck (different inputs were built with different boundary and initial conditions, following the different experimental set-ups). This optimization comes from the need of having a large amount of runs needed to reduce the noise generated by statistical sampling, against the computational time needed for each run. The figures of merit that were chosen to compare to the experimental results were the mass flow rate in the primary loop and the temperature in cold leg and hot leg. To those parameters there were associated the following uncertainties:

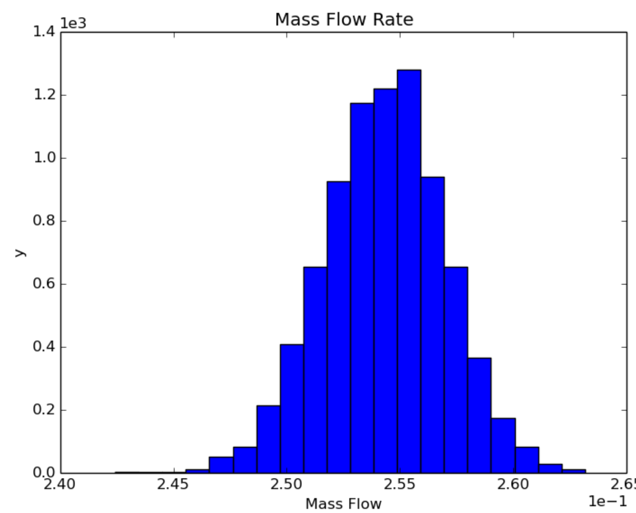

Fig. 2. Histogram after optimized binning.

Mass Flow	Temperature
±0.033 kg/s	±2 K

In Eq. (1), using $n \simeq 8000$ for the number of Monte Carlo samples, they are calculated a number of bins $k = 15$. In Figure 3 it is shown the plot of the relative error between the linearized output of mass flow rate as function of core power. Since the mass flow rate is proportional to the cube root of the core power [9], the error between the cube root behaviour and the linear regression is shown in Figure 3.

CONCLUSIONS

As it is easily inferable, the RAVEN code has just started to face the fundamental topic of the validation activities. The approach, here briefly proposed, represents a comprehensive methodology to assess the ability, of a generic system code, to

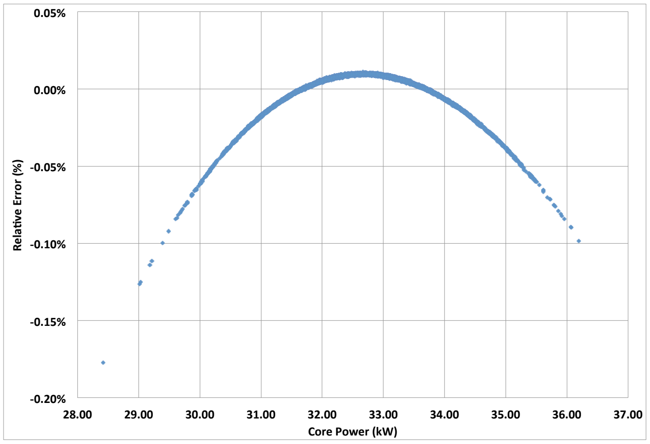

Fig. 3. Relative error as function of core power.

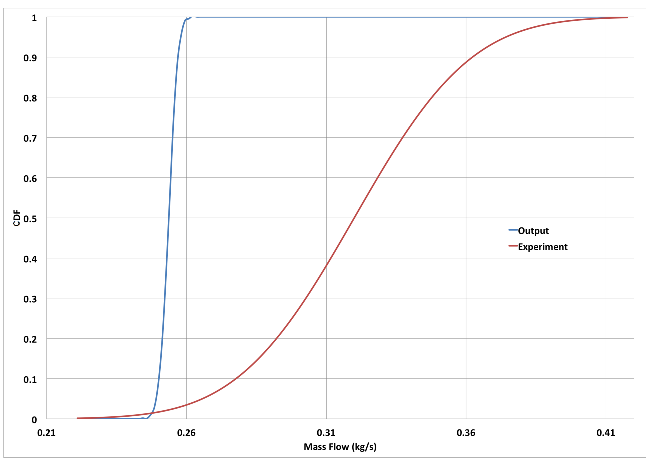

Fig. 4. CDF Comparison.

model the reality. All the validation metrics currently available in RAVEN can provide a much better and clearer assessment of the accuracy boundaries of the code under investigation then a single metric approach. Obviously, the here reported methodology only represents the starting point on which all the future development is going to be based on. In the near future, multi-dimensional distribution (Multivariate and Custom) are going to be finalized. These distributions are crucial in modeling the correlations, in the input and output space, among different uncertain parameters (FOMs).

ACKNOWLEDGMENTS

This work is supported by the U.S. Department of Energy, under DOE Idaho Operations Office Contract DE-AC07-05ID14517. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

In addition, a fundamental support has been provided by "La Sapienza", University of Rome.

REFERENCES

1. A. ALFONSI, C. RABITI, D. MANDELLI, J. COGLIATI, and B. KINOSHITA, "RAVEN AS A TOOL FOR DYNAMIC PROBABILISTIC RISK ASSESSMENT: SOFTWARE OVERVIEW," in "Proceedings of International Conference of mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2013), Sun Valley (Idaho)," (2013), pp. 1247–1261.
2. A. ALFONSI, C. RABITI, D. MANDELLI, J. COGLIATI, and B. KINOSHITA, "RAVEN and Dynamic Probabilistic Risk Assessment: Software Overview," in "Proceedings of European Safety and Reliability Conference (ESREL 2014), Wroclaw (Poland)," (2014).
3. C. RABITI, A. ALFONSI, D. MANDELLI, J. COGLIATI, and B. KINOSHITA, "RAVEN, a New Software for Dynamic Risk Analysis," PSAM 12 Probabilistic Safety Assessment and Management, Honolulu, Hawaii (June 2014).
4. W. L. OBERKAMPF and C. J. ROY, *Verification and Validation in Scientific Computing*, Cambridge University Press, 2rd ed. (2010).
5. C. RABITI, D. M. A. ALFONSI, J. COGLIATI, and B. KINOSHITA, "MATHEMATICAL FRAMEWORK FOR THE ANALYSIS OF DYNAMIC STOCHASTIC SYSTEMS WITH THE RAVEN CODE," in "Proceedings of International Conference of mathematics and Computational Methods Applied to Nuclear Science and Engineering (M&C 2013), Sun Valley (Idaho)," (2013), pp. 320–332.
6. D. ANDERS, R. BERRY, D. GASTON, R. MARTINEAU, J. PETERSON, H. ZHANG, H. ZHAO, and L. ZOU, "RELAP-7 Level 2 Milestone Report: Demonstration of a Steady State Single Phase PWR Simulation with RELAP-7," Tech. Rep. INL/EXT-12-25924, Idaho National Laboratory (INL) (2012).
7. G. G. CLOOMIS and K. SODA, "Results of the Semiscale MOD-2A Natural Circulation Experiments," Tech. Rep. NUREG/CR-2335, Idaho National Engineering Laboratory (1982).
8. C. RABITI, A. ALFONSI, J. COGLIATI, D. MANDELLI, and R. KINOSHITA, "REACTOR ANALYSIS AND VIRTUAL CONTROL ENVIRONMENT (RAVEN) FY12 REPORT," Tech. Rep. INL/EXT-12-27351, Idaho National Laboratory (INL) (2012).
9. "Natural circulation in water cooled nuclear power plants," Tech. Rep. IAEA-TECDOC-1474, International Atomic Energy Agency (IAEA) (2005).