Orthogonal-Array based Design Methodology for Complex, Coupled Space Systems

IAC-10.B5.2.4

ORTHOGONAL-ARRAY BASED DESIGN METHODOLOGY FOR COMPLEX, COUPLED SPACE SYSTEMS

G.Ridolfi
Politecnico di Torino, Italy, guido.ridolfi@polito.it
Delft University of Technology, The Netherlands, g.ridolfi@tudelft.nl

E. Mooij
Delft University of Technology, The Netherlands, e.mooij@tudelft.nl

D. Cardile
Politecnico di Torino, Italy, diego.cardile@polito.it

S. Corpino
Politecnico di Torino, Italy, sabrina.corpino@polito.it

F. Stesina
Politecnico di Torino, Italy, fabrizio.stesina@polito.it

The process of designing a complex system, formed by many elements and sub-elements interacting between each other, is usually completed at a system level and in the preliminary phases in two major steps: design-space exploration and optimization. In a classical approach, especially in a company environment, the two steps are usually performed together, by experts of the field inferring on major phenomena, making assumptions and doing some trial-and-error runs on the available mathematical models. To support designers and decision makers during the design phases of this kind of complex systems, and to enable early discovery of emergent behaviours arising from interactions between the various elements being designed, the authors implemented a parametric methodology for the design-space exploration and optimization. The parametric technique is based on the utilization of a particular type of matrix design of experiments, the orthogonal arrays. Through successive design iterations with orthogonal arrays, the optimal solution is reached with a reduced effort if compared to more computationally-intense techniques, providing sensitivity and robustness information. The paper describes the design methodology in detail providing an application example that is the design of a human mission to support a lunar base.

NOMENCLATURE

ANOVA	Analysis Of Variance
CCD	Central Composite Design
Isp	Specific Impulse
MDF	Multi Disciplinary Feasible
NHD	Non Hierarchical Decomposition
OA	Orthogonal Array
SOS	System of Systems

I. INTRODUCTION

In this paper we discuss a method for the design of complex systems using orthogonal arrays. In an earlier paper, ref. [1], we described a decomposition approach for a category of complex systems called "*System-of-Systems*". A system-of-systems is, by definition, a system composed of multiple elements and sub-elements. For instance, a liquid fuel rocket engine can be considered an SOS composed of the structure, fuel system, and control system, each system further decomposed into components such as fuel tanks, nozzle(s) etc.

At a much higher level, a planetary exploration mission architecture can also be considered an SOS, which entails the system of Earth infrastructure, inter-planetary transportation, and the planet surface habitat, each of these systems comprising components such as the launch vehicles, transfer vehicles, and lander in the interplanetary transportation system.

The elements of the System-of-Systems have multiple interactions between each other and the aggregate activity is typically more than the simple interactions of the various parts. The analysis and comprehension of the emergent behaviours arising from these interactions is of a crucial importance for the proper design of such kind of complex systems; in ref. [2] this aspect has already been assessed and discussed in detail.

The system-of-systems used as a case study in this paper is composed of the mathematical models for elements and subsystems belonging to a hypothetical human mission to support the return of mankind on the Moon with a permanent outpost. In particular, the models that have been developed allow for the design of a manned re-entry capsule, a service module for the capsule, a lander system, an ascent vehicle, and an Earth-Moon transfer vehicle. The transfer vehicle brings all the other elements from a low Earth orbit to a low lunar orbit. The astronauts are transferred on the lunar surface by the lander and the ascent module, which will bring them on orbit again at the end of the mission. The service module provides the necessary ΔV to inject the capsule with the astronauts in the transfer orbit from the Moon towards the Earth. For more detailed information on the mission architecture see ref. [1]

The main goal of this paper is to describe the design methodology that has been formulated and applied to the System of Systems just described, providing a detailed description of the obtained results. The design methodology has been developed to cope with the evolving environment of engineering companies and agencies. The adopted design processes are radically changing, and the tools that support the engineers and the decision makers are evolving consequently. In a traditional approach, even with the experience achieved in the last decade by applying the paradigm of Concurrent Engineering, companies are organized in engineering groups that are usually aligned with disciplines, parts, and/or processes concurring for the development of a product/system. In most of the cases the various groups have full authority on design issues belonging to their inherent domain only. Each engineering group uses its own design methods and software tools. Most important, the interfaces are usually determined by the experts and provided manually, therefore not fulfilling the requirement of concurrency in the design process. The experts are also responsible for achieving compromises between engineering groups and all the most important design choices are taken together, using all the available information.

The problem of designing using highly integrated mathematical models is on an open debate among the research groups dealing with complex systems[17]-[26]. As will be explained in more detail in section II, three different classes of methodology can in principle be used to solve a system-of-systems, but the stochastic algorithms, e.g., genetic algorithms, are those who are more widely applied. The main disadvantages of those methodologies is that they are "*simulation intensive*", causing long run times, providing almost no insight in the problem of interest and leading to poor convergence of the solution in some cases.

The authors are investigating a different and possibly more efficient way to deal with this kind of systems, based on parametric design. We think, according to what has been addressed in previous works[2],[3], that supporting the design team with graphical information, instead of providing a ready solution, with no clear insight in the behavior of the single elements and in the interactions, is crucial for System Engineering processes and tools. The scope of systems engineering is to make sure that the development process leads to the most cost-effective final product. Before every decision is made, especially for those that are hard to undo in an advanced phase, the alternatives should be carefully assessed, understood and discussed. This can be achieved only if there is an efficient communication between the various disciplines/systems/elements that determine the performance(s) of the system. A system engineering tool should help in this direction.

Thus, the main objective of the methodology presented in this paper is to support designers and decision makers during all the design phases of a complex system by keeping the concurrency of the design process and providing information on the behaviour of the system and its components, in terms of interactions, robustness and sensitivity.

The remaining part of the paper is organized as follows. In section II, the main aspects and potentials of the design methodology are described. In section III, a brief overview of the applied test case to verify the methodology is provided. The results of the design iterations are described and commented in section IV. Finally, in section V, conclusions and recommendations are provided.

II. DESIGN SPACE EXPLORATION USING ORTHOGONAL ARRAYS

It is commonly accepted that one of the main objectives of the design process of systems of any complexity is to predict the behavior of the system in its operative environment, and to set all the physical and functional characteristics of the system such that it performs as required. The cheapest and maybe fastest way to predict the behavior of a system is to create a (n analytical) model and to extract information about its behavior and its performance by executing multiple experiments with different levels of the design variables, i.e., the input to the system. It is typically said, especially since the last few years, that the objective is to find the levels of the design variables that "*optimize*" the system, under certain boundary conditions and constraints. The "*optimization*" of a mathematical model is usually obtained using so-called gradient-based methods. These methods use first-derivatives and, sometimes, second derivatives of the mathematical equations that describe the physical phenomenon of interest. These derivatives are used to determine the search direction in the design space that should lead to the optimum. The gradient-based methods perform very well for problems in which the mathematical functions do not present discontinuities or plateaus, i.e., flat regions in the design space. For this kind of problems, in recent years many so-called stochastic methods, such as those based on genetic algorithms and particle swarms[12],[13], have been developed and applied, even to problems of a certain complexity[4]-[8]. One of the drawbacks of such methodologies, if observed from a company perspective, is that the analyses are not fully traceable, due to some level of randomness that all those methods exploit. Further, the computation of sensitivities is still much dependent on the type of equations describing the system. Mistree *et al*. successfully applied a parametric design technique based on orthogonal arrays to study the trajectory of a re-entry vehicle through the Earth atmosphere[9]-[11]. The objective of the work was to substitute the Monte Carlo technique with a more efficient methodology to reduce the number of simulations for the design of space vehicle trajectories. Their work is based on the experimental design approach developed by Dr. Taguchi[16] in the 1960s. Using Taguchi's orthogonal arrays, Mistree achieved a very high level of efficiency in the engineering effort consumed in conducting experiments to obtain the information needed to guide decisions related to a particular product.

II.1 Factorial Design and Orthogonal Arrays

Orthogonal arrays are special matrices used for factorial design and result in so-called fractional factorial designs, as only a subset of all possible combinations is addressed. Factorial design is a very efficient and systematic approach used to study the effect of several factors on certain responses of interest. The term "*factors*" is used to address the design variables, while the term "*responses*" indicates the output variables (the objectives) that describe the performance of the system. Performing a factorial design means that the mathematical model of the system of interest, considered as a *black box*, is experimented with all (or a subset in case of a fractional factorial design) combinations of factors levels. The levels of a factor are the values that the factor can assume, e.g., a minimum, nominal and maximum value. Studying the effect of a factor on the response means studying the variation in the response caused by a change in the level of the factor itself. This variation is computed, for instance, by averaging out the responses obtained from the simulations with the factor of interest at a certain level, irrespectively of the levels of the other factors.

. This is also called the main effect of the factor. The factorial design allows also the study of interactions between factors. The interaction effect between two or more factors can be computed by applying two or more times the same procedure used for computing the main effect. For instance, to compute the interaction between two factors one can select the levels of the factors of interest and averaging out the responses obtained from the simulations irrespectively of the levels of the other factors. To obtain a full picture of the interaction of the parameters, this procedure shall be repeated for all the available combinations of factor levels.

There are many alternative types of factorial designs. They differ from each other based on the number of design evaluations needed to provide information relative to factor effects and the effects of their interactions. The number and types of factors effects and interactions that is possible to distinguish, given a certain factorial design, is closely related to the number of design evaluations. As a general rule (necessary, but not sufficient), the number of design points must be at least equal to the number of factors and interactions effects that one is willing to estimate. If the number of design points is too low, not all main and interaction effects can be distinguished, causing the so-called confounding effect, which arises when we want to study an interaction effect with the same combination of variable levels of another interaction or a main effect. The confounding effect, typical when using orthogonal arrays, will be explained in more detail later in this section.

Full factorial designs require the largest number of design points since all combinations of variables levels are tested. Consider for instance having three design variables each one at two levels. According to a full factorial design the number of design points needed is 2^3 (= 8), as shown in Fig. 1.

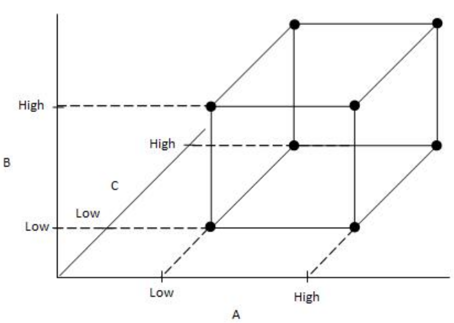

Fig. 1: Full factorial design: three factors at two levels.

The other class of factorial designs is called fractional factorial. In this case the required data points are only a fraction of those required by a full factorial design. There are three main subclasses of fractional factorial designs. The class called *Resolution III* is the least demanding in terms of required design points. For a *Resolution III* class no main effects are confounded with any other main effect, but main effects are confounded with two-factors interactions (and higher order) that may also be confounded with each other. In Fig. 2, a *Resolution III* design with 3 factors at 2 levels is shown. As can be seen, already with only 3 factors the number of design points required is half if compared to the relative full factorial design presented in Fig. 1.

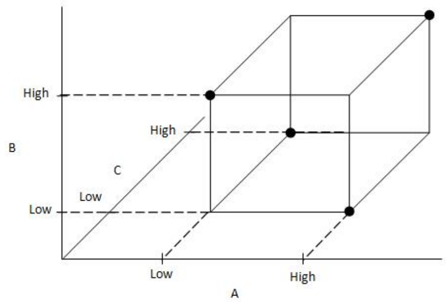

Fig. 2: Fractional factorial design (Orthogonal Design): three factors at two levels.

Orthogonal Arrays are *Resolution III* designs expressed in a tabular form. In Table 1, we present an L_8 orthogonal array, i.e., 8 experiments with a maximum of 7 factors at 2 levels. Actually, it is not necessary to assign a factor to all the columns of the orthogonal array. Columns can be left empty to be able to study interaction effects of interest, avoiding confounding.

Table 1: L8 (27) orthogonal array, adapted from ref. [15]

Experiment Number	Column						
	1	2	3	4	5	6	7
1	-1	-1	-1	-1	-1	-1	-1
2	-1	-1	-1	1	1	1	1
3	-1	1	1	-1	-1	1	1
4	-1	1	1	1	1	-1	-1
5	1	-1	1	-1	1	-1	1
6	1	-1	1	1	-1	1	-1
7	1	1	-1	-1	1	1	-1
8	1	1	-1	1	-1	-1	1
	A	B	C	D	E	F	G
	Factor Assignment						

The level *-1* represents the minimum level of the design variable of interest, *1* represents the maximum. With a full factorial design the number of required experiments would have been 2^7, 128.

To represent the orthogonal array of Table 1 in a graphical form, as shown in Fig. 2, a hypercube of 7 dimensions (because 7 factors are taken into account) would be required.

Besides significantly reducing the computational effort in exploring the design space, orthogonal arrays enable an efficient determination of the effects of the design variables (or control parameters) on the outputs, (or performance parameters). The term "*orthogonality*" has to be interpreted in the combinatorial sense. To explain this concept Phadke introduces the so-called balancing property[15], which means that for any pair of columns, all combinations of factor levels are present an equal number of times. For instance, in Table 1, in the first two columns all the factor combinations (i.e., -1 -1, -1 1, 1 -1, and 1 1) are present two times. The same happens for any other pair of columns. This property assures that once all the simulations are completed, for any couple of factors all the factor-level combinations are tested.

The utilization of orthogonal arrays for experiment planning is straightforward. Referring to the OA presented in Table 1, in each column we read the level that has to be assigned to the design variables for each experiment. In each row, we read the variable sets for every single experiment to be performed. Therefore, the number of rows represents the total number of experiments. The orthogonal property of the arrays developed by Dr. Taguchi, causes the already mentioned confounding effect. Suppose we want to study the effect of the interaction of two of the parameters assigned to the columns of the OA on a performance parameter that we call *y*. For each experiment we obtain a certain value of the performance, so that at the end of all the experiments we will obtain a performance vector $\overline{y} = [y_1, y_2, ..., y_n]$, with *n* equal to the number of experiments. The interaction of factors A and B, for instance, can be studied with the factor analysis technique, ref. [15]:

$$AxB = \left(y_{A_1B_1} - y_{A_{-1}B_1}\right) - \left(y_{A_1B_{-1}} - y_{A_{-1}B_{-1}}\right) = \left(y_{A_1B_1} - y_{A_{-1}B_{-1}}\right) - \left(y_{A_1B_{-1}} - y_{A_{-1}B_1}\right) \quad [1]$$

Looking at Table 1, we see that when factor *C* is at level *-1* the combinations of A and B are $A_{-1}B_1$ and A_1B_{-1}, while when the factor *C* is at level *1* the combinations of A and B are $A_{-1}B_{-1}$ and A_1B_1. Thus, it is not possible to distinguish the effect of factor C from the interactions of factor A and factor B; the effect of factor C is confounded with the effect of the interaction AxB. The 2-factor confounding effect can be avoided by not assigning any parameter to the column for which we want to study the interaction, with the drawback of a reduced number of main factor effects that is possible to study.

The subclass of fractional factorial designs called *Resolution IV* requires more design points than a *Resolution III* with the benefit of having no main effects confounded with any other main effect or with any two-factor interaction, but two-factor interactions can be confounded with each other and with higher-order interactions. Last, the *Resolution V* subclass allows for experimentation with no main effect or two-factor interaction confounded with any other main effect or two-factor interaction, although two-factor interactions can be confounded with higher-order interactions, e.g., three-factor interactions (AxBxC).

Using a full factorial design or one of the fractional factorial design subclasses (e.g., orthogonal arrays), depends on many aspects. Low-resolution fractional factorial designs allow for a faster experimentation (lower computational effort), but less information on main effects and interactions can be collected. High resolution designs, or full factorial in the extreme case, are advised when no *a priori* information is available on the model of interest, so that fewer assumptions can be made regarding which interactions are negligible to obtain a unique interpretation of output data.

Orthogonal arrays have been used in the design session presented in this paper according to an incremental approach. Starting from the analysis of factors main effects and few two-factor interactions, more and more design points have been added, if confounding was suspected to be present, increasing the basic design to a *Resolution IV, Resolution V* or even full factorial design, if necessary. The possibility of pursuing an incremental approach is one of the characteristics of orthogonal arrays that motivated us to use them for the design of complex, coupled systems. Initially, few variables at two levels can be used for screening. When more information about the design region of interest and the behaviour of the system became available after the initial screening, a more specific analysis could be performed focusing the attention on the relevant parameters and interactions only. This approach provides the design team with the possibility to gain more insight in the system of interest and is much less computationally intense, if compared to Monte Carlo analyses[9] or evolutionary algorithms[12],[13].

II.2 Higher order factorial designs

Two-level factorial designs allows for the determination of linear factor main effects and linear interaction effects. In other words, the models that can be identified experimenting with factors at only two levels are of the following form:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j \qquad [2]$$

where Y is the response of interest, β_0, β_i and β_{ij} are the coefficients for the factors x_i.

There are many cases in which the curvature of the design space is very important, especially when optima are inside the design space of interest, not on the border. When curvature is present the two-level factorial design does not provide reliable results since only limited curvature is detectable by the two-factor product terms. This is the main motivation that leads us to take techniques into account that are suitable for analyzing and identifying models with curvature. Amongst the higher-order models, the second-order models of the following form are the most widely used, typically for most engineering problems:

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j \qquad [3]$$

The parameter β_0 is the mean response of the system, the parameters β_i are the coefficients for the linear main effects, the parameters β_{ii} are the coefficients for the quadratic effects, and the parameters β_{ij} represent the coefficients for the linear interactions.

A second-order model like the one described by [3] will not represent a reasonable approximation on the whole design space in general, but for relatively small regions the results are usually very satisfactory. There are many designs, which allow fitting a second-order regression model: Central Composite Design (CCD), three-level factorial design, Box-Behnken design, D-optimal design, and so on. The CCD is the most widely used class of factorial designs used for identifying second order models, and we choose the CCD also because of its heritage from the two-level factorial design[27]. This particularly efficient design is formed by a two-level factorial design (full or fractional) for linear effects, plus axial points and a central point for curvature effects. The design reduction using *Resolution III*, *IV*, or *V* designs and the incremental approach for determining priorities and factor importance fully apply to the CCD due to the heritage from the two-level design mentioned before. Indeed, in Fig. 3, we show a full factorial CCD (left) and a CCD obtained using an Orthogonal Array (right).

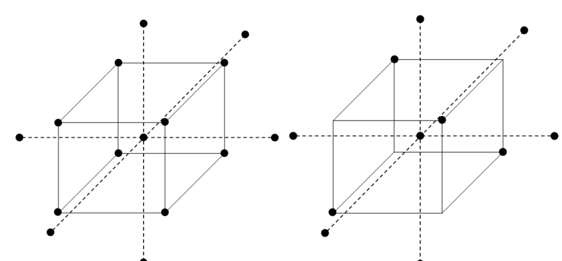

Fig. 3: Central Composite Design (CCD). Full factorial (left), Orthogonal Array (right).

II.3 Analysis of variance

As already mentioned in the introduction of this paper, the main objective of the methodology is to support the design team providing information on the behaviour of the system and its components, i.e., main factor influence, interaction effects, robustness and design sensitivities. To obtain all the required information from the results of the simulations, executed according to the variable settings planned with the orthogonal arrays and the CCD, the ANalysis Of VAriance, ANOVA, and regression analysis are tools we use synergistically for studying the system response.

As the name suggests, ANOVA is a technique used to identify the relative effect of different factors on the overall variability of the performance detected during the simulations. To do so, the overall variance is partitioned in its components determined by the effect of the factors and the included interactions taken into account. The larger the contribution to the overall variance the larger is the effect on the analysed performance. Suppose we obtain a performance vector $\mathbf{y} = [y_1, y_2, ..., y_n]$, with n equal to the number of experiments, as a result from the simulations with the variable settings planned with the CCD. The total variation of the performance *y*, also called the total sum of squares, can be computed as follows:

$$SS_T = \sum_{i=1}^{n} (y_i - \overline{y})^2 \quad [4]$$

with $\overline{y} = \sum_{i=1}^{n} y_i$ being the mean response.

The total sum of squares is computed as a function of the observations of the real model, i.e., the original mathematical model. The best we can do is to identify another model, a regression model, out of the observations of the real model, like the model presented in equation [3]. We would like to obtain the total sum of squares of the regression model as closely as possible to the total sum of squares of the real model, to be as confident as possible that the global variability of the performance detected on the real model is also explained in the regression model. The total sum of squares of the regression model can be expressed as follows:

$$SS_R = \sum_{i=1}^{n} (\hat{y}_i - \overline{y})^2 \quad [5]$$

where the vector $\hat{\mathbf{y}}$ contains the responses computed with the regression model. The difference between the two sums of squares is called error sum of squares:

$$SS_E = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e^2 \quad [6]$$

Eq. [6] indicates how much of the global variability is not explained by the regression model. It already provides information on factors or interactions that are missing from the regression model and that have a certain contribution in determining the global variability of the real model. For quadratic models and limited regions of the design space (limited intervals of the design variables) the error sum of squares is usually zero or very small.

The total variability of the data is not sufficient for a deep understanding of the importance of each single factor of the regression model. The total sum of squares shall be partitioned in its components to gain an insight into the influence of each term of the regression model on the total variability.

In literature three main methodologies for partitioning the total of sum of squares can be identified: the sequential sum of squares decomposition, the classical sum of squares decomposition and the partial sum of squares decomposition[28]. The following terminology is used in the sum of squares decomposition: $SS(A)$ is the sum of squares associated with factor A; $SS(A|B)$ is the sum of squares of the factor A given that factor B is already in the model, i.e., the variability added by factor A to the total variability computed with only factor B in the model.

The classical sum of squares decomposition gives an indication of the change in variability of the data due to adding an extra term to the model, given that all the other terms have been added except for the terms that contain the effect under test. For instance, the sum of squares of factor C, with A and B already in the model, with all the interactions (two and three variables) can be computed as follows:

$$\begin{aligned} SS(C) &= SS(C|A,B,AB) = \\ &SS(A,B,C,AB) - SS(A,B,AB) \end{aligned} \quad [7]$$

It has to be noted that the interactions involving factor C (i.e., AC, BC, ABC) have been left out the analysis, to avoid confounding the partial sum of squares due to factor C with the partial sum of squares due to the interactions involving C. The Classical sum of squares is the favourite method of computing the sum of squares, since it gives a clear indication of the "*snow-ball*" effect of excluding or including a term into the model. This is why we prefer this decomposition over the others mentioned before.

The sum of squares of each term of the model will allow the design team to understand what the main factors that produce a significant variation of the performance are, in the design region of interest. As we will see in the example in section 0, the partial sums of squares are also easier to show irrespectively of the dimensionality of the design space. Notably, if there are more than 2 design variables multi-dimensional representations of the design space will prove difficult to show, whereas bar plots clearly indicating factor importance are much easier to manage with high dimensionality design spaces.

II.4 Regression analysis

In the previous subsection we mentioned several times the process of identifying a model by using the observations at the design points planned with the CCD, see equation [3]. The coefficients β_i and β_{ij} are somehow linked to the partial sums of squares of the design factors (β_0 represents the mean response). Also these coefficients give an idea on the sensitivity of the performance to a change in the design parameters. Indeed, they represent the partial derivatives of the design factors with respect to the performance of interest. The difference is that the partial sum of squares also takes into account the ranges of the design parameters, but neglects the sign of the effect (since it is computed using squared values).

The computation of the coefficients β_0, β_i and β_{ij} allows for the computation of response surfaces and contour plots. For the visualization of the results and presentation to the design team this is a fundamental step in the methodology.

The regression coefficients are computed using a linear least square interpolation. Expressing the model of equation [3] in a matrix notation as shown in equation [8], we obtain a least square estimation $\hat{\boldsymbol{\beta}}$ of the regression coefficients $\boldsymbol{\beta}$ as shown in equation [9].

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{21} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \qquad [8]$$

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X'X}\right)^{-1}\mathbf{X'y} \qquad [9]$$

The matrix **X'X** is often singular and is easiest inverted by, for instance, Singular Value Decomposition or QR decomposition.

Response surfaces are very powerful in presenting the shape of the design space to the design team, for fast modification of the design-variable settings and neighbourhood analysis of the design-point conditions. Further, contour plots allow for a fast and effective boundaries and constraints analysis, even with more objectives, i.e., more contour plots superimposed.

II.5 Summary: the mixed hypercube

Given the mathematical model of the complex system to design, the objective is to select the best combination of the levels of the design variables to optimize the performance(s) while satisfying boundaries and constraints. Typically, continuous and discrete, or architectural, variables are of interest for the analysis. Using Orthogonal Arrays coupled with a fractional factorial CCD to plan the simulations, we collect the performance(s) from the model. ANOVA is used to obtain information on factor and interaction importance, thus the sensitivity of the performance(s) to the variation of the design variables' levels and the robustness of the design space of interest. Regression analysis is used to efficiently show the performance(s) behaviour as a function of the variation of the design parameters, thus to show the design region to the design team with contour plots, boundaries and constraints analysis.

In the title of this subsection we introduce the "*mixed hypercube*" as the name we give to the design methodology presented in this paper. Hypercube, because when the design factors are more than three, the geometrical representation of all the design dimensions (with each one being a design variable) is a hyperspace. When we put boundaries to these dimensions we obtain a hypercube. Mixed is mentioned because we can use both continuous and discrete (nominal or architectural) design variables. The need of separating the design variables in these two classes arises from the fact that the computation of response surfaces is meaningless when discrete variables are involved (there are no intermediate levels that the variables can assume). Thus, for each design combination of architectural variables, a CCD analysis is performed on the continuous variables providing information on multiple objectives, trends and shape of the design region, best settings of the design variables, robustness and sensitivities, as shown in Fig. 4.

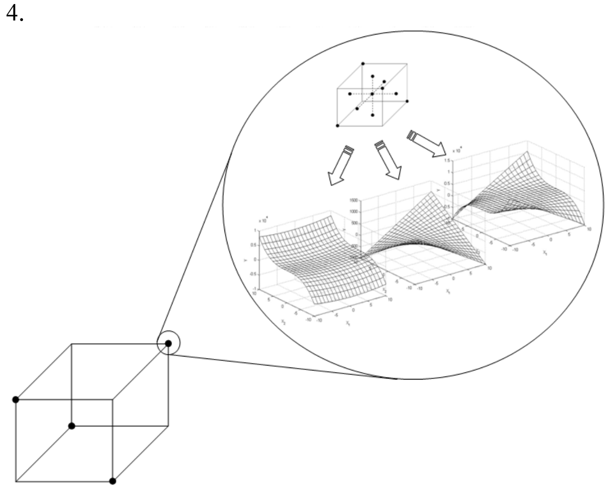

Fig. 4: Hypercube design with multiple objectives and mixed continuous-discrete variables

The lower left hypercube of Fig. 4 represents the design of experiments with architectural variables. For each point of that hypercube another hypercube is build with the CCD experiment design with the continuous variables only.

III. DESIGN CASE

In this section we provide the information needed to capture the rationale behind the results provided in the next section. A hypothetical human mission to the Moon to support a human outpost for a minimum of fifteen years has been taken into account. Mathematical models for a re-entry capsule, a service module, a lander system, an ascent vehicle and an Earth-Moon transfer vehicle have been developed and implemented. In

Table 2 the most relevant design variables that affect the design of the mission are listed and the rationale behind the various possible architectures is described. More details on the mathematical models are provided in ref. [1]. The objective of the design session is to minimize the number of launches to support the lunar outpost for the required time while minimizing the dimensions and the mass of the capsule and the service module. Due to the fact that cost models for such systems are not available in literature, minimizing the number of launches and the mass of the capsule and the service module is reasonably similar to minimizing the cost of the mission as a whole, at least as a first approximation. In the future, a more detailed analysis will be performed to include relationships between mass, technology level and cost in the design process. In Table 3, the requirements and the design variables' levels taken into account as a baseline design are summarized. These settings are similar to those considered in the ESAS document[30] and discussed in ref. [1]. The baseline design represents the first tentative design-variable set used to solve the problem of minimizing the objectives while not violating the constraint. The baseline design is the central point of the hypercube analysis. The baseline design leads to the following performances. The mass of the capsule and the service module is equal to 24 tons and the number of (human) launches to support the lunar base for 15 years is 113.

The values of the design variables and requirements have been assigned as a first guess in order to begin with the design process. In Table 4, the ranges of the design variables are described and the nomenclature used in the graphs shown in the next section is indicated.

Table 2: System of systems model summary

	Rationale
Mission Requirements	
Lunar outpost operative life	The foreseen operative life of the lunar base to be supported by the astronauts. As a baseline we consider 15 years.
Maximum time of the crew in the outpost	The maximum time that each astronaut can spend in the outpost in nominal conditions. As a baseline we consider two months.
# Crewmembers in the outpost	The minimum number of astronauts present in the lunar base at the same time. The baseline is six astronauts.
# Capsule crewmembers	The number of astronauts that the capsule is able to host. This number affects the dimensions and the mass of the capsule, thus the dimensions and mass of the service module and all the other elements of the system of systems architecture.
Capsule	
Crewmembers comfort level	The comfort level is related to the available volume per astronaut within the pressurized compartment of the capsule. It ranges between a minimum of 1(tolerable limit) and a maximum of 4 (optimal), ref. [31]
Mission duration	The time foreseen from the astronauts' departure from Earth to the capsule re-entry on Earth. It affects the amount of provisions to support the astronauts, thus mass and unpressurized volume. In cascade, mass and dimensions of the service module and the other elements are affected. The baseline is 13.5 days.
Sidewall angle	The angle of the capsule's conic structure. It affects the volume (pressurized and unpressurized) of the capsule and it represents an interface variable with the re-entry trajectory module (not implemented for the results presented in this paper) for heat fluxes, decelerations and foot print computation. The baseline is 32.5 deg.
Isp	The specific impulse of the propellant for attitude control and de-orbiting manoeuvres. This propellant characteristic affects the volume, the mass and the complexity of the propulsion system, thus volume and mass of the whole capsule.
Other elements	
Isp	The design of the other elements of the system of systems architecture presented in this paper is tightly coupled with the capsule mass and volume characteristics. Besides other system-specific design parameters, already described in ref. [1], the specific impulse of the propellant of each element plays a fundamental role in the determination of masses and volumes, especially for those elements whose main function is transportation (energy change) like the service module and the transfer module.
Launcher Class	
Mass in LEO	The class of the launcher limits the on-orbit capabilities-per-launch. For this design session, we consider four launcher classes from 25 tons (Ariane5 class) to 150 tons (future hypothesized heavy launcher class).

Table 3: Baseline design. Adapted from ref. [1].

Requirements	
Lunar outpost operative life	15 [years]
Maximum time of the crew in the outpost	0.2 [years]
# Crewmembers in the outpost	6
Design Variables	
Capsule Crewmembers comfort level	2
Capsule Mission duration	13.5 [days]
# Capsule crewmembers	4
Capsule Sidewall angle	32.5 [deg]
Capsule Isp	274 [s]
Service Module Isp	364 [s]
Lander System Isp	435 [s]
Ascent Module Isp	364 [s]
Transfer Vehicle Isp	451 [s]
Launcher Class	25[tons]

Table 4: Design variables taken into account in the design process. Type A: Architectural variable. Type C: Continuous variable.

Design Variables	Factor	Type	Levels Level -1	 Level 0	 Level 1
Launcher Class	A	A	25 [tons]	50 [tons]	75 [tons]
Outpost Operative Life	B	A	10 [years]	15 [years]	20 [years]
# Capsule crewmembers	C	A	3	4	5
Capsule Crewmembers comfort level	D	A	1	2	3
Capsule Isp	E	C	200 [s]	264 [s]	375 [s]
Capsule Sidewall angle	F	C	28 [deg]	32.5 [deg]	37 [deg]
Service Module Isp	G	C	200 [s]	350 [s]	500 [s]
Ascent Module Crewmembers comfort level	H	A	1	2	3

IV. RESULTS AND DISCUSSIONS

The experiments designed with the mixed hypercube approach, considering the architectural and continuous variables indicated in Table 4 have been analysed with the ANOVA. The total number of simulations that have been performed is 145. This allowed a reduction of the computational effort if compared to a full factorial design. Indeed, a full factorial design with 8 design variables would have required 6561 (=3^8) simulations.

The results concerning the factor contribution to the performances and constraint violation, i.e., sensitivity of performances and constraints to the design factors and their interactions, are shown in Fig. 5, Fig. 6, and Fig. 7.

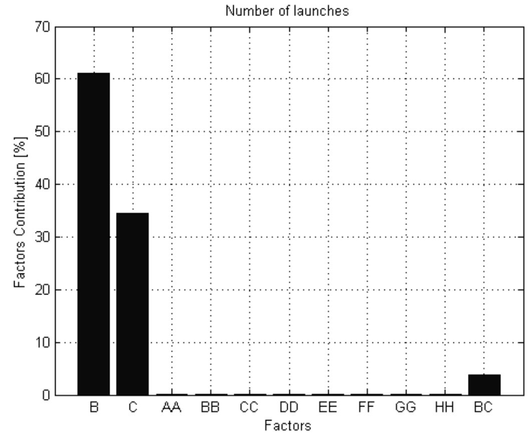

Fig. 5: Factor Contribution to the Number of (human) launches

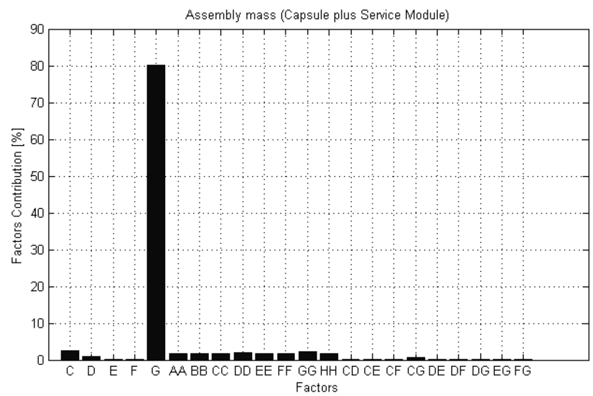

Fig. 6: Factor contribution to the mass of Capsule plus Service module.

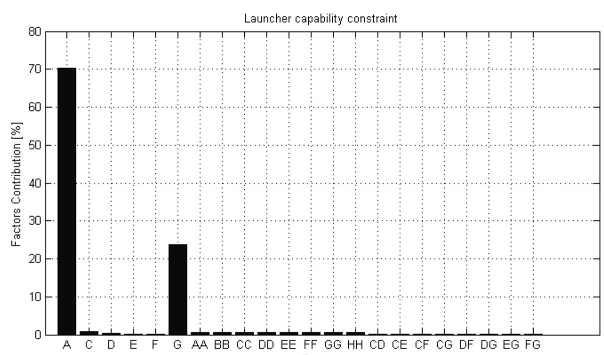

Fig. 7: Factor contribution to the Constraint Violation

In the previous figures only the relevant factors and interactions are shown, when the same letter is repeated two times, it indicates that we refer to a quadratic effect of that factor. As we can see in Fig. 5, the number of (human) launches is most affected by the lunar base operative life, that is a requirement, and by the crewmembers that the capsule can host. Also their interaction contributes to the determination of the performance, but with a reduced importance. This means that the number of launches needed to support the lunar base is very sensitive to the duration of the nominal life of the base itself and quite sensitive to the dimensions of the capsule. Also the launcher class, thus the maximum payload capability of the launcher, does play an important role. Indeed, in Fig. 7 we can read that the launcher class affects the constraint violation up to 70%. The constraint is violated each time that the capsule/service module assembly mass exceeds the launcher payload capability. In Fig. 6, we consider the capsule/service module assembly mass, and we can observe that the design variables directly linked to the design of the capsule contribute the most, together with the specific impulse of the service module propellant. The specific impulse of the capsule does not contribute much to the mass since the ΔV to be delivered by the capsule itself is much lower than the ΔV to be delivered by the service module (in the simulation we used 50 m/s for the capsule and 1700 m/s for the service module[1]). To gain more insight in the behaviour of the system of systems, the information gathered with the ANOVA is used for the regression analysis to plot the response surfaces of the performances and constraints as a function of the most relevant parameters. In Fig. 8 and Fig. 9, the trends of the number of launches needed to support the outpost, as a function of the outpost operative life and the maximum number of crewmembers that can be hosted in the capsule is shown, for different levels of other design variables.

As already mentioned before, the surfaces involving architectural variables do not have physical relevance, since only few of the points on the surface are valid. However, they are very useful to understand the trends of the performances and to provide visual information to the design team.

In Fig. 8, we can read that with the operative life of the outpost that increases and the number of crewmembers hosted in the capsule that decreases, the number of launches increases consequently. This is probably an expected result, given the problem of interest; maybe less expected is the mild interaction between the two parameters identified with the ANOVA (see Fig. 5) and corroborated by the trends of Fig. 8. Indeed, it can be clearly read that the sensitivity to variations of factor B increases as the factor C decreases; this is a clear indication of the presence of an interaction between the two parameters. The influence of the number of crewmembers increases as the outpost operative life increases, i.e., when the total number of astronauts to be landed on the Moon increases.

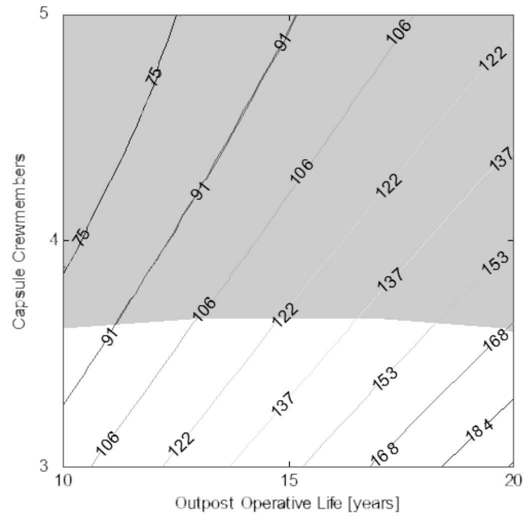

Fig. 8: Number of launches as a function of the outpost operative life and number of capsule crewmembers. The other parameter levels are as follows A -1, D 1, E -1, F 0, G 0, H 0. Gray area is infeasible.

The grey area represents the infeasible region of the design space. In that region the assembly mass, i.e., capsule mass plus service module mass, exceeds the mass deliverable by the selected launcher.

The fact that the other design parameters do not affect this performance much is also testified by the trends described in Fig. 9. The trends in Fig. 9 have been obtained as a function of factor B and factor C, with the other factors at a level that shall increase the dimension of the feasible region, especially because factor G is increased. The trend of the performance is not significantly different if compared to Fig. 8, and this result was already anticipated from the factor importance analysis performed with the ANOVA and reported in Fig. 5. The most relevant consequence is related to the fact that the launcher class is at 50 tons, thus causing the shift of the design region into an area that is far from the constraint.

In Fig. 10 and Fig. 11, the trend of the capsule and service module assembly mass, expressed in kg, is shown as a function of the number of crewmembers hosted in the capsule and the Service module propellant Isp, with the capsule crewmembers comfort level at 1 and 3 respectively.

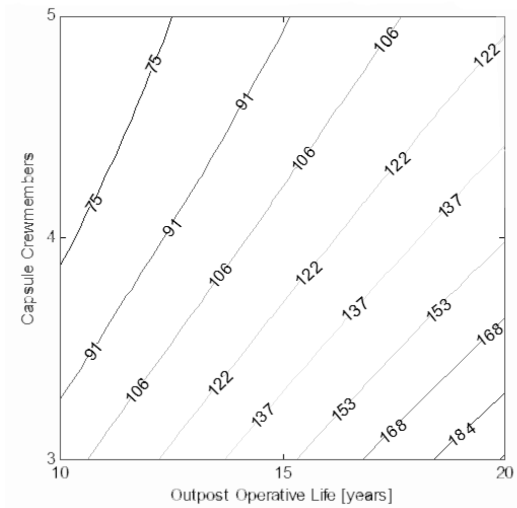

Fig. 9: Number of launches as a function of the outpost operative life and number of capsule crewmembers. The other parameter levels are as follows A 0, D -1, E 1, F 0, G 1, H 0.

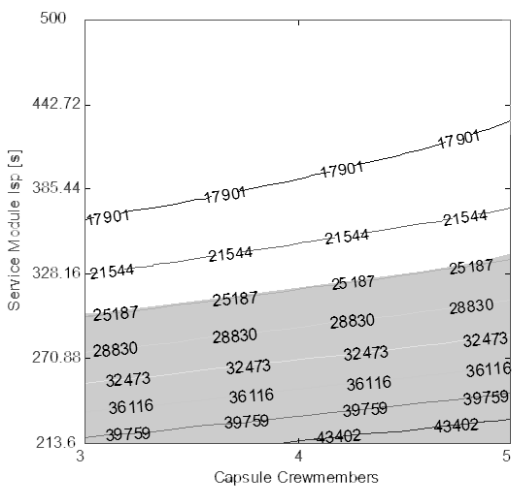

Fig. 10: Capsule and service module assembly mass [kg], as a function of the service module propellant Isp and number of crewmembers hosted in the capsule. The other parameter levels are as follows A -1, B -1, D -1, E 0, F 0, H 0.

The results presented in Fig. 6, helped us in the identification of the design parameters to which this performance is most sensitive to capture most of the variability detected during the simulations. Factor G clearly dominates the performance if compared to the influence of factor C. In particular we read that with the specific impulse that increases and the number of crewmembers hosted in the capsule that decreases, the assembly mass decreases.

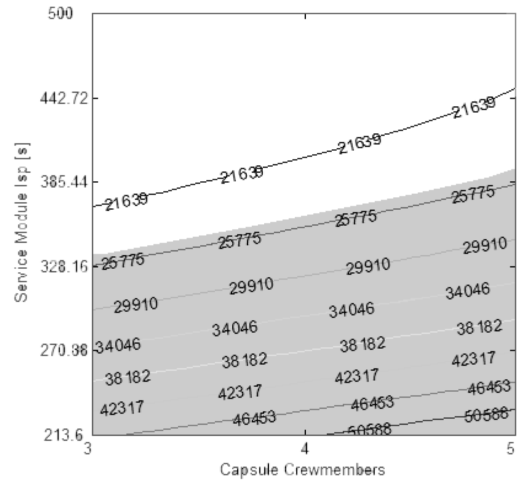

Fig. 11: Capsule and service module assembly mass [kg], as a function of the service module propellant Isp and number of crewmembers hosted in the capsule. The other parameter levels are as follows A -1, B -1, D 1, E 0, F 0, H 0.

The difference between Fig. 10 and Fig. 11 is to be attributed to the variation of factor D. Indeed, in Fig. 6 certain relevance in the determination of the assembly mass had already been discovered. In particular, with the comfort level that increases, the assembly mass increases as well, leading closer and closer to the constraint. With factor G at lowest level, there is no possible way to not violate the constraint by only playing with the number of crewmembers hosted by the capsule, with the other design parameters set as specified.

The specific impulse of the service module propellant is much relevant concerning the determination of the constraint violation, but is not the only one, as can be observed in Fig. 7. Indeed, in Fig. 12, the constraint is plotted as a function of launcher class and service module propellant Isp. As we can read, the only possible approach to obtain the feasibility of the mission, not violating the constraint, while taking into account a hypothetical Ariane5 launcher class, with 25 tons of available mass, is to increase the propellant Isp. In Fig. 12 we also read that with the launcher mass availability that increases the sensitivity of the constraint to the propellant Isp decreases consequently, thus providing more flexibility concerning this particular design choice. Of course, the launcher class is almost never an option, it is rather a given reality, but this type of analysis showed that using the design methodology proposed in this paper, the design team has the power of identifying the impact of the requirements on the performances, in such a way to adjust the free parameters to meet requirements and performances.

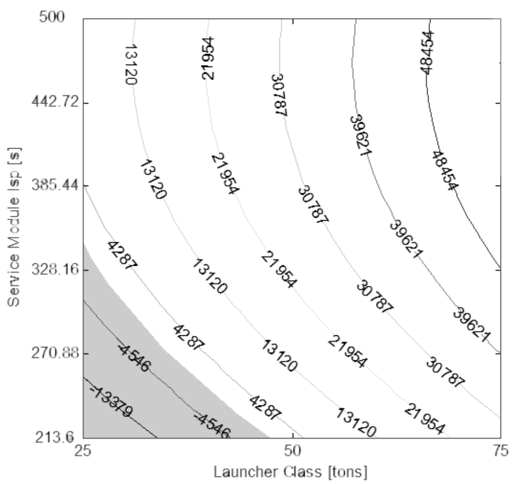

Fig. 12: Constraint as a function of the launcher class and service module propellant Isp. The other parameter levels are as follows B 0, C 0, D 0, E 0, F 0, H 0.

Further, as showed in the example in which we considered the outpost operative life as a design variable, see Fig. 8 and Fig. 9, this methodology can be used as a tool to bargain on requirements. With considerations on affordability, reliability, flexibility, and cost in mind (not taken into account in the mathematical model, but clear in the design team background) requirements can be adjusted in an informed way, considering their effect on the system of systems as a whole.

The combinations of design variables to be plotted in pairs to study the objectives and constraints are much more than those reported in the contour plots shown in this section, and they increase when the number of design parameters increases. The ANOVA, allowed us to reduce the number of plots actually needed to plot all the design variables pairs to a few graphs, and still retaining most of the variability experienced during the simulations.

V.1 Design session close-up

To close the design session, we collect all the information gained by analysing the ANOVA and the response surfaces graphs to select the best combination of design-variable levels to determine the baseline for successive design iterations, maybe at a deeper level of detail.

Suppose that the launcher class is selected as baseline. It seems to be the most reasonable choice, especially because we discovered that there can be feasible architectures choosing this launcher class. Unless programmatic and cost analyses are performed, given the information we derived from the mathematical models there is no particular motivation of applying for a requirement change request, thus we cope with the value of 15 years. Concerning the dimensions of the capsule it seems advantageous to strive for a larger capsule. The effect on the reduction of launches is larger than the effect on the increase in mass, and the mass constraint can be met anyway. The crewmembers comfort level depends on the choice of the service module propellant Isp. Choosing an Isp of 500 s will bring the design point to be far enough from the constraint, thus allowing the comfort level to be high. With an intermediate level of Isp, that will cause the mass of the propellant to increase (but probably the cost to decrease), a high comfort level can still be chosen but with the warning of putting the design point close to the constraint. On the other hand, the combination of factor G at the intermediate level and factor C at the high level, with a high capsule comfort level would lead to unfeasibility anyway, see Fig. 8. Thus, it seems wiser to set the service module propellant Isp at the highest level. Concerning the Capsule Isp and sidewall angle, we did not experience any major impact on the objectives and the constraints. The selection of the levels of these parameters would need to be taken considering other issues not included in the current version of the mathematical models, like cost and re-entry conditions and requirements, for instance, thus for the time being we could cope with the values of the baseline. The ascent module crewmembers comfort level can be set equal to the baseline value for the same reason.

With these settings of the design variables, the design point of the system of systems shows to be placed in an interesting position on the objective space, with 90 launches and an assembly mass of almost 22 tons, see Fig. 13.

In Fig. 13, all the design points computed with the mixed hypercube method are shown. As we can see, the ANOVA and response surface analyses, with constraint and sensitivity analyses allowed us to improve the baseline design. Indeed, the selected design point dominates the baseline design point, i.e., it is better considering all the objectives at the same time. Further, there are not design points on the graph that are better than the one selected, besides those obtained with a relaxed requirement on the lifetime of the lunar outpost.

In Fig. 14, the non feasible solutions obtained during the design session are shown.

The design points computed with the mixed hypercube approach strongly depend on the initial design point chosen as baseline. If the baseline point would have been far away from the actual one, the solution we found would probably not have been discovered, especially with reduced dimensions of the hypercube. This is the main drawback of working with a local design methodology; it provides information in the region within the hypercube only, thus not suitable to find global optimum solutions.

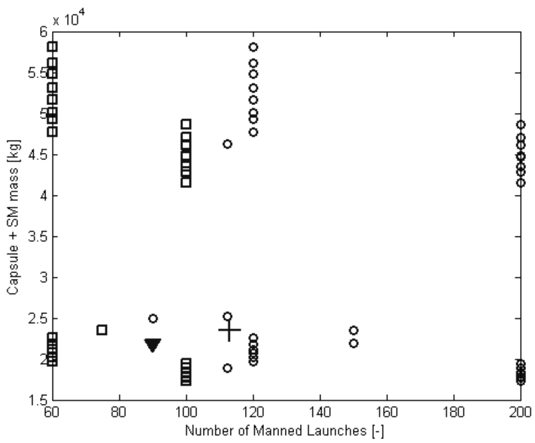

Fig. 13: Objective space, feasible solutions

(☐ Feasible solutions with 10years outpost operative life, ○ Feasible solutions,▼ Selected design point, ┼ Initial baseline design point)

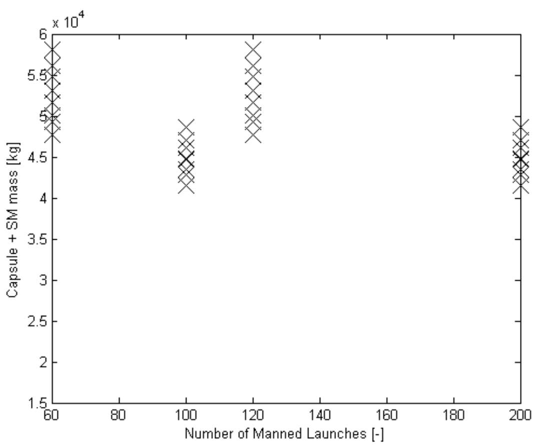

Fig. 14: Objective space, infeasible solutions

Concluding the design session, we can say that the most relevant results of the analysis are two. The first is that the launch of a capsule and a service module to support a human mission is still possible with currently available launcher classes (even though some changes will have to be done to be able to host humans). The second is that, for successive design phases, the design team must be very careful in handling the service module Isp, since it is the main driver concerning performances and constraint.

V. CONCLUSIONS AND RECOMMENDATIONS

In the present paper a design methodology for complex systems has been described and applied to design a human mission to support a base on the moon.

The mixed hypercube approach, based on the utilization of orthogonal arrays, enabled a reduction of the computational effort in sampling the mathematical models for successive data analysis and identification. The advantages of using orthogonal arrays to plan the experiments are that the results coming from the experiments are valid over the entire experimental region spanned by the control factors. Further, we consider the characteristic of the deterministic planning of the experiments very important; it translates in traceability and repeatability of the results. Continuous or discrete variables may be used in the methodology without major complication; particularly useful when dealing with architectural configurations of a complex system.

The analysis of variance technique allowed us to perform a sensitivity analysis and a factor-importance analysis over the entire design region within the hypercube. The response-surface analysis based on the results coming from ANOVA resulted faster and more effective since only the most relevant parameters were taken into account.

The visualization of the results in the form of bar plots for factor sensitivities, contour plots with constraints for the objectives trends as a function of the design variables is crucial for a methodology to support the design team to make informed decisions. "*What-if*" scenarios can be easily analysed, so that fast decisions on heterogeneous systems and architectures can be made in a reasonable amount of time with all the needed information available. The customer can be actively involved in the design process, since requirements can be treated as design variables with all the advantages mentioned before.

The mixed hypercube approach allowed us to determine a design point that is better than the baseline but it is maybe not the optimal one, also because not all the design variables have been taken into account for this particular example. However, the results provide indications on the region of the design space in which to invest most of the effort for the search of the optimum solution.

The methodology is currently being coupled with global optimization techniques in such a way to obtain global optima on top of which local refinement, factor importance and response surfaces analysis can be performed as described in the paper.

VI. ACKOWLEDGMENTS

The research is performed in cooperation with Thales Alenia Space Italy, in the framework a regional research programme called STEPS. The authors wish to thank Giorgio Luigi Ferrari, from Thales Alenia Space Italy, and Sergio Chiesa, Politecnico di Torino, for the support in the development of the case study analyzed in this paper.

VII. REFERENCES

[1] Ridolfi, G., Cardile D., Mooij, E., Chiesa, S., and Ferrari, G., "*A System of Systems Approach for the Concurrent*

Design of Space Missions". IAC 2010, Prague. (to be published)

[2] Ridolfi, G., Mooij, E., and Chiesa, S., "*A Modeling Framework for the Concurrent Design of Complex Space Systems*". AIAA MST conference proceedings. August 2010, Toronto.

[3] Ridolfi, G., Mooij, E., and Corpino, S., "*A System Engineering Tool for the Design of Satellite Subsystems*". AIAA MST conference proceedings. August 2009, Chicago. AIAA 2009 – 6037

[4] Sobieszczanski-Sobieski, J., and Haftka, R.T., "*Multidisciplinary Aerospace Design Optimization: Survey of Recent Developments*".

[5] Sobieszczanski-Sobieski, J., *Multidisciplinary Optimization for Engineering Systems: Achievements and Potential*, NASA TM 101566, 1989

[6] Sobieszczanski-Sobieski, J., "*Optimization by decomposition: a step from hierarchic to non hierarchic systems*". NASA CP 3031

[7] Alexandrov, N.M., *Multidisciplinary Optimization: State of the Art,* Proceedings of the ICASE/NASA Langely Workshop on Multidisciplinary Design Optimization, March 1995

[8] Alexandrov, N.M., and Kodiyalam, S., "*Initial results of an MDO method evaluation study*". AIAA-1998-4884

[9] Mistree, F., Lautenshlager, U., Erikstad, S.O., and Allen, J.K., *Simulation Reduction Using the Taguchi Method*, NASA Contractor Report CR-93-4542.

[10] Lautenshlager, U., Erikstad, S.O., Allen, J.K., and Mistree, F., *Experimental Satellite Trajectory Analysis Using Decision-Based Robust Design*. Journal of Guidance, Control, and Dynamics. Vol. 18, No 5, Sept.-Oct. 1995.

[11] Erikstad, S.O., Lautenshlager, U., Bras, B., and Allen, J.K., Mistree, F. *Integrating Robustness into Multiovbjective Space Vehicle Design Process*. Journal of Guidance, Control, and Dynamics. Vol. 18, No 5, Sept.-Oct. 1995.

[12] Coello Coello, C. A., and Van Veldhuizen, D. A. and Lamont, G. B. "*Evolutionary Algorithms for Solving Multi-Objective Problems*". Kluwer Academic Publishers. 2002.

[13] Kennedy, J., and Eberhart, R. C. "*Swarm Intelligence*". Morgan Kaufman Publisher. 2001.

[14] Chemla, G., Jézéquel, L., and Troclet, B. *Robust Design Optimization by First Design M*ethod, Journal of Spacecraft and Rockets, Vol. 46, No.4, July-August 2009

[15] Phadke, M.S., *Quality Engineering Using Robust Design*, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[16] Taguchi,G., *Systems of Experimental Design*, Kraus International Publications, New York, 1987

[17] Sobieszczanski-Sobieski, J., Hongman, K., Mullins, J., Ragon, S., and Soremekun, G. "*A parallel trade study Architecture for Design Optimization of Complex Systems*". AIAA ASME, April 2005, Texas. AIAA 2005 – 2201

[18] Sobieszczanski-Sobieski, J., Bairstow, B.K., and de Weck, O. "*System-level Optimization and Concurrent Engineering: A comparative Experiment*". AIAA ASME, April 2007, -Hawaii. AIAA 2007 – 1879

[19] Sobieszczanski-Sobieski, J., Altus, D.T., Phillips M., and Sandusky, R., "*Bi-Level System Synthesis (BLISS) for concurrent and distributed processing*". AIAA paper AIAA-2002-5409

[20] Sobieszczanski-Sobieski, J. *Integrated System-of-Systems Synthesis,* AIAA Journal, Vol. 6, No. 5, May 2008

[21] Tedford, N.P., Martins, J.R., *On the common Structure of MDO problems: A comparison of architectures*, AIAA 2006-7080

[22] Yi, S.I., Shin, J.K., and Park, G.J. *Comparison of MDO methods with mathematical examples*, Structural and Multidisciplinary Optimization 2008 35: 391-402. DOI 10.1007/s00158-007-0150-2

[23] Cramer, E.J., Dennis Jr., J.E., Frank, P.D., Lewis, R.M., and Shubin, G.R., *Problem Formulation for Multidisciplinary Optimization*, CRPC-TR93334

[24] Olds, J.R. *Multidisciplinary Design Techniques Applied to Conceptual Aerospace Vehicle Design*, 1993, NASA Technical Report, NASA-CR-104409.

[25] Kodiyalam, S. "Evaluation *of Methods for Multidisciplinary Design Optimization (MDO),* Part I". NASA / CR-2000-210313

[26] Kodiyalam, S., and Yuan, C. "*Evaluation of Methods for Multidisciplinary Design Optimization (MDO),* Part II". NASA / CR-2000-210313

[27] Montgomery, D.C., "*Design and Analysis of Experiments*", John Wiley and Sons, New York, 2001. ISBN 0-471-31649-0.

[28] Draper, N.R., and Smith, H., "*Applied Regression Analysis*". John Wiley & Sons, New York 1998.

[29] Khuri, A.I., and Cornell, J.A., "Response Surfaces. Design and Analysis", Second Edition. Marcel Dekker Inc., New York, 1996. ISBN 0-8247-9741-8

[30] Stanley, D., Cook., S., and Connolly, J., *"NASA's Exploration Systems Architecture Study"*. NASA, TM 214062, 2005

[31] Anonymous. *"Man-System Integration Standards"*. Nasa, STD 3000, 1995