An innovative nature-inspired heuristic combined with Response Surface Methodology to find the optimal region in Discrete Event Simulation Models

Cassettari Lucia, Mosca Marco, Mosca Roberto
DIME, University of Genoa
Genoa, ITALY

Giribone Pier Giuseppe
CARIGE Bank - Financial Administration
Genoa, ITALY

***Abstract*—** The search for a stationary point in the study of Discrete Event Simulation models is a complex problem. This is because the equation of the objective function is never known a priori to the experimenter. In the case of restricted investigation domains the Response Surface Methodology typically provides, through the use of Central Composite Design, the experimental design most suitable for the construction of first and second order regression meta-models. The problem becomes more complex when the domain to be investigated is larger because, in that case, it becomes impossible to identify a meta-model regression able to describe the behaviour of the objective function on the entire domain. Such a limitation can be overcome by an appropriate use of research techniques such as gradiental or direct search methods. However, the presence in the domain of local stationary points may affect their effectiveness and forces the experimenter to track the investigation starting from several points of the domain, with a consequent increase in the number of function evaluations and computational time. In more recent times global research techniques have been developed, often inspired by natural processes. However they generally not perform well applied to Discrete Event Simulation models. For this reason the Authors have developed a new search algorithm called Attraction Force Optimization (AFO). The proposed approach applied to industrial problems up to 10-dimensional, offers significant advantages in terms of both exploration capacity and convergence speed. An application of the proposed technique to a real industrial case completes the discussion.

***Keywords*—**Discrete Event Simulation, Optimization, Optimal region, Nature-Inspired Heuristic, Response Surface Methodology

I. Introduction

The problem of discrete event simulation-based optimization is the finding of an optimal configuration for a stochastic function with an unknown structure [1]. This kind of problem in the past was approached through the use of different classes of optimization techniques, possibly adapted to the simulation-based applications [2]. In general it can be stated that each of the available techniques in the literature presents both advantages and limitations. Therefore a more powerful search algorithm in absolute terms is not identifiable but, from case to case, some techniques are more effective than others. Below there is a brief description of the principal algorithms available in the literature.

BFGS algorithm is a quasi-Newton local search method and it is considered one of the most efficient gradient methodologies [3]. This technique uses information derived from the calculation of partial derivatives in order to locally approximate an objective function through a second order quadratic form without the explicit estimation of Hessian matrix.

In addition to gradient techniques (such as Steepest Descent/Ascent, Newton and quasi-Newton methods), Direct search techniques can be implemented in order to find points of minimum/maximum. The main idea of these algorithms is to reach optima using only function evaluations, without estimating partial derivatives. Nelder-Mead simplex is one of the most popular direct search techniques for unconstrained non-linear multi-dimensional minimization [4,5].

Pattern Search methods [6] are a class of numerical algorithms which does not use gradient calculation and, consequently, they can optimize highly non-continuous target functions.

At each step, the procedure selects a set of points, called Mesh, in the neighborhood of the current point, called centroid. Mesh is built by a set of vectors, called Pattern. If the search finds a point in the Mesh which is better than the current point, it becomes the new centroid for the next iteration.

Among the Pattern Search methodologies, the most popular are: Generalized Pattern Search (GPS) and Mesh Adaptive Direct Search (MADS). Given that the main loop of the first one doesn't use random numbers, GPS is a deterministic search, while the second one is a stochastic algorithm. They can also be classified in function of the vector basis implemented for the Mesh building in N+1 or 2N search.

The most well-known global optimization methodologies are the so-called "nature inspired" heuristics. The main feature of these approaches is to simulate the behavior of a group of individuals (agents) that cooperates in order to reach a common objective (Genetic Algorithm and Particle Swarm Optimization) or a natural law (Simulated Annealing)[7].

Genetic Algorithms are a global heuristic based on the metaphor of natural selection of a species [8]. These methodologies dynamically modify a population of individuals (a set of points in the feasible domain) in order to form a better group in the next generation (iteration). The goodness of the solutions is directly linked with the topology of the target function: individuals that fit the environment has a higher probability to belong to the next generation. Usually, the best members pass without undergoing a change (elites children), the others are selected after some modification of their chromosomes (coordinates of the independent variables). The classic operators applied to these classes of individuals are mutation and cross-over. The first operation simulates the changes caused by the environment on a single member (the coordinates of one point are modified through a stochastic perturbation), while the second operator represents the coupling (the coordinates of two points are mixed randomly).

Simulated Annealing is a global heuristic which models the physical process of heating a material and the following gradual cooling till a minimum-energy particle configuration has been reached [9].

At each iteration of the main loop, the algorithm generates randomly a new point in the domain of the feasible solutions: its distance from the previous one is proportional to the current temperature. The technique accepts all the new points which improve the solution of the objective function and, with a certain probability, linked proportionally to the temperature, the points for which the system increases the energy. The expedient to accept solution that doesn't improve the current one, allows to avoid local minima and to explore better the domain, especially when the temperatures are high. A cooling schedule decreases the temperature: the main control parameter tends to be zero at the end of the optimization. Iteration after iteration, a smooth decrease of temperature leads to a lower acceptance probability of not-good solutions and a smaller search space. This process converges to the individuation of the optimal region of the objective function.

Kennedy and Eberhart, inspired by the social collaborative models used by groups of animals (flocks of birds, shoal of fishes) in order to find food, proposed an agent-based method which allows to optimize highly non-linear functions, called Particle Swarm Optimization (PSO) [10]. The main idea consists in modeling the exchange of information between agents, oriented to the common objective of finding the best source of food (the optimum of the target function). In the process of finding good solutions for the problem two operators, characteristic of social sciences, are simulated: the individual influence on the group, called simple nostalgia, and the group influence on the single agent, called publicized knowledge.

Many other researchers have extended their studies to test other routines inspired mainly by natural processes and social behaviors [11].

Among the most recent, are described briefly below the Harmony Search (HS), Artificial Bee Colony (ABC) and Continuous Ant Colony Optimization (CACO).

Z.W. Geem, J. H. Kim and G. V. Loganathan in 2001 proposed the HS: a music-inspired metaheuristic for optimization [12-14]. The process of finding optimal solutions is compared to the so called "improvisations" performed by a musician tuning his instrument in order to achieve the standard harmony, considered aesthetically perfect (the optimum of the target function).

The Artificial Bee Colony (ABC) Algorithm has been developed in 2005 by Dervis Karaboga to solve optimization problems [15]. Such Metaheuristic simulates the behavior adopted by the honeybees to exchange information within the hive leading to the identification of the most promising food source (ie the optimal region) in the surrounding environment (the objective function).

The Ant Colony Optimization, designed in discrete context by Marco Dorigo in 1992 [16] and later extended in the continuous [17], aims to simulate the transmission method of the information adopted by the ants (based on the distribution of pheromone). The regions in which it will be deposited the greater amount of pheromone will be those in which you will have a greater probability of finding the global optimum of the target function.

Among the techniques available in the literature, in stochastic simulation-based optimization problems, both DES both Monte Carlo, the Authors have so far preferred to use the Simplex Method and the Steepest Descent [18,19]. However, both methods have clear limitations. On the one hand both techniques are likely to find a local minimum and not the global one and, on the other hand, the local search techniques require, to avoid this problem, the utilization from more starting points of the domain, necessitating, consequently, a very high number of function evaluations.

Since in complex models such as those of industrial type many variables are often involved and being generally the execution time of each single simulation run very high, also the global search heuristics are not efficient and too expensive from the computational point of view. For this reason in recent times the Authors developed a new heuristic in order to overcome the limitations of traditional techniques, ensuring the convergence with a reduced number of iterations [20]. Such technique was only evaluated using the main benchmark functions suggested in the literature, and comparing its performances with those of the search algorithm mentioned above.

In this paper, the authors applied AFO technique in a real industrial optimization problem, comparing the results with those achieved by the other consolidated techniques available in the Optimization Toolboxes developed by MathWorks and other more recent techniques before described.

II. AFO

The technique "Attraction Force Optimization", AFO, is a new algorithm inspired by the attractive forces existing in Nature (for instance gravitational or electrostatic ones).

A general description of this heuristic, which is applied to an industrial simulator with the aim of finding the optimum area, is presented in this section.

This technique tries to inherit the advantages of both classic and modern philosophies. From the traditional approach, AFO allows a greater analytical tractability, tracking of convergence to the solution, the ability to always get the same result, starting from the same initial conditions. From the modern approach, AFO is able to present a thorough exploration of the experimental domain, the use of rather simple calculation operators and, also, the presence of a population of points interacting between them.

The algorithm starts placing a particle for each vertex, defined by the lower-bound and upper-bound of the problem and the scheme is then completed by adding *m-2* particles for each parallel direction to the generatrix base of the vectorial space.

The attractive force associated to each particle $f(x_i)$ is proportional to the potential assumed by the vectorial field (fitness function to maximize) in that point, x_i and the number of particles is the population of agents which remains constant for every iterations.

In the current scheme, the particle that has the best function evaluation, called base particle, x_{BASE}, is related to the others, identifying the distance at which the attractive forces are counterbalanced. The coordinates for these equilibrium points are estimated applying the following classic physic formula:

$$\vec{x}_{EQ,i} = \frac{\|r\|_2}{1+\omega\sqrt{\frac{f(x_i)}{f(x_{BASE})}}} \quad (1)$$

where:

$\|r\|_2$ the 2-norm distance between the i-th particle and the base particle

ω calibration parameter of the heuristic (typically $\omega = 1$)

AFO generates the scheme for the next iteration starting from the base and the less attracted equilibrium particle, i.e. the x_{EQ} with the bigger distance from x_{BASE}. The measure of the next scheme will be less than the previous one and the experimenter will be confident to find the optimal region inside the next scheme or on its frontier. The heuristic ends when the convergence has been reached.

III. The Industrial Test Case

The problem that has been studied in this work is related to a manufacturing system used for the production of transmission shaft for rotors of hydraulic pumps at high pressure. The goal of the study is to identify the optimal configuration of the industrial plant, namely the number of machines for each work station, in order to maximize the profit on a ten-years operating period.

Figure 1 shows the steps of the manufacturing process under examination. For each work station some different machines are placed in parallel.

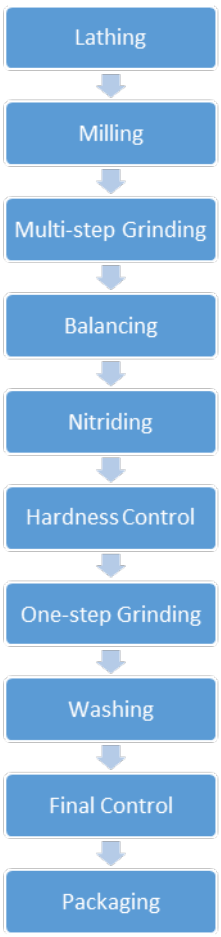

Fig. 1. Phases of the manufacturing process

The simulation software that was used to model the real system is SIMUL8, a particularly flexible software that allows to implement models based on Discrete-Event Simulation logic (DES). Appropriate probability distributions have been associated with each work station, with the aim of reproducing the stochasticity characterizing the timing set-up of the machine, the inter-arrival time of the materials, failures, etc. Table I shows the distributions used for each work station.

After studying the significance of the variables involved in the considered real case, based on the application of Design of Experiments techniques (DOE) [21-24], it was considered appropriate to limit the optimal configuration of the study analyzing the two most significant variables on the objective function, that is the number of machines used for the Nitriding (x_1) and the number of machines for the Final Control (x_2). Both variables have been considered varying in a range between 4 and 20.

TABLE I. WORKING TIME (HOURS/LOT SIZE)

Working Station	*Type of Probability Distributions (parameters)*
Lathing	Normal (3.8,0.5)
Milling	Uniform (3,5)
Multi-step Grinding	Normal (2.5, 1.1)
Balancing	Uniform (1,2)
Nitriding	Normal (4.5,1)
Hardness Control	Triangular (1,2,3)
One-step Grinding	Normal (2.4,0.5)
Washing	Fixed value 2.2
Final control	Normal (7.6, 1.5)
Packaging	Normal (1.7, 0.4)

Once the simulation model has been validated from the analogic point of view, the Authors determined the minimum run time in order to obtain statistically reliable results from the model [25-28]. From the analysis of the graph of Figure 2 the Authors decided to use a simulation run of 1000 days.

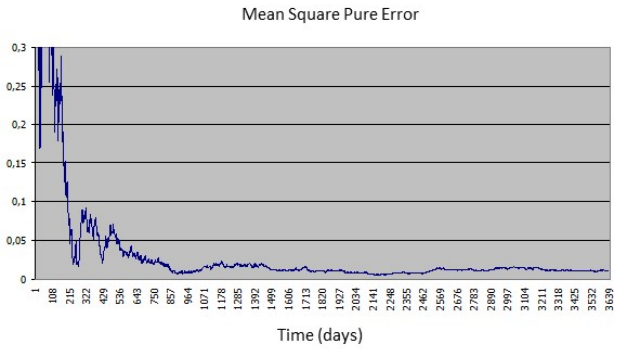

Fig. 2 Experimental error evolution curve of the simulation model

IV. AFO APPLICATION

The just mentioned heuristic has been applied in order to find the area of the optimal response surface obtained by the simulation approach described previously.

In particular, in this section the results that have been obtained from an application of AFO are outlined, pointing out that the goodness of the proposed solution does not depend on a specific calibration parameters at the discretion of the investigator, but rather is almost independent from it.

In the first case, the parameters are set as follows:

- number of particles along x-axis: 4;
- number of particles along y-axis: 4;
- number of AFO iterations: 3;
- parameter $\omega = 1$.

The algorithm proceeds as well in identifying the optimum area needing the calculation of 36 function evaluations. Figure 3, starting from the investigated experimental domain, highlights the progressive narrowing of the schemes during the three iterations of the technique. It is worth noting the rapid convergence of the schemes, with the final identification of the optimum point of coordinates (10.8756 , 17.1862).

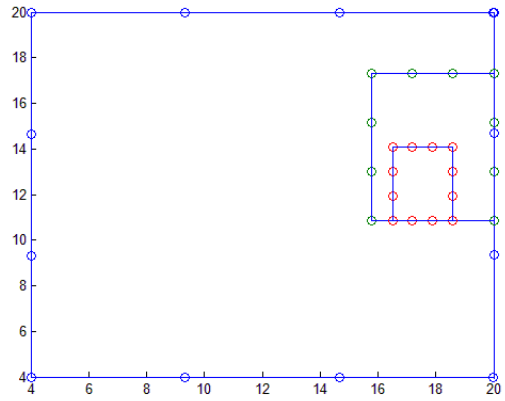

Fig. 3. AFO results: progressive narrowing of the schemes around optimum area

The scheme in Figure 4 shows the progressive value of the objective function, i.e. the optimal value, computed by AFO as a function of the number of iterations: it can be underlined that, given the set of parameters that have been fixed in this scenario, only two iterations would be sufficient to converge to the optimum zone.

The graphs in Figure 4 shows, always in function of the progressive number of experiments, the value of the basis-particle (graph on the top), which is the numeric value of the potential associated with the same particle and the value of the area subtended the schemes (graph on the bottom).

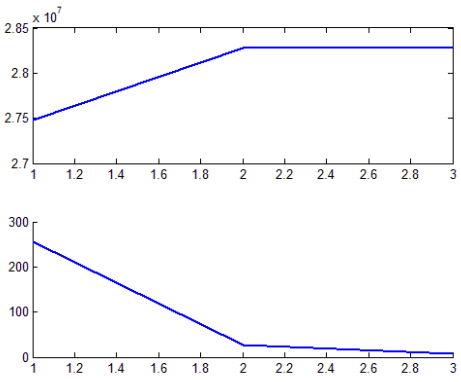

Fig. 4 Analysis of results: output of Matlab code after AFO implementation

It has been proceeded varying sequentially the parameters that can be set a-priori by the modeler, in order to evaluate if the technique is sensitive to the variation of any of them. Four different cases have been identified, whose parameters are configured as follows:

Case 1:

- number of particles along x-axis: 4;
- number of particles along y-axis: 4;
- number of AFO iterations: 4;
- parameter ω=1.

Case 2:

- number of particles along x-axis: 5;
- number of particles along y-axis: 5;
- number of AFO iterations: 3;
- parameter ω=1.

Case 3:

- number of particles along x-axis: 3;
- number of particles along y-axis: 3;
- number of AFO iterations: 4;
- parameter ω=1.

Case 4:

- number of particles along x-axis: 4;
- number of particles along y-axis: 4;
- number of AFO iterations: 4;
- parameter ω=1,5.

In Figures 5-8 the graphs obtained with Matlab for each of the four cases considered are reported. The first (Figure 5) is the one related to the case 1, namely that in which it was left unchanged the amount of particles placed along the sides of the schema, but there is an additional iteration than previously presented: the number of function evaluations required in this case is equal to 48 (12 more than before) and this additional computational effort does not justify an improvement of the solution, since the algorithm finds the same point than before. The second graph (Figure 6), with the addition of a particle on each side of the scheme, for a total of 48 function evaluations, identifies a very "near" point on the domain than they did in the first case (10.8805 ; 17.8818). Figure 7 shows the case 3, i.e. one in which a particle has been removed from each side, leaving the number of experiments set at 4, for a total of 32 function evaluations: the candidate point has coordinates (10,9572 ; 17.9892). Figure 8 shows the optimization results carried varying the contraction parameter that, if it is set to a higher value, allows a slower and more accurate investigation of the domain: in this case, the optimum can be found in (10.9104 , 17.8870) .

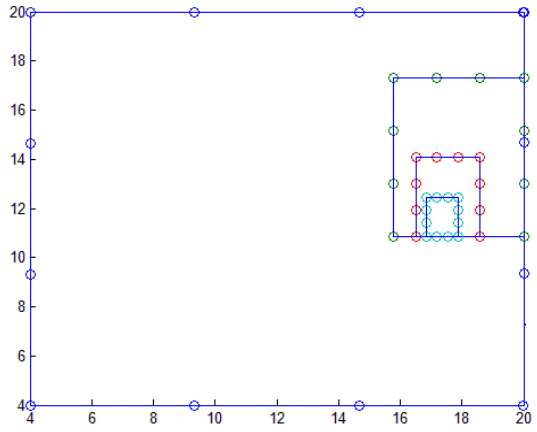

Fig. 5. AFO results for Case 1: progressive narrowing of the schemes around optimum area

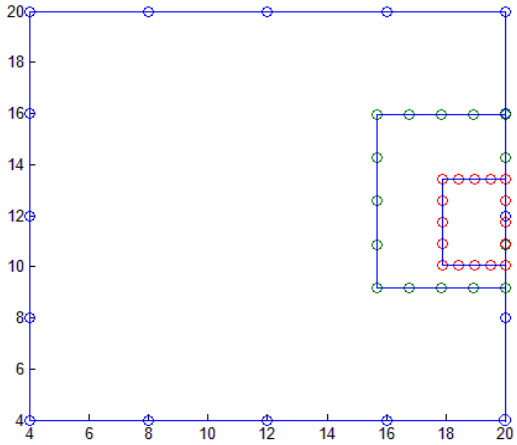

Fig. 6. AFO results for Case 2: progressive narrowing of the schemes around optimum area

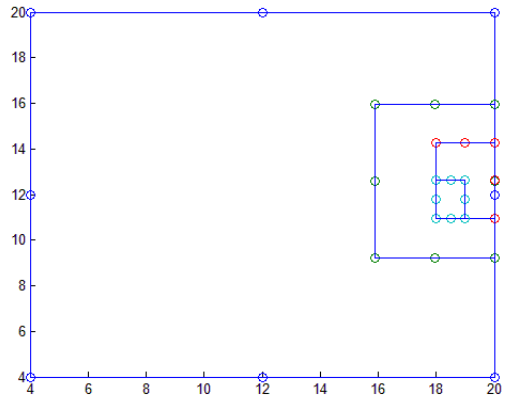

Fig. 7. AFO results for Case 3: progressive narrowing of the schemes around optimum area

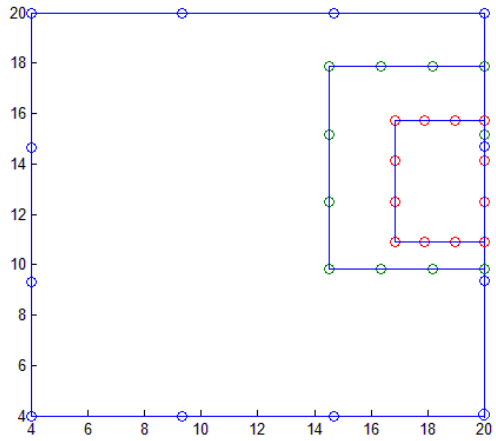

Fig. 8. AFO results for Case 4: progressive narrowing of the schemes around optimum area

When studying the graphics that have been just shown, it is worth underlining the robustness of Attraction Force Optimization, since it is capable of performing optimally independently from how the set of parameters is varied. It is possible to examinate figures 9-12 in which, for each one of the four cases, the value of the optimal objective function and the value of the area subtended the schemes are shown, both as function of the progressive number of AFO iterations.

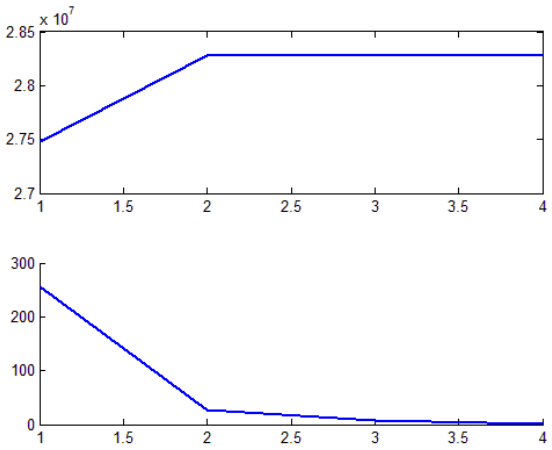

Fig. 9. Analysis of results for Case 1: output of Matlab code after AFO implementation

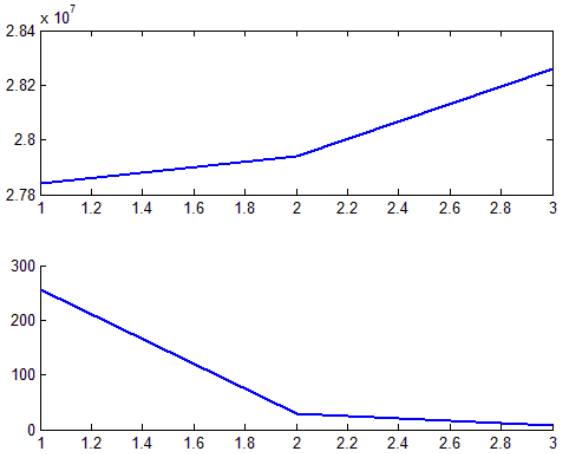

Fig. 10. Analysis of results for Case 2: output of Matlab code after AFO implementation

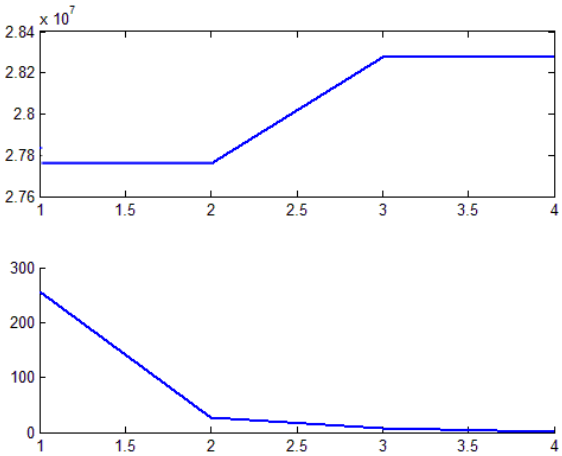

Fig. 11. Analysis of results for Case 3: output of Matlab code after AFO implementation

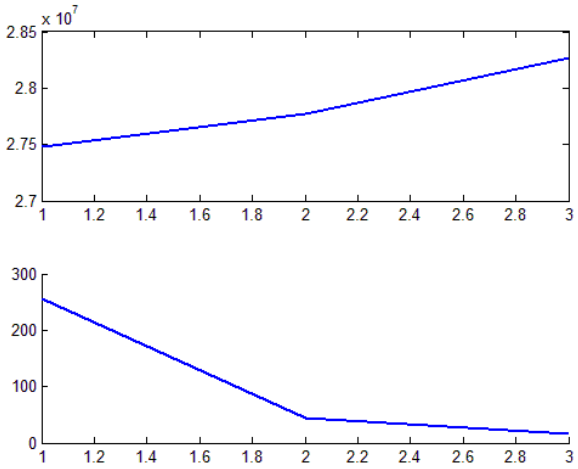

Fig. 12. Analysis of results for Case 4: output of Matlab code after AFO implementation

V. BENCHMARK AMONG OPTIMIZATION TECHNIQUES

After finishing the tests with AFO, the Authors have compared AFO's performances with those of some other traditional optimization techniques.

In order to get a comparison in terms of performance efficiency, all optimization techniques have been set with the same external conditions and a limit of 50 function evaluations has been put.

At the end of the experimental tests, all techniques have shown moderate-high efficiency and have correctly identified the optimum region.

In the grid of Figure 13 there is a summary of the coordinates of the optimum points found by the different techniques and the corresponding values of the objective function.

	16	17	18	19	20
8	21267980	21111980	20955980	20799980	20643980
9	24794030 C	24638105	24482105	24326105	24170105
10	26407055	28166855 H	28010780 B	27854705 A	27698705
11	26270480 M	28297280	28287905 D L	28131980 F	27975980
12	26133980	28160855	28151480	27995480	27839480
13	25997480	28024430 E	28014980	27858980	27702980
14	25860980	27887930 N O	27878480	27722480	27566480
15	25724480	27751430 I	27741980	27585980	27429980
16	25587980	27614930 G	27605480	27449480	27293480
17	25451480	27478430	27468980	27312980	27156980
18	25314980	27341930	27332480	27176480	27020480

	Legend
A	BFGS method
B	Nelder - Mead Simplex
C	GPS N+1 method
D	GPS 2N method
E	MADS N+1 method
F	MADS 2N method
G	Algoritmi genetici
H	Simulated Annealing
I	Particle Swarm Optimization
L	AFO
M	Harmony Search (HS)
N	Artificial Bee Colony (ABC)
O	CACO

Fig. 13. Grid of the objective function values for each combination of the two independent variable. The upper letter indicates the convergence point from the different techniques.

The grid has been built in order to assess in comparative terms the performance of the implemented optimization methods. From the knowledge of the objective function values obtained for each possible combination of the investigated variables, it has identified the optimum area of the objective function, i.e. the profitability of the company in terms of future cash flows over a period of 10 years (yellow area highlighted in Figure 12). The procedure for the grid generation has required the calculation of the 256 possible combinations, by varying at each iteration the number of the machines of the two work centers (Nitriding and Final Control) from 4 to 20 units. The computational cost for the grid construction has been particularly high and can be estimated at 25 hours of simulation time.

The construction of the grid, however, allowed an immediate and objective comparison of the results obtained.

It is useful to underline that AFO has identified the optimum zone needing just the explicit calculation of 36 evaluations of the objective function, while for all the other techniques the maximum limit of fval is rather higher, since it was fixed equal to 50. This highlights the goodness of AFO heuristic, being able to find an optimal solution requiring a greatly reduced computational burden.

VI. STATIONARY POINT ANALYSIS

Once identified the supposed stationary point by AFO algorithm, given the stochastic nature of the problem studied, the authors have analyzed the behavior of the response in the region around the identified optimum point. There may be in fact other coordinates within the optimal region equally interesting from an industrial point of view, for example by engaging less resources compared to the optimal point identified. These problems can be addressed through the proper construction of the confidence region centered on the location of the stationary point [29].

Let consider a second order metamodel equation:

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^2 + \sum_{i<j}^{k} b_{ij} x_i x_j = b_0 + \vec{x}^T \vec{b} + \vec{x} \hat{B} \vec{x} \quad (2)$$

The stationary point of the regression model can be computed setting the derivatives to zero as follows:

$$\frac{\partial \hat{y}(\vec{x})}{\partial \vec{x}} = \frac{\partial}{\partial \vec{x}} [b_0 + \vec{x}^T \vec{b} + \vec{x} \hat{B} \vec{x}] = \vec{b} + 2\hat{B}\vec{x} \quad (3)$$

The j-th derivative $d_j(\vec{x})$ can be written as:

$$d_j(\vec{x}) = b_j + 2\hat{B}_j^T \vec{x} \quad (4)$$

where vector $\hat{B}_j^T$ represents the j-th row of the second order coefficients matrix $\hat{B}$.

These derivatives are simple linear functions in $x_1, x_2, ..., x_k$.

Let denote the vector of derivatives as the k-dimensional $\vec{d}(\vec{x})$ and let suppose to calculate the derivatives in $\vec{d}(\vec{t})$, obtaining the coordinates of the true stationary point of the system, unknown a priori.

If the errors of the model of the equation (2) are i.i.d., then

$$d(t) \approx N\left(\vec{0}, \mathrm{Var}[\vec{d}(\vec{t})]\right) \quad (5)$$

where $\mathrm{Var}[\vec{d}(\vec{t})]$ is the variance-covariance matrix of $\vec{d}(\vec{t})$.

It follows that

$$\frac{\vec{d}^{\,T}(\vec{t})\left\{\hat{\mathrm{Var}}[\vec{d}(\vec{t})]\right\}^{-1}\vec{d}(\vec{t})}{k} \approx F_{k,N-p} \quad (6)$$

where $F_{k,N-p}$ is a F-Distribution having k and $N-p$ degrees of freedom and $\mathrm{Var}[\vec{d}(\vec{t})]$ is a $k \times k$ matrix, function of $\vec{t}$.

In case of two independent variables the formula are so defined:

$$\begin{cases} d_1(\vec{t}) = b_1 + 2\left(b_{11}t_1 + \dfrac{b_{12}}{2}t_2\right) \\ d_2(\vec{t}) = b_2 + 2\left(\dfrac{b_{12}}{2}t_1 + b_{22}t_2\right) \end{cases} \quad (7)$$

$$\mathrm{Var}[\vec{d}(\vec{t})] = \begin{bmatrix} \mathrm{Var}[d_1(\vec{t})] & \mathrm{Cov}[d_1(\vec{t}), d_2(\vec{t})] \\ \mathrm{Cov}[d_1(\vec{t}), d_2(\vec{t})] & \mathrm{Var}[d_2(\vec{t})] \end{bmatrix} \quad (8)$$

It can be noticed that the elements of $\mathrm{Var}[\vec{d}(\vec{t})]$ has been derived from the variance-covariance matrix of the regression coefficients $\Psi_{VAR-COV} = \left(X^T X\right)^{-1}\sigma^2$.

The following equalities are indeed satisfied:

$$\mathrm{Var}[d_1(\vec{t})] = \mathrm{Var}[b_1] + 4t_1^2\,\mathrm{Var}[b_{11}] + t_2^2\,\mathrm{Var}[b_{12}] \quad (9)$$

$$\mathrm{Var}[d_2(\vec{t})] = \mathrm{Var}[b_2] + t_1^2\,\mathrm{Var}[b_{12}] + 4t_2^2\,\mathrm{Var}[b_{22}] \quad (10)$$

$$\mathrm{Cov}[d_1(\vec{t}), d_2(\vec{t})] = 4t_1t_2\,\mathrm{Cov}[b_{11}, b_{22}] + t_1t_2\,\mathrm{Var}[b_{12}] \quad (11)$$

Based on relationship (6), we have:

$$\Pr\left\{\vec{d}^{\,T}(\vec{t})\left\{\hat{\mathrm{Var}}[\vec{d}(\vec{t})]\right\}^{-1}\vec{d}(\vec{t}) \le k\,F_{\alpha,k,N-p}\right\} = 1-\alpha \quad (12)$$

where $F_{\alpha,k,N-p}$ is the upper α-th percent point of the F-Distribution having k and $N-p$ degrees of freedom.

It follows that the unknown values of $\vec{t} = t_1, t_2, \ldots, t_k$, that fall inside the $100(1-\alpha)\%$ confidence region for the stationary point, satisfy the inequality below:

$$\vec{d}^{\,T}(\vec{t})\left\{\hat{\mathrm{Var}}[\vec{d}(\vec{t})]\right\}^{-1}\vec{d}(\vec{t}) \le k\,F_{\alpha,k,N-p} \quad (13)$$

Concerning the application presented in this paper, the Authors built a Central Composite Design centered in the stationary point $\vec{x}_S = (11;17)$, as shown in Figure 14.

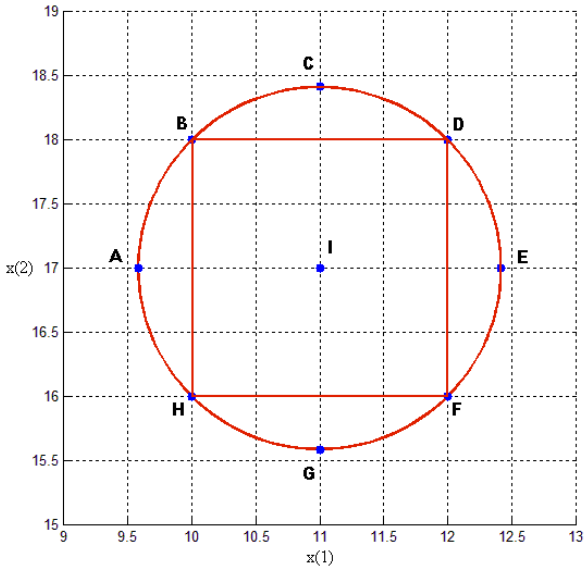

Fig. 14. Central Composite Design centered in the stationary point

The experimental responses obtained are reported in Table 2.

TABLE II. CENTRAL COMPOSITE DESIGN EXPERIMENTAL RESPONSES

Nitriding	Final Control	Simulation response
X1	X2	f(X1, X2)
9,586	17	26400335
10	18	28010780
11	18,4142	28124470
12	18	28151480
12,414	17	27667715
12	16	26133980
11	15,5858	25267715
10	16	26407055
11	17	28297280
11	17	28444730
11	17	28437205

The regression meta-model equation obtained is reported below.

$$\begin{aligned} \hat{y} = {} & -2.65229871073019E+008 \\ & +1.17263972064171E+007 * x_1 \\ & +2.58595089050437E+007 * x_2 \\ & +1.03443750000056E+005 * x_1 * x_2 \\ & -6.03520208333318E+005 * {x_1}^2 \\ & -7.65733958333299E+005 * {x_2}^2 \end{aligned} \quad (14)$$

ANOVA table reported in Fig. 15 shows that the regression model suitably fits the experimental data.

ANOVA for Response Surface Quadratic model

Analysis of variance table [Partial sum of squares - Type III]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	1.199E+013	5	2.399E+012	17.44	0.0035	significant
A-A	*3.444E+011*	*1*	*3.444E+011*	*2.50*	*0.1744*	
B-B	*7.410E+012*	*1*	*7.410E+012*	*53.87*	*0.0007*	
AB	*4.280E+010*	*1*	*4.280E+010*	*0.31*	*0.6010*	
A^2	*2.057E+012*	*1*	*2.057E+012*	*14.95*	*0.0118*	
B^2	*3.311E+012*	*1*	*3.311E+012*	*24.07*	*0.0045*	
Residual	6.878E+011	5	1.376E+011			
Lack of Fit	*6.740E+011*	*3*	*2.247E+011*	*32.58*	*0.0299*	*not significant*
Pure Error	*1.379E+010*	*2*	*6.896E+009*			
Cor Total	1.268E+013	10				

Fig. 15 ANOVA table for the quadratic regression model

Also the $R^2 = 0.9453$ and $R^2_{adj} = 0.8905$ parameters confirm the goodness of fit of the quadratic form identified, whose graphical representation is shown in Figure 16.

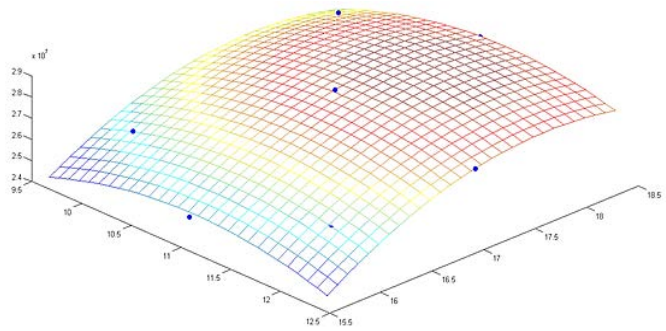

Fig. 16 Response Surface of the optimal region

The authors then proceeded with the construction of the optimal region through the elements $\vec{d}(\vec{t})$ and $\mathrm{Var}[\vec{d}(\vec{t})]$.

For the calculation of $\vec{d}(\vec{t})$, having an estimation of $\vec{b}$, you can easily find an expression of the system starting from (6).

According with (8), it is necessary the elements $\mathrm{Var}[d_1(\vec{t})]$, $\mathrm{Cov}[d_1(\vec{t}), d_2(\vec{t})]$ and $\mathrm{Var}[d_2(\vec{t})]$.

These elements derive from the variance-covariance matrix of regression coefficients $\Psi_{VAR-COV} = (X^T X)^{-1}\sigma^2$, where σ^2 is replaced by the magnitude of experimentally obtained s^2, whose best estimator is *MSE*.

Along the main diagonal of $\Psi_{VAR-COV}$ can be derived $\mathrm{Var}[b_1], \mathrm{Var}[b_2], \mathrm{Var}[b_{12}], \mathrm{Var}[b_{11}], \mathrm{Var}[b_{22}]$ and the covariance $\mathrm{Cov}[b_{11}, b_{22}]$.

According with (8), (9) e (10) it is possible to calculate $\mathrm{Var}[d_1(\vec{t})]$, $\mathrm{Var}[d_2(\vec{t})]$ and $\mathrm{Cov}[d_1(\vec{t}), d_2(\vec{t})]$, consequently the matrix $\mathrm{Var}[\vec{d}(\vec{t})]$ is determined.

Being known all elements the Author proceed to the confidence region construction using $\alpha = 0.05$.

Applying the (11) to the industrial problem, we obtain

$$\vec{d}^T(\vec{t})\left\{\hat{\mathrm{Var}}[\vec{d}(\vec{t})]\right\}^{-1}\vec{d}(\vec{t}) \le k\,F_{\alpha=0.05,k=2,N-p=11-6=5} = 2\cdot 5.7861 = 11.5722$$

The region highlighted in yellow in Fig. 17, represents the confidence region for the stationary point $\vec{x}_S$.

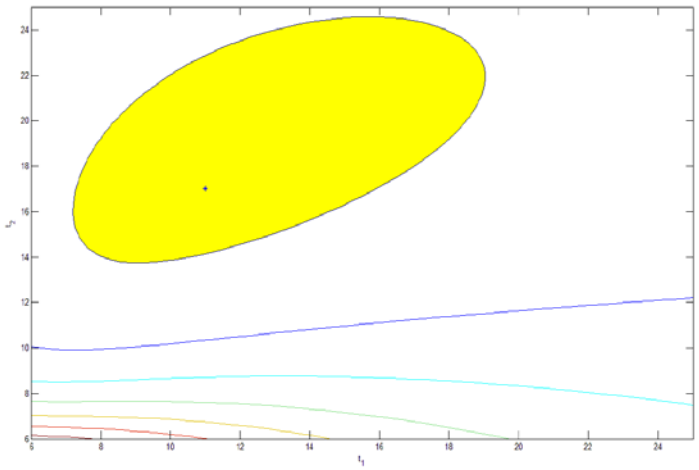

Fig. 17 Confidence region of the stationary point

The procedure identifies an optimal region consistent with the results obtained by the search algorithms and the grid reported in Fig. 13. The domain portion detected is more skewed to the right of the stationary point. This is reasonable since the response surface actually has values close to those of stationarity in this experimental region.

The technique used was therefore suitable to describe the confidence region of the analyzed industrial problem.

VII. Conclusions

The aim of the study has been to test the effectiveness and the efficiency of a new nature-inspired heuristic applied in the simulation-based optimization problems. The technique, called AFO, has been compared with other consolidated optimization techniques in a real industrial optimization problem. The goal was to determine the optimal plant configuration in a manufacturing line in order to obtain the maximum profit over a 10 years period.

The results demonstrate how the analyzed heuristic reaches the optimum area with a considerable saving in terms of number of function evaluations, compared to traditional techniques. It would be interesting to test its effectiveness in other optimization problems with a higher number of independent variables in order to understand if and under what conditions the algorithm can be more powerful than the other techniques available in literature.

REFERENCES

[1] P. Buchholz, "Optimization of Stochastic Discrete Event Simulation Models", Dagstuhl Seminar Proceedings 09261, Models and Algorithms for Optimization in Logistics

[2] L.A. Riley, "Discrete-event simulation optimization: a review of past approaches and propositions for future direction, Proceedings of the 2013 Summer Computer Simulation Conference, 2013

[3] R. Fletcher, "Practical methods of optimization", John Wiley and Sons, 1987.

[4] R. R. Barton, J. S. Ivey, "Modifications of the Nelder-Mead Simplex Method for stochastic simulation response optimization", proceedings of the Winter simulation conference, 1991.

[5] J. C. Lagarias, "Convergence properties of the Nelder Mead simplex method in low dimensions", SIAM J. Optim., Vol.9, No.1, Society for Industrial and applied Mathematics, 1998.

[6] R. M. Lewis, V. Torczon, "Pattern Search Methods for Linearly Constrained Minimization", SIAM Journal on Optimization, Vol.10, No 3, pp 917–941, 2000.

[7] S. Fioribello, P. G. Giribone, S. Ligato, "Design of a robust calibration for the Hull-White stochastic tree by means of global heuristics of research", AIFIRM, 2016.

[8] J. H. Holland, "Adaptation in Natural and Artificial Systems", MIT Press, 1975.

[9] S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, "Optimization by Simulated Annealing", Science, 1983.

[10] R. Eberhart, J. Kennedy, "Particle Swarm Optimization", Proceedings of ICNN' 95 - International Conference on Neural Networks, vol. 4, pp. 1942-1948, 1995.

[11] X-S. Yang, "Nature-Inspired Metaheuristic Algorithms - Second Edition", Luniver Press (Univ. of Cambridge), 2010.

[12] Z. W. Geem, "Music Inspired harmony search algorithm: theory and applications" , Springer, Berlin (2009).

[13] Z. W. Geem, "Harmony Search algorithms for structural design optimization",Springer, Berlin (2009).

[14] Z. W. Geem, J. H. Kim, G.V. Loganathan, "A new heuristic optimization algorithm: Harmony Search", Simulation, 76, 60-68, 2001.

[15] D.Karaboga, "An idea based on honey bee swarm for numerical optimization", Technical Report – TR06 Ercyes University – Kayseri, Turkey, October 2005.

[16] M. Dorigo, "Optimization, Learning and Natural Algorithms" Ph.D Thesis, Dipartimento di Elettronica, Politecnico di Milano 1992.

[17] M. Dorigo, K. Socha, "Ant Colony Optimization for continuous domains" , European Journal of Operational Research, 2006.

[18] L. Cassettari, R. Mosca, R. Revetria, F. Rolando, Sizing of a 3,000,000t bulk cargo port through discrete and stochastic simulation integrated with response surface methodology techniques. Recent Advances in Signal Processing, Computational Geometry and Systems Theory - ISCGAV'11, ISTASC'11, 2011.

[19] P. Lonardo, D. Anghinolfi, M. Paolucci, F. Tonelli, "A stochastic linear programming approach for service parts optimization" , CIRP Annals - Manufacturing Technology, Volume 57, Issue 1, 2008, Pages 441–444.

[20] I. Bendato, L. Cassettari, P. G. Giribone, S. Fioribello, "Attraction Force Optimization: a deterministic nature-inspired heuristic for solving optimization problems", Applied Mathematical Sciences, Vol. 10, 2016, no. 20, 989 – 1011.

[21] I. Bendato, L. Cassettari, F. Rolando, M. Mosca, R. Mosca, "Stochastic techno-economic assessment based on a Monte Carlo simulation and the response surface methodology: The case of an innovative linear Fresnel CSP system, Energy, Vol. 101, 2016, pp. 309–324.

[22] I. Bendato, L. Cassettari, M. Mosca, R. Mosca, "New Markets Forecast and Dynamic Production Redesign Through Stochastic Simulation", International Journal of Simulation Modelling, 14(3), 2015.

[23] I. Bendato, L. Cassettari, M. Mosca, R. Mosca, "A Design of Experiments/Response Surface Methodology Approach to Study the Economic Sustainability of a 1 MWe Photovoltaic Plant", Renewable and Sustainable Energy Reviews, 51, 2015, 1664-1679.

[24] E. Briano, C. Caballini, P. Giribone, R. Revetria R., "Design of experiment and montecarlo simulation as support for gas turbine power plant availabilty estimation", 12th WSEAS International Conference on Automatic Control; Modelling and Simulation, ACMOS '10, 2010.

[25] L. Cassettari, R. Mosca, R. Revetria, "Experimental Error Measurement in Monte Carlo Simulation". In: In: Abu-Taieh, E.M.O., El Sheikh, A.A.R. (Ed.), Handbook of Research on Discrete Event Simulation Environments: Technologies and Applications, IGI Global, 2009, pp 92-142.

[26] I. Bendato, L. Cassettari, P.G. Giribone, R. Mosca, "Monte Carlo method for pricing complex financial derivatives: an innovative approach to the control of convergence", Applied Mathematical Sciences, 9(124), 2015, 6167-6188.

[27] R. Mosca, A.G. Bruzzone, L. Cassettari, M. Mosca, "Risk analysis for industrial plants projects: An innovative approach based on simulation techniques with experimental error control", 21st European Modeling and Simulation Symposium, EMSS 2009.

[28] L. Cassettari, R. Mosca, R. Revetria, "Monte Carlo Simulation Models Evolving in Replicated Runs: A Methodology to Choose The Optimal Experimental Sample Size", Mathematical Problems in Engineering, Volume 2012 (2012).

[29] R.H. Myers, D.C. Montgomery, C.M. Anderson, "Response Surface Methodology: Process and Product Optimization Using Designed Experiments",Wiley, 2009 .