75th International Astronautical Congress (IAC) - Milan, Italy, 14-18 October 2024.

IAC-24,B6,IP,27,x89504

Methods for Optimizing a Monte Carlo Campaign for an Aerospace Model: Sampling and Representativeness Considerations

Mr. Rocco Larocca[†], Dr. Andrea Uccella[‡], Dr. Simone Battistini[††], Dr. Filippo Maggi[‡‡]

[†] Main Author, *Politecnico di Milano,* rocco.larocca@mail.polimi.it
[‡] Co-Author, *MBDA Italia,* andrea.uccella@mbda.it
[††] Co-Author, *MBDA Italia,* simone.battistini@aerospace.unb.br
[‡‡] Co-Author, *Politecnico di Milano,* filippo.maggi@polimi.it

Abstract

Computer-based experimental design is nowadays a key factor in any industrial engineering process, as it allows to verify and validate the conceived product even when the operative conditions are very hard to reproduce in a laboratory environment, as in the case of space applications. As a matter of fact, computer-based simulations minimize the design risks and the overall required temporal and economic costs. During the concept phase of a product, Monte Carlo (MC) Methods represent the core of the simulation step, guaranteeing the randomness behaviour that characterizes real experiments. Unfortunately, a big computational load might be needed to guarantee the representativity of the problem.
This work aims at defining a tool to estimate the minimum number of runs by which a MC process can be said to be representative. The investigation started from looking at sampling generation methods (i.e., the process of generating the uncertain inputs) for retrieving the best techniques in producing more even spaced points (e.g., no clusters or not-covered zones). The parameters of the obtained sampling performance were used to develop a tool that takes in input the number of uncertain variables, and returns the minimum number of runs needed to be representative.
It is shown that Quasi-MC Methods, in particular the Sobol's sampling technique, perform better than frequently used Pseudo-Random series in reaching representativeness by a very small number of runs. The tool presented in this paper was tested on a generic 3DOF model of a rocket to retrieve and compare the number of runs needed to represent the dynamics of the system with a number of different sampling methods. An analysis of the representativeness and the role of uncertainties in the problem by *a posteriori* considerations was also carried out, allowing to validate the results of the previous investigation and to tune the overall algorithm for several different space applications.
Keywords: Experimental Design, Quasi-Monte Carlo Methods, Sobol, Simulation, Validation

1 Introduction

In the context of industrial engineering, the testing phase represents a key moment for evaluating all the trade-offs, assumptions and crucial choices made from the very first instants of a conceptual design. Nevertheless, in order to deeply understand how a certain system is actually working, the mere observation becomes partially meaningless; as a matter of facts, the initial conditions of any test result fundamental in determining its outcomes. Therefore, changing the related boundary factors could mean increasing the acquired information, highlighting the cause-and-effect relationship on which the system is based on.
As a consequence, experimentation becomes a vital part of any engineering development, even more so in the aerospace field in which high capitals and efforts can be assigned to a very single mission. This translates in the impossibility to test "on field" a certain aerospace product for multiple times, both for economical and temporal reasons; moreover, it is frequent to face systems whose environmental working conditions can be difficult to be reproduced on ground (e.g., rockets), not guaranteeing the maximum precision in the outputs in front of important investments. With the advent of computing technology and numerical methods, engineers and scientists have had the possibility of simulating real systems on computers by creating proper mathematical models able to represent the most deep physical behaviours. This means including linear, nonlinear, ordinary, and partial differential equations, so producing an high-complexity environment whose fidelity to reality appears as a primary requirement. In the hypothesis of working with a very accurate computer-based model, it would be at this point possible to vary the boundary inputs given to that, simulating multiple runs and so retrieving how the desired performances are conditioned by initial random parameters.

Nowadays, the above-mentioned process of modeling a real system on a computer-based form and of testing it virtually is something quite diffused in the aerospace sector. Being more specific, the idea of simulating a model by different uncertain inputs with the aim of reproducing a real case scenario is commonly known as Monte Carlo (MC) Methods: their advantages are the possibility of directly computing the uncertainties even when the model contains discontinuities or nonlinearities, its simplicity in the implementation, its robustness, as well as its transparency. Nevertheless, simulating an huge number of uncertain cases in order to gain in randomness and representativeness of the "on field" experiment can be very time consuming and inefficient, although it can be justified by the absence of any criteria for selecting a proper quantity of runs given a certain number of uncertain inputs.

The objective of this paper is to focus on those drawbacks, trying to find possible solutions to optimize the overall MC computation. As a consequence, it is worth to go in depth on the two cited aspects:

- Inefficiency in the computations: in the assumption of clear and performing code, it depends on redundant and non-representative cases that are simulated "guessing" the correct number of runs;
- Absence of predictive criteria: it links to the previous statement, and underlines the need of an instrument for individuating the optimal number of simulations to be representative.

Indeed, if any *a priori* information about the sampling method performances in producing better sequences will be available and associated to a given number of uncertain variables, the very same information can be used to find the most proper number of runs and so to radically increase the possibility of being representative with the lowest effort possible. In this sense, multiple sampling methods (i.e., Pseudo-Random, Latin Hypercube and Quasi-Monte Carlo series) have been introduced, trying to compare by proper algorithms their performances and ability to reproduce a real scenario by the lowest number possible of runs.

By following the previous reasoning, it has been shown that Quasi-MC Methods, in particular the Sobol's sampling technique, perform better than frequently used Pseudo-Random series in achieving representativeness with a significantly fewer number of runs. A 3DOF aerospace model of a rocket has been used to practically discuss representativeness and uncertainties' role in the complexity of the problem, allowing to corroborate the results and to tune the overall instruments also for very different engineering systems.

All the implementations, computations and simulations have been performed in the MATLAB environment.

1.1 Paper organization

This paper has been organized in four sections: in Sec. [2], a brief mathematical introduction about Monte Carlo Methods and available sampling techniques is reported, presenting proper instruments for evaluating the samples' performances and laying the foundations for the successive applications (i.e., main assumptions and methods); indeed, Sec. [3] exploits the previous concepts for structuring an iterative *a priori* selective function to find the most proper number of runs to be representative, combining then fundamental *a posteriori* considerations for defining the achieved representativeness; Sec. [4] applies the obtained algorithm to a 3DOF rocket model, corroborating the above theoretical outcomes on a real-shaped computer-based experiment; finally, Sec. [5] resume all the above results to deduce the final conclusions about the different sampling techniques, enhancing the convenience of more efficient simulation process in the context of the aerospace industry.

2 Theoretical Backgorund

General steps of a Monte Carlo Method can be resumed as [1]:

1. Determine initial statistical properties of inputs (uniform distribution, Gaussian distribution, etc.);
2. Generate a set of N (number of MC iterations) random values resembling the probability distributions for each input quantity;
3. Perform calculations (mathematical model) with these sets;
4. Analyze the results, that are represented by a probability distribution.

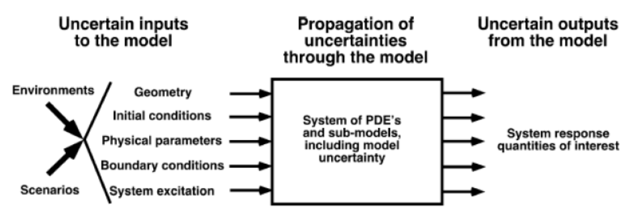

Fig. 1: Reproduction of Real Experiments in Simulated Models by MC Approach

Clearly, in the idea of a optimal computer-based experimental design, the entire process would be simulated trying to represent in the best way and by the lowest N possible the randomness and every condition characterizing a potential real test, in other words two aspects that are dictated by the sampling process (i.e., Step 2).

For this reason, it makes sense to investigate all the aspects characterizing the sampling operations, so the available sampling techniques. Excluding all the options requiring *a posteriori* information since highly-dependant on the model (e.g., ANOVA [1] methods), three sampling possibilities have been selected:

- Pseudo-Random (PRS) logic, based on stochastic considerations and without any criteria of optimization [2];
- Latin Hypercube (LHS) logic, exploiting permutations on a PRS to better cover the (stratified) sampling grid [3];
- Quasi-MC logic (QMC), working on optimized quasi-random numerical sequences able to avoid repetitions/clusters, and defined as completely deterministic (i.e., Sobol series [4]).

All these methods have been deeply analyzed by focusing on their capacity to produce (or not) the most uniform sampling possible, thus trying to identify the one having the most even spaced points. Moreover, repeatability and easiness to be implemented in MATLAB have been two additional discriminant factors in neglecting further sampling logic.

Judging *a priori* the uniformity of the outputted sampling grids has required different performance parameters, working both by mathematical and statistical considerations. In particular:

- Discrepancy has been defined as the mean distance between two points of the domain made by two consecutive dimensionalities (i.e., uncertain variables): the more the points are distributed uniformly, the lower the discrepancy is. It can been shown that, accordingly to the outcomes of the Koksma-Hlawka Theorem [5], QMC Sobol's results to have the lower discrepancy possible for most of the (n,N) combinations (where n is the number of dimensions and N the number of runs);
- Considering the definition of Voronoi's polygons applied to the sampling bidimensional domain (where the two selected dimensions are the last ones since typically suffering the most clusters or correlation effects), and so observing how the related areas can be a figure of merit of the points' uniformity (refer to Fig.2), the "Voronoi's coefficient" has been introduced as an index of areas' regularity for a further comparison of the three methods. Voronoi's coefficient is calculated as:

$$\varkappa = \min(\max(A_i)) - \max(\min(A_i)), \quad i = 1,\dots,N \quad (1)$$

A better sampling would lead to a minor $\varkappa$, in other terms in a slighter difference between maximum and minimum areas. The calculation of Voronoi's coefficient is punctual, in the sense of requiring a precise number of dimensionalities and runs: this actually leads to a "local" comparison of the three methods, and it can be used as check for the discrepancy values;

[1] i.e., Analysis of the Variance, based on Bayesian statistics

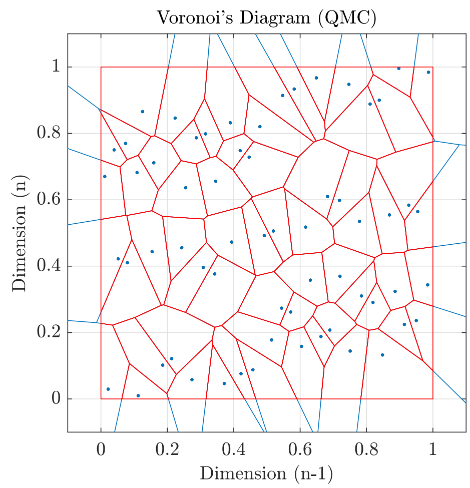

Fig. 2: Voronoi's Diagram for QMC Sobol's Sampling ($n = 15$, $N = 63$)

- Kurtosis is a statistical parameter quantifying how outlier-prone a distribution is [6]. In particular, a perfect uniform distribution has a Kurtosis equal to 1.8 (3 for a Gaussian one). A sampling technique aiming for the most uniform possible result should produce a Kurtosis the nearest to the ideal value.

The theoretical summary exposed above, if applied to the individuated sampling methods, would reward QMC Sobol's as the most promising method to obtain more uniform samplings' productions, finding in LHS a valid alternative and dismissing PRS due to the absence of any performance criteria; indeed, the low discrepancy character of Quasi-MC Methods reflects in the best Voronoi's coefficient value between the three sampling logic, and so these mathematical properties allow to produce the most uniform possible sampling grid for a very large number of (n,N) combinations, reaching the Kurtosis values near to the ideal ones.

Nevertheless, all the above theoretical considerations have to be considered as only-comparative ones, in the sense that there is not an absolute best sampling method but always one that is performing better than the others. This concept has justified the idea of an investigating function, looking for the best number of runs and sampling method possible given as input the number of uncertain variables for the model of interest .

3 The *a priori* Investigator

The *a priori* investigation has been operated by a function that has been called in the MATLAB environment as `MCreducer`; it exploits all the concepts anticipated in the above theoretical background. This function, in its predictive part, works without any information coming from *a posteriori* considerations and it is able to handle both uniform or normal distributions, automatically excluding constant inputs or unrecognized probability distributions' trends.

1. The function is started by inputting a certain number of dimensionalities n and a guess range of number of runs N_{guess}. The first step consists of a comparison between discrepancies for all the provided N_{guess}, looking for the method and the first candidate solution N_1 providing the lower discrepancy with the maximum difference with respect to the other possibilities;

2. The candidate solution N_1 proceeds by a further check calculating the Voronoi's coefficient for the exact (n, N_1) couple: if the minimum result is associated to the method outputted by Step 1, N_1 becomes the final solution N_{end};

3. This step varies depending on the previous one. If the check at Step 2 has been satisfied, the function is stopped and the solution N_{end} with the best performing method is released, otherwise the iterative process goes on selecting at Step 1 the second solution N_2, repeating the cycle until a valid output is found.

As anticipated, the provided *a priori* investigator is designed to work without any information coming from *a posteriori* considerations, and so its results can be identified as optimal in the sense of uniformity (i.e., best sampling performances) but not in the sense of representativeness for which the MC simulation becomes unavoidable. Nonetheless, the assured even spacing sampling method with N samples for the assigned n faces the simulation of the model with much more acquired information with respect to typical MC processes, therefore also in the case in which the suggested solution results not to be representative, it can be used as lower bound of the guess for a successive iteration of `MCreducer` that it is, at this point, started with a more and more significant information load. The above observation clears the way for a larger optimization algorithm that, although requiring some *a posteriori* considerations, enhances the *a priori* ones step by step, allowing to reach representativeness in very few iterations of the overall code.

Representativeness can be traced on the basis of the Central Limit Theorem applied to MC campaigns, by which the definition of root mean square error ε_N can be retrieved:

$$\varepsilon_N = |E[f(\boldsymbol{y})] - E_N[f(\boldsymbol{x_i})]| = \frac{\sigma(f(\boldsymbol{x_i}))}{\sqrt{N}} \qquad (2)$$

The operator $E(\cdot)$ stays for the expectation (or mean value) of the random distribution $f(\cdot)$, so a MC simulation can be said as representative whenever the nominal expectation function $E(f(\boldsymbol{y}))$ is well approximated by the sampled random distribution one $E(f(\boldsymbol{x_i}))$, with a convergence rate of $O(1/\sqrt{N})$; in other words, whenever convergence is achieved, the model and the entire MC analysis figures as representative of the real experiment.
Exploiting this concept and defining the nominal trend as the one coming by a precise knowledge of the uncertainties acting on the model, a first error can be introduced as:

$$\texttt{err1} = \left|\frac{\mu_{nom} - \mu_i}{\mu_{nom}}\right|, \; \mu_i = \frac{1}{i}\sum_{j=1}^{i}\mu_j \qquad (3)$$

i goes from 1 up to N, while μ_{nom} stays for the nominal mean value and μ_i derives by a continuous updating of the overall mean by summation of the means retrieved by each single MC run.
Nevertheless, in order to avoid possible misleading solutions due to false convergences of the trend (i.e., passages through the nominal mean), the complete flattering of the curve has been assured by a second error (i.e., second derivative) that shall be the nearest to zero:

$$\texttt{err2} = \frac{\partial^2(\mu_{nom} - \mu_i)}{\partial N^2} \qquad (4)$$

In this way, assigning proper tolerances to both the errors, a larger iterative scheme has been structured, combining *a priori* theoretical concepts with *a posteriori* information, so retrieving uniformity of the sampling but most of all representativeness of the problem.

Algorithm 1 The overall *a priori* Investigator

```
Define guess interval
Given n and guess, operate MCreducer retrieving N1
  and method
Calculate μnom and μi
if err1 < tol1 then
    Calculate err2
    if err2 < tol2 then
        N1 valid, convergence is achieved
    else
        N1 not valid, new guess = N1^(1+err1)
        Restart from Step 1
    end if
else
    N1 not valid, new guess = N1^(1+err1)
    Restart from Step 1
end if
Output N, method and highlight the reached represen-
  tativeness
```

4 The *a posteriori* Analysis: a 3DOF Aerospace Model

In order to corroborate the overall *a priori* investigation made by Algorithm 1, and consequently the theoretical discussion of Sec.2, the very same process has been applied to a engineering model. In particular, it corresponds to a simplified scheme representing a rocket vehicle whose aim is to complete a flying mission elapsing into prescribed position's requirements: therefore, an external environment plus dynamics sub-model is interfaced with a guidance block in the idea of a closed loop cycle.
More specifically, the engineering model can be indicated as 3DOF point-mass since including the three aero-forces (normal, axial, side), as well as thrust and weight [7]. This actually classifies the current framework in a conceptual design phase of the system's modeling.

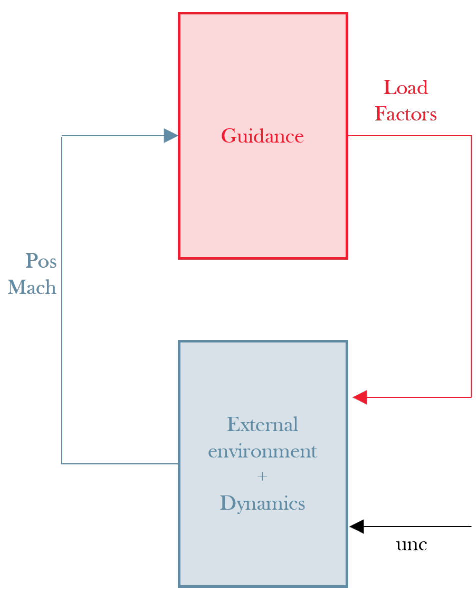

Fig. 3: Synthetic Scheme of the 3DOF Rocket Model

A synthetic scheme of the overall 3DOF framework is shown in Fig.3, by which it can be understood how the uncertain content acts on the quantities involved into the external environment and dynamics' modeling, and so it appears right and proper to focus on it exploiting how any uncertainty weights on the final outputs of the block.
Nevertheless, before going deep inside the 3DOF architecture, it is important to state some hypotheses and mission's conditions that have been imposed on the aerospace system in analysis:

- The system shall be started by a boosting phase, followed by the activation of an air-breathing engine. Therefore, the thrust is conditioned firstly by the booster, then by the engine;
- The system receives a commanded Mach of 0.8, so the outputted value shall not be too far from the assigned requirement;
- For the sake of simplicity, no wind components have been taken into account in modeling the atmospheric conditions and so the dynamics of the system;
- Position and velocity are retrieved in a Cartesian scheme.

The characteristics of the model as well as the kind of mission that it has to accomplish have been the main two reasons for which range, Mach and mass have been chosen as reference output parameters for evaluating both the flight performances and the validity of the *a priori* analysis. Indeed:

- Range is a typical figure of merit in a 3DOF aerospace model since it indicates the final distance travelled by the airframe, and so the point of arrival of the same. It is highly dependant of the thrust and of the aerodynamic forces (thus, coefficients);
- Mach can help in evaluating the response of the overall model to uncertain conditions in achieving the required commanded value till the end of the flight. Also in this case, it results as highly correlated to the thrust and to the aerodynamic performances;
- Mass is fundamental for observing the nominal behaviour of the propulsive unit (i.e., fuel consumption). It depends on its own uncertainty (adopted on the MTOW[2]), less by aerodynamic aspects.

As anticipated in the previous lines, the uncertainties acting on the vehicle are related to the main aspects of the airframe's flight: propulsion, aerodynamics, attitude, etc. A total number of 15 dimensionalities has been identified, therefore the `MCreducer` first input is defined as $n = 15$. Nevertheless, for setting all the needed conditions for beginning the iterative process, tolerances have to be assigned:

- `tol1` has been fixed to 10^{-3} for assuring a convergence rate up to the 99.9%, although for practical reasons a less demanding requirement can be imposed;
- `tol2` has been chosen as equal to 10^{-1} for ensuring a almost complete flattering of the convergence trend.

Finally, the initial lower bound of the `guess` has been set to 15 (as the dimensionality of the problem) for speeding up the computation, excluding trivial solution in the one-unit interval.
The obtained iteration sequences for each of the three selected sampling methods are reported in Tab.1: it can be

[2]MTOW: Maximum Take Off Weight

observed how the cycle is stopped at $N = 63$ for the Quasi-MC option, with a non-negligible saving with respect to PRS or LHS. Nonetheless, in order to qualitatively and quantitatively visualize the convergence trend as well as Sobol's capacity to better produce well-shaped distributions, for each of the chosen outputs it will be shown:

- **Convergence of the mean**: for each parameter and for each sampling method, the trend of the mean is calculated, comparing step-by-step it to the nominal value at the suggested outputs of `MCreducer` to effectively confirm the satisfaction of the tolerance requirements. In this way it will be possible to have a numerical idea of the convergence's grade reached by the three sampling techniques, and to have a check on the quantities obtained by the *a priori* investigator;
- **Histograms of frequency**: for each parameter and for each sampling method the frequency histogram[3] at $N_{QMC} = 63$, the (theoretical) smallest number of runs required by Sobol's sampling, is represented. These plots can help evaluating the advantages of this sampling technique in terms of uniformity of the obtained values at a very (theoretical) lower N with respect to the other two sampling algorithms' ones. Qualitative conclusions retrieved by the mere observation of the histograms will be confirmed by the Kurtosis parameter that, as seen in Sec.2, corresponds to a measure of how outlier-prone a distribution is.

Moreover, for the sake of completeness and for guaranteeing the validity of the obtained results, the simulation is extended up to $N_{max} = 600$ such to observe the uselessness in performing more runs than the ones suggested by `MCreducer` and to graphically appreciate the convergence properties of a representative campaign. As a consequence, for the selected outputs it will be possible to answer the following questions:

1. Is the suggested number of run (i.e., N_{tot}) large enough to reach representativity?
2. If yes, is the suggested N_{tot} too large with respect to convergence?
3. Is the mean diverging from the 3σ area, i.e. is the model inaccurate in retrieving the specific output?
4. What is the sampling method suggesting the lowest representative number of runs?

It is worth to have a focus on the third question; as a fundamental hypothesis of this paper, it has been assumed the correctness of the model on which the MC is applied in order to neglect additional errors and computational efforts given by inaccuracies or mathematical loops inside the code. Nevertheless, this idea needs to be confirmed: an inexact modeling of the real phenomenon would mislead representativity considerations. For this reason, the concept of **3σ area** is introduced: it is obtained by calculating the standard deviation at each run up to N_{tot}, then by multiplying it for ± 3. In this way, the region in which the 99.73% of cases will fall in is identified, a value considered reliable enough for aerospace systems. If one or more runs fall outside this zone, it would indicate the presence of errors inside the model blocks, and so a fallacious code.

[3] The number of bins for each plot has been determined following the relation of Bendat & Piersol by which $n_{bins} = 1.87 \cdot (N_{QMC} - 1)^{0.4}$

At this point, all the instruments are set for discussing the simulation's results:

Range

Range figures as the outcome of an integration process made on the Equations of Motion of the model, so it conveys almost all the uncertainties characterizing the dynamics of the system (i.e., the most subject to correlation effects with respect to uncertain inputs). In particular, by a sensitivity analysis it would possible to show its highly positive dependence with the thrust, while a negative correlation would be registered with the axial aerodynamic coefficient.
The latter causes a missed convergence around the nominal mean since, apart for the nominal run, c_x will act always in worsening the vehicle performances.

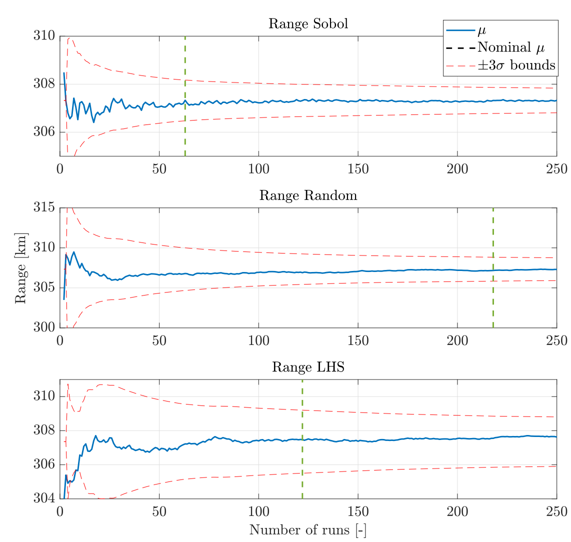

Fig. 4: Convergence Trends: Range

Nevertheless, from Fig.4 it is evident the flattering trend around the values suggested by `MCreducer` as well as the accuracy of the model stated by a mean always inside the 3σ boundaries.

The true advantage of QMC Sobol's sampling can be appreciated by Fig.5 in which the histograms of frequency and the Kurtosis values are reported: it is clear that, both in a qualitative and quantitative sense, at $N = 63$ the best bell shape is assumed by Sobol's one.

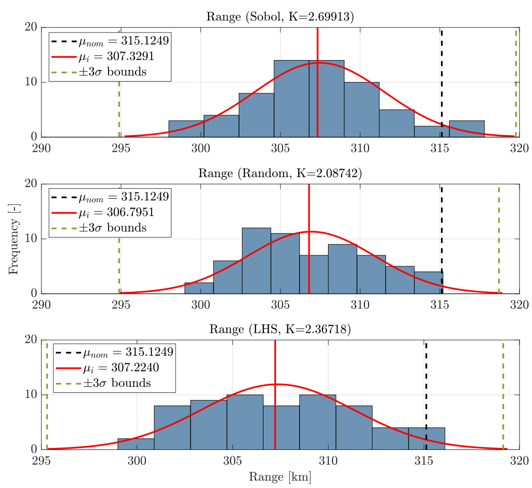

Fig. 5: Frequency Histograms: Range

Furthermore, looking at Tab.1 it is evident how the errors, overcoming the assigned tolerance for the c_x effects, are all in the same order of magnitude albeit they are obtained for very different numbers of runs, again underlining Sobol's performances at a very lower N.

Summing up, it is at this point possible to answer to the four check questions previously illustrated: limiting the analysis only to the range, the suggested $N_{tot} = 63$ can be said as representative whenever Sobol's sampling is applied, as well as it results as coherent with the convergence trend and the outcomes of the *a priori* investigator. The mean stays inside the prescribed safety boundaries.

Mach

Mach figures as a fundamental parameter for evaluating the accuracy of the model since it exists a strict requirement about a commanded Mach value that, in the case of a performing code, has to be maintained till the end of the mission. Discussing its dependencies, they are similar to those of range so the aerodynamic effects due to the detrimental character of c_x are still present.

Therefore, looking at Fig.6, the nominal value is not approached by the updated mean's trend that, in any case, results as convergent around the values given by the overall investigation. More importantly, the curve stabilizes around 0.85, positively accomplishing the prescribed requirement and assuring the accuracy of the modeling phase, something further underlined by the 3σ boundaries.

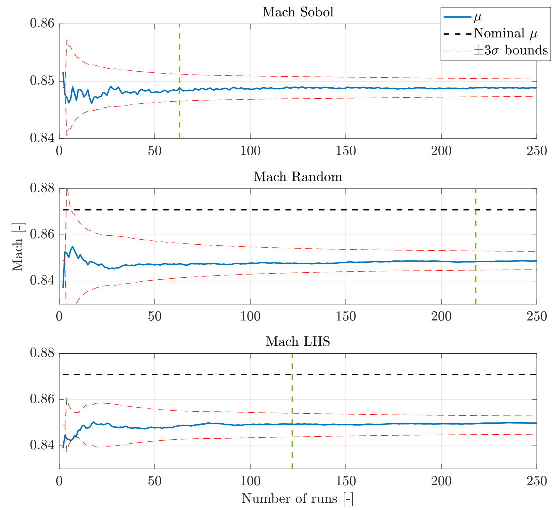

Fig. 6: Convergence Trends: Mach

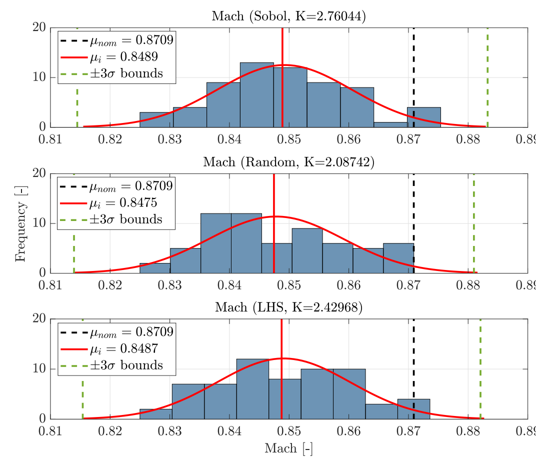

Fig. 7: Frequency Histograms: Mach

Analogously to the range, considerations about sampling performances have been derived mainly by looking at Fig.7 in which the histograms of frequency related to the Mach final values have been reported. Sobol's results again in the best normal shape, well symmetrically distributed and with its peak around the updated mean; on the contrary, PRS and LHS at $N_{QMC} = 63$ begin to assuming the Gaussian shape manifesting the necessity of more simulations.

Tab.1 corroborates the conclusions retrieved by convergence trends and frequency histograms: looking at `err1`, QMC Sobol's performs better than the other methods by a very lower number of runs (i.e., $N_{tot} = 63$), with the second derivative's value staying well behind the assigned tolerance. Looking at the Kurtosis, and remembering how a perfect shaped normal bell outputs a Kurtosis equal to 3, it is possible to appreciate the better bell

	Sequence	**`err1` (Range)**	**`err1` (Mach)**	**`err1` (Mass)**
QMC (Sobol)	15 → 32 → 63	2.4738 %	2.5274 %	0.0169 %
PRS	17 → 40 → 153 → 218	2.5233 %	2.5858 %	0.0589 %
LHS	17 → 30 → 52 → 95 → 122	2.4754 %	2.4673 %	0.0845 %

Table 1: Iterative Sequences and Convergence Errors for QMC, PRS and LHS

shape of Sobol's, as well as comparing to the range's ones, it can be observed how Mach is slightly better producing a normal distribution due to a less important correlation to the more influential and uncertain input variables.

To summarize, for the Mach case, $N_{tot} = 63$ results large enough to guaranteeing representativeness by a Sobol's sampling process, largely overcoming the other two methods and coherently with the predictions of `MCreducer`. More importantly, the mean trend remains inside the 3σ boundaries and behind the commanded value, guaranteeing a correct performance under this point of view.

Mass

Mass is a key parameter for evaluating the overall performances of the system from the propulsive, aerodynamic and attitude point of view. Despite this fact, it comes by simpler considerations with respect to range or Mach, therefore its dependence on the involved uncertainties is different. By a sensitivity analysis it would possible to show a correlation coefficient equal to +1 (i.e., positive linear trend) with the mass uncertainty itself, while the link with c_x is limited, avoiding the effects seen for the previous outputs.

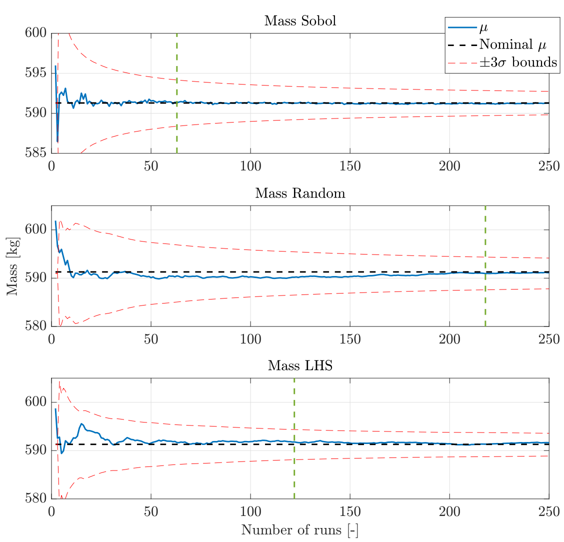

Fig. 8: Convergence Trends: Mass

Fig.8 confirms the above-mentioned characteristics for the convergence trend of the mass: an exact matching between resulting and nominal means is finally achieved at the values provided by Algorithm 1, with a complete flattering of the tendency that justifies the non-necessity of further runs. Moreover, the converging trend stays always in the 3σ region.

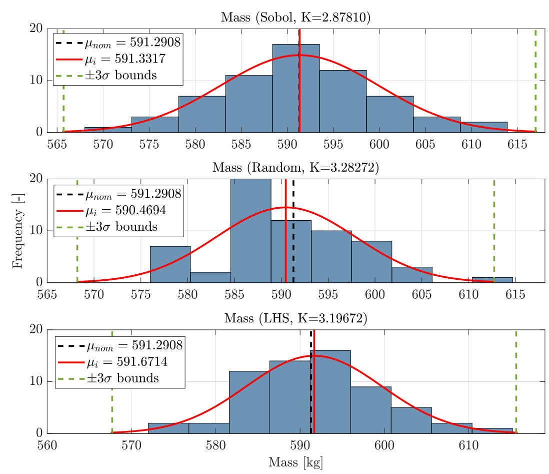

Fig. 9: Frequency Histograms: Mass

Mass allows to judge QMC Sobol's sampling method as the most appropriate for the problem in analysis; indeed, Fig.9 shows a almost perfect normal bell for the outputted mass values produced by Sobol, while at $N = 63$ the results of PRS are not even remotely comparable.
The advantages explained by the histograms of frequency are confirmed by the numerical results included in Tab.1. `err1` finally returns into the prescribed tolerance requirement and shows its lower value for QMC that, at a number of runs N much lower with respect to the other methods, can be said to reach representativeness.

Answering the four questions about representativeness appears at this point immediate for the mass: it is reached by a very high precision at the suggested N_{tot} without unnecessary successive runs. The model for retrieving the mass results accurate.

5 Conclusions

MC Methods have been identified as a fundamental step in any conceptual design aiming at reducing the overall costs. The absence of predictive criteria for determining the correct number of runs N to achieve representativeness clearly limits their potential. The paper' objective has been to focus on all the factors inducing representative results, discovering how in reality the *a priori* operations in producing the needed sampling are fundamental for fully exploiting every single run extracting the greatest amount of information from them.

At this point, the very same idea has justified the research of the most performing sampling method by an *a priori* predictive function that, given a certain n, outputs the most appropriate sampling method and the related N for gaining the most uniform possible sampling grid. By multiple applications of this algorithm, QMC Sobol's has been identified as the most performing method for a great number of n cases.

Nonetheless, although the obtained uniformity has been optimized, representativeness can not be said yet as reached: indeed, *a posteriori* considerations are still needed. This lack has pushed for a larger algorithm able to retrieve representativity observations by convergence analyses and eventually to start a new complete iteration if convergence would not be correctly achieved. In this way, the *a priori* investigator has been observed to adapt itself to the model under study, for example the 3DOF aerospace system discussed in Sec.1. By this practical case, it has been confirmed the optimal behaviour of QMC Sobol's:

- Representativeness can be said as achieved in a very less number of runs with respect to PRS or LHS;
- No additional runs have been operated for achieving representativeness, providing the user of a mathematical based N.

All these aspects translate in correctly achieving the aim of this paper as exposed in Sec.1, thus estimating the correct minimum number of simulations for a MC campaign reaching representativeness. Most importantly, these results provide a significant advantage in terms of reducing computational effort, as well as they can be used to verify and validate the built model and its processes.

References

[1] Montgomery D. C., *Design and Analysis of Experiments*, John Wiley & Sons Inc., 8^{th} edition, 2013, ISBN: 978-1-118-14692-7

[2] Rubinstein R.Y., Kroese D.P., *Simulation and the Monte Carlo method*, Wiley, Third ed., 2017, Hoboken, New Jersey (USA)

[3] McKay M.D., Beckman R.J., Conover W.J., *A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code*, Taylor & Francis Gruop LLC, Technometrics, May 1979

[4] Bratley P., Fox B.L., *ALGORITHM 659 Implementing Sobol's Quasirandom Sequence Generator*, Université de Montréal, 1988

[5] Niederreiter H., *Quasi-Monte Carlo methods and pseudo-random numbers*, Bull. Am. Math. 1978, Vol. 84, No. 6

[6] Gentile C., *Using the Kurtosis Measure to Identify Clusters in Wireless Channel Impulse Response*, IEEE Transactions on Antennas and Propagation, Vol.61, No.6, June 2013

[7] Fleeman E. G., *Tactical Missile Design*, Second Edition, American Institute of Aeronautics and Astronautics, Inc., Virginia, USA