We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800

Open access books available

122,000

International authors and editors

135M

Downloads

154

Countries delivered to

Our authors are among the

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

CLARIVATE ANALYTICS

BOOK CITATION INDEX

INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Sensitivity Analysis in Discrete-Event Simulation Using Design of Experiments

José Arnaldo Barra Montevechi,
Rafael de Carvalho Miranda and Jonathan Daniel Friend

Additional information is available at the end of the chapter

1. Introduction

The use of discrete-event simulation as an aid in decision-making has grown over recent decades [1, 2, 3, 4]. It is already used as one of the most utilized research techniques for many sectors due to its versatility, flexibility and analysis potential [5, 6].

However, one of simulation's greatest disadvantages is that, on its own, it does not serve as an optimization technique [7]. This forces simulation practitioners to simulate multiple system configurations and choose the one which presents the best system performance. Computational development has helped alter this scenario due to the increasing availability of faster computers and ever improving search and heuristic optimization techniques.

Simulation optimization can be defined as the process of testing different combinations of values for controllable values, aiming to find the combination of values which offers the most desirable output result for simulation models [8].

In support of this claim, [1, 9, 10, 11] assert that using optimization along with simulation has been continuously increasing due to the emergence of simulation packages which possess integrated optimization routines.

The overarching idea of including these routines is to search for improved definitions for the system parameters in relation to its performance. However, according to [10], at the end of optimization, the user has no way of knowing if an optimal point was truly reached.

Despite the fact that simulation has been around for more than half a century, until quite recently, the scientific community was reluctant to use optimization tools in simulation. The first time the subject emerged in two renowned simulation books, [12] and [13], was at the close of the 20th century [9]. This resistance has begun to diminish with the convent of meta-heuristic research, along with strides being made in statistic analysis [14].

According to [15], verification of system performance for a determined set of system parameters with reasonable precision using simulation, demands a considerable amount of computational potential. In order to find the optimal or near-optimal solution, one needs to verify a large number of parameter values, thus, optimization via simulation is normally exhaustive from a computational standpoint.

Having highlighted the computational strains, [8] states that despite the evolution of optimization software, a common criticism made about such commercial packages is that, when more than one variable is manipulated at a time, the software becomes very slow.

Considering that not all decision variables are of equal importance in respect to the response variable that one desires to optimize [16, 17], a sensitivity analysis may be carried out on the simulation model in order to select the variables which will compose the optimization search space in order to limit the number of variables and, in turn, make the search faster.

Thus, in order to proceed to the variable selection, screening can be done in order to separate the most important variables from those which may be eliminated from consideration [16, 17]. The same author presents some examples of experimental design utilized in screening experiments:

- 2^n factorial design;
- 2^{n-p} fractional factorial design;
- Supersaturated designs;
- Groups screening designs.

The current chapter presents an application of Design of Experiments (DOE), specifically fractional factorial design, in order to select the significant input variables in a simulation model, and thus accelerate the optimization process. For information about experimental design, the reader can consult [1, 4, 18, 19].

Fractional factorial design is a DOE technique in which only a fraction of the total number of experiments is executed, thus realizing fewer experiments than full factorial design. Throughout this chapter, it is shown that the use of such a design serves to reduce the search space in the optimization phase of simulation studies.

In this chapter, real examples of how to conduct sensitivity analysis with factorial design are given. In order to reach this goal, two study objects are presented, comparing the optimization carried out without previous investigation of input variable significance, with the optimization carried out in reduced search space. Finally, a comparison is made of the results of the optimization, with and without the sensitivity analysis.

2. Simulation optimization

A simulation model generally includes n input variables (x_1, x_2,...,x_n) and m output variables (y_1, y_2, ..., y_m) (Figure 1). The optimization of this simulation method implies finding the optimal configuration of input variables; that is, the values of x_1, x_2, …, x_n which optimize the response variable(s) [20].
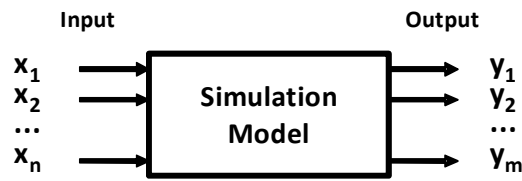
Figure 1. Simulation Model [20]

Optimization helps respond to the following questions: What are the optimal adjustments to the input variables (x) which maximize (or minimize) a given simulation model output? The objective is to find an optimal value which maximizes or minimizes a determined performance indicator [11].

According to [21], simulation optimization is one of the most important technologies to come about in recent years. These authors recall that previous methodologies demanded carrying out complex changes to the simulation model, thus consuming time and computational potential and, in many cases, not even being economically viable for real cases due to the large number of decision variables.

A traditional simulation optimization problem (minimization of a single objective) is given in Eq. 1 [22]:

$$\min f(\theta) \tag{1}$$

$$s.t\ \theta \in \Theta$$

Where $f(\theta) = E[\psi(\theta, \omega)]$ is the system's expected performance, estimated for $\hat{f}(\theta)$, which is obtained using the simulation model samples $\psi_j(\theta, \omega)$, observed according to the discrete or continuous input parameters, restricted by θ within a viability set $\Theta \subset \mathfrak{K}^d$.

According to [23], the optimization method which serves for the problem presented in Eq. (1) depends on whether the simulated variables are discrete or continuous. There are many methods for resolving this problem in the literature, like the one presented in Eq. (1), and unfortunately depending on the model being optimized, some methods cannot guarantee that an optimal solution is found [24].

Table 1 shows the main optimization software packages which are both on the market and cited in academic literature, as well as the simulation packages with which they are sold. The optimization techniques utilized in each software package is also shown.

As shown in Table 1, different optimization software packages utilize different search methods, such as: Evolutionary Algorithms [25], Genetic Algorithms [26], Scatter Search [27], Taboo Search [28], Neural Networks [29] and Simulated Annealing [30].

According to [31] and [32], the simulation optimization's greatest limitation is the number of variables being manipulated, as the software's performance is considerably reduced in models with a great number of variables. Thus, [33] asserts that convergence time is the most significant restriction in reaching computational efficiency for optimization algorithms.

Optimization Software	Simulation Package	Optimization Technique
AutoStat®	AutoMod®, AutoSched®	Evolutionary and Genetic Algorithms
OptQuest®	Arena®	Scatter Search, Taboo and Neural Networks
Optimiz®	Simul8®	Neural Networks
Optimizer®	Witness®	Simulated Annealing e Taboo Search
SimRunner®	ProModel®	Evolutionary and Genetic Algorithms

Table 1. Optimization software packages [4, 7, 9]

In order to ease this process, the use of fractional factorial design can be used to conduct sensitivity analysis on a simulation model in order to select the input variables which truly impact the response variable and enable the elimination of variables which are not statistically significant. In terms of the simulation, sensitivity analysis may be interpreted as a systematic investigation of the model's outputs, in accordance with the model's input variables [19].

By using DOE techniques, it is possible to reduce the number of experiments executed, determine which independent variables affect the dependent variable, and identify the amplitude or intensity of this effect. For optimization purposes, identification of the most significant variables is important, as the greater the number of variables in the search space, the longer the optimization process will take.

Thus, by using sensitivity analysis in simulation optimization problems, one can work with those input variables which actually have a significant effect over the determined response variable, thus reducing the number of experiments necessary and the computational potential involved in this process.
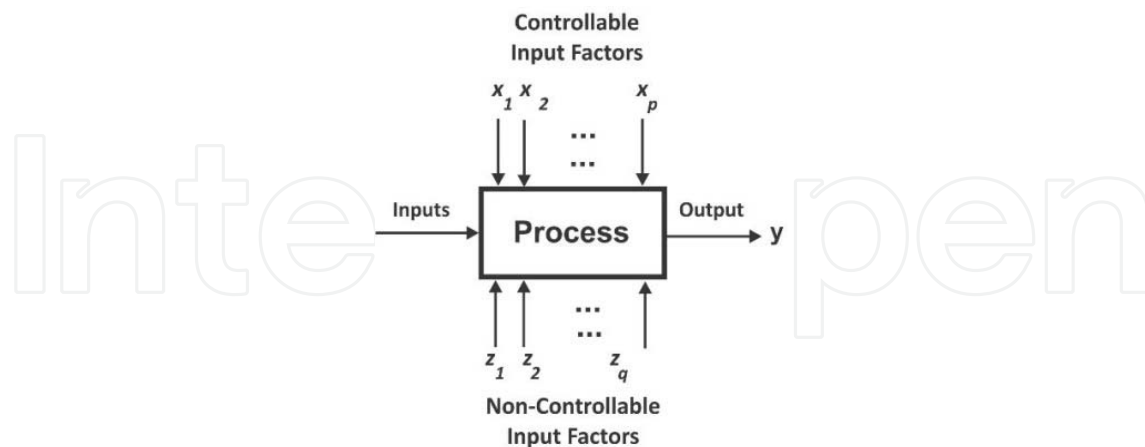
3. Experimentation strategies

An experiment can be defined as a test or series of tests in which purposeful changes are made to input variables, with the objective of observing and identifying the form in which the system responses are affected, in function of the changes carried out on the input variables [18].

According to [34], there are two types of process variables (Figure 2): controllable variables (x_1, x_2, ..., x_P), and non-controllable variables (z_1, z_2, ..., z_q), which are many times called "sound". The same author states that the experiment's objectives can be:

- Determine the variables which have the most influence over the response (y);
- Determine the values of x (significant variables) in order that the response is close to the nominal demand;
- Determine the values of x (significant variables) in order that the variability in y is small;

- Determine the values of x (significant variables) in order that the effect of the non-controllable variable effects are minimized.

Figure 2. General process model [34]

The experimentation strategy is the method of designing and conducting experiments [18]. According to this author, there are many methods which can be used for the realization of experiments. Some examples are listed below:

Best-guess: This strategy is based on the specialists' technical or theoretical knowledge, which alters the value of one or two variables for the test in function of the previous result. This procedure presents at least two disadvantages: The first disadvantage occurs when the initial configuration does not produce the desired result and then the analyst must search for another input variable configuration. These attempts may continue indefinitely and certainly take a long time without guaranteeing success. The second disadvantage is that, supposing an acceptable initial configuration, the analyst will be tempted to stop testing, even though there is no guarantee that the best results has been obtained.

One factor at a time: This strategy involves selecting the starting configuration for each input variable and then successively varying each variable within a given range, while simultaneously maintaining the other variables constant. The greatest disadvantage of this strategy is its inability to detect interaction between variables; nonetheless, many analysts disregard this fact and it is often used [18].

Factorial Design: According to [18], when an experiment involves the study of two or more factors, the most effective strategy is factorial design. In using this strategy, factors are altered at the same time, instead of one at a time. That is, in each complete attempt or experimental replica, all possible combinations are investigated [35]. This strategy is more efficient than the one previously mentioned, as it allows for the effects of a single factor to be estimated across various factor levels, thus leading to valid conclusions within the experimental conditions, [18] and is the only way to discover interaction between factors [18, 35], avoiding incorrect conclusions when there is interaction between factors. The main problem with factorial design is the exponentially increasing number of combinations with each increase in the number of factors [19].

Response Surface Methodology (MSR): This method consists in a set of mathematical and statistical techniques that are used for modeling and analysis, in which the response of interest is influenced by multiple variables, and the objective is to optimize this response [35]. According to these authors, the relation between the independent and dependent response variables is unknown in most problems. [35] states that the first step of MSR is finding an accurate approximation for the true relation between the response (y) and the independent variables. In general, polynomials with a low degree are used to model a given region of the independent variables.

4. Simulated experiment design

According to [36], although classic experimental design methods were developed for real world experiments, they are perfectly applicable to simulated experiments. In fact, according to the same author, simulated experiment design presents many opportunities for improvements which are difficult or impossible to carry out using actual experiments.

[37] asserts that research related to experimental design are frequently found in specialized publications, but they are rarely read by simulation practitioners. According to these same authors, most simulation practitioners can get more from their analyses by using DOE theory developed specifically to experiment with computational models.

The benefits of DOE enable the possibility of improving simulation process performance by avoiding trial-and-error searches for solutions [38]. More specifically, the use of factorial design can minimize or even eliminate the disadvantages brought about by experimenting with simulated systems instead of the real system.

According to [36], in order to facilitate the understanding of simulation's role in experimental execution, it is necessary to imagine a response value (Y) or a dependent value variable can be represented in the following equation:

$$Y = f(x_1, x_2, ..., x_n) \tag{2}$$

Where:

- x_1, x_2, x_n represent the input variables, factors or independent variables;
- f represents the simulation model's transformation function.

[39] declares that simulation is a black box which transforms inputs variables into simulated outputs, which imitate the real system's perspective output. For each scenario, the analyst carries out one or more runs and registers the average output values.

In simulation models, the levels chosen for each factor must enable the effects to be programmed in the model. In order to exemplify this question, the following situation is proposed: a determined factor which is desired to be optimized corresponds to the possibility of using an experienced employee (upper level) or a new hire (lower level), thus verifying, what the impact would be on daily throughput. In simulation models, the modeler must be familiar with each variable to be affected by the change in levels. Thus, the modeler must decide which distribution to use for each variable time in the operation.

Experimentation using simulation presents some special advantages over using physical or industrial systems [4]:

- By using simulation, it is possible to control factors that, in reality, are uncontrollable, such as client arrival rate;
- By using simulation, it is possible to control the basic origin of variation, which is different from physical experiments, thus avoiding use of blocks.

Another experimental design characteristic is that commercial simulators come with random number generators and therefore, from an experimental point of view, the trouble of randomizing the experimental replicas is eliminated. Randomization is a problem with physical experimentation [36].

5. Design and analysis of experiments

According to [18], DOE can be defined as the process of designing experiments in order that the appropriate data are collected and analyzed by statistical methods, thus resulting in valid and objective conclusions. Any experimental problem must contain two elements: experimental design and statistical data analysis.

DOE techniques are seen with a broad range of application in many knowledge areas, thus showing itself as a set of tools of great importance for process and product development.

Those involved in the research should have a previous idea of the experiment's objective, which factors will be studied, how the experiment will be conducted and a comprehension of how the data will be analyzed [34].

According to [18], DOE should consider the following stages:

1. **Problem recognition and definition:** Completely develop all ideas about the problem and the objectives to be attained through the experiment, thus contributing to greater comprehension of the process and eventual problem solution;
2. **Choice of factors and working levels:** Choose the factors to undergo alterations, the intervals of these factors and the specific levels for each run to be carried out;
3. **Selection of the response variables:** Determine the response variables which really supply useful information about the performance of the process under study;
4. **Selection of the experimental design:** Consider the sample size (number of replications), selection of the correct order of runs for the experimental attempts, or the formation of blocks and other randomization restrictions involved;
5. **Realization of experiments:** Monitor the process to guarantee that everything is being completed according to the design – errors in this stage can destroy the experiment's validity;
6. **Statistical data analysis:** Analyze the data using statistical methods, given that results and conclusions are objective and not the outcome of opinions – residuals analysis and verification of model validity are important to this phase;

7.   **Conclusions and recommendations:** Provide practical conclusions based on the results and recommend a plan of action. Accompanying sequences and confirmation tests must be conducted in order to validate the experiment's conclusions.

Stages 1 – 3 are commonly called the pre-experimental design and, for the experiment's success, it is important that these steps are carried out in the most appropriate manner possible [34].

5.1 DOE: Main concepts

There are three basic principles to DOE [18]:

- **Randomization:** Execution of experiments in a random order in order that the phenomenon's unknown effects are distributed among the factors, thus increasing the investigation's validity. According to the author, randomization is the base for the use of statistical methods in experimental design;
- **Replication:** Repetition of the same test multiple times, thus creating a variation in the response variable which is utilized to evaluate experimental error. With the use of each replication, it is possible to obtain an estimate of experimental error, allowing for the determination of whether the differences observed in the data are statistically different, as well as obtaining a more accurate estimate of an experimental factor's effect.
- **Blocking:** Design technique used to increase precision with the comparisons between the factors of interest. It is frequently utilized to reduce or eliminate variability transmitted by factors of sound. It should be utilized when it is not possible to maintain homogeneity of experimental conditions.

Now that the basic principles of DOE have been defined, the following list presents some of the fundamental terms which are used when dealing with DOE techniques:

- **Factor:** According to [37], factors are input parameters and the structural considerations which compose an experiment. Factors are altered during experimental conduction. According to [40], a factor may assume at least two values during an experiment, being quantitative or qualitative;
- **Levels:** The variations possible for each factor [41];
- **Main effect:** According to [36], the main effect for a factor may be defined as the average of the differences in the response variable, when the factor changes from an inferior to superior level;
- **Response variable:** The response variable is the performance measure for the DOE. The response variables describe how the system responds under a certain configuration of input factors [8];
- **Interaction:** There is interaction between the factors when the response difference between the levels of a given factor is not the same as for the rest of the factors.

Aside from these commonly utilized experimental design terms, two further important concepts should be presented: Analysis of variance (ANOVA) and residuals analysis.

According to [35], in order to test if the alteration in one of the levels or interaction is significant, a hypothesis test for the average can be used. In the case of DOE, this test can be conducted using ANOVA. The statistical test ANOVA is utilized to accept or reject hypotheses investigated with DOE. Its objective is to analyze the average variation of results and demonstrate which factors actually produce significant events over the system response variables [42].

However, according to [18], it is not advisable to trust solely in ANOVA since the validity of its suppositions may be unreliable. Problems with these results may be identified using residual analysis.

Residual analysis is an important procedure to guarantee that the models developed by means of experimentation adequately represent the responses of interest. [18] defines residuals as the difference between the predicted value and the observed experimental value; the same author also asserts that residues should be normal, random and non-correlated.

5.2. Full factorial design

Full factorial design with two levels or factorial 2^k is a type of design in which two levels are defined for each factor, an upper and lower level, and combinations of factors are tested [8]. 2^k factorial design is one of the most important types of factorial design, according to [35], and can be particularly useful in the initial phases of experimental work, especially when many factors are being investigated. Full factorial design offers the fewest executions for the k factors to be studied.

In full factorial design, the number of experiments is equal to the number of experimental levels, elevated to the number of factors. In the case of factorials with two levels, the number of experiments (N) in order to evaluate k factors is given by $N = 2^k$. These designs possess a simplified analysis and form the base of many other experimental designs [34].

In using this strategy, the factors are altered simultaneously and not just one at a time, which indicates that, for each run or complete replica, all possible combinations are investigated [35]. For example, if there are *a* levels for factor *A* and *b* levels for factor *B*, then each replica will contain *ab* combinations [18].

One aspect to be considered is that, as there are only two levels per factor, it must be supposed that the response is approximately linear within the range of levels chosen [35]. Another important aspect is that, for experiments with a great number of factors, full factorial design results in an extremely large number of combinations. In this situation, fractional factorial planning is used in order to select a subset of combinations within the full factorial design, aiming to identify the significant factors in system performance [8].

According to [39], many studies in operational research use full factorial design due to its simplicity and because the technique allows the analyst to identify interactions between factors as well as their main effects.

Factorial designs are more efficient than the one-at-a-time approach, as they allow for the factors' effects to be estimated via the other factors' levels, thus leading to valid conclusions about the experimental scope; they are also the only manner to discover interactions among the variables, thus avoiding erroneous conclusions when interactions between the factors are present [18].

5.3. Fractional factorial designs

When there is little interest in interaction behavior among the factors which compose the system, this can be disregarded [35]. Instead, fractional factorial design can be used.

For example, consider a factorial design of 2^5. In this planning, five degrees of freedom correspond to the main effects, 10 degrees of freedom correspond to second order interactions and 16 degrees of freedom correspond to the highest order of interactions. In initial system or project studies, there is little interest in the highest level of interactions [35].

If interactions can be disregarded, fractional factorial design involving fewer executions for a complete set of 2^k executions can be used in order to obtain information about the main effects and lower order interactions [35].

Thus, fractional factorial design provides a means by which to obtain estimates of main effects and, perhaps, second order interactions, with a fraction of the computational force required for full factorial design [4].

According to [18], the greatest application of fractional factorial designs is in screening experiments, where many factors are present in a system and the objective is to identify which factors indeed exercise a significant effect over the given response of interest. For the factors identified as significant through the use of fractional designs, the author recommends a more careful analysis with the use of other designs, such as full factorial design.

In fractional factorial design, a subset of 2^{k-p} is constructed from a set of all of the possible points for a 2^k design, and a simulation is executed for only the chosen points [4].

For this type of factorial design, the analyst must be attentive to its resolution. According to [35], the concept of resolution design is the form in which the fractional factorial designs are related in accordance with the associative standards which they produce. A design's resolution is represented by a Roman subscript number; for example, $2_{III}^{(3-1)}$ represents a factorial design with resolution III, with half of the experiments used in full factorial design [18]. The designs with resolution of III, IV and V are particularly important; they are listed in detail below [35].

- **Design resolution III:** These are the designs in which the main effect is associated with another main effect, but these main effects are associated with second order interactions and second order interactions may be associated.
- **Design resolution IV:** These are the designs in which the main effect is associated with any other main effect or any other second order interaction, but second order interactions are associated with each other.

- **Design resolution V:** These are the designs in which no main effect or second order interaction is associated with any other main effect or second order interaction, but the second order interactions are associated with third order interactions.

6. Sensitivity analysis development stages

As a way of simplifying DOE application in sensitivity analysis of simulation models, the following sequence of steps is proposed.

The flowchart proposed in Figure 3 presents the necessary steps for conducting sensitivity analysis in discrete-event simulation models. Four stages are defined for the proposed method:

1. Simulation;
2. Fractional Factorial Design;
3. Full Factorial Design;
4. Optimization.

In the first step, the analyst should define the optimization objectives as well as verify if the simulation model is verified and validated. In doing so, the model will be apt to proceed to the following step, fractional factorial design. In this phase, the analyst should determine the model's input factors, their levels and select the response variables for analysis.

Once these initial steps have been completed, fractional factorial design can be applied. During execution of the experiments, the analyst should return to the simulation step and carry out the experiments in the simulation package. With the experiments done, the data should be analyzed using statistical means, determining the factors' significance level as well as its lower order interactions. At the end of this phase, the non-significant factors can be removed from analysis.

The third stage defined in the sensitivity analysis phase may be omitted by the analyst, depending on the degree of precision desired or for the cases in which the simulation models demand a lot of computational time in order to be processed. In this stage, a full factorial design is generated with experimental data, and only factors which show to be significant in the previous steps are tested. Depending on the necessity for executing more experiments in order to comply with full factorial design, the analyst will have to return to the simulation phase again to execute new experiments. In this stage, the residues should be analyzed in order to validate the results. Statistical analysis should be conducted once again in order to finalize this stage.

In the following stage, optimization simulation is utilized again. The significant factors after full factorial design application are utilized to configure the optimization tool, which is then executed. Many different configurations are tested for the input parameters until the optimizer converges on a solution. It is up to the analyst to evaluate the results and generate his or her conclusions and recommendations.

In order to demonstrate the utilization of this method in sensitivity analysis of discrete-event simulation models, two simulation models will be used as study objects in this chapter.
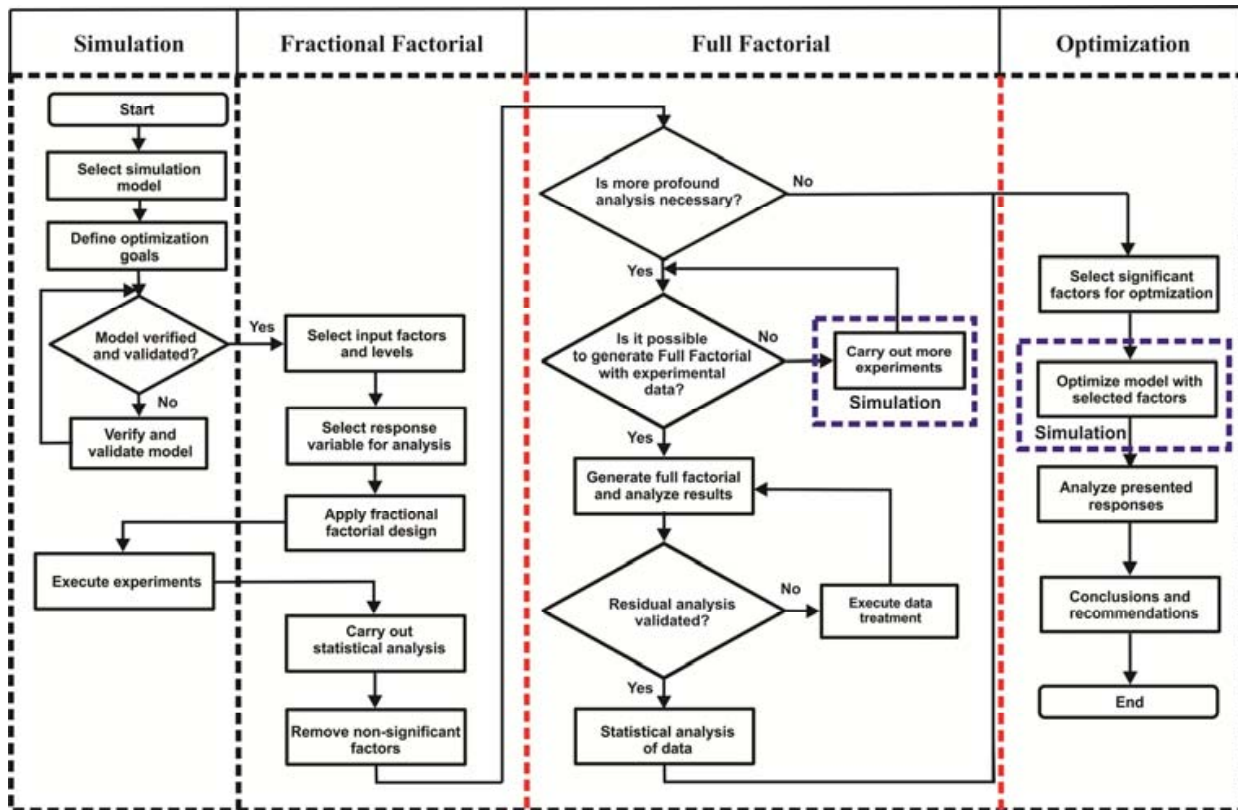
Figure 3. Sequence of steps in order to conduct sensitivity analysis

It should be highlighted here that, in spite of the proposed modeling being applied to great variety of discrete-event simulation models, this approach could be unviable for models which demand a greater amount of computational time to be processed. In these cases, other types of experimental design which involve a smaller number of experiments could be used, such as Plackett-Burman Design; however, according to [18], such designs should be used with great care, as they possess limitations which should be evaluated carefully. An example of this shortcoming is the inability of certain designs to analyze main effects.

7. Modeled systems

The simulation models presented in this chapter were implemented in the software ProModel® and optimized in the package's optimization software, SimRunner®. However, it should be highlighted that the results presented could have been obtained using other commercial simulation packages. Likewise, a commercial statistical software package was utilized to analyze the data.

7.1. Study object 1

The first simulation model represents a quality control station from a telecommunications components factory. The cell is responsible for testing a diverse range of products before shipping them to final clients. This cell receives approximately 75% of the products from the six production lines in the company.

The model in question was verified and validated, thus being ready for study. In order to verify the computational model and correct possible occurrences, some simulator resources such as counters, variables and signals were used, besides the conventional animation. Once the model was validated and verified, simulation was carried out. Statistical tests were utilized which compared the results obtained in the simulation model with data from the real system. The model was considered valid based on statistical tests which did not indicate a statistical difference between real and simulated data.

These conditions are indispensable for conducting sensitivity analysis. The utilization of a non-validated model would lead to erroneous conclusions and undesirable decision-making. For more information about validation and verification, readers are recommended to consult [3]. Figure 4 presents an image of the model implemented in the simulation software.
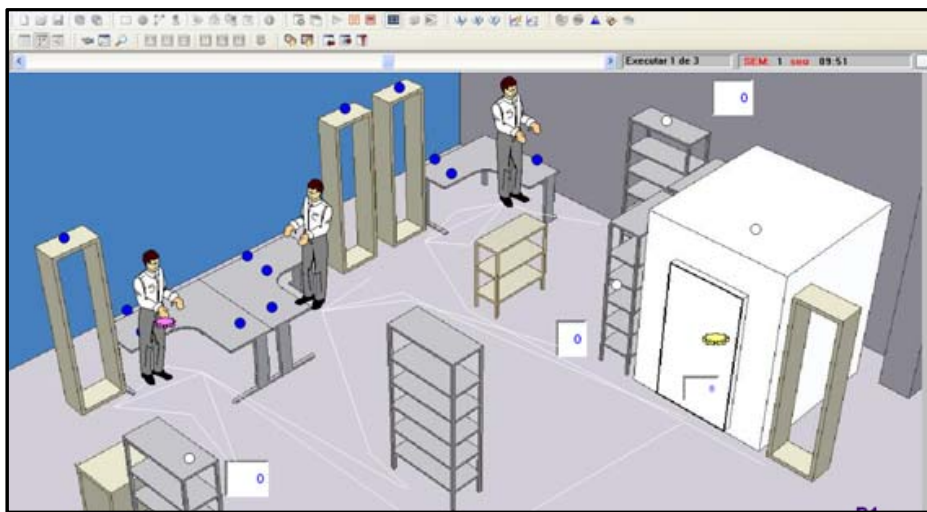
Figure 4. Representation of the real system implemented in the simulation software

The quality control station possesses the following characteristics:

- 7 inspection posts;
- 3 operators;
- 19 types of products to be tested;
- 31 types of operations possible to be carried out depending on the type of product.

For the case in question, discrete variables were defined, with little variation in the lower and higher levels. This fact is justified due to the fact that the majority of simulation optimization problems work with such conditions; however, the experimentation can be conducted using other variable types and a greater variation between the upper and lower limits. Other types of applications can be seen in [43].

For this study object, two levels were defined for each factor. For [18], when the experiment's objective is factor screening or process characterizations, it is common to utilize a small number of variables for each factor. According to the author, lower and higher levels are the most sufficient to obtain valid conclusions about the factors' significance. For example, even if the factor "Type 1 operators" exhibited the possibility of

hiring 1 to 4 operators, for the purposes of the experimental matrix, only two levels would be considered: the lower level being 1 and the upper level being 4.
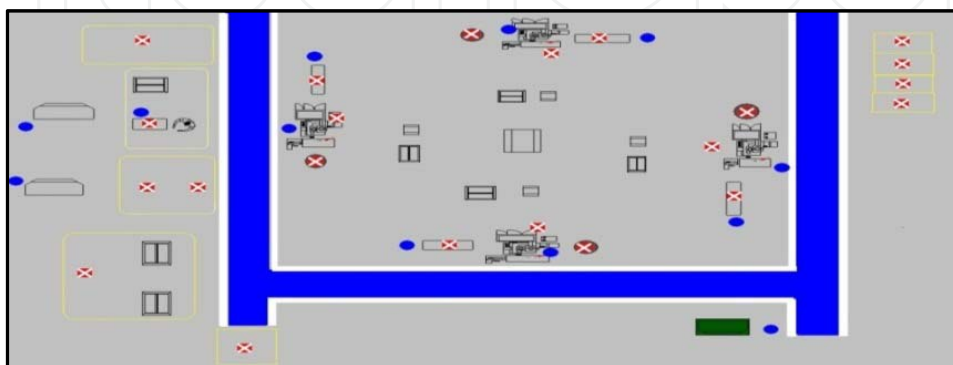
Variable	Factor	Lower level (-)	Upper level (+)
A	Type 1 operators	1	2
B	Type 2 operators	1	2
C	Type 3 operators	1	2
D	Type 1 inspection posts	1	2
E	Type 2 inspection posts	1	2
F	Type 3 inspection posts	1	2
G	Type 4 inspection posts	1	2
H	Type 5 inspection posts	1	2
J	Type 6 inspection posts	1	2
K	Type 7 inspection posts	1	2

Table 2. Experimental factors for the first study object

The optimum set of variables will be determined using three approaches. The first one performs several experiments to identify the main factors. After identifying the statistically significant simulation factors by using a two sample t hypothesis test (a usual procedure from any statistic package), the original fractional factorial design can -be converted to full factorial, eliminating the non-statistically significant terms. As these parameters are also important to the simulation arrangement, despite not being statistically significant, they can be kept constant in proper levels. Comparatively, a second approach can be established by using the main factors identified at the experiments DOE as input for the optimization via SimRunner. Finally, the third approach is performed using all ten factors in optimization via Simrunner.

7.2. Study object 2

The second study object represents an automotive components production cell. The objective in this study object is to find the best combination of input variables which maximizes cell throughput. As with the previous case, the model was verified and validated, being ready for sensitivity analysis and optimization. Figure 5 shows an image of the model implemented in the simulation software.

Figure 5. Representation of the real system implemented in the simulation software

The cell presents the following characteristics:

- 41 machines;
- 3 operators;
- 8 different types of products;
- 46 types of possible processes throughout the system.

Variable	Factor	Lower level (-)	Upper level (+)
A	Type 1 operators	1	2
B	Type 2 operators	1	2
C	Type 3 operators	1	2
D	Type 1 machines	1	2
E	Type 2 machines	1	2
F	Type 3 machines	1	2
G	Type 4 machines	1	2
H	Type 5 machines	1	2
J	Type 1 inspection posts	1	2
K	Type 2 inspection posts	1	2
L	Type 3 inspection posts	1	2
M	Type 4 inspection posts	1	2

Table 3. Experimental factors for the second study object

The optimum set of parameters is determined by three approaches similar to the first application.

8. Experimentation

8.1. Identification of significant factors

According to [4], in simulation experimental designs provide a way to decide which specific configurations to simulate before the runs are performed so that the desired information can be obtained with the fewest simulation runs. For instance, considering the second application where there are 12 factors, if a full factorial experiment were chosen, $2^{12} = 4096$ runs would be necessary. Therefore, a screening experiment must be considered. Screening or characterization experiments are experiments in which many factors are considered and the objective is to identify those factors (if any) that have large effects [18]. Typically, screening experiment involves using fractional factorial designs and is performed in the early stages of the project when many factors are likely considered to have little or no effect on the response [18]. According to this author, in this situation, it is usually best to keep the number of factors levels low.

8.2. Study object 1

For the first study object, ten experimental factors, each with two levels, were defined, as seen in Table 2. When considering full factorial design, a total number of $2^{10} = 1024$

experiments would be necessary. In order to reduce the acceptable number of experiments, fractional factorial design is used.

Table 4 presents four factorial designs for 10 factors and their resolutions. As the objective of this analysis was to identify the model's sensitivity to certain factors, resolution IV was chosen. Resolution IV indicates that no main effect is associated with any other main effect or second order interaction, but there is interaction between certain second order interactions [18].

Fraction	Resolution	Design	Executions
1/8	V	$2^{(10-3)}$	128
1/16	IV	$2^{(10-4)}$	64
1/32	IV	$2^{(10-5)}$	32
1/64	III	$2^{(10-6)}$	16

Table 4. Factorial designs for 10 factors and their resolutions

Experiment	A	B	C	D	E	F	G	H	J	K	WIP
1	-	-	-	-	-	-	+	+	+	+	99
2	+	-	-	-	-	-	+	-	-	-	95
3	-	+	-	-	-	-	-	+	-	-	101
4	+	+	-	-	-	-	-	-	+	+	104
5	-	-	+	-	-	-	-	-	+	-	94
6	+	-	+	-	-	-	-	+	-	+	98
7	-	+	+	-	-	-	+	-	-	+	100
8	+	+	+	-	-	-	+	+	+	-	99
9	-	-	-	+	-	-	-	-	-	+	93
10	+	-	-	+	-	-	-	+	+	-	95
11	-	+	-	+	-	-	+	-	+	-	97
12	+	+	-	+	-	-	+	+	-	+	98
13	-	-	+	+	-	-	+	+	-	-	101
14	+	-	+	+	-	-	+	-	+	+	101
15	-	+	+	+	-	-	-	+	+	+	99
16	+	+	+	+	-	-	-	-	-	-	98
17	-	-	-	-	+	-	+	+	-	-	101
18	+	-	-	-	+	-	+	-	+	+	95
19	-	+	-	-	+	-	-	+	+	+	100
20	+	+	-	-	+	-	-	-	-	-	98
21	-	-	+	-	+	-	-	-	-	+	94
22	+	-	+	-	+	-	-	+	+	-	95
23	-	+	+	-	+	-	+	-	+	-	93
24	+	+	+	-	+	-	+	+	-	+	99

25	-	-	-	+	+	-	-	-	+	-	93
26	+	-	-	+	+	-	-	+	-	+	98
27	-	+	-	+	+	-	+	-	-	+	99
28	+	+	-	+	+	-	+	+	+	-	98
29	-	-	+	+	+	-	+	+	+	+	95
30	+	-	+	+	+	-	+	-	-	-	101
31	-	+	+	+	+	-	-	+	-	-	98
32	+	+	+	+	+	-	-	-	+	+	100
33	-	-	-	-	-	+	-	-	+	+	99
34	+	-	-	-	-	+	-	+	-	-	98
35	-	+	-	-	-	+	+	-	-	-	97
36	+	+	-	-	-	+	+	+	+	+	99
37	-	-	+	-	-	+	+	+	+	-	98
38	+	-	+	-	-	+	+	-	-	+	100
39	-	+	+	-	-	+	-	+	-	+	98
40	+	+	+	-	-	+	-	-	+	-	99
41	-	-	-	+	-	+	+	+	-	+	98
42	+	-	-	+	-	+	+	-	+	-	97
43	-	+	-	+	-	+	-	+	+	-	100
44	+	+	-	+	-	+	-	-	-	+	96
45	-	-	+	+	-	+	-	-	-	-	99
46	+	-	+	+	-	+	-	+	+	+	100
47	-	+	+	+	-	+	+	-	+	+	98
48	+	+	+	+	-	+	+	+	-	-	103
49	-	-	-	-	+	+	-	-	-	-	99
50	+	-	-	-	+	+	-	+	+	+	96
51	-	+	-	-	+	+	+	-	+	+	101
52	+	+	-	-	+	+	+	+	-	-	100
53	-	-	+	-	+	+	+	+	-	+	99
54	+	-	+	-	+	+	+	-	+	-	100
55	-	+	+	-	+	+	-	+	+	-	98
56	+	+	+	-	+	+	-	-	-	+	102
57	-	-	-	+	+	+	+	+	+	-	96
58	+	-	-	+	+	+	+	-	-	+	97
59	-	+	-	+	+	+	-	+	-	+	95
60	+	+	-	+	+	+	-	-	+	-	99
61	-	-	+	+	+	+	-	-	+	+	97
62	+	-	+	+	+	+	-	+	-	-	97
63	-	+	+	+	+	+	+	-	-	-	99
64	+	+	+	+	+	+	+	+	+	+	98

Table 5. The $2_{IV}^{(10-4)}$ design matrix for principal fraction and results

Among the resolution IV designs presented in Table 4, the fractional factorial design $2_{IV}^{(10-4)}$ was chosen. This design, despite possessing a greater number of executions than the design $2_{IV}^{(10-5)}$, allows for the reduction of full factorial designs for six or less factors, without requiring new experiments; that is, the results of this design can be used in full factorial designs with six or less factors, without needing to conduct additional experiments. However, if preliminary studies show that there are a significant number of fewer variables (5 or less), the factorial design $2_{IV}^{(10-5)}$ could be chosen with no problems.

It is worth mentioning that factorials with a resolution less than IV should be omitted because, in these types of design, the main effects are associated with second level interactions, and second level interactions could possess interaction between each other, as well, thus making these designs undesirable.

Table 5 shows the design matrix for the principal fraction and the results obtained for the WIP. Wip represents the total number of pieces in quality control inspection; its result is shown by the variable introduced in the simulation model which subtracts the pieces which leave the system (inspected pieces) from the total number of entities which enter the system (pieces to be inspected). This value is then offered at the conclusion of the simulation in which the report is generated. In the table, the best results attained with the experimentation are shown. In Table 5, the symbols – and + indicate the lower and upper levels shown in Table 2, respectively.

As an example, the number of operator types 1, 2 and 3 and the number of inspection posts types 1, 2 and 3 (A B C D E F) were defined in the simulator as being equal to the lower level (1); for the inspection posts types 4, 5, 6 and 7 (G H J K), the upper level was defined. A replica utilizing this configuration was run in the simulation software, and work in process (WIP) statistics were stored for analysis. This process was repeated 63 other times until all of the experimental matrix's configurations were run.

The $2_{IV}^{(10-4)}$ fractional factorial design used in this research was not replicated. Therefore, it is not possible to assess the significance of the main and interaction effects using the conventional bilateral t-test or ANOVA. The standard analysis procedure for a non-replicated two-level design is a normal plot of the estimated factor's effects. However, these designs are so widely used in practice that many formal analysis procedures have been proposed to overcome the subjectivity of normal probability [18]. [44], for instance, recommend the use of Lenth's method, a graphical approach based on a Pareto Chart for the error term. If the error term has one or more degrees of freedom, the line on the graph is drawn at t, where t is the (1 - α/2) quartile of a t-distribution with a number of degrees of freedom equal to the number of effects/3. The vertical line in the Pareto Chart is the margin of error, defined as ME = t x PSE. Lenth's pseudo standard error (PSE) is based on sparseness of the effects principle, which assumes the variation in the smallest effects is due to the random error. To calculate PSE the following steps are necessary: (a) calculates the absolute value of the effects; (b) calculates S, which is 1.5 x median of the step (c); calculates the median of the effects that are less than 2.5 x S and (d) calculates PSE, which is 1.5 x median calculated in step (c).

With the aid of statistical software, it was possible to perform quantitative analysis of the stored data. Figure 6 presents the Pareto Chart for $2_{IV}^{(10-4)}$ fractional design with significance level of 5%.
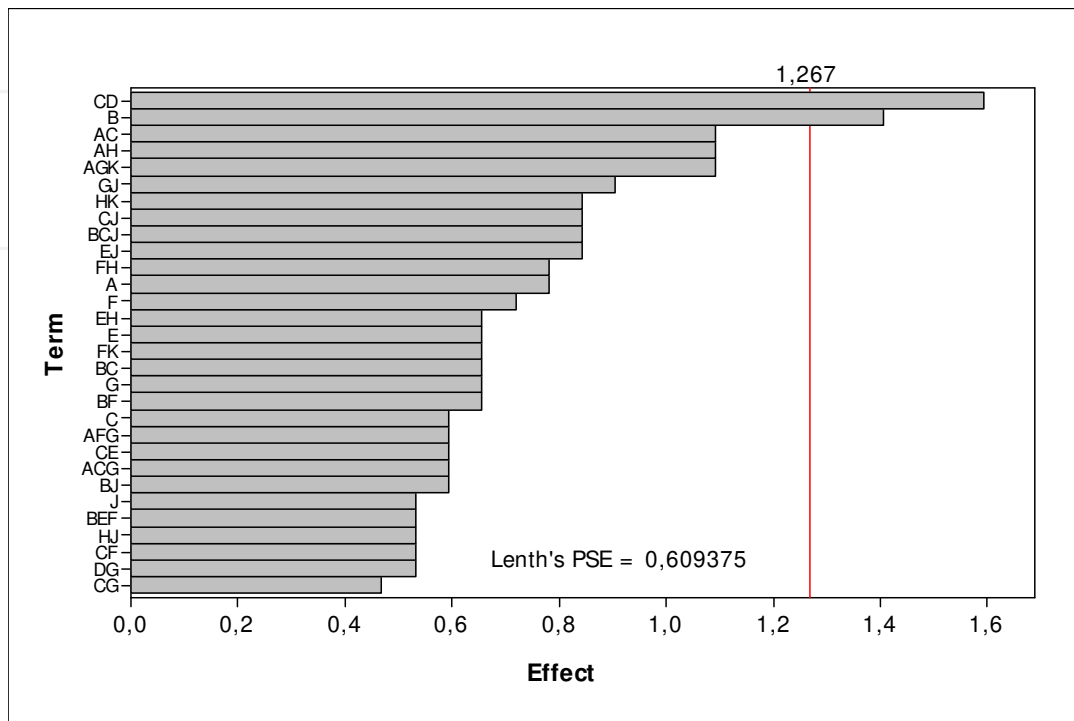
Figure 6. Pareto chart for $2_{IV}^{(10-4)}$ fractional design with a significance level of α = 5%

By analyzing the figure, it can be seen that factor B (number of type 2 operators) and interaction CD (number of type 3 and type 1 inspection posts) are significant. According to [18], if the experimenter can reasonably assume that certain high-order interactions are negligible, information on the main effects and low-order interactions may be obtained. Otherwise, when there are several variables, the system or process is likely to be driven primarily by some of the main effects and low-order interactions. For this reason, it is reasonable to admit that factors A, E, F, G, H, J and K are not significant, although they are still necessary for the simulation model and must be kept at the lower level (-).

Factors C and D may be considered significant, seeing as the interaction between the two factors are quite significant. In Figure 7, it can be seen that B and C exercise a positive effect over the WIP; shifting from the lower to upper level causes an increase in the WIP. Inversely, factor D exercises a negative effect; shifting from the lower to upper level causes WIP to fall. Analysis of interaction behavior in fractional factorial design is not recommended, since the effects possess the property of aliases. That is, according to [18], two or more factors are aliased when it is not possible to distinguish between the effects of overlapping factors. Only three main factors may be considered significant (B, C, D), and full factorial with these factors can be carried out with the data from the 64 experiments.

Factorial design's own structure helps explain why only factors B, C and D were chosen to compose the new factorial design. The main reason B is aliased with the triple interaction

AGH, which can be disregarded according to the chosen resolution for factorial design and by the sparsity of effects principle [18]. In turn, interaction CD is aliased with two triple interactions, AFH and BFG, which can be disregarded, just as in the last case. It can also be aliased with the double interaction in JK; however, as the main factors J and K are not significant, this interaction may also be discarded. The alias structure used in this analysis is available in many statistical packages.

Fractional factorial design $2^{(10-4)}_{IV}$ was converted into a full factorial design 2^3 with replicas. Residual analysis is also possible, since now there are replicas, seeing that the experimentation changed from fractional factorial design $2^{(10-4)}_{IV}$ to full factorial design 2^3 (8 experiments).
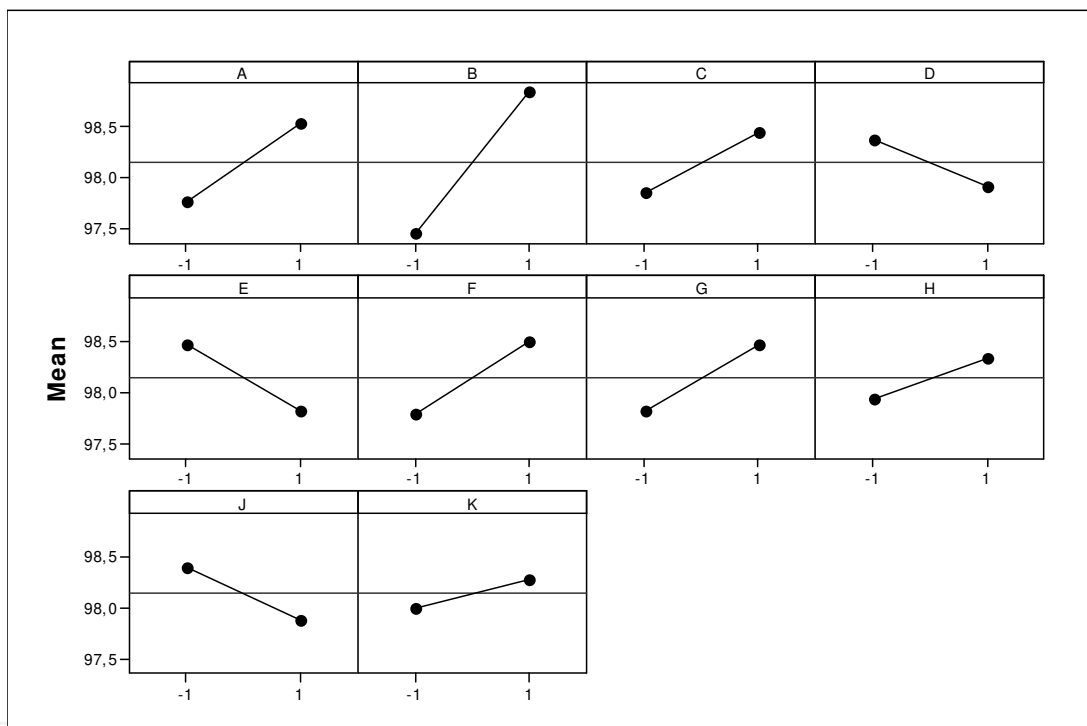
Figure 7. Main effects plot for WIP

According to [18], the residues need to be normal, random and not correlated in order to validate the experimental values obtained. Figure 8 shows the verification of the residues normality.

Evaluating the normality probability graph, one can see that the data are adjusted to a normal distribution, as is evidence by the way the points fall over the line in the graph as well as analysis of the P-value. One can see the data points follow the straight line, and the P-value for the normality test was less than 0.05, leading to the conclusion that the data are normally distributed. Figure 9 shows the verification of the residues independence. The standardized graphs versus the observed values do not present any random patterns of grouping or bias.
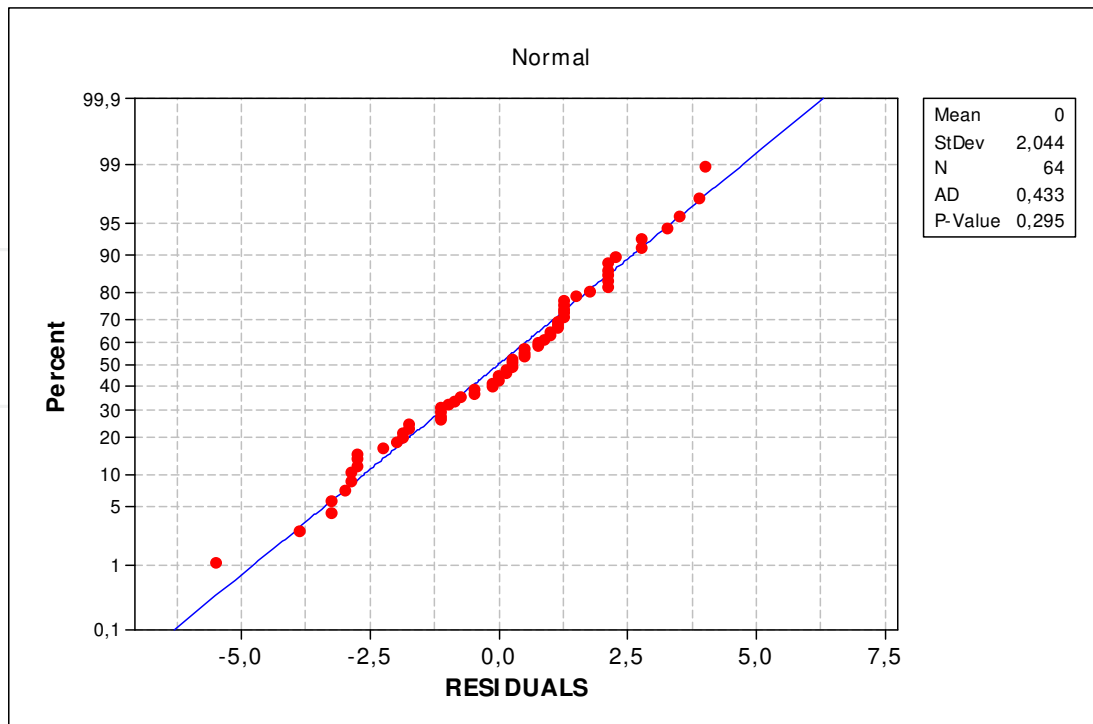
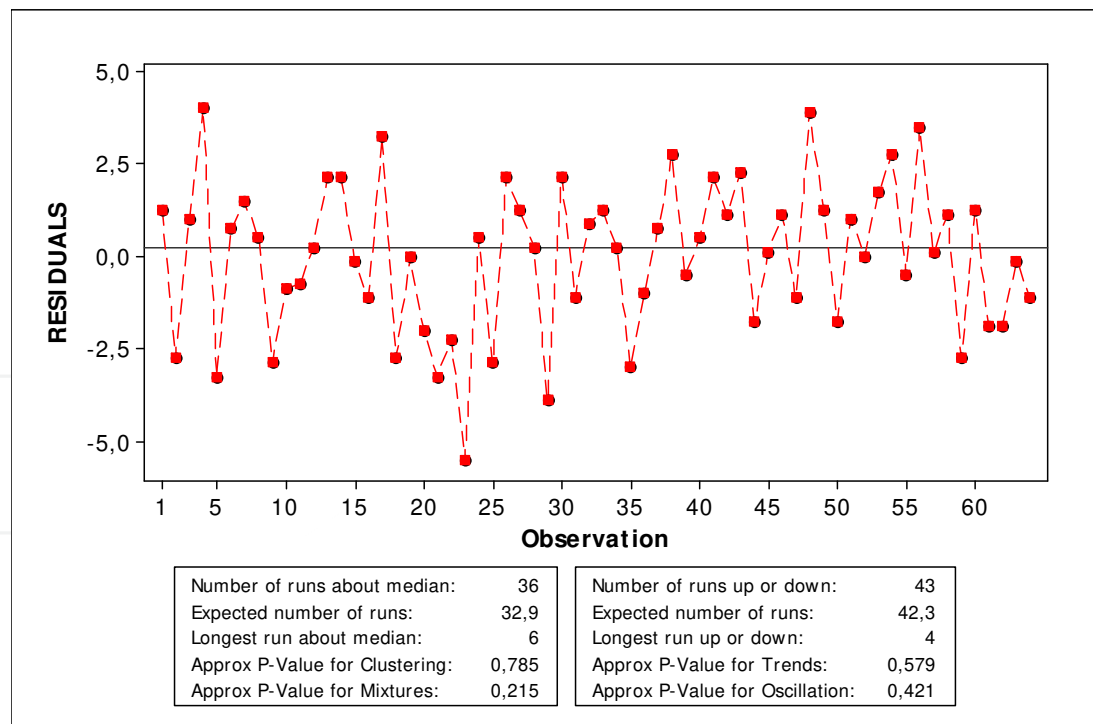**Figure 8.** Verification of the residues normality

Figure 9. Verification of the residues independence

Once the residual validity was verified, the results could be analyzed using DOE. The analyses continued to be carried out via graphical analysis due to its ease of comprehension.

Figures 10 and 11 present the analysis for the new design. By analyzing figure 10, it can be verified that factor B (number of type 2 operators) and the interaction CD (number of type 3 operators and number of type 1 inspection posts) remained significant. In this new design, no other main factor or interaction demonstrated a significance level greater than 5%.
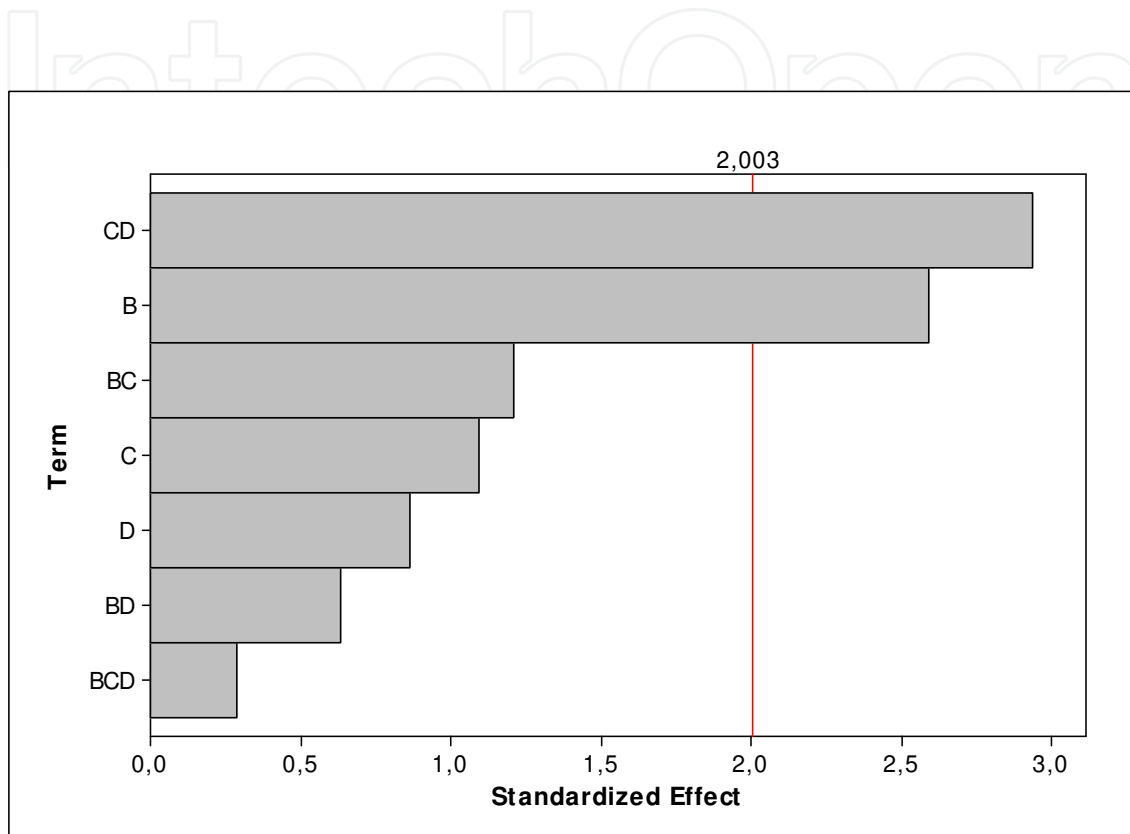
Figure 10. Pareto Chart for full factorial design with significance level $\alpha = 5\%$

Analysis of Figure 11 shows that factors B and C exercise a positive effect over the WIP; that is, they should be kept at the lower level in order to minimize the WIP count. Factor D should be kept at the upper level, since it exercises a negative effect on the WIP count. By observing Figure 11, it can be seen that the CD interaction has a strong effect on diminishing the WIP when the main effects C and D remain at their own respective lower and upper levels.

There are strong indications about an improved configuration for the input variables in order to minimize the WIP count. However, these suppositions will be tested using commercial optimization software. First, 10 input variables will be utilized; afterwards, only the three variables which showed to be significant according to the study in this section will be evaluated, and the seven other factors will be fixed in the lower level, seeing as they are not significant.
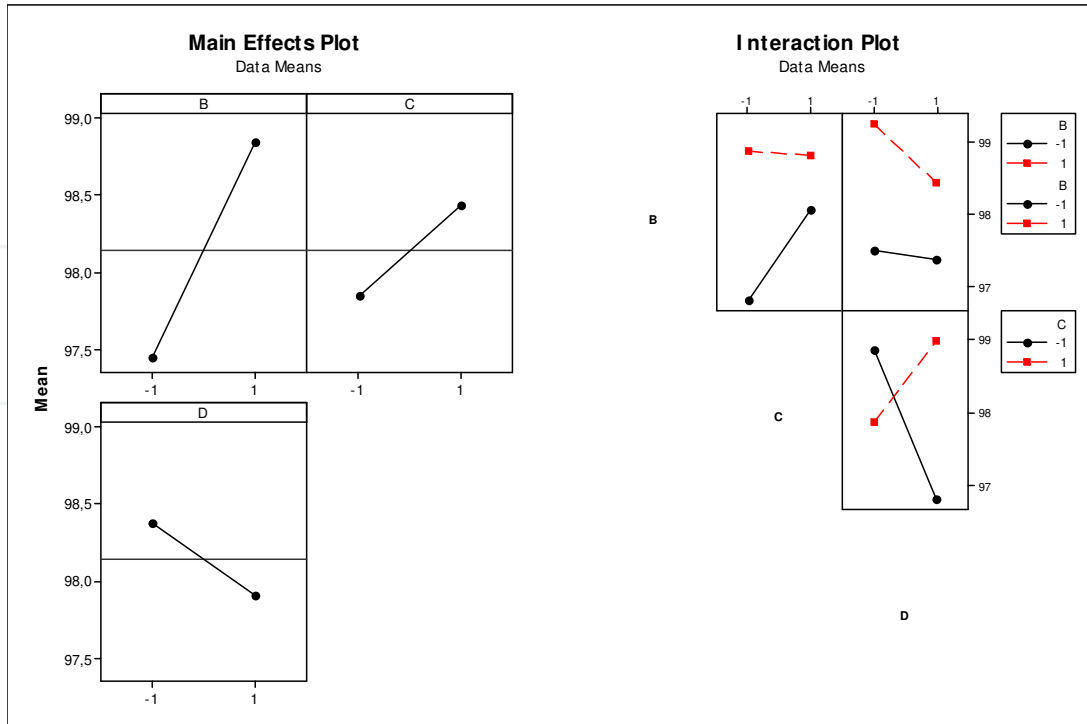
Figure 11. Factorial and interaction plots for 2^3 full factorial design

8.3. Study object 2

For the second study object, 12 experimental factors were defined, as presented in Table 3. Each factor possesses two levels. Unlike the previous case, for this study the objective is to maximize the manufacturing cell's throughput. Thus, the significance of the 12 factors will be analyzed. Considering full factorial design, $2^{12} = 4096$ experiments would be necessary. Similar to the previous case, fractional factorial design was used to reduce the number of experiments.

Table 6 presents four factorial designs for 12 factors and their resolutions. As the analysis objective in this case is to identify the model's performance sensitivity to the factors, a resolution IV design was chosen.

Fraction	Resolution	Design	Executions
1/256	III	$2^{(12-8)}$	16
1/128	IV	$2^{(12-7)}$	32
1/64	IV	$2^{(12-6)}$	64
1/32	IV	$2^{(12-5)}$	128

Table 6. Factorial designs for 10 factors and their resolutions

Out of the resolution IV plans presented in Table 6, fractional factorial design $2_{IV}^{(12-6)}$ was chosen. Researchers opted for this design due to its location between the fractional factorial designs $2_{IV}^{(12-7)}$ and $2_{IV}^{(12-5)}$, thus enabling a reduction to full factorial design for six or less

factors, without having to carry out new experiments. It should be noted here that if more than six factors are shown to be significant, another factorial design could be done while taking advantage of the data already acquired from the fractional factorial design $2_{IV}^{(12-6)}$ and realizing only the non-tested experiments.

Table 7 presents the experimental design matrix for the principal fraction and throughput. Throughput represents the number of pieces produced by the manufacturing cell. For this case, researchers needed to create a variable to store the number of pieces produced and present this value at the end of the simulation. The greatest value produced is highlighted.

Experiment	A	B	C	D	E	F	G	H	J	K	L	M	Throughput
1	-	-	-	-	-	-	-	-	+	+	+	+	369200
2	+	-	-	-	-	-	-	+	+	+	-	-	408200
3	-	+	-	-	-	-	-	+	-	-	-	+	392600
4	+	+	-	-	-	-	-	-	-	-	+	-	390000
5	-	-	+	-	-	-	-	+	-	-	+	-	392600
6	+	-	+	-	-	-	-	-	-	-	-	+	390000
7	-	+	+	-	-	-	-	-	+	+	-	-	400400
8	+	+	+	-	-	-	-	+	+	+	+	+	416000
9	-	-	-	+	-	-	+	-	-	-	+	+	403000
10	+	-	-	+	-	-	+	+	-	-	-	-	413400
11	-	+	-	+	-	-	+	+	+	+	-	+	429000
12	+	+	-	+	-	-	+	-	+	+	+	-	429000
13	-	-	+	+	-	-	+	+	+	+	+	-	413400
14	+	-	+	+	-	-	+	-	+	+	-	+	408200
15	-	+	+	+	-	-	+	-	-	-	-	-	421200
16	+	+	+	+	-	-	+	+	-	-	+	+	429000
17	-	-	-	-	+	-	+	-	-	+	-	-	390000
18	+	-	-	-	+	-	+	+	-	+	+	+	429000
19	-	+	-	-	+	-	+	+	+	-	+	-	426400
20	+	+	-	-	+	-	+	-	+	-	-	+	431600
21	-	-	+	-	+	-	+	+	+	-	-	+	416000
22	+	-	+	-	+	-	+	-	+	-	+	-	410800
23	-	+	+	-	+	-	+	-	-	+	+	+	410800
24	+	+	+	-	+	-	+	+	-	+	-	-	434200
25	-	-	-	+	+	-	-	-	+	-	-	-	392600
26	+	-	-	+	+	-	-	+	+	-	+	+	408200
27	-	+	-	+	+	-	-	+	-	+	+	-	400400
28	+	+	-	+	+	-	-	-	-	+	-	+	395200
29	-	-	+	+	+	-	-	+	-	+	-	+	405600
30	+	-	+	+	+	-	-	-	-	+	+	-	397800
31	-	+	+	+	+	-	-	-	+	-	+	+	395200
32	+	+	+	+	+	-	-	+	+	-	-	-	429000

33	-	-	-	-	-	+	+	-	+	-	-	-	400400
34	+	-	-	-	-	+	+	+	+	-	+	+	423800
35	-	+	-	-	-	+	+	+	-	+	+	-	421200
36	+	+	-	-	-	+	+	-	-	+	-	+	421200
37	-	-	+	-	-	+	+	+	-	+	-	+	400400
38	+	-	+	-	-	+	+	-	-	+	+	-	397800
39	-	+	+	-	-	+	+	-	+	-	+	+	405600
40	+	+	+	-	-	+	+	+	+	-	-	-	444600
41	-	-	-	+	-	+	-	-	-	+	-	-	371800
42	+	-	-	+	-	+	-	+	-	+	+	+	410800
43	-	+	-	+	-	+	-	+	+	-	+	-	403000
44	+	+	-	+	-	+	-	-	+	-	-	+	382200
45	-	-	+	+	-	+	-	+	+	-	-	+	384800
46	+	-	+	+	-	+	-	-	+	-	+	-	387400
47	-	+	+	+	-	+	-	-	-	+	+	+	395200
48	+	+	+	+	-	+	-	+	-	+	-	-	421200
49	-	-	-	-	+	+	-	-	-	-	+	+	397800
50	+	-	-	-	+	+	-	+	-	-	-	-	423800
51	-	+	-	-	+	+	-	+	+	+	-	+	382200
52	+	+	-	-	+	+	-	-	+	+	+	-	413400
53	-	-	+	-	+	+	-	+	+	+	+	-	395200
54	+	-	+	-	+	+	-	-	+	+	-	+	405600
55	-	+	+	-	+	+	-	-	-	-	-	-	392600
56	+	+	+	-	+	+	-	+	-	-	+	+	416000
57	-	-	-	+	+	+	+	-	+	+	+	+	395200
58	+	-	-	+	+	+	+	+	+	+	-	-	429000
59	-	+	-	+	+	+	+	+	-	-	-	+	429000
60	+	+	-	+	+	+	+	-	-	-	+	-	434200
61	-	-	+	+	+	+	+	+	-	-	+	-	410800
62	+	-	+	+	+	+	+	-	-	-	-	+	423800
63	-	+	+	+	+	+	+	-	+	+	-	-	418600
64	+	+	+	+	+	+	+	+	+	+	+	+	449800

Table 7. The $2_{IV}^{(12-6)}$ design matrix for principal fraction and results

Similar to the last case, fractional factorial design $2_{IV}^{(12-6)}$ was utilized without replicas. With the help of statistical software, the data were analyzed. Figure 12 shows the Pareto Chart for $2_{IV}^{(12-7)}$, fractional design with a significance level of 5%. By analyzing the first figure, it can be seen that factors G (number of type 4 machines), A (number of type 1 operators), H (number of type 5 machines), B (number of type 2 operators), E (number of type 2 machines) and the interaction BG (number of type two operators and number of type 4 machines) are significant according to the adopted significance level. It can be said that the factors C, D, F, J, K, L and M are not significant, although they are necessary for the simulation and their values may be fixed at the lower level

(assuming a value of 1), since they do not exercise a significant effect over the model's throughput.

By analyzing Figure 13, it can be seen that the main factors A, B, E, G, and G exercise a positive effect over throughput; that is, by altering the lower level to the upper level, there is an increase in throughput.

Interaction behavior analysis in fractional factorial design is not recommended, seeing that aliasing between effects tends to emerge. Thus, full factorial design with the significant factors will be carried out using the data from the 64 experiments carried out. As was the previous case, the alias structure for factorial design $2_{IV}^{(12-6)}$ enables the explanation for why A, B, E, G and H were chosen to make up the full factorial design.

The main factor A is aliased with the two triple interactions BCH and HLM. Factor B is aliased with two other triple interactions, ACH and CLM. Factor E is aliased with the interactions DFG and FJK. Factor G is aliased with DEF and DKJ. Factor H is aliased with the interactions ABC and ALM. Finally, the interaction BG is aliased with four triple interactions, ADL, CEK, CFJ and DHM. All these interactions can be disregarded according to the chosen level of resolution for factorial design and the principle of sparsity of effects principle [18].

It can be concluded that, although the simulation model possesses 12 input variables which may be arranged in order to maximize total throughput, only 5 variables significantly contribute to increased throughput. In following, an optimization of this simulation model will be executed, using a commercial software package, first optimizing all 12 input variables, and then with only the 5 input variables which are statistically significant.
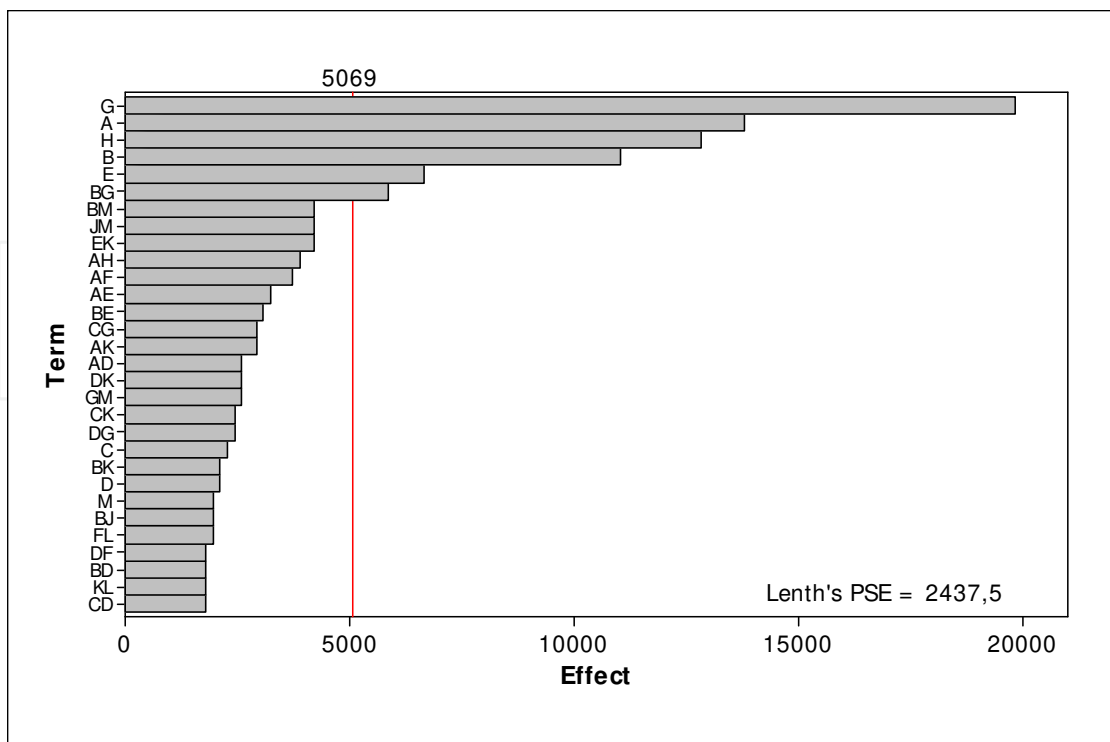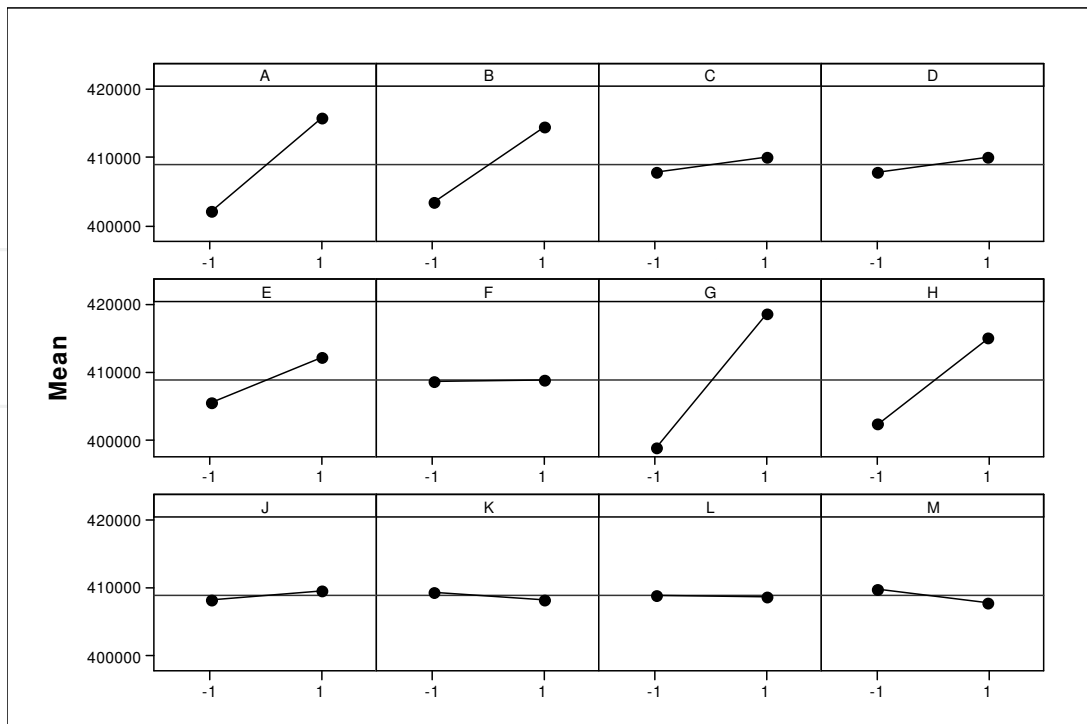

Figure 12. Pareto Chart for $2_{IV}^{(12-7)}$ fractional design with significance level of 5%

Figure 13. Main effects plot for throughput

Fractional factorial design $2_{IV}^{(12-6)}$ was converted into a full factorial design 2^5 with replicas. Before analyzing the new design's results, the validity of the results was tested, as was the case with the previous experiment. Once the validity of the new design's residues' was verified, it was then possible to statistically analyze the results with DOE.

With the new design, the main factors A, B, E, G and H, the interactions BG, AGH, ABEG and AH showed to be significant, according to the 5% significance level (Figure 14). All of the main factors presented positive effects on the throughput, according to Figure 15; that is, shifting from the lower (–) and upper (+) level, the production cell's throughput increases.

It can then be concluded that, although the simulation model possesses 12 input variables which can be arranged in order to maximize throughput, only five variables significantly contribute to increased production. In following, a comparison will be performed between the commercial optimization software; first, all 12 input variables will be optimized, and then only the five variables which are statistically significant will be optimized.

It is worth mentioning here another optimization approach which is commonly utilized in simulation optimization. By using full factorial design with replicas, it is possible to generate a metamodel for the response variable under analysis. With a mathematical model on hand, traditional optimization tools such as Microsoft Excel's Solver may be utilized in place of simulation optimization tools. An example of such a technique can be seen in [43].

Another approach which is commonly employed in the literature is Response Surface Methodology. As was cited in the previous strategy, a mathematical model, generally non-linear of second-order, is generated through the experimental data and then is optimized. The shortcoming of this strategy is that the model must possess a robust fit, which allows

for a satisfactory representation of the response factor. If the model is not robust, experimental strategies should be employed in order to redefine the experimental region, which many times is not applicable for simulation optimization problems.
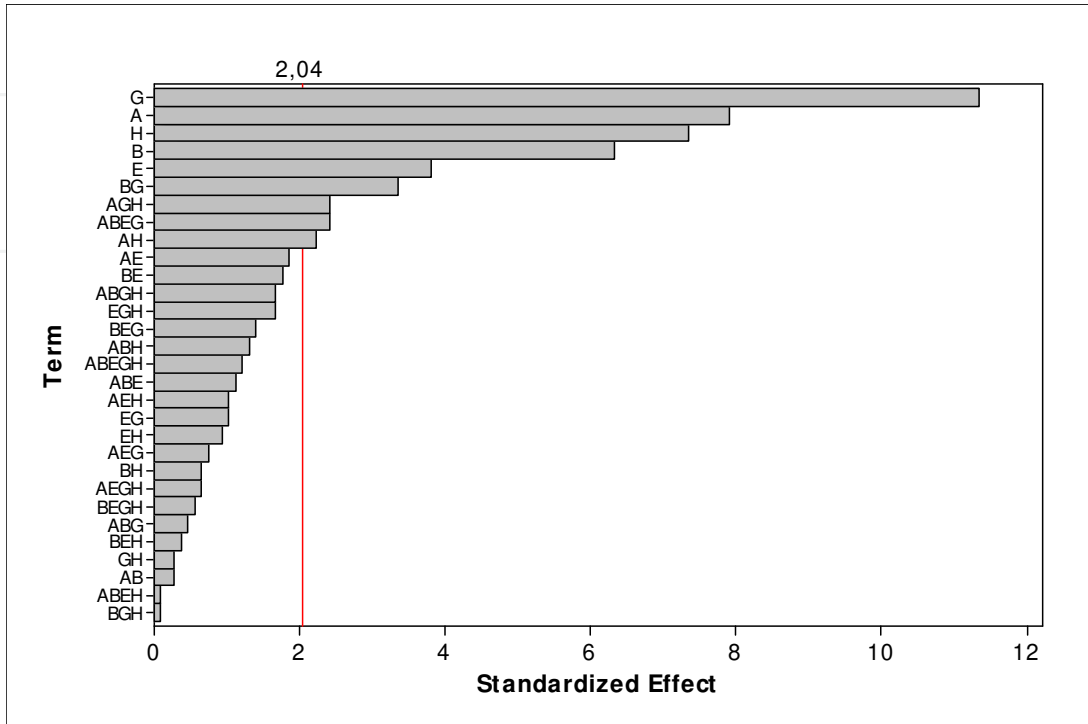

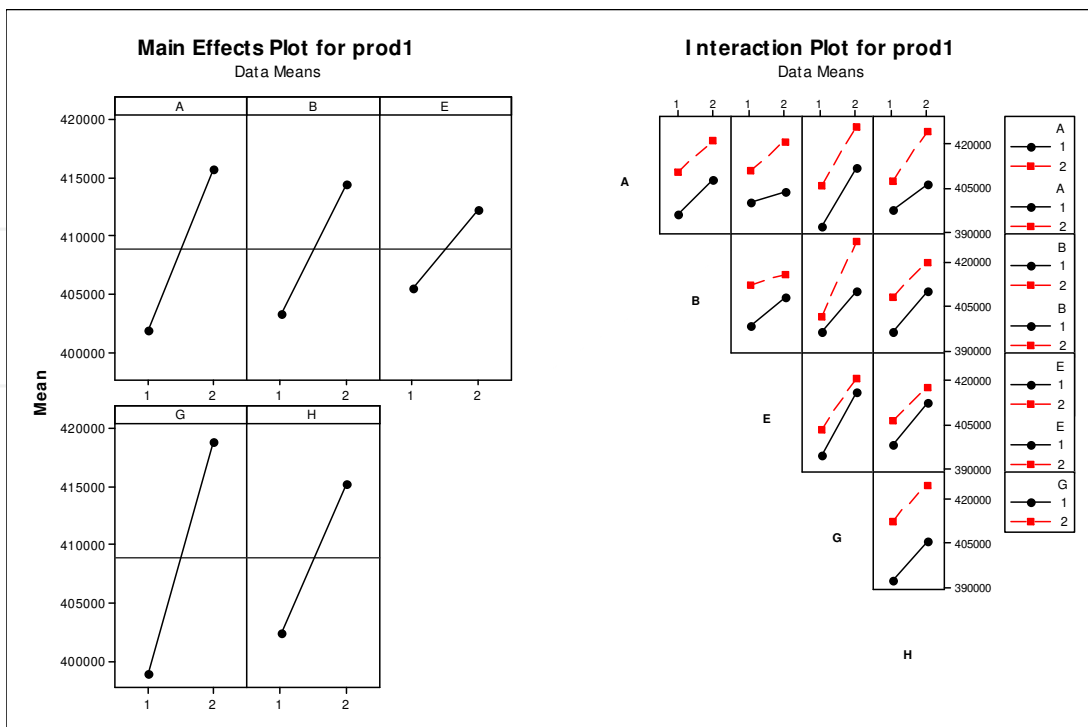**Figure 14.** Pareto Chart for 2^5 full factorial design with significance level $\alpha = 5\%$

Figure 15. Factorial and interaction plots for 2^5 full factorial design

9. Simulation model optimization

Through the sensitivity analysis, each simulation model's significant variables may be identified. Considering only these results, the best combination of input variables can be inferred in order to optimize the simulation models; however, there is no way of guaranteeing this affirmation based in only a sensitivity analysis.

One way of confirming these results is through optimization. An example application of simulation optimization will be employed as a means of evaluating the efficiency of fractional factorial design in the execution of sensitivity analysis.

The adopted procedure is to optimize the study objects in two different ways. In the first case, all input variables will be optimized and, in the second case, only the factors selected in the sensitivity analysis will be optimized. Finally, the results attained will be compared in order to verify if the design techniques were advantageous to the process. The time involved in the process will not be the basis for comparison; thus it is evident that the number of experiments necessary will be reduced for the model to arrive at a solution.

The simulation software package SimRunner® from the ProModel Corporation will be utilized for the execution of experiments; however, there are other simulation optimization software packages that could have been chosen for this investigation. SimRunner® integrates resources to analyze and optimize simulation models through multivariable optimization. This type of optimization tests multiple factor combinations in search of the system input variable configuration which leads to the best objective function value [20].

SimRunner® is based in a genetic algorithm and possesses three optimization profiles: Aggressive, Moderate and Cautious. These profiles are directly related to the confidence in the solution and the time necessary to find this solution. The cautious profile was chosen for this study in order to consider the greatest possible number of solutions and in turn, guarantee a more comprehensive search and present better responses [45].

9.1. Optimization of the first study object

The optimization objective for the first simulation model was to find the best input variable combination in order to minimize the system's work in process count. As presented in Table 2, this model possesses 10 input variables, being varied at the lower level (1) and the upper level (2).

In the first optimization stage, 10 input model variables were selected and the optimizer was configured. The results found can be seen in Figure 16.

The optimizer converged with 296 experiments. The best result obtained was 92, which was attained during experiment 261, as seen in Figure 16. The values found for the factors are shown in Table 8.

The sensitivity analysis for the first study object identified three factors with significant effects which can be utilized for simulation (Table 9). The other variables were maintained at their original values, defined as the lower level (*).
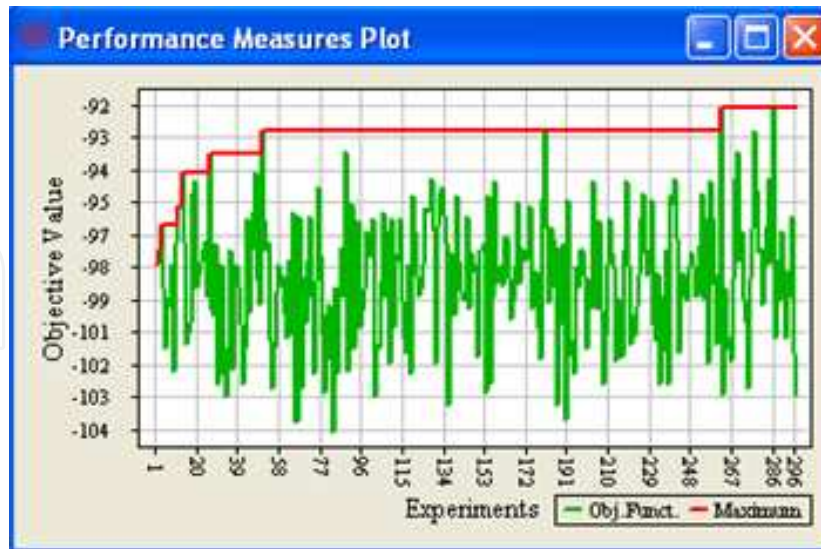
Figure 16. Performance measures plot for optimization using all factors

Factor	Variable	Value
A	Type 1 operators	2
B	Type 2 operators	2
C	Type 3 operators	1
D	Type 1 inspection posts	2
E	Type 2 inspection posts	1
F	Type 3 inspection posts	1
G	Type 4 inspection posts	1
H	Type 5 inspection posts	2
J	Type 6 inspection posts	1
K	Type 7 inspection posts	1

Table 8. The best solution for optimization using all factors

Factor	Variables	Value Range
B	Type 2 operators	1 - 2
C	Type 3 operators	1 - 2
D	Type 1 inspection posts	1 - 2

Table 9. Significant factors for the first study object

The results found can be seen in Figure 17.

Simrunner® converged after 8 experiments. The best result was 93, which was obtained in the seventh experiment, as shown in Figure 17. The values are shown in Table 10.
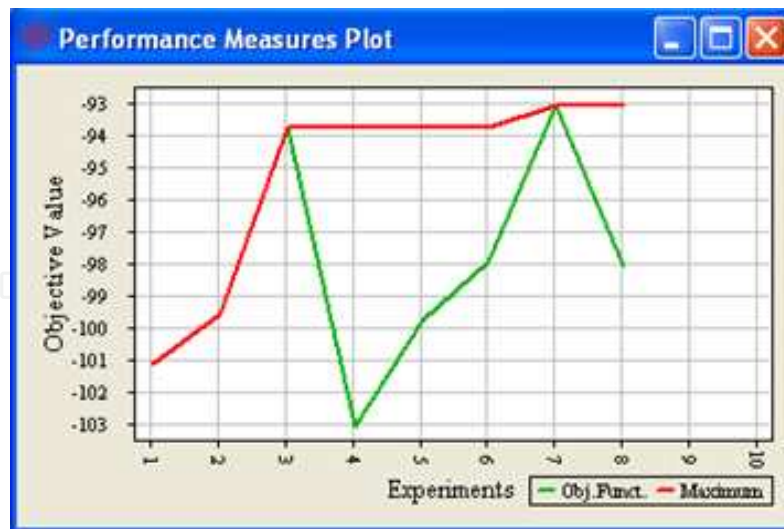
Figure 17. Performance measures plot for optimization using significant factors

Factor	Variable	Value
A	Type 1 operators	1*
B	Type 2 operators	1
C	Type 3 operators	1
D	Type 1 inspection posts	2
E	Type 2 inspection posts	1*
F	Type 3 inspection posts	1*
G	Type 4 inspection posts	1*
H	Type 5 inspection posts	1*
J	Type 6 inspection posts	1*
K	Type 7 inspection posts	1*

Note: The parameters identified with (*) were not used as input for optimization. They were kept at lower level (1).

Table 10. The best solution for optimization using significant factors

9.2. Optimization of the second study object

The optimization objective for the second simulation model was to find the best combination of model input variables in order to maximize the manufacturing cell's throughput. As presented in Table 3, the model possesses 12 input variables which are varied from the lower level (1) to the upper level (2).

In the first optimization phase, the 12 model input variables were selected and the optimization software was set up for experimentation. The results found can be seen in Figure 18.

The optimizer converged with 173 experiments. The best result found was 452,400, which was obtained in experiment 10 (Figure 18). The obtained values are shown in Table 11.
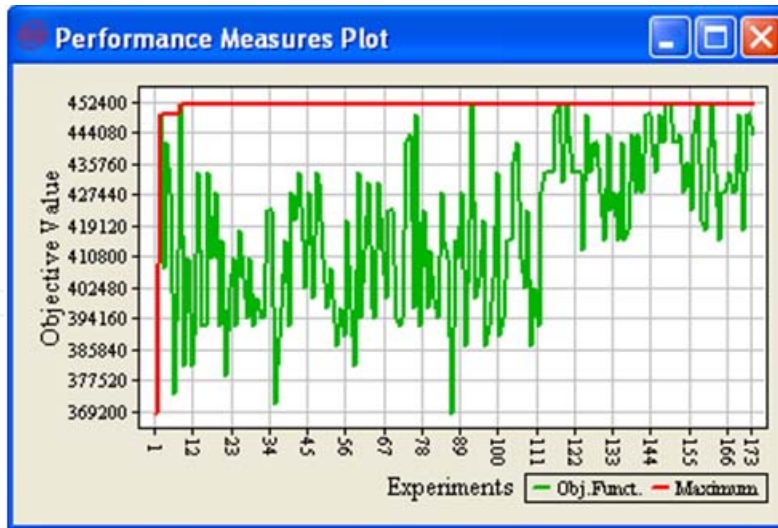
Figure 18. Performance measures plot for optimization using all factors

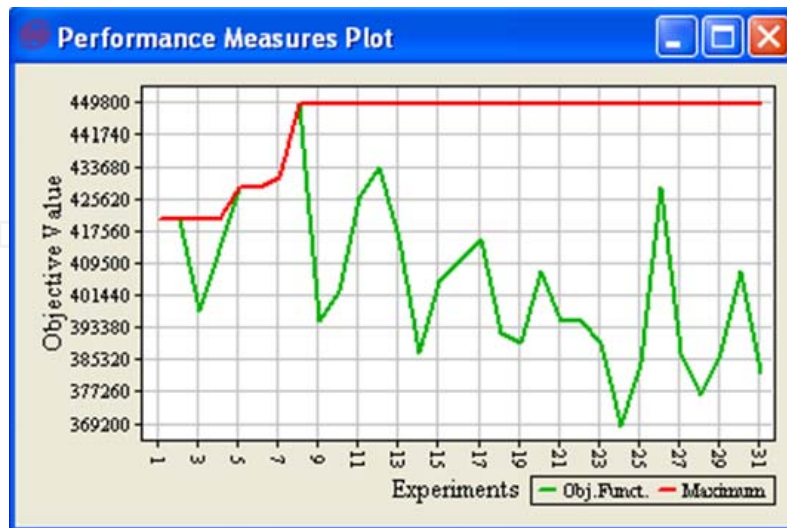| Factor	Variable	Value
A	Type 1 operators	2
B	Type 2 operators	2
C	Type 3 operators	1
D	Type 1 machines	2
E	Type 2 machines	2
F	Type 3 machines	2
G	Type 4 machines	2
H	Type 5 machines	1
J	Type 1 inspection posts	2
K	Type 2 inspection posts	2
L	Type 3 inspection posts	2
M	Type 4 inspection posts	2

Table 11. Best solution for optimization using all factors

The sensitivity analysis for the second study object identified five factors with significant effects which will be used for simulation optimization inputs (Table 12). The other model input variables were kept at their lower level (*).

Factor	Variables	Value range
A	Number of type 1 operators	1 - 2
B	Number of type 2 operators	1 - 2
E	Number of type 2 machines	1 - 2
G	Number of type 4 machines	1 - 2
H	Number of type 5 machines	1 - 2

Table 12. Significant factors for the second study object

The results are shown in Figure 19.

Figure 19. Performance measures plot for optimization using significant factors

Simrunner® converged after 31 experiments. The best value found was 449,800, which was attained in the eighth experiment carried out using the optimizer. The factors' values can be seen in Table 13.

Factor	Variable	Value
A	Type 1 operators	2
B	Type 2 operators	2
C	Type 3 operators	1*
D	Type 1 machines	1*
E	Type 2 machines	2
F	Type 3 machines	1*
G	Type 4 machines	2
H	Type 5 machines	2
J	Type 1 inspection posts	1*
K	Type 2 inspection posts	1*
L	Type 3 inspection posts	1*
M	Type 4 inspection posts	1*

Note: The parameters identified with (*) were not used as input for optimization. They were kept at lower level (1).

Table 13. Best solution for optimization using significant factors

10. Results analysis

10.1 First study object

Table 14 presents a comparison of the results attained utilizing the three methods for the first study object. As far as the number of experiments executed, the advantage of using

sensitivity analysis to identify the significant factors becomes obvious. The commercial optimizer carried out 296 experiments when all of the input variables were chosen; when only the significant factors were utilized, merely 8 experiments were executed. Summing up the 64 experiments utilized with fractional factorial design, the result (72) is still four times smaller than the number of experiments executed with the optimizer when all 10 input variables were utilized.

Parameter	Optimization using all factors	Optimization using significant factors	Design of experiments
A	2	1*	1
B	2	1	1
C	1	1	1
D	2	2	2
E	1	2	2
F	1	1*	1
G	1	1*	1
H	2	1*	1
J	1	1*	2
K	1	1*	1
Result (WIP)	92	93	93
Confidence Interval (95%)	(83 – 100)	(86 – 99)	-
Number of runs	296	8	64

Note: The parameters identified with (*) were not used as input for optimization. They were kept at the lower level (1).

Table 14. Optimization results for the three procedures of the first study object

In respect to the responses found, it should be highlighted that, due to the simulation model's stochastic character, the response presented by the optimizer should be analyzed with care while considering both the average value and confidence interval of each result found.

Analyzing only the average optimization result found, it can be noted that the result found by the optimizer, when all 10 decision variables were manipulated was greater, reaching a WIP result of 92. However, when the response's confidence interval is analyzed, it can be said that the optimization's responses, when considering only the significant factors and their respective confidence intervals, are equal. The advantage of the response found using the sensitivity analysis is that only two factors (D and E) had to remain at the upper level, while the rest were kept at the lower level in order to minimize WIP.

The results in Table 5 were found with only the use of fractional factorial design and were selected using the best results during the experimentation process. This approach, however, does not take into consideration any simulation optimization approach and should be viewed with caution. The result using only DOE shows the possibility of using experimental

design with optimization. This possibility was not explored in detail here, as it was not this chapter's objective; however, many authors [1, 4, 18, 19] present optimization techniques using only experimental design.

10.2. Second study object

Table 15 shows a comparison of the results obtained using the three methods for the second study object.

Parameter	Optimization using all factors	Optimization using significative factors	Design of experiments
A	2	2	2
B	2	2	2
C	1	1*	2
D	2	1*	2
E	2	2	2
F	2	1*	2
G	2	2	2
H	1	2	2
J	2	1*	2
K	2	1*	2
L	2	1*	2
M	2	1*	2
Result (WIP)	452400	449800	449800
Confidence Interval (95%)	(445182 – 459617)	(440960 – 458639)	-
Number of runs	173	31	64

Note: The parameters identified with (*) were not used as input for optimization. They were kept at lower level (1).

Table 15. Optimization results for the three procedures analyzing the second study object

In relation to the number of experiments executed, once again, the sensitivity analysis showed itself to be efficient. Along with the reduction in the number of factors, the number of experiments fell from 173 to 31. With the addition of 64 experiments using fractional factorial design, the method was efficient with 95 experiments, a little more than half of the experiments when all factors were considered.

In relation to the optimization result, once again, it was necessary to perform a more detailed analysis of the responses. In spite of the average difference between the solutions presented with the optimizer (using all 12 and then only 5 significant factors) being around 2,600 pieces, it was verified that the solutions were within the same confidence interval. Thus the post-sensitivity analysis optimization solution's quality again showed itself to be efficient when comparing the optimization of all of the input variables.

As was true with the last case, the results in Table 7 were found with only the use of fractional factorial design and were selected using the best results during the experimentation process.

11. Conclusion

The objective of this chapter was to present how experimental design and analysis techniques can be employed in order to identify significant variables in discrete-event simulation models, thus aiding simulation optimization searches for optimal solutions.

To develop this application, the main concepts of simulation optimization were presented during this chapter. In following, two applications were developed to verify how fractional factorial design can be used in sensitivity analysis for simulation models, identifying its advantages, disadvantages and effects on model optimization.

For optimization, the identification of the significant variables was extremely important, as it enabled the reduction of the search space and computational potential necessary to perform the search for an optimal solution. Optimization was performed for the two applications in two distinct forms, utilizing simulation optimization.

The first simulation optimization approach was to use all of the models' input variables. No previous studies were performed to determine if the models' variables exercised significant effects on overall system performance. The second approach relied on sensitivity analysis in order to identify the variables which influenced system performance. After identifying the significant variables, model optimization was utilized while using the reduced search space. The third approach involved using only the experimentation's results without using any simulation optimization procedure.

By analyzing the results, the advantages of using sensitivity analysis become evident, not only due to the reduction in necessary computational potential for the optimization process, but also for the greater level of detail and knowledge acquired about the process under study. By using this experimentation, it is possible to verify the process's variables which exercise the greatest effects on overall system performance, thus being able to determine the effect each variable has on the process as well as their interactions. These interactions would be very difficult to define and easy to disregard in simulation projects without the use of DOE.

It should not go without noting, however, that one must take extreme caution during execution of these experiments, as just one experiment realized under incorrect conditions or out of the correct matrix order can lead to erroneous results. In order to analyze the results obtained during experimentation, the user should have a solid understanding of DOE and statistics. Those erroneous conclusions in simulation could possibly lead to incorrect implementation, which could generate very tangible costs in the real world. Thus, it is recommended that one research further the concepts shown in this chapter.

One approach that is commonly used for simulation model optimization which was not explored in great detail in this chapter is the development of a mathematical metamodel, which represents a determined model output, according to optimization. This approach has a vast field of application and could have been applied in the problems presented. It is also recommended that the reader study this topic further as well [18, 43, 46]. In this sense, the

use of Kriging metamodeling for simulation has established itself in the scientific simulation community, which can be seen in [46, 47, 48, 49], thus demonstrating its promising research field.

The use of discrete-event simulation along with optimization is still scarce; nonetheless, in the last decade, important studies about this area of operational research have started to be realized, supporting the wider acceptance of this approach while also investigating the barriers to its continuous improvement. The use of sensitivity analysis in DOE enables a reduction in search space while increasing the optimization process's efficiency and speed.

Simulation optimization helps take simulation from being merely a means of scenario evaluation to a much greater solution generator. In doing so, sensitivity analysis plays a crucial role in this process, as it helps overcome the time and computational potential barriers presented by simulation models with large numbers of variables, thus making optimization an even greater tool for aiding decision-making.

In respect to future research possibilities, a potentially rich area in terms of investigation would be the examination of experimental designs which would reduce the number of experiments needed in order to identify the significant factors in pre-optimization phases. Another point which could be investigated more profoundly is the inclusion of qualitative techniques, such as brainstorming, cause and effect diagrams and Soft Systems Methodology, in order to select the factors to be utilized in experimentation. A field which is little-explored is sensitivity analysis in the optimization of multiple-objective models.

Author details

José Arnaldo Barra Montevechi, Rafael de Carvalho Miranda and Jonathan Daniel Friend
Universidade Federal de Itajubá (UNIFEI), Instituto de Engenharia de Produção e Gestão (IEPG), Itajubá, MG, Brazil

Acknowledgement

12. References

[1] Banks J, Carson II JS, Nelson BL, Nicol DM (2005) Discrete-event Simulation. New Jersey: Prentice-Hall.

[2] Bruzzone AG, Bocca E, Longo F, Massei M (2007) Training and recruitment in logistics node design by using web-based simulation. Int. J. Internet Manuf. Serv. 1(1): 32-50.

[3] Sargent RG (2009) Verification and validation of simulation models. In: Winter Simulation Conference, Proceedings... Austin, TX, USA.

[4] Law AM (2007) Simulation modeling and analysis. New York: McGraw-Hill.

[5] Jahangirian M, Eldabi T, Naseer A, Stergioulas LK, Young T (2010) Simulation in manufacturing and business: A review. Eur J Oper Res. 203(1):1-13.

[6] Ryan J, Heavey C (2006) Process modeling for simulation. Comput Ind. 57(5): 437-450.

[7] Law AM, Mccomas, MG (2002) Simulation-Based Optimization, In: Winter Simulation Conference, Proceedings... San Diego, CA, USA.

[8] Harrel CR, Mott JRA, Bateman RE, Bowden RG, Gogg TJ (1996) System improvement using simulation. Utha: Promodel Corporation.

[9] Fu MC (2002) Optimization for Simulation: Theory vs. Practice. J. Comput. 14(3): 192-215.

[10] Fu MC, Andradóttir S, Carson JS, Glover F, Harrell CR, Ho, YC, Kelly JP, Robinson SM (2000) Integrating optimization and simulation: research and practice. In: Winter Simulation Conference, Proceedings... Orlando, FL, USA.

[11] Harrel CR, Ghosh BK, Bowden R (2004) Simulation Using Promodel. New York: McGraw-Hill.

[12] Law AM, Kelton WD (2000) Simulation modeling and analysis. New York: McGraw-Hill.

[13] Banks J, Carson II JS, Nelson BL, Nicol DM (2000) Discrete event system simulation. New Jersey: Prentice-Hall.

[14] April J, Better M, Glover F, Kelly JP, Laguna M (2005) Enhancing business process management with simulation optimization. In: Winter Simulation Conference, Proceedings... Monterey, CA, USA.

[15] Andradóttir S (1998) Simulation optimization. In: Banks J, editor, Handbook of Simulation. New York: John Wiley & Sons. pp. 307-333.

[16] Biles WE (1979) Experimental design in computer simulation. In: Winter Simulation Conference, Proceedings... San Diego, CA, USA.

[17] Biles WE (1984). Experimental design in computer simulation. In: Winter Simulation Conference, Proceedings... Dallas, TX, USA.

[18] Montgomery DC (2005) Design and Analysis of Experiments. New York: New York: John Wiley & Sons, Inc.

[19] Kleijnen JPC (1998). Experimental design for sensitivity analysis, optimization, and validation of simulation models. In: Banks J, editor. Handbook of Simulation. New York: John Wiley & Sons. p. 173-223.

[20] Carson Y, Maria A (1997) Simulation optimization: methods and applications. In: Winter Simulation Conference, Proceedings... Atlanta, GA, USA.

[21] Azadeh A, Tabatabaee M, Maghsoudi A (2009) Design of Intelligent Simulation Software with Capability of Optimization. Aust. J. Basic Appl. Sci. 3(4): 4478-4483.

[22] Fu MC (1994) Optimization via simulation: A review. Ann Oper Res. 53:199-247.

[23] Rosen SL, Harmonosky CH, Traband MT (2007) Optimization of Systems with Multiple Performance Measures via Simulation: Survey and Recommendations. Comput. Ind. Eng. 54: 327-339.

[24] Bettonvil BWM, Castillo E, Kleijnen JPC (2009) Statistical testing of optimality conditions in multiresponse simulation-based optimization. Eur J Oper Res. 199(2): 448-458.

[25] Coello CAC, Lamont GB, Van Veldhuizen DA (2007) Evolutionary Algorithms for Solving Multi-Objective Problems (Genetic and Evolutionary Computation). New York: Springer.

[26] Holland JH (1992) Adaptation in Natural and Artificial Systems. Cambridge: MIT Press.

[27] Martí R, Laguna M, Glover F (2006) Principles of Scatter Search. Eur J Oper Res.169(2): 359-372.

[28] Glover F, Laguna M, Martí R (2005) Principles of Tabu Search, In: Gonzalez T, editor, Approximation Algorithms and Metaheuristics. London: Chapman & Hall/CRC.

[29] Ripley B (1996) Pattern Recognition and Neural Networks. Cambridge: University Press.

[30] Aarts EHL, Korst J, Michiels W (2005) Simulated Annealing. In: Burke EK, Kendall G, editors. Introductory tutorials in optimisation, decision support and search methodologies. New York: Springer. pp. 187-211.

[31] April J, Glover F, Kelly JP, Laguna M (2003) Practical introduction to simulation optimization. In: Winter Simulation Conference, Proceedings... New Orleans, LA, USA.

[32] Banks, J. Panel Session: The Future of Simulation. In: Winter Simulation Conference, Proceedings... Arlington, VA, USA, 2001.

[33] Tyni T, Ylinen J (2006) Evolutionary bi-objective optimization in the elevator car routing problem. Eur J Oper Res. 169(3):960-977.

[34] Montgomery DC (2009) Introduction to Statistical Quality Control. New York: John Wiley & Sons, Inc.

[35] Montgomery DC, Runger GC (2003) Applied Statistics and Probability for Engineers. New York: John Wiley & Sons, Inc.

[36] Kelton WD (2003) Designing simulation experiments. In: Winter Simulation Conference, Proceedings... New Orleans, LA, USA.

[37] Kleijnen JPC, Sanchez SM, Lucas TW, Cioppa TM (2005) State-of-the-Art Review: A User's Guide to the Brave New World of Designing Simulation Experiments. J. Comput. 17(3): 263–289.

[38] Montevechi JAB, Pinho AF, Leal F, Marins FAZ (2007) Application of design of experiments on the simulation of a process in an automotive industry. In: Winter Simulation Conference, Proceedings... Washington, DC, USA.

[39] Sanchez SM, Moeeni F, Sanchez PJ (2006) So many factors, so little time…. Simulation experiments in the frequency domain. Int J Prod Econ. 103: 149–165.

[40] Kleijnen JPC (2001) Experimental designs for sensitivity analysis of simulation models. In: Eurosim, Proceedings… Delft, Netherlands.

[41] Chung CA (2004) Simulation Modeling Handbook: a practical approach. Washington, DC: CRC Press.

[42] Landsheer JA, Wittenboer GVD, Maassen GH (2006) Additive and multiplicative effects in a fixed 2 x 2 design using ANOVA can be difficult to differentiate: demonstration and mathematical reason. Soc Sci Res. 35: 279–294.

[43] Montevechi JAB, Almeida Filho RG, Paiva AP, Costa RFS, Medeiros AL (2010) Sensitivity analysis in discrete-event simulation using fractional factorial designs. J. Simulat. 4: 128-142.

[44] Ye KQ, Hamada M (2001) A step-down Lenth method for analyzing unreplicated factorial designs. J Qual Technol. 33(2):140–153.

[45] Simrunner User Guide. ProModel Corporation: Orem, UT. USA. 2002.

[46] Kleijnen JPC, van Beers W, van Nieuwenhuyse I (2010) Constrained optimization in simulation: A novel approach. Eur J Oper Res. 202 (1): 164-174.

[47] Kleijnen JPC (2009) Kriging metamodeling in simulation: A review. Eur J Oper Res. 192(3): 707–716.

[48] Ankenman B, Nelson BL, Staum J (2010) Stochastic Kriging for Simulation Metamodeling. Oper. Res. 58(2):371–382.

[49] Biles WE, Kleijnen JPC, van Beers WCM, van Nieuwenhuyse I (2007) Kriging metamodeling in constrained simulation optimization: an explorative study. In: Winter Simulation Conference, Proceedings... Washington, DC, USA.