

EMPIRICAL MODEL BUILDING FOR THE OPTIMUM DESIGN OF ROTATING MACHINERY

Sergio Butkewitsch
EMBRAER – Empresa Brasileira de Aeronáutica S.A.
Av. Brigadeiro Faria Lima, 2170, 12227-901 – Sâo José dos Campos – SP – Brazil
sergio.choze@embraer.com.br

Elaine G. Assis
Uniminas – Pça. Prof. Duarte, 99, 38400-156 – Uberlândia – MG - Brazil
elainega@ netsite.com.br

Valder Steffen, Jr.
School of Mechanical Engineering – Federal University of Uberlândia
Campus Santa Mônica, Bloco 1M, 38400-089 – Uberlândia – MG - Brazil
vsteffen@mecanica.ufu.br

Abstract. *This work shows the application of statistical modeling techniques to the optimum design of rotating machinery. The fundamentals of rotor dynamics are reviewed in order to present the physical quantities involved in the statement of an optimization problem aiming at improving the configuration of a rotor-bearing system, whose behavior is evaluated in terms of its critical speeds and strain energy. Objective and constraint functions belonging to this optimization procedure are represented by means of response surface meta-models, whose basic concepts are also briefly reviewed. The optimum values of the design variables (bearing stiffness/damping coefficients and inertia values/positions) are determined by several different optimization methods, spanning gradient based and heuristic strategies. The results obtained are analyzed, leading to the conclusions and some suggestions for future research work.*

Keywords. Rotor dynamics, response surface modeling, heuristic optimization

1. Introduction

Opposite to general models such as differential equations and the like, empirical models are intended to represent the specific behaviour of a given system with respect to a pre-defined set of parameters. These models are constructed from the data gathered by observing the system of interest. So, as far as data provided by numerical procedures and/or physical experiments are processed and used to create symbolic models of a physical reality, empirical modeling (also called meta-modeling) techniques are being applied. Generally, meta-modeling techniques are developed in four steps:

(1) Experimental design: a design space, including a range of design possibilities, is sampled in order to reveal its contents and tendencies. This can be understood as a structured observation of the phenomenon to be modeled;

(2) Choice of a model: the nature of the empirical model itself is determined, tacking into account that the relations contained in the data gathered in the previous step have to be mathematically represented, with the highest possible accuracy;

(3) Model fitting: the model whose shape is defined in (2) is fitted to the data collected in (1). Differences in fitting schemes may affect the efficacy of meta-modeling techniques in the solution of a given problem;

(4) Verification of model accuracy: the three precedent steps are sufficient to build a first tentative model, whose overall quality and usefulness have to be evaluated by adequate sets of metrics and tests.

Each combination of design space sampling (1), model choice (2) and fitting procedure (3) leads to the use of specific verification procedures (4). A general overview of combination possibilities spanning the four major steps of empirical model building is resented in Fig (1), adapted from Simpson et all (1997).

Figure (1) also mentions some meta-modeling approaches, which are very popular in a variety of applications. Response Surface Methods (RSM) are global analytical meta-models. This means they are intended to represent physical relationships found in a design space by means of a unique closed form equation whose coefficients have to be estimated through statistical (least squares regression) techniques. The analytical form is a considerable advantage of RSM over other types of meta-models in terms of physical insight and ease of use, but its global nature can be a handicap in the case of highly non-linear design spaces.

If one searches for more symbolic/abstract empirical models, Bayesian or "krigging" are of a kind that no longer offer analytical representation of the functional relationships pertaining to the design space. When compared to RSM, they are more difficult to implement and costly to run, but can cope better with non-linear design spaces due to their inherent structure intended to model local behaviour along design spaces.

Increasing the level of abstraction, highly symbolic, heuristic models such as neural networks operate with transformation matrices that lead to the estimate of an output, given the corresponding input. Neural networks, in particular, exhibit a high degree of robustness (Rao et all, 1995; Bishop, 1996) with respect to eventual noise collected during the "Experimental Design" phase.

From the brief comparison outlined in the latter paragraphs, it can be stated that each of the different kinds of empirical models have its own advantages and drawbacks, and the choice for one of them will depend upon the particular problem to be solved and the resources available for the solution. On the other hand, all meta-modeling techniques, regardless of abstraction level, offer two distinguished positive characteristics for simulation and optimization purposes:

- *Low computational cost:* if the meta-model is a response surface, a low order polynomial equation has to be solved for a set of inputs. For the case of neural networks, a matrix multiplication operation has to be performed. Once they are constructed, meta-models become more and more inexpensive to use in long term basis;

- *Superior numerical conditioning:* this is a key characteristic in many fields of engineering. For example, if one intends to optimize a structure subject to impact loadings, it is virtually impossible to directly couple a numerical optimizer with a finite element solver due to the highly non-linear nature of the analysis. Instead, a response surface based on analysis results can be easily optimized (Yang et all, 1999). With the availability of low cost and well conditioned predictive tools, sophisticated design approaches can be adopted;

In this paper the authors intend to take advantage of these features in order to obtain the optimal design of a rotating machine, as described in sections 4 and 5. Details of the optimization algorithms (gradient based and heuristic search methods) used for achieving this goal appear in section 3. Response surface empirical models of the optimality metrics are developed, and the data needed for their generation are obtained through a finite element model of the rotor, according to the mathematical foundations presented at section 2.

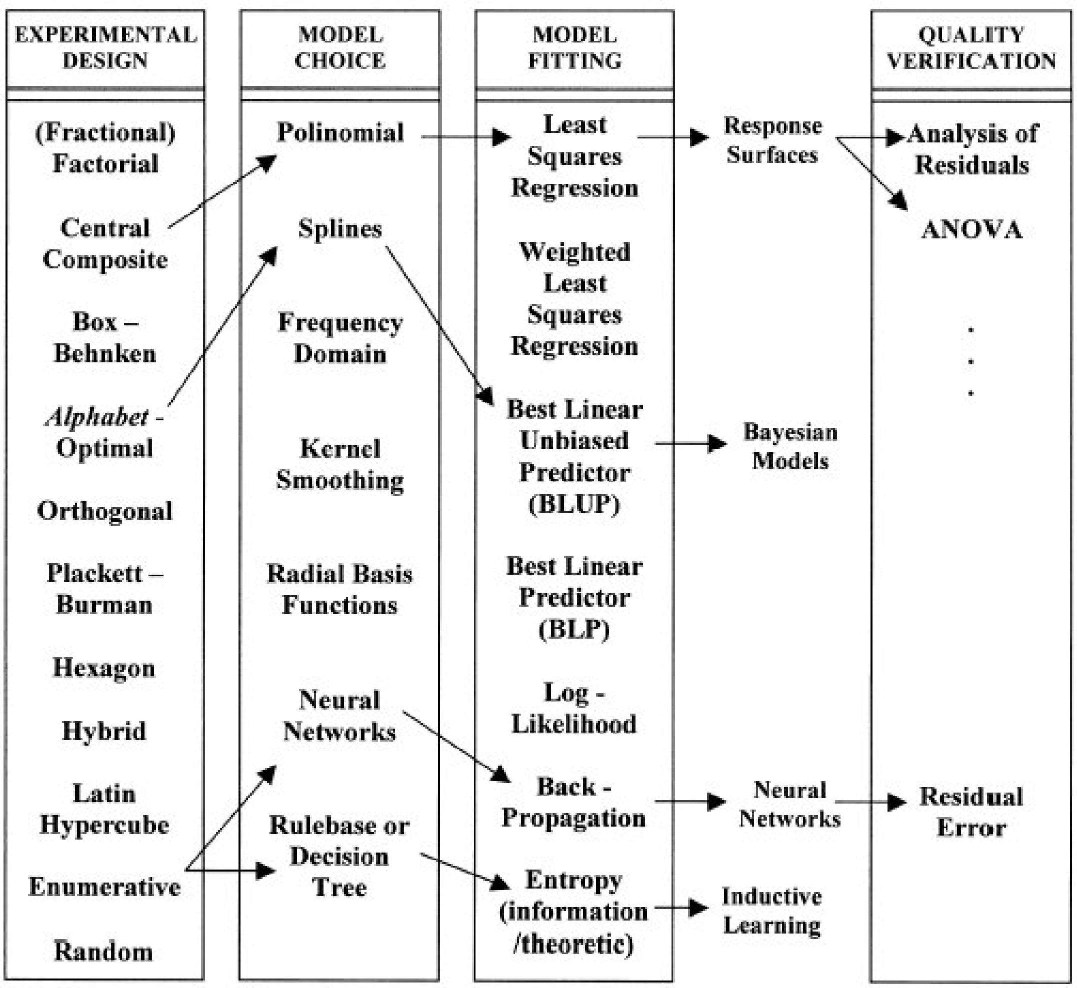

Figure 1. Empirical modeling schemes.

2. Mathematical model for rotating machinery

The rotor model is obtained by using the Finite Element Technique and is composed of three basic elements that are: disk, shaft and bearings. Figure (2) shows the inertial frame (X,Y,Z) and the frame (x,y,z) that is fixed to the disk.

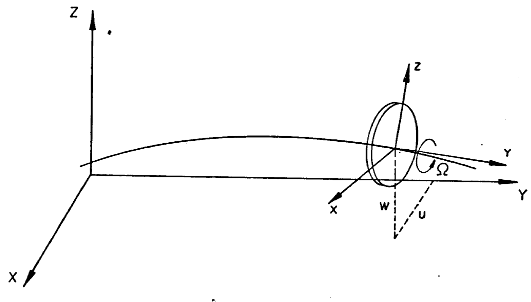

Figure 2. Rotor reference frames.

To obtain the equations of motion Lagrange's equations are used and for that purpose the kinetic energy, strain energy and the virtual work of the external forces have to be calculated. Damping is included in the bearing model. Besides, it is possible to include modal damping in the equations by analogy with a single-degree-of-freedom system. According to Lalanne and Ferraris (1998) the system equation of motion is written as:

$$M\ddot{q} + C\dot{q} + Kq = F_1 + F_2 \sin \Omega t + F_3 \cos \Omega t + F_4 \sin a\Omega t + F_5 \cos a\Omega t \tag{1}$$

where q = N order generalized coordinate displacement vector; K = stiffness matrix which takes into account the symmetric matrices of the beam and nonsymmetrical matrices of the bearings; C = matrix consisting of skewsymmetrical matrices due to gyroscopic effects and nonsymmetrical matrices due to bearing viscous damping; F_1 = constant body force such as gravity; F_2, F_3 = forces due to unbalance; F_4, F_5 = forces due to nonsynchronous effects; and a = coefficient.

The modal base of the associated non-gyroscopic system is used to reduce the number of degrees-of-freedom in order to calculate the system eigenvalues and eigenvectors and to obtain the response to force excitations. Natural frequencies and critical speeds are obtained from Eq. (1), which is rewritten for the homogeneous case. In the present contribution only unbalance forces are taken into account.

For optimization purposes, dynamic characteristics of the rotor-bearing system are considered, namely: critical speeds, unbalance responses and strain energy. Strain energy is considered as given by:

$$U = 0.5\{\delta\}^T K\{\delta\} \tag{2}$$

3. Optimization overview

Numerical optimization techniques have been widely used in general design of mechanical systems. More specifically, in the field of gyroscopic mechanical systems dynamics, automatic design modifications have led to many applications: resonance and critical speed avoidance, vibration level reduction, model updating and the like.

Such methods take advantage of computer automation capabilities through a set of mathematical methods. The standard mathematical formulation of the optimization problem is as follows (Vanderplaats, 1998):

$$\max, \min\left[F(\{X\})\right] \tag{3}$$

that is, find the best possible (minimum or maximum) value of a function that represents a performance criterion, subject to

$$G_j(\{X\}) \leq 0 \tag{4}$$

standing for a set of threshold values to j aspects of system performance

$$H_k(\{X\}) = 0 \tag{5}$$

that is, a set of target values to k aspects of system performance, and

$$\{X\}^{LB} \leq \{X\} \leq \{X\}^{UB} \tag{6}$$

which are bounds to the values of the elements contained by the vector {X}. These elements are called design or decision variables (whose initial values are denoted as X 0), and all the functions (F,G and H) involved in the optimization problem depend upon these variables.

3.1. Gradient based optimzation methods

These are the most traditional and widely used design optimization methods, due to their reliability and efficiency in a wide range of engineering applications. Three fundamental steps are usually necessary to their implementation:

- *Definition of the search direction* – This procedure is the optimization algorithm itself. Gradients (∇) of the objective function (in the sequential methods) and both objective and constraint functions (in the direct methods) are manipulated in order to establish search directions along the design space.

- *Definition of the step in the search direction* – Once a search direction {S} is defined in the previous step, the general optimization problem is restricted to a one – dimensional search. The quantity α in Eq. (7) is the size of the optimizer's move along the search direction in order to update the design configuration from $\{X\}i$ to $\{X\}i+1$.

$$\{X\}^{i+1} = \{X\}^{i} + \alpha \cdot \vec{S} \tag{7}$$

- *Convergence verification* – Convergence is achieved for the design variable set $\{X^*\}$ upon the satisfaction of the Kuhn – Tucker conditions, expressed by Eqs. (8) to (10):

$$G_j(\{X^*\}) \leq 0 \tag{8}$$

$$\lambda_j \cdot G_j(\{X^*\}) = 0 \tag{9}$$

$$\nabla F(\{X^*\}) + \sum_j \lambda_j \cdot \nabla G_j(\{X^*\}) = 0 \tag{10}$$

where λj are the (non-negative) Lagrange multipliers.

3.2. Heuristic optimization methods

Also known as "random" and "intelligent" optimization strategies, this group of optimization methods varies the design parameters according to probabilistic rules. It is common to resort to random decisions in optimization whenever deterministic rules fail to achieve the expected success.

On the other hand, however, heuristic techniques tend to be more costly, sometimes to the point that certain applications are not feasible unless alternative formulations, designed to spare computational resources, are introduced. Such formulations comprise the response surface meta-modeling method that is used in this paper to represent system responses to be optimized by heuristic methods, as well as the deterministic (gradient based) ones.

3.2.1. Genetic algorithms

Genetic Algorithms are random search techniques based on Darwin's "survival of the fittest" theories, as presented by Goldberg (1989). Genetic algorithms were originated with a binary representation of the parameters and have been used to solve a variety of discrete optimization problems. A basic feature of the method is that an initial population evolves over generations to produce new and hopefully better designs. The elements (or designs) of the initial population are randomly or heuristically generated.

A basic genetic algorithm uses four main operators, namely *evaluation*, *selection*, *crossover* and *mutation* (Michalewicz, 1996), which are briefly described in the following:

- *Evaluation* – the genetic algorithms require information about the fitness of each population member (fitness corresponds to the objective function in the classical optimization techniques). The fitness measures the adaptation grade of the individual. An individual is understood as a set of design variables. No gradient or auxiliary information is required, only the fitness function is needed.

- *Selection* - the operation of choosing members of the current generation to produce the prodigy of the next one. Better designs, viewed from the fitness function, are more likely to be chosen as parents.

- *Crossover* – the process in which the design information is transferred from the parents to the prodigy. The results are new individuals created from existing ones, enabling new parts of the solution space to be explored. This way, two new individuals are produced from two existing ones.

- *Mutation* – a low probability random operation used to perturb the design represented by the prodigy. It alters one individual to produce a single new solution that is copied to the next generation of the population to maintain population diversity.

3.2.2. Simulated annealing

Annealing is a term from metallurgy used to describe a process in which a metal is heated to a high temperature, inducing strong perturbations to its atoms positions. Providing that the temperature drop is slow enough, the metal will eventually stabilize into an orderly structure. Otherwise, an unstable atom structure arises.

Simulated annealing can be performed in design optimization by randomly perturbing the decision variables and keeping track of the best resulting objective value. After many tries, the most successful design is set to be the center about which a new set of perturbations will take place.

In an analogy to the metallurgical annealing process, let each atomic state (design variable configurations) result in an energy level (objective function value) *E*. In each step of the algorithm, the atoms positions are given small random displacements due to the effect of a prescribed temperature *T* (standard deviation of the random number generator). As an effect, the energy level undergoes a change *.E* (variation of the objective function value). If *.E* δ 0, the objective stays the same or is minimized, thus the displacement is accepted and the resulting configuration is adopted as the starting point of the next step. If *E* > 0, on the other hand, the probability that the new configuration is accepted is given by Eq. (11):

$$p(\Delta E) = e^{-\left(E/T\cdot k_b\right)} \tag{11}$$

where k_b is the Boltzman constant, set equal to 1.

Since the probability distribution in Eq. (11) is chosen, the system evolves into a Boltzman distribution. The random numbers *r* are obtained according to a uniform probability density function in the interval (0,1). If $r < P(\Delta E)$ the new configuration is retained. Otherwise, the original configuration is used to start the next step.

The temperature *T* is simply a control parameter in the same units as the objective function. The initial value of *T* is related to the standard deviation of the random number generator, whilst its final value indicates the order of magnitude of the desired accuracy in the location of the optimum point.

Thus, the annealing schedule starts at a high temperature, which is discretely lowered (using a factor $0 < rt < 1$) until the system is "frozen", hopefully at the optimum, even if the design space is multi-modal.

4. Description and development of an illustrative case study

As for demonstrating the application of the techniques outlined in the previous sections, this work develops a case study based on the optimization of the rotor depicted in Fig. (3). Its left hand part is a picture of the rotor prototype (used for the experimental identification of structural parameters, such as described in Oliveira, 1999), while its right counterpart is a dimensional sketch highlighting the most important dynamical elements of the rotor structure.

This rotor system is supposed to operate at an angular speed of 200 rpm. Finite element analysis of an initial proposed design point out that the first and second critical speeds are equal to 223.00 rpm and 246.67 rpm, respectively. Therefore, the following optimization problem can be devised:

Maximize: first and second critical speeds (so that they are safely above the typical operating speed of 200 rpm)
Subject to the constraint: strain energy at the operating speed is minimum

by manipulating design variables listed in Tab. (1). It should be noted that, regarding the design objective and constraint, the optimization problem could have been stated conversely:

Minimize: the strain energy when operating at 200 rpm
Subject to: safe separation between each of the critical speeds and the operating speed

All in all, this means that the task of optimally designing this rotor system is a multi-objective problem. Its optimum solution can be influenced by the formulation, so special measures are to be taken to address its multi-criterion nature.

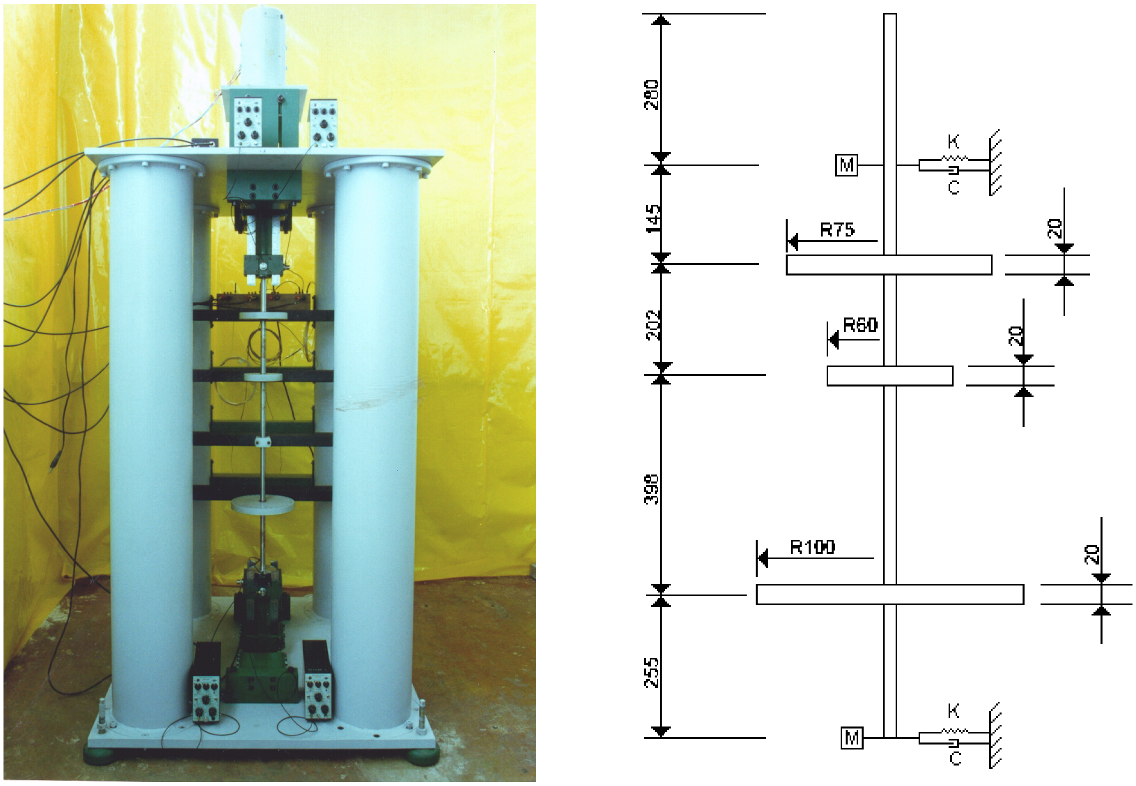

Figure 3. Rotor chosen for the case study development.

Table 1. Design variables descriptions and bounds.

Variable	Description	Lower Bound	Upper Bound
V1	Damping of Bearing 1 [N.s/m]	66.4	83
V2	Damping of Bearing 2 [N.s/m]	66.4	83
V3	Stiffness of Bearing 1 [N/m]	6880	8600
V4	Stiffness of Bearing 2 [N/m]	6880	8600
V5	Diameter of Disc 1 [m]	0.16	0.2
V6	Thickness of Disc 1 [m]	0.0016	0.002
V7	Diameter of Disc 2 [m]	0.096	0.12
V8	Thickness of Disc 2 [m]	0.0016	0.002
V9	Diameter of Disc 3 [m]	0.12	0.15
V10	Thickness of Disc 3 [m]	0.0016	0.002

The first step for solving the optimization problem thus stated is to build empirical models that describe the variation of the key physical quantities (objective and constraints) with respect to the ten design variables of interest. If the response surface method is chosen, objective and constraint functions will be available in low order polynomial closed form, allowing for the use of gradient based and/or heuristic optimization algorithms in an automatic design process.

In the development of empirical models, it is crucial to screen which design variables really influence the responses involved. The importance of this action is twofold: a) For better understanding and judgment regarding the design process and b) For sparing the use of computational/experimental resources aimed at collecting data for the construction of higher fidelity empirical models.

Table (2) is an excerpt of the $2^{10-3}_{R=V}$ two level fractional factorial design implemented to select which design variables are statistically significant with respect to the responses considered. It should be noted that the combination of ten design variables in two levels gives rise to up to 2^{10} = 1024 different configurations. The fraction of $2^{10-3} = 2^{7}$ = 128 combinations strongly reduces the computational effort necessary for the design variable screening procedure, while still having sufficient resolution ("R = V" stands for "Resolution = 5") to allow for the regression of adequate linear models (since each factor is varied in only two levels). Box and Drapper (1987) provide in-depth explanation on how to balance the experimental design effort in view of important aspects such as resolution and model accuracy.

Table 2. Excerpt of $2_{R=V}^{10-3}$ two level fractional factorial design for decision variables screening.

Combination	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
1	99.6	99.6	10320	10320	0.24	0.0024	0.144	0.0024	0.18	0.0024
2	99.6	99.6	10320	10320	0.24	0.0024	0.096	0.0016	0.18	0.0024
.						.				
.						.				
.						.				
127	66.4	66.4	6880	6880	0.16	0.0016	0.144	0.0016	0.18	0.0024
128	66.4	66.4	6880	6880	0.16	0.0016	0.096	0.0024	0.18	0.0024

Table (3) shows the statistically significant design variables for each of the responses of interest.

Table 3. Statistical significance table, at the 5% significance level [(S) – Significant; (NS) – Not Significant].

Response	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
First Critical Speed	NS	NS	S	S	S	S	S	S	S	S
Second Critical Speed	NS	NS	S	S	S	S	S	NS	S	S
Strain Energy	NS	S	NS	S	NS	NS	S	NS	NS	NS

And the linear response surface empirical models resulting from the linear standard least squares regression analysis of this $2_{R=V}^{10-3}$ two level fractional factorial experimental design are given by Eqs. (12) to (14):

$$318.79 + 0.0023 \cdot V3 + 0.0031 \cdot V4 - 200.88 \cdot V5 - 10218.40 \cdot V6 - 161.87 \cdot V7 - 4117.47 \cdot V8 - 177.11 \cdot V9 - 5876.30 \cdot V10 \tag{12}$$

$$307.18 + 0.0037 \cdot V3 + 0.0037 \cdot V4 - 207.45 \cdot V5 - 10336.80 \cdot V6 - 159.51 \cdot V7 - 147.28 \cdot V9 - 5417.08 \cdot V10 \tag{13}$$

$$0.03 - 0.00011 \cdot V2 - 2.5 \cdot 10^{-6} \cdot V5 + 0.023 \cdot V7 + 0.030 \cdot V9 \tag{14}$$

with adjusted squared multiple correlation coefficients equal to 81.22%, 75.88% and 73.90% respectively. Based on these explained variances values, linear empirical models can be assumed as satisfactory representations of the responses of interest for the purpose of optimizing the rotor system of Fig. (3). Although better accuracy could possibly be achieved with higher order response surface equations, this potential gain has to be balanced against two relevant issues: a) the additional computational effort needed to generate higher fidelity models and b) Eventual numerical difficulties for the optimization of models based on non-linear equations.

Particularly with respect to the critical speeds empirical models, the signs of the coefficients present an additional evidence of consistency with the expected physical behaviour. Positive coefficients for the stiffness terms indicate that the critical speeds assume higher values when they are increased. Conversely, critical speeds drop when the inertias augment, as indicated by the negative coefficients associated to disk thickness and diameters. Besides, it should be noted that the grand average value of the second critical speed (307.18) is smaller than that of the first critical speed (318.79). This fact serves as a warning with respect to optimization procedures, since it indicates that within the design space being considered the first and second critical speeds have the potential to overlap or even shift their positions along the angular speed spectrum.

The responses represented by equations (12) to (14) reflect the three criteria that ultimately measure the optimality of any given rotor design. In order to simultaneously account for the three of them, Eq. (15) is defined as a global optimality functional whose minimization ensures the best compromise solution among all of the participating optimality criteria:

$$\Phi(\{X\}) = \sqrt{\sum_k \left[\frac{W_k \cdot \left(F_k - F_k^*\right)}{F_{pk} - F_k^*} \right]^2} \tag{15}$$

where:

- *$\Phi(\{X\})$* is a compromise objective function
- *Fk* is the *k-th* response of interest, in a total of K . K = 3 in the present case, and each of the F functions is replaced by its corresponding response surface model among Eqs. (12) to (14)
- *F*k* is the target value for the *k-th* response

- *Fpk* is the worst value accepted for the *k-th* response
- *Wk* is the weighting factor applied for the *k-th* response of interest

This formulation is well accepted because it considers engineering specifications through *F*k* and *Fpk*, which helps in keeping a practical insight over the optimization problem. It should also be noted that the optimization problem defined through Eq. (15) is unconstrained because the *K* responses encompass both objective and constraint functions. This is particularly useful when heuristic optimization techniques are used because most of the times it is not trivial to implement the handling of explicit constraints when such methods are used.

5. Case study results and discussion

The minimization of Eq. (15) is pursued by means of the three optimization methods outlined in section 3, that is, gradient based (BFGS "Quasi-Newton"), Genetic Search (G.S.) and Simulated Annealing (S.A.). The results yielded by these different approaches are summarized in Tabs. (4), (5) and (6) which contain design variable optimal values, response surface estimates and finite element results, respectively. The relative differences between the data in Tabs. (5) and (6), which ultimately measure the accuracy of the response surface empirical models with respect to the finite element original data, are presented in Tab. (7)

Table 4. Optimal design variables sets.

Optimization Method	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
BFGS	99.6	99.6	10320	10320	0.16	0.0016	0.096	0.0016	0.14044	0.0016
G.S.	84.44	84.23	8601	8599.8	0.1621	0.0016	0.098	0.0023	0.1223	0.0018
S.A.	78.82	80.84	8846	8303.0	0.1619	0.0016	0.1115	0.0016	0.1321	0.0018

Table 5. Design optimization results obtained with gradient based and heuristic methods (response surface predictions).

Optimization Method	First Critical Speed [rpm]	Second Critical Speed [rpm]	Strain Energy at 200 rpm [J]
BFGS	270.02	289.35	9.60E-03
G.S.	259.12	277.47	5.57E-03
Simulated Annealing	257.7247	273.6743	7.31E-03

Table 6. Design optimization results obtained with gradient based and heuristic methods (finite element predictions).

Optimization Method	First Critical Speed [rpm]	Second Critical Speed [rpm]	Strain Energy at 200 rpm [J]
BFGS	263.40	289.80	3.26E-03
G.S.	261.60	273.60	6.82E-03
Simulated Annealing	262.20	277.20	5.83E-03

Table 7. Response surfaces *versus* finite element predictions for optimum design performance (relative differences).

Optimization Method	Relative (%) Differences		
	First Critical Speed	Second Critical Speed	Strain Energy at 200 rpm
BFGS	2.51%	-0.16%	194.48%
G.S.	-0.95%	1.41%	-18.29%
Simulated Annealing	-1.71%	-1.27%	25.44%

Table (7) shows that the empirical models derived for the critical speeds manage to predict them within a very narrow uncertainty range.

On the other hand, the same does not happen with the strain energy at 200 rpm and this calls for a discussion about two aspects. In the first place, the inherent physical nature of the quantity to be meta-modeled has to be considered. With the general quadratic form expressed in Eq. (16), where {X} and [K] are the displacement vector and stiffness matrix respectively, it is unlikely to obtain a reasonable representation for the strain energy by means of a simpler linear model. The same is expected for the critical speeds, but Eqs. (12) and (13) happen to behave properly within the particular range defined by the design variables side constraints. Therefore, the accuracy of the linear approximations for the critical speeds tends to deteriorate if this range is broadened.

$$\{X\}^T \cdot [K] \cdot \{X\} \tag{16}$$

The second issue is related to the lack of homocedasticity and normality within the strain energy data, as illustrated by the residual plot in Fig (4). Since the strain energy values are very small, an adequate scaling could have improved the predictive capabilities of the response surface associated to this physical quantity, acting in this case as a variance stabilizing transformation. Alternatively, the weighted least squares approach (with or without data transformation) could also result in a better empirical model.

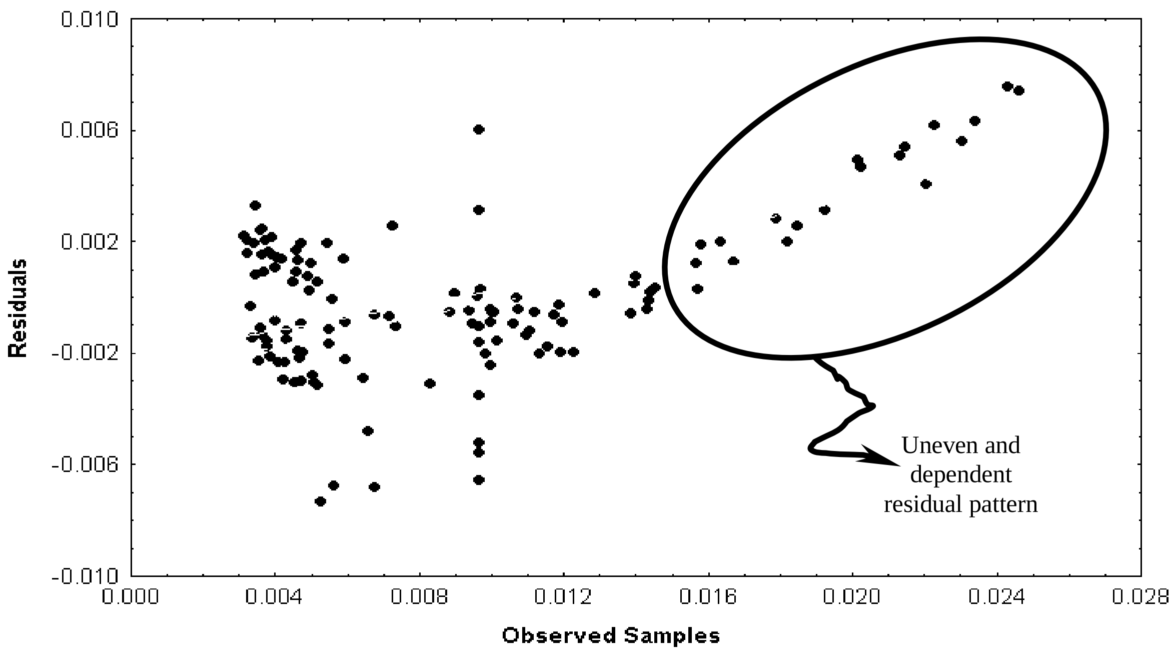

Figure 4. Uneven and dependent residual distribution indicating lack of homocedasticity and normality within the strain energy experimental design data.

Still based on the data contained in Tabs. (4), (5), (6) and (7), the following issues arise:

(1) There is not enough statistical certainty to comment the strain energy evolution during the optimization process;

(2) It is safe to evaluate the critical speeds evolution during the empirical model based optimization procedure;

(3) The comparison of the three algorithms employed during the optimization procedure reveal that the numerical conditioning of the functional given by Eq. (15) is outstanding, since the BFGS method managed to get to the global optimum of the available design space by exhausting the ranges defined between the design variables side constraints. On the other hand, the heuristic methods seem to have been trapped in local optima, which is probably due to inconvenient setup of sensitive parameters such as the initial population size and number of generations (in the case of the genetic search) and temperature (in the case of the simulated annealing). Indeed, response surface empirical modeling enable the removal of several of the difficulties for gradient based optimization methods, reducing or even inverting their eventual handicaps relative to heuristic algorithms.

6. Conclusions and perspectives for future research work

This paper has demonstrated the use of empirical modeling techniques aimed at representing the behaviour of a gyroscopic system so that optimization algorithms can be efficiently applied for design improvement purposes. Empirical modeling allowed for considerable flexibility in defining the multi-criterion optimization problem associated with the automatic design of a rotor system, enabling the optimization of a functional with useful characteristics: inherent good numerical conditioning, implicit constraint handling and possibility to explicitly account for engineering design specifications.

With regard to numerical conditioning and local optima, better results could have been obtained with the heuristic optimization strategies, and it would be useful to spend some effort in trying to define better values to parameters such as initial population size, number of generations and temperature, as well as rules for their choice.

Multidisciplinary Design Optimization (MDO) is also an interesting field opened by the use of empirical modeling. In the case of the rotor system optimized in this paper, one could imagine other design criteria: construction cost, heat transfer issues and others. Of course, these quantities could not be calculated in the same finite element based module used to obtain the critical speeds and the strain energies. However, data related to them could be easily generated in other platforms, and then transformed in response surface equations that could be plugged into Eq. (15) as new criteria to be satisfied in the optimization process.

Another aspect to be further investigated is the robustness of optimal solutions with respect to uncontrollable random perturbations that may occur. To illustrate this subject one can consider the data in Tab. (8), which shows how the response to imbalance at a given node of the finite element model vary when assembly imperfections shift the disk positions of ± 20.00 mm with respect to their "theoretical" locations, considering the optimal rotor configurations operating at 200 rpm.

Table 8. Response to unbalance (m) of rotor optima considering the influence of assembly imperfections (200 rpm).

Optimization Method	"Theoretical" Position	+ 20.00 mm offset	- 20.00 mm offset
BFGS	3.49E-06	3.16E-06	3.79E-06
Genetic Algorithm	4.35E-06	3.98E-06	4.71E-06
Simulated Annealing	4.36E-06	3.98E-06	4.74E-06

The data in Tab. (8) shows that the effect of assembly imperfections is, at least, noticeable (simulations with additional offset values have to be performed in order to verify if the variations are indeed statistically significant). The extent of this influence can be checked by an additional factor to the experimental design (i.e., the offset from the "theoretical" location). If this influence is found to be relevant, it can be controlled through the addition of the corresponding response surface to a multi-criterion formulation such as Eq. (15).

Finally, one of the most critical issues of empirical modeling for the sake of simulation and optimization is the equilibrium between the experimental (numerical or physical) cost needed to construct the models and their predictive accuracy. Vast data sets do not necessarily result in adequate models, and model sophistication does not always result in improved accuracy.

Considerable research effort (Jin et all, 2000) is being devoted to establish the benefits and drawbacks of several different meta-modeling strategies, and the rotor optimization problem developed in this article is suitable for such an investigation, specifically if the following approaches are taken into account:

- Use of higher order response surfaces, as already suggested in section 4;
- Modification of the least squares approach from "standard" to "weighted", due to the frequent occurrence of outliers among the data generated by means of the experimental designs;
- Use of other metrics capable of capturing the multi-criterion essence of this kind of engineering problem;
- Use of various types of empirical models, such as those mentioned in Fig. (1): neural networks, radial basis functions, krigging models, splines and others.

7. References

Bishop, C., 1996, "Neural Networks for Pattern Recognition", Oxford University Press, Oxford, U.K., 504 p.

Box, G. E. P. and Drapper, N. R., 1987, "Empirical Model Building and Response Surfaces", Wiley_Interscience, New York, USA, 688 p.

Goldberg, D., 1989, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, USA, 412 p.

Jin, R., Chen, W. and Simpson, T., 2000, "Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria", Proceedings of the 8th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Long Beach, CA, USA.

Lalanne, M. and Ferraris, G., 1998, "Rotordynamics Prediction in Engineering", John Wiley & Sons, USA, 266 p.

Michalewicz, Z., 1996, "Genetic Algorithms + Data Structures = Evolution Programs", Springer Verlag A.I. Series, New York, USA, 397 p.

Oliveira, A. G., 1999, "Techniques for Excitation Characterization in Rotating Machinery", PhD Dissertation, UFU, Uberlândia, Brazil (in portuguese).

Rao, V. B. And Rao, H. V., 1995, "C++ Neural Networks and Fuzzy Logic", M and T Books, New York, USA, 512 p.

Simpson, T. W., Peplinsky, J. D., Koch, P. N. and Allen, J. K., 1997, "On the Use of Statistics in Design and the Implications for Deterministic Computer Experiments", Proceedings of the ASME Design Engineering Technical Conferences, Sacramento, CA, USA.

Vanderplaats, G. N., 1998, "Numerical Optimization Techniques for Engineering Design", VR&D Inc., Colorado Springs, Colorado, USA, 417 p.

Yang, R. H., Tho, C. H., Wu, C. C., Johnson, D. and Cheng, J., 1999, "A Numerical Study of Crash Optimization", Proceedigns of the ASME Design Engineering Technical Conferences, Las Vegas, NE, USA.