

## *Research Article*

# **Monte Carlo Simulation Models Evolving in Replicated Runs: A Methodology to Choose the Optimal Experimental Sample Size**

**Lucia Cassettari, Roberto Mosca, and Roberto Revetria**

*DIPTeM, University of Genoa, 16145 Genova, Italy*

Correspondence should be addressed to Roberto Revetria, roberto.revetria@unige.it

Received 1 October 2011; Revised 29 February 2012; Accepted 17 March 2012

Academic Editor: Kwok W. Wong

Copyright © 2012 Lucia Cassettari et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The idea of a methodology capable of determining in a precise and practical way the optimal sample size came from studying Monte Carlo simulation models concerning financial problems, risk analysis, and supply chain forecasting. In these cases the number of extractions from the frequency distributions characterizing the model is inadequate or limited to just one, so it is necessary to replicate simulation runs many times in order to obtain a complete statistical description of the model variables. Generally, as shown in the literature, the sample size is fixed by the experimenter based on empirical assumptions without considering the impact on result accuracy in terms of tolerance interval. In this paper, the authors propose a methodology by means of which it is possible to graphically highlight the evolution of experimental error variance as a function of the sample size. Therefore, the experimenter can choose the best ratio between the experimental cost and the expected results.

## **1. Introduction**

According to the universally accepted definition, a Monte Carlo simulator is a teller of the possible histories of the object system. The model “goodness” depends not only on its constructor ability (i.e., system analysis, data survey, and logic transcription) but also on a correct experimental activity, which should include, among its main targets, experimental error measurement, which is generally distributed as a normal distribution  $(0, \sigma^2)$ , affecting the model [1–3].

The  $\sigma^2$  entity, which can be estimated according to Cochran’s theorem [3] through the measurement of the mean square pure error (MSPE), its unbiased estimator, is an intrinsic characteristic of each model.  $\sigma^2$  is strictly connected to the investigated reality, since it is directly dependent on the overall stochasticity by which this reality is affected.

In other words any object system displays its own level of stochasticity affecting the behaviours of the output variables and entering in the simulation model, by producing a characteristic “noise,” which cannot be set aside. In the experimental phase, the real problem is not the background noise entity, which cannot be eliminated since it is inherent in any stochastic system, but the possibility to add another major source of noise to it. This error component, on the contrary, can be controlled and, if necessary, even eliminated. It is associated with a number of extractions from distributions of random variables in the model. However, said number is inadequate to obtain, in terms of statistical parameters, total adherence in the simulation phase to the probability distributions describing real system behaviour.

The study of the MSPE trend in the simulated time makes it possible to solve the problem through a graph whose examination clearly points out, without any particular difficulty in interpretation, the total model noise fraction in each moment of the simulated time [4, 5]. Therefore, while examining the graph, it is possible, if necessary, to separate from the total noise the real system noise. So all the positive consequences that can arise from it in terms of reliability analysis on the model output results can be appreciated.

On the contrary, there are object systems that cannot be managed in the experimental phase according to the MSPE evolution scheme in the simulated time.

This occurs each time the number of extractions from the frequency distributions is limited to a single value or, in any case, to a limited number of samples that is not large enough to obtain an effective description of the aforementioned distributions.

A typical case—yet not the only one as shown further below—is represented by the Corporate Models used for the construction of possible future economic scenarios. In these models some input variables, characterising any following accounting period, are assigned under the form of frequency distributions displaying a character of uncertainty that grows in the time (costs of raw materials, personnel, services; sale proceeds, transfers, investments, etc.). In the experimental phase, a single value will be selected, at worst, from these distributions to characterise any specific activity [6, 7].

The main difference between this methodology and the evolution time methodology is that, in this case, both the variance of the mean response (called by the authors  $MSPE_{MED}$ ) and the variance of the standard deviation (called by the authors  $MSPE_{STDEV}$ ) must be monitored. These two parameters make it possible to choose the optimal number of runs needed to obtain an unbiased evaluation of the experimental error affecting the objective function.

As it is well known, the larger the sample, the better the description of the population. With this methodology it is possible to graphically highlight the evolution of the variance experimental error as a function of the sample size. So the experimenter will be able to choose the best ratio between the experimental cost and expected results.

In conclusion, the proposed methodology allows the determination of the number of replicated runs capable of minimising, according to the experimenter’s needs, the noise generated by an inadequate overlapping of the probability density functions of the variables concerned.

## 2. Theoretical Development of the Methodology

The technique for MSPE study in replicated runs, which is conceptually similar to the methodology used for the MSPE study in simulated time, can be divided into the following phases:

- (1) to set a number  $K > 2$  of simulation runs, carried out in parallel, in which the independent model variables are maintained always at the same level, modifying

**Table 1:** Collection of experimental data.

Experimental table					
Runs					
	$y_{1,1}$	$y_{1,2}$	...	...	$y_{1,K}$
	$y_{2,1}$	$y_{2,2}$			...
$y_{ij}$	$y_{3,1}$				...
	...				...
	...				...
	$y_{N,1}$	...	...	...	$y_{N,K}$

only the triggering seeds of the random numbers. In the case of a single replication factorial experiment or central composite design application,  $K$  will be equal to the central runs used in the experimental design (in this regard, it should be borne in mind that the variance of the pure experimental error must be constant in each point of the operability region and hence at the centre as well as along the boundary);

- (2) to determine a number  $N \gg 1$  of replications  $y_{ij}$  with  $i = 1, \dots, N$ ,  $j = 1 \dots K$  for each simulation run (see Table 1);
- (3) to calculate for each of the  $K$  runs  $N$  means, defined as  $\overline{y}_{ij}$  with  $i = 1, \dots, n \dots N$ , where

$$\overline{y}_{nj} = \frac{\sum_{i=1}^n y_{ij}}{n}; \quad (2.1)$$

- (4) to calculate  $N$  means of means, defined as  $\overline{Y}_i$  with  $i = 1 \dots n \dots N$ , given by

$$\overline{Y}_n = \frac{\sum_{j=1}^K \overline{y}_{nj}}{K}; \quad (2.2)$$

- (5) to calculate  $N$  values of  $\text{MSPE}_{\text{MED}}$ , with  $1 \leq j \leq K$  and with  $1 \leq i \leq N$ , as shown in Table 2

$$\text{MSPE}_{\text{MED}}(i) = \frac{\sum_{j=1}^K (\overline{y}_{ij} - \overline{Y}_i)^2}{K - 1}. \quad (2.3)$$

These values, transferred on the plane  $(i, \text{MSPE}_{\text{MED}})$ , highlight the trend of the mean square pure error curve in the replicated runs, thus showing, step by step, the entity of the error variance affecting the simulation trial by impacting each objective function.

As mentioned above, according to Cochran's theorem,  $\text{MSPE}_{\text{MED}}$  is the best estimator of the experimental error variance  $\sigma^2$  and, consequently, allows for the measurement of the experimental error affecting the mean value of the means distributions.

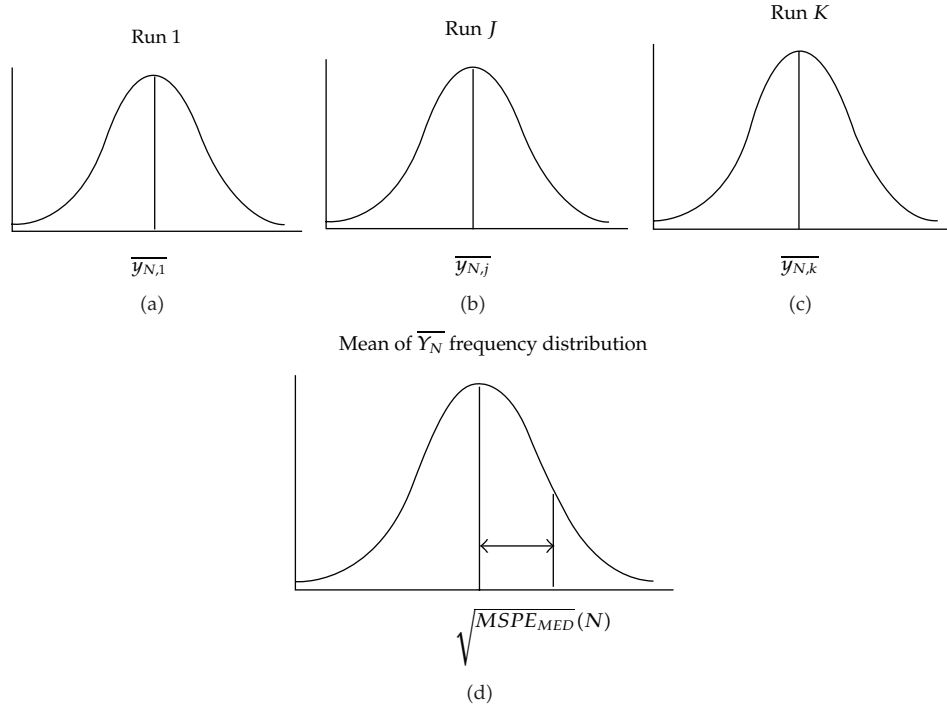


Figure 1:  $MSPE_{MED}$  generation display.

Table 2: Means calculation.

Means table						
Runs						Row mean
	$\bar{y}_{11} = y_{11}$	$\bar{y}_{12}$	...	...	$\bar{y}_{1K}$	$\bar{Y}_1 = \frac{\sum_{j=1}^K \bar{y}_{1j}}{K}$
	$\bar{y}_{21} = \frac{y_{11} + y_{21}}{2}$	$\bar{y}_{22}$			...	
	...				...	
	...				...	
	$\bar{y}_{N1} = \frac{\sum_{i=1}^N y_{i1}}{N}$	...	...	...	$\bar{y}_{NK} = \frac{\sum_{i=1}^N y_{iK}}{N}$	$\bar{Y}_n = \frac{\sum_{j=1}^K \bar{y}_{Nj}}{K}$

Figure 1 effectively shows the  $MSPE_{MED}$  concept as a dependent variable dispersion measure:

- (i) having chosen the number  $N$  of replications, for each of the  $K$  runs, we can obtain a frequency distribution having a mean of  $\bar{y}_{Nj}$ ;
- (ii) the duly sampled  $K$  means,  $\bar{y}_{Nj}$  with  $1 \leq j \leq K$ , produce the mean frequency distribution whose mean is  $\bar{Y}_N$  and whose variance has  $MSPE_{MED}$  as an unbiased estimator.

**Table 3:** Standard deviation calculation.

		Stdev table		
		Runs		Row mean
stdev <sub>ij</sub>	stdev <sub>11</sub> = 0	...	stdev <sub>1k</sub> = 0	stdev <sub>1</sub> = 0
	stdev <sub>21</sub> = $\sqrt{\frac{(y_{11} - \bar{y}_{21})^2 + (y_{21} - \bar{y}_{21})^2}{2}}$	...	...	...
	...	...	...	...
	...	...	...	...
stdev <sub>N1</sub> = $\sqrt{\frac{(y_{11} - \bar{y}_{N1})^2 + \dots + (y_{N1} - \bar{y}_{N1})^2}{N}}$	...	stdev <sub>kN</sub>	stdev <sub>N</sub> = $\frac{\sum_{j=1}^K \text{stdev}_{Nj}}{K}$	

The same approach is also valid for the standard deviation (see Table 3). For each of the  $K$  runs it is possible to calculate  $N$  standard deviation, defined as  $\text{stdev}_{ij}(y_{1j}, y_{2j}, \dots, y_{ij})$  with  $1 \leq i \leq N$ , from which we can obtain  $N$  means of standard deviation, given by

$$\overline{\text{stdev}_n} = \frac{\sum_{j=1}^K \text{stdev}_{nj}}{K}. \quad (2.4)$$

Now we can calculate  $N$  MSPE<sub>STDEV</sub> as follows:

$$\text{MSPE}_{\text{STDEV}}(i) = \frac{\sum_{j=1}^K (\text{stdev}_{ij} - \overline{\text{stdev}_i})^2}{K - 1} \quad (2.5)$$

with  $1 \leq j \leq K$  and with  $1 \leq i \leq N$ .

The knowledge of the MSPE values in each point makes it possible to obtain important inferences on the behaviour of the real experimental responses. In fact, always by the effect of Cochran's Theorem, it makes it possible to know the interval in which there is a 99.7% probability that the value  $y^*$  from a single simulation run lies in it.

While in the time-based evolution systems, having generally to manage small samples (in fact the  $K$  parallel runs in the simulation problems seldom overcome ten elements) the generic expression of this interval is given by

$$\bar{y} - 3\sqrt{\text{MSPE}_{\text{MED}}} \leq y^* \leq \bar{y} + 3\sqrt{\text{MSPE}_{\text{MED}}}, \quad (2.6)$$

in the runs-based evolution systems, what happens is

$$\bar{y} - 3\sqrt{\text{MSPE}_{\text{MED}}} - 3\sqrt{\overline{\text{VAR}} + \text{MSPE}_{\text{STDEV}}} \leq y^* \leq \bar{y} + 3\sqrt{\text{MSPE}_{\text{MED}}} + 3\sqrt{\overline{\text{VAR}} + \text{MSPE}_{\text{STDEV}}}, \quad (2.7)$$

where  $\overline{\text{VAR}}$  is the square of  $\overline{\text{stdev}_N}$ .

Moreover, it should be noted that when each sample of experimental responses, resulting from  $K$  parallel runs, would be broad enough to provide an exhaustive description

of the population behaviour, the two MSPEs evolving in the runs would crash on the  $X$ -axis (MSPE = 0). This way the totality of the stochastic description of the real system, and, hence, in the model, is represented by the experimental response variance. Put otherwise

$$\lim_{N \rightarrow \infty} \text{MSPE}_{\text{MED}} = \lim_{N \rightarrow \infty} \text{MSPE}_{\text{STDEV}} = 0 \quad (2.8)$$

for which

$$\bar{y}_{\text{MED}} - 3\sqrt{\text{VAR}} \leq y^{*\infty} \leq \bar{y}_{\text{MED}} + 3\sqrt{\text{VAR}}, \quad (2.9)$$

For the experimenter, the problem is not to obtain a theoretical MSPE = 0 but to limit the number of runs  $N$  through a careful check of the experimental error evolution in terms of both of magnitude and adjustment, so as to limit also its impact on  $y^*$  to acceptable values.

As previously shown, the parameter  $N$  influences, for each replication, the number of runs that needs to be carried out in the statistical parameters calculation (mean, etc.) of the dependent variable and the number of  $R$  survey points of the dependent variable MSPE calculation.

As concerns the number of  $K$  runs carried out in parallel, it is clear that the interest to choose a high  $K$  value can be correct. In fact, the higher the  $K$ , the wider the sample used for the operation. Therefore, the  $K$  size necessarily affects the accuracy of the mean of the dependent variable mean/variance distribution. In many cases, despite the computational power available, it may happen that, as  $K$  grows, the time to calculate MSPE quickly becomes heavy.

It is an obvious consequence that the study of the experimental error evolution, and the resulting search for the characteristic background noise, can be particularly burdensome, due to the lack of a careful setting of the various parameters.

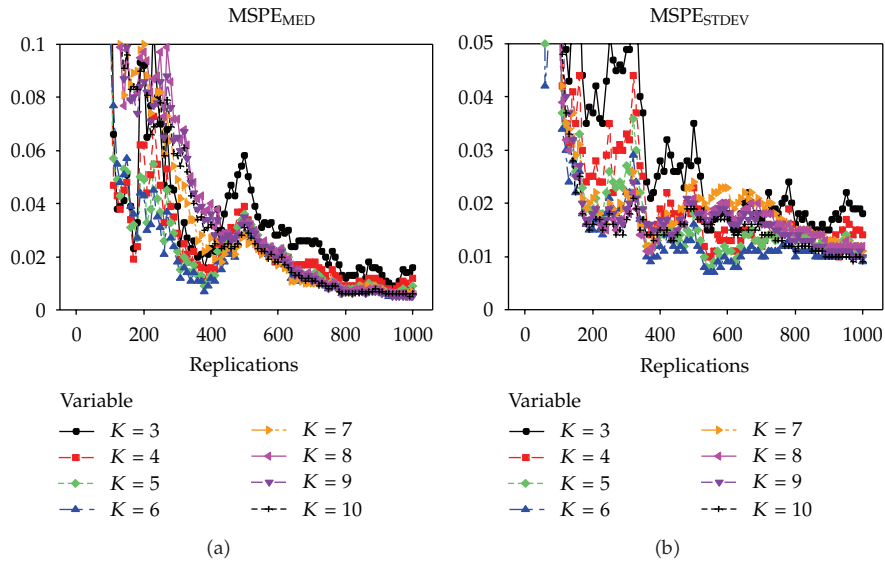
While, for the purpose of a correct evaluation of both the stochastic effect on the experimental response and the characteristic noise, it is important for  $N$  to choose values having a size of  $10^4$  or greater.

Figure 2 displays the comparison of the behaviours of the mean and standard deviation of the mean squares pure error, with reference to the simulation model described further below in paragraph 3, for each  $K$  value in the chosen 3–10 inquiry range.

The figure confirms the poor influence of the  $K$  replication number. It should be noted that, both for  $\text{MSPE}_{\text{MED}}$  and  $\text{MSPE}_{\text{STDEV}}$ , the curves tend to overlap after about 600 runs, with the only exception of  $K = 3$  which maintains a trend of about a hundredth greater than the others.

This leads to the conclusion that it is not necessary to operate with high values of  $K$  since, as already shown, it is a linear multiplication factor of the computational effort [3]. On the contrary, as it will be seen further below in other case studies, significant changes of statistical parameters can be obtained by increasing the  $N$  replications of each run (with  $K$  limited to 5). This can be logically correct as this operation is equivalent to an increase of the extraction number from the frequency distributions and, hence, to a more careful identification of them using experimental samples having a greater width.

Finally, it should be remembered that, in multiobjective problems, the curve stabilization point to be used as the “optimal run length” is that of the curve that reaches the stabilization



**Figure 2:** Influence of the number of replications on the behaviour of  $MSPE_{MED}$  and  $MSPE_{STDEV}$ .

phase with the greatest number of runs (in this case the greater between the  $MSPE_{MED}$  and the  $MSPE_{STDEV}$  as shown in Figure 2).

### 3. Applications to Two Simulation Models

In order to better understand the application of the methodology under consideration, two applications addressing two different types of business management problems are illustrated below.

#### 3.1. The IT Corporate Model

The application is related to an engineering company, operating internationally in the sale to third parties and performing design only or design plus supply for the realization of engineering works with the “turn-key” formula.

The company model [8] is made up of three fundamental subsystems as shown in Figure 3:

- (1) marketing and production: fed with exogenous variables (market, management), it is bound to generate both the necessary personnel and the other management variables as an output,
- (2) personnel: fed with the number of personnel provided as well as by exogenous phenomena such as overtime and production bonus, it is possible to estimate the costs related to the said personnel,
- (3) profit and loss: fed with the output resulting from the two previous models, it gives the economic parameters (i.e., EBTDA) in the future accounting periods as an output.

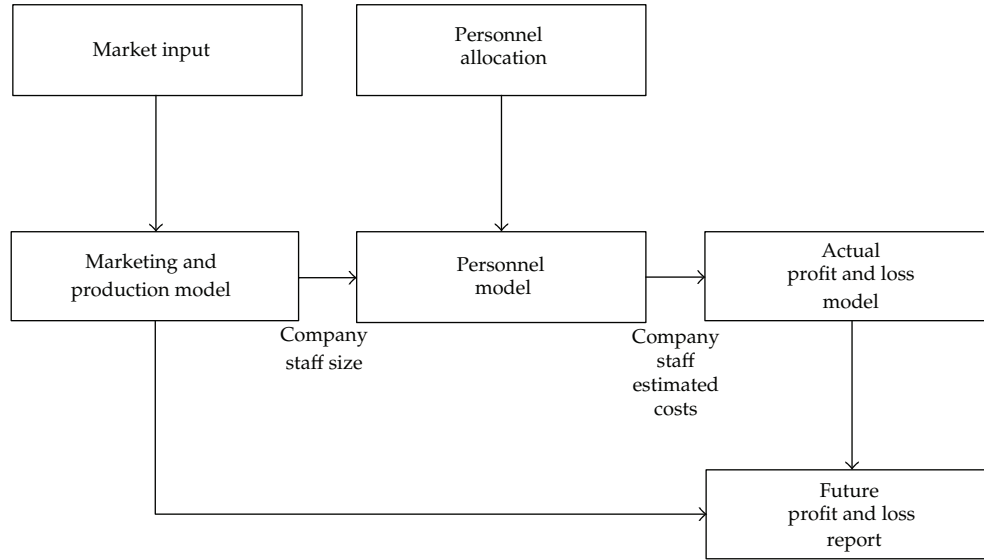


Figure 3: Model structure scheme.

By setting the four years following the last accounting period as maximum temporal forecast horizon, the model allows to carry out yearly forecast of

- (i) order situations (acquired and acquirable),
- (ii) in-house necessary resources (management variables),
- (iii) career movements, seniority, and personnel scholarship level,
- (iv) incomes and EBTDA.

Uncertainty has been applied to all the significant variables of the marketing and production model and the profit and loss one, by means of  $\eta(t)$  coefficients that multiply the basis relationships of the former deterministic model. Each variable takes the form of a normal-type distribution  $(\mu(\eta(t)), \sigma(\eta(t)))$  with a growing variance in the following years ( $t$ ) in order to take into consideration the increasing levels of uncertainty with the projection in the future (Figure 4).

The information provided about the IT model architecture and about the nature of frequency distributions clearly underscores how difficult it is to establish a link between traditional experimentation and the reliability of the output results. Indeed, it is absolutely wrong to believe that it is possible to obtain statistically reliable results with a single extraction or just a few extractions from each of the 52 normal distributions of the input variables.

At this point, the necessity to carry out an experimental strategy capable of overcoming this drawback is clear. It can be achieved by using an  $N$  number of replications for each simulation run:

- (i) coverage suitably, the input distributions.
- (ii) the same distributions are affected by an experimental error rate which is under the experimenter's control and wholly visible in its evolution as the  $r_p$  survey points grow.



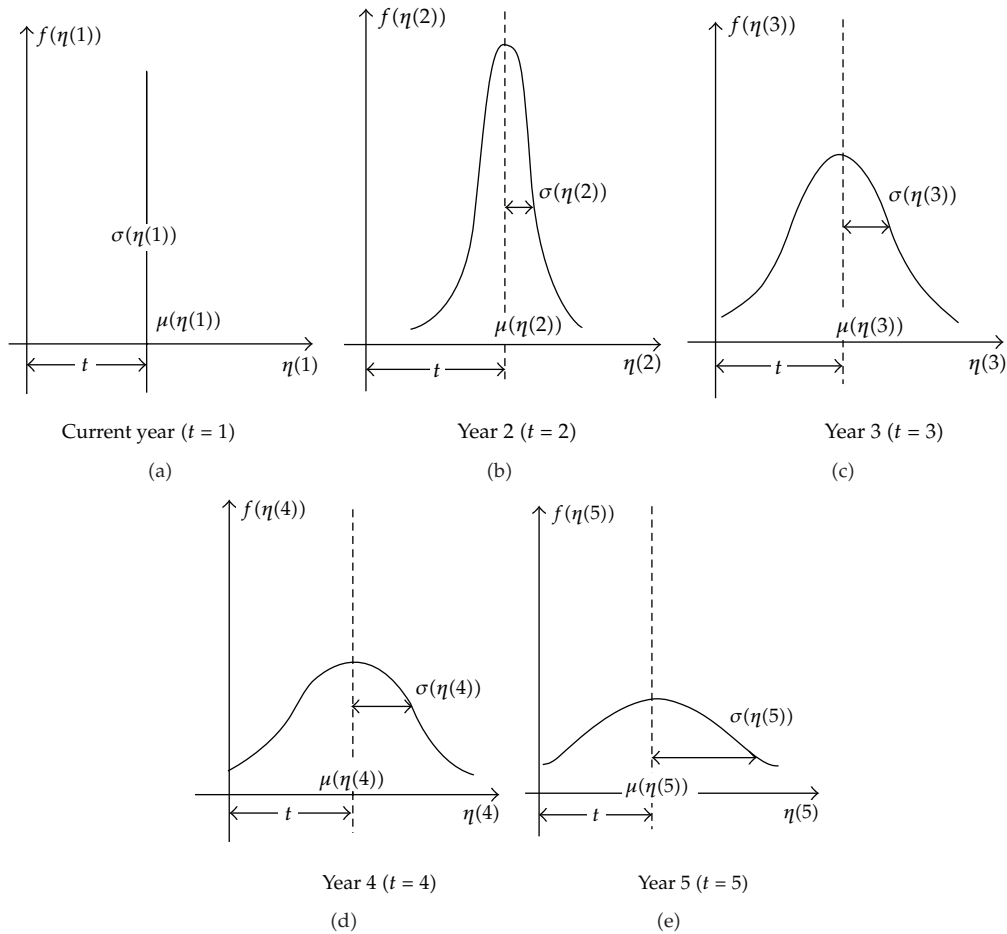


Figure 4: Scheme of variable stochasticity.

In order to evaluate the methodology efficacy, the authors have carried out a series of tests on an objective function, EBITDA, expressing the maximum model sensitivity. The EBITDA is calculated with the “profit and loss” model whose input is the results from the “marketing and production” and “personnel” submodels and accurately represents the ultimate synthesis of the entire action of the forecasting model.

It is important to note in this regard that the response obtained from the IT model on the EBITDA shows the following features:

- (i) the  $MSPE_{MED}$  and  $MSPE_{STDEV}$  of the two population distributions, at the changing of  $K$ , have values that are undoubtedly low ( $10^{-3}$ ). This points out, differently from many other cases, a strong centripetal trend with consequent confidence intervals width on the mean responses that are almost negligible for  $N > 200$  (Figures 2 and 5);
- (ii) the mean variance  $\overline{VAR}$  is significantly high if compared to the  $\overline{Y}_N$  (the value is about 3,5 versus 23.30) The oscillation field of the dependent variable EBITDA at the Year 5 gives a range of expected responses, compared to 23.30 included, between

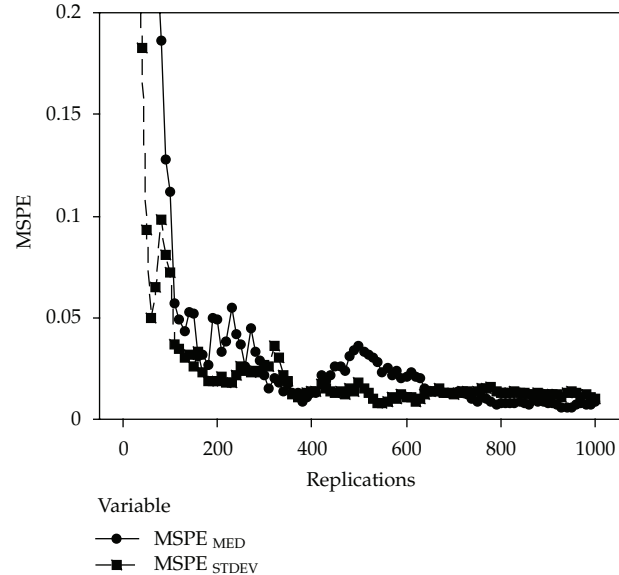


Figure 5: IT Model:  $MSPE_{MED}$  and  $MSPE_{STDEV}$  evolution trend.

29 and 17.6. This happens although, for  $N > 200$ , both MSPEs are almost negligible.  $\overline{VAR}$  is in fact the parameter connected to the real system stochasticity. The ability of the experimenter, thanks to the proposed methodology, is to not add to model outputs a noise not characteristic of the real system but strictly connected with the experimental phase.

In these cases, by extending  $N$  up to values that are apparently out of a common experimental practice (such as 25,000/30,000/50,000 and more) it will be generally possible to observe important drops both in the MSPE of the standard deviation and, in some cases of systems affected by endogenous stochasticity of a particular type, in the magnitude of the mean variance.

It is duty to remind that, when optimisation problems are faced using Monte Carlo simulators, the three variance sources ( $\overline{VAR}$ ,  $MSPE_{MED}$ , and  $MSPE_{STDEV}$ ) must be strongly controlled since they can generate, with their magnitudes, significant deviations of the punctual mean responses by putting into evidence, in consequence, distorted and/or inexistent superficial response behaviours.

### 3.2. The Project Management Model

The objective of this study was the evaluation of the job delivery date of a company order relative to the revamping-reburning-S.A.B. of a 320 MW unit of a thermoelectric power plant [9]. For this scope the authors developed a tool to

- (i) deal with the problem of forecasting in a stochastic regime,
- (ii) provide Project Managers with an online control methodology (fed by data coming from the job site) to rebalance the effects of the stochastic nature and/or incorrect initial forecasts.

The application of stochasticity to the deterministic duration values supplied by the estimators takes place by assuming the duration of each activity equal to appropriate frequency distributions.

Nine delay scenarios were defined for the activities comprising the order:

- (i) Gaussian, with a mean equal to the value hypothesised by the deterministic approach and standard deviations equal to 33%, 20%, and 10% of the activity duration,
- (ii) negative exponential, with the same mean and standard deviations as indicated above,
- (iii) mixed, part Gaussian and part negative exponential, with the same mean and standard deviations as indicated above.

Undoubtedly, the scenarios will be characterised by an uncertainty peak at the estimation time with a resulting effect on the delivery dates. This uncertainty is bound to decrease with job progress under the effect of the progressive transformation of the various stochastic elements into deterministic ones (i.e., completed operations or, in phase of completion, corrective interventions carried out in progress) and, hence, with the possibility for the model to anticipate the delivery date with reasonable accuracy.

Thanks to the transformation of increasing portions of activities that switched from the stochastic to the deterministic regime and to the improved knowledge of the possible delays for the future activities, it became possible, during the experimentation phase, to reduce the global uncertainty level of the object system. As a consequence, it is possible to gradually approximate, with increasing accuracy, the maximum and minimum durations of the possible plant delivery delay.

Starting from the nine hypothesised scenarios, the authors decided to use the Monte Carlo method for each scenario to find the probabilistic delivery range compared to the job final delivery date hypothesised by the company.

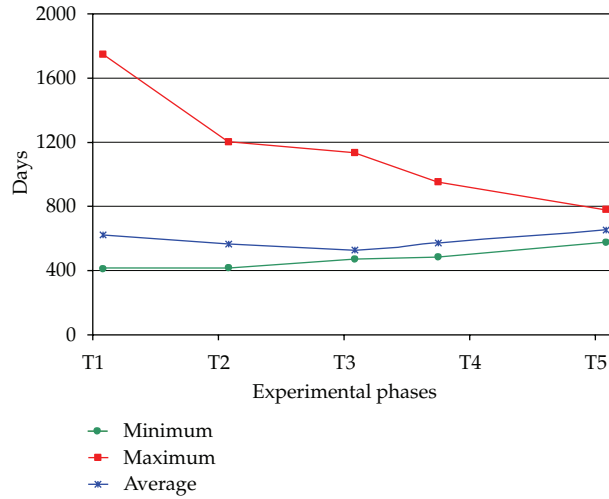
Figure 6 shows the output of the simulation in terms of job completion time obtained by the simulation model where there is a noticeable variation (such as maximum, minimum and average value) in relation to the stochasticity and time progress.

One of the authors' targets was to underscore how changes in the endogenous stochasticity can induce changes in the related  $MSPE_{MED}$  and  $MSPE_{STDEV}$  behaviour and consequently in the reliability of the model output.

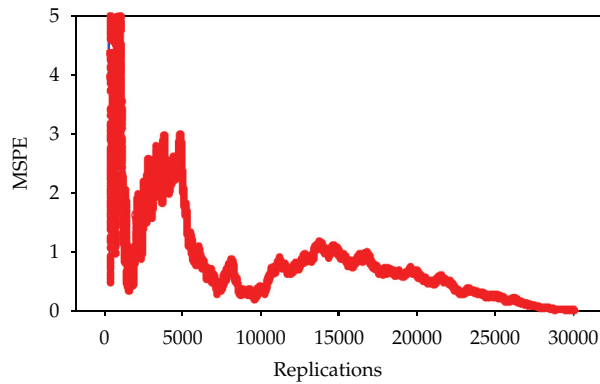
Among the stochastic typologies considered, it has been chosen by way of example to select two different scenarios characterised, respectively, by

- (1) negative exponential having a variance of 33% compared to the deterministic value of each operation, being assumed, by the estimators, as the minimum completion time of each operation,
- (2) negative exponential having a variance of 10%.

Figures 7 and 8 show the behaviour of  $MSPE$  mean evolution in the case of the 33% negative exponential distribution. As it can be noted, the mean of the  $MSPE_{MED}$  shows a beginning phase, up to 5000 replications, characterised by broad oscillations with a first settlement phase around  $0.6 \text{ days}^2$  between 5,000 and 15,000 runs. It follows a decreasing phase, with a more or less constant slope of about  $45^\circ$  compared to the  $x$ -axis. This phase is characterised by punctual oscillations having a very small width, bringing the  $MSPE_{MED}$  value at 30.000 runs within an interval of  $\pm 0,002 \text{ days}^2$ .



**Figure 6:** Project delivery date with stochasticity.

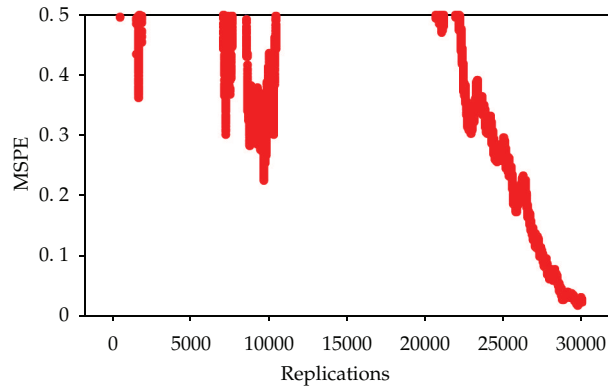


**Figure 7:**  $MSPE_{MED}$  exp neg 33%.

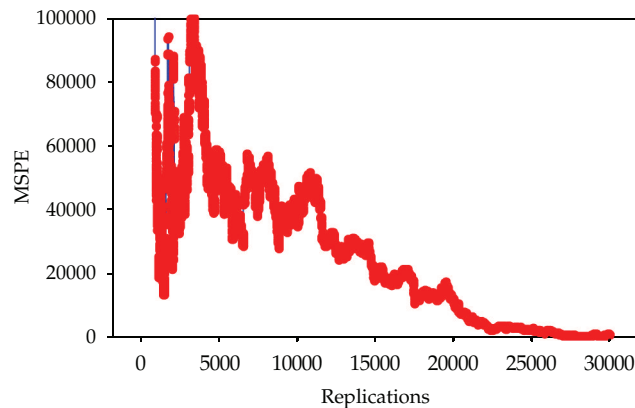
It is interesting to note how the  $MSPE_{STDEV}$  contribution on the total delay of the job delivery is, in any case, of great relevance at 30,000 runs (Figure 9). It involves a standard deviation on the mean variance distribution of about 30 days, with the increase of the maximum completion date to about 330 days due to the two variances, in terms of the related standard deviations. This analysis phase on the stochasticity resulting from the 33% negative exponential points out, with respect to the budgeted time equal to 380 days, that the maximal cumulative delay is about 500 days.

The analysis of the “10% negative exponential” case underlines an  $MSPE_{MED}$  evolution trend characterised by strong oscillations between 0 and 5,000 runs with a rapid fall of the MSPE value to  $5E-2$ , followed by a second phase of further descent between 5,000 and 15,000 runs up to  $E-3$  with a final settlement trend around this value in the last curve part (Figure 10).

As concerns  $MSPE_{STDEV}$  (Figures 11 and 12), it is possible to observe a widely oscillating phase up to 10,000 runs followed by a sudden decrease from  $100 \text{ days}^2$  to  $20 \text{ days}^2$  between 10,000 and 15,000 runs to achieve a settlement around to  $20 \text{ days}^2$  at 30,000 runs after a small peak at  $50 \text{ days}^2$  at 22,500 runs.



**Figure 8:**  $MSPE_{MED}$  exp neg 33% (zoom).



**Figure 9:**  $MSPE_{STDEV}$  exp neg 33%.

By translating the influence of the 10% negative exponential in terms of maximum delay, it is possible to note a possible increase of 117 days compared to 380 days scheduled in the budget.

Considering the behaviour of the variance evolution, the comparative analysis of the  $MSPE_{MED}$  and  $MSPE_{STDEV}$  curves in the 10% negative exponential case (Figure 13) shows the need to use a value of  $N$  greater than 30,000.

The comparative analysis of the mean and variance MSPE curves in the 33% negative exponential case (Figure 14) also shows, considering the behaviour of the variance evolution, the need to use an  $N$  value greater than 30,000.

The endogenous stochastic variation is highlighted by the model in a strong way both in terms of mean and variance of the mean populations and in terms of MSPE of both parameters. In particular, in both case studies, with reference to the MSPE, it is the standard deviation that assumes an important role in terms of impact on the final delivery date, while the  $MSPE_{MED}$  in both cases (10%, 33%) leads to a stabilization at values, which cannot affect, in a significant way, the mean of the means value.

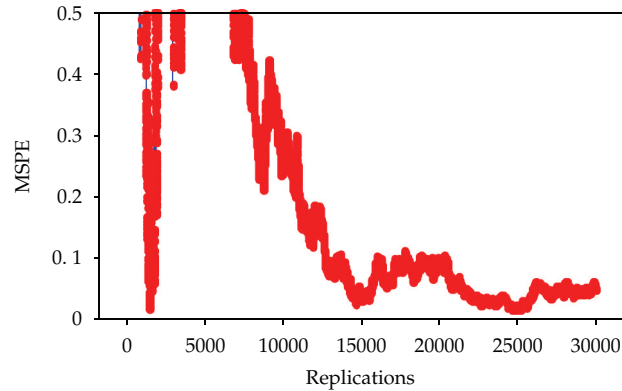


Figure 10:  $MSPE_{MED}$  exp neg 10% (zoom).

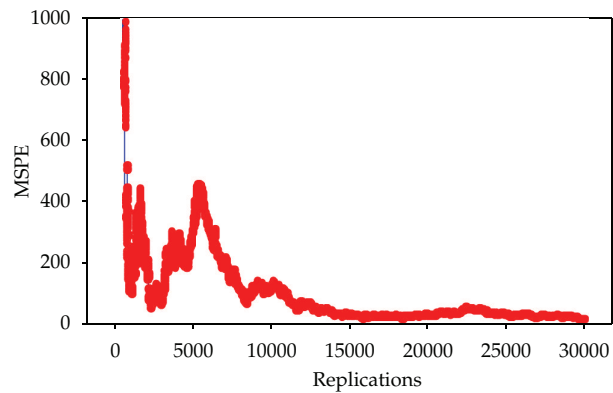


Figure 11:  $MSPE_{STDEV}$  exp neg 10%.

As shown above, the fundamental role of the MSPE evolution curves lies in indicating the value of the optimal  $N$  at which it is necessary to stop the experimental phase, with the certainty that the final result, in a ratio of costs/benefits, will not be influenced by the inner stochasticity of the model.

From this point of view, the experimentation on the project model confirms the validity of the methodology that in both cases, characterised by an extremely different magnitude of the duration distribution variances, identifies 30,000 as the lowest number of runs necessary to achieve a suitable stabilisation of the  $MSPE_{STDEV}$ .

The experimentation underscores once again that the role of the experimental response mean population variance cannot be set aside. This is well known in elementary statistics, but it is forgotten too often by simulation model builders.

#### 4. Conclusions

The aim of the authors was the theoretical systemisation of the mean square pure error evolution methodology in the replicated runs, not only through the in-depth study of some

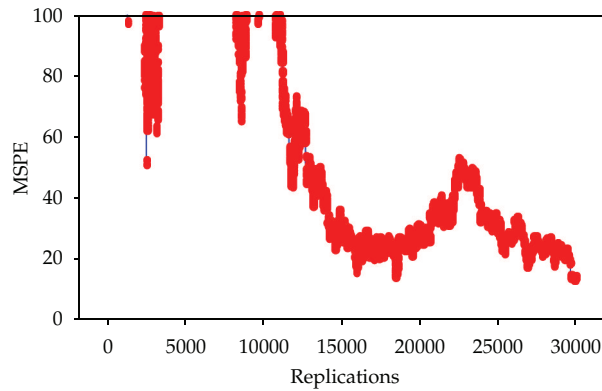


Figure 12:  $MSPE_{STDEV}$  exp neg 10% (zoom).

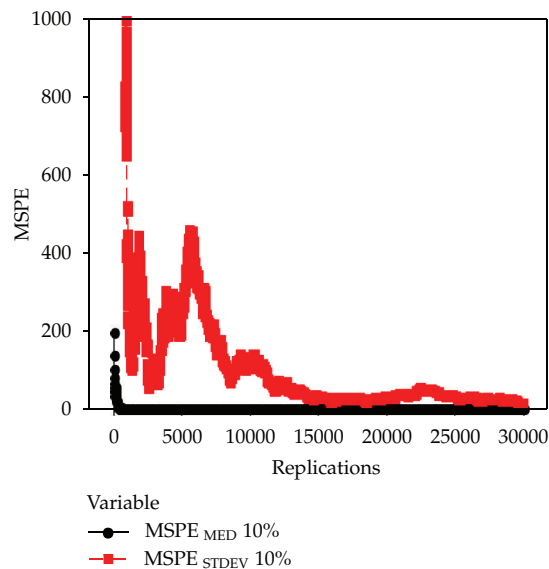


Figure 13: Comparison of  $MSPE_{MED}$  versus  $MSPE_{STDEV}$  exp neg 10%.

theoretical fundamental steps but also through their application to some appropriately chosen simulation models.

The methodology, outlined in the mid-80s, had never been studied in an exhaustive way, because it was considered for use only in few particular operating situations studied through the use of Monte Carlo simulation models [8].

The observations proposed in this work point out, on the contrary, a wide range of model situations to which the tool can be applied to obtain a correct experiment. Among these, the authors recall the models of systems operating on jobs, which are temporally closed and nonrepetitive, the stochastic models for project management, models for the description of repetitive operations characterized by a demand of high variability performances, and the financial models with an evolution in future time [10].

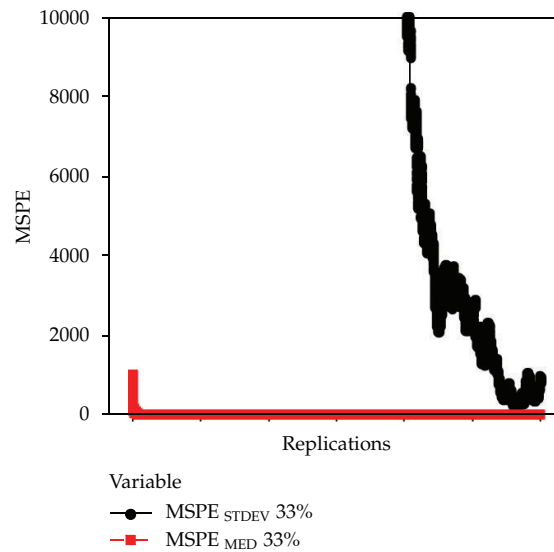


Figure 14: Comparison of  $MSPE_{MED}$  versus  $MSPE_{STDEV}$  exp neg 33%.

The key point of this discussion is the acknowledgment of the need, in these kinds of models, of the simultaneous control not only of the two traditional output values, as in the objective function study of temporal evolution models (response and related MSPE), but also of four parameters, such as mean, variance, and the related MSPE, with the variance being the element with the greatest capacity of affecting punctual response.

Based on these assumptions, the role of the  $MSPE_{MED}$  and  $MSPE_{STDEV}$  is fundamental in determining the run number capable of providing the experimenter the most suitable ratio between run time and punctual response accuracy.

## References

- [1] G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York, NY, USA, 1987.
- [2] R. H. Myers and D. C. Montgomery, *Response Surface Methodology*, John Wiley and Sons, 1995.
- [3] D. C. Montgomery, *Design and Analysis of Experiments*, John Wiley and Sons, 1997.
- [4] R. Mosca and P. Giribone, "Optimal length in O.R. simulation experiments of large scale production system," in *Proceedings of the IASTED International Symposium on Applied Modelling and Simulation (AMS '82)*, pp. 78–82, Paris, France, 1982.
- [5] L. Cassettari, R. Mosca, and R. Revetria, "Experimental error measurement in Monte Carlo simulation," in *Handbook of Research on Discrete Event Simulation Environments*, E. M. O. Abu-Taieh and A. A. R. El Sheikh, Eds., chapter 6, pp. 92–141, Information Science, 2010.
- [6] R. Mosca, P. Giribone, M. Scotti, and M. Mondani, "Dynamic management of a DSS applied to a factual company economy," in *Proceedings of the 7th IASTED International Symposium on Modelling, Identification and Control (MIC '88)*, pp. 16–20, Grindelwald, Switzerland, 1988.
- [7] R. Mosca, A. G. Bruzzzone, L. Cassettari, and M. Mosca, "Risk analysis for industrial plants projects: an innovative approach based on simulation techniques with experimental error control," in *Proceedings of the European Modelling and Simulation Symposium (EMSS '09)*, Tenerife, Spain, September 2009.
- [8] R. Mosca, P. Giribone, S. Casciani, and G. Ghislanzoni, "A concrete application of a modular decision support system," in *Proceedings of the 6th IASTED International Symposium on Modelling, Identification and Control (MIC '87)*, pp. 320–328, Grindelwald, Switzerland, 1987.



- [9] M. Schenone, A. Frigato, and M. Mosca, "Risk analysis applied to project management and project control," in *Proceedings of the 19th IASTED International Symposium on Modelling, Identification and Control (MIC '99)*, Innsbruck, Austria, 1999.
- [10] L. Cassetari, P. G. Giribone, M. Mosca, and R. Mosca, "The stochastic analysis of investments in industrial plants by simulation models with control of experimental error: Theory and application to a real business case," *Applied Mathematical Sciences*, vol. 4, no. 73–76, pp. 3823–3840, 2010.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

