19th Brazilian Congress of Thermal Sciences and Engineering
November 06th-10th, 2022, Bento Gonçalves - RS - Brazil

ENC-2022-0051
LOW-COST IDENTIFICATION OF TWO-FACTOR INTERACTIONS THROUGH FRACTIONAL FACTORIAL METHOD APPLIED TO 1D MODEL OF A S-CO_2 CENTRIFUGAL COMPRESSOR

Elóy Esteves Gasparin
Gabriel Inácio Rodrigues Silva
São Paulo State University (UNESP), Av. Brasil Sul, 56 – Centro, Ilha Solteira – SP, 15385-000, Brazil
eloygasparin@gmail.com
gabriel-inacio.silva@unesp.br

Daniel Dezan
Federal University of ABC, Av. dos Estados, 5001 – Bangu, Santo André - SP, 09210-580, Brazil
daniel.dezan@ufabc.edu.br

Fábio Saltara
Paulo Mello
Jurandir Yanagihara
University of São Paulo, Av. Prof. Luciano Gualberto, 380 – Butantã, São Paulo - SP, 05508-010, Brazil
fsaltara@usp.br
padu.mello@gmail.com
jiy@usp.br

Leandro Oliveira Salviano
São Paulo State University (UNESP), Av. Brasil Sul, 56 – Centro, Ilha Solteira – SP, 15385-000, Brazil
leandro.salviano@unesp.br

Abstract. *The Sensitivity Analysis (SA) of high-dimensional and computational expensive models is able to provide important insights into model behavior. The so-called screening methods are known for their low-cost and good proxy for main effects identification, which is essential for SA of computationally expensive models, such as the Computational Fluid Dynamic (CFD) model of centrifugal compressors. If utilized with factor fixing purposes, it can significantly diminish computational effort for a subsequential optimization procedure. Morris' screening method has been widely applied to engineering problems and provided excellent main effects prediction and robust factor ranking. Moreover, Morris' Design of Experiment (DoE) demonstrated to provide good sample space coverage for Response Surface (RS) training. However, the two-factor interactions are not assessed by Morris' method, since its one-factor-at-a-time nature is not fit for that purpose. Therefore, this work intended to assess if the Fractional Factorial (FF) screening method is able to provide a low-cost first assessment of two-factor interactions on computational expensive engineering models before the factor fixing is performed, which could avoid the fixing of an interacting factor. Thus, an already developed and computationally fast one-dimensional (1D) model of s-CO2 centrifugal compressor was considered herein, allowing the comparison with more expensive and robust SA methods for proper evaluation of this strategy. Axial height (ΔZ), number of blades (Z_{fb}) and blade thickness (t) were considered as inputs. Overall, this combined new strategy was able to properly rank variables influence and its two-factor interactions, which reduces the computational cost of factor fixing SA of high-dimensional models (for 8 input variables this methodology is able to reduce the number of model runs in more than 80%).*

Keywords: *Sensitivity Analysis, Fractional Factorial, Centrifugal Compressor, Supercritical CO2*

1. INTRODUCTION

The Sensitivity Analysis (SA) of large and high-dimensional models is a powerful tool for gaining insight on variables importance and correlations (SALTELLI, 2008), this statistical analysis studies the behavior of input variables toward an interesting output, describing the importance of each input in determining the output variability. There are two main kinds of SA methods: qualitative screening methods and quantitative variance-based methods (CAMPOLONGO *et al.*, 2011). The Elementary Effects (EE) Method, developed by Morris (1991), is considered a good practice in qualitative SA for large and expensive computational models, since it uses only a few model executions to assess a good proxy of variables importance (FENG *et al.*, 2019). This screening method has been refined to identify the non-influential input variables and has been effective (CAMPOLONGO; CARIBONI, 2007). Moreover, its One-At-a-Time (OAT) Design of Experiment (DoE) was improved in order to maximize trajectories distance and better represent the sample space (GE *et al.*, 2014). However, due to its OAT nature, the EE method is unable to properly identify interactions between input variables (SALTELLI *et al.*, 2019).

On the other hand, the quantitative SA methods estimate a sensitivity index that represents a percentage of the output variance that each variable is responsible for, due to its main component or to its interactions with order variables. However, quantitative methods require a high number of model evaluations, becoming unfeasible for computational expensive models. Therefore, the Fractional Factorial (FF) sampling is a promising screening method that arises as an alternative to expensive quantitative SA methods, since it should be able to identify two-factor interactions, maintaining the low-cost assessment of Morris' method (SALTELLI, 2008). This screening method has been used as a preliminary evaluation of factors influence (NAGEEB EL-HELALY *et al.*, 2018) in other areas, as a preliminary assessment of the model. However, it is still underused for the identification of two-factor interactions on engineering problems.

In order to properly evaluate the combination strategy of EE and FF methods, the SS-ANOVA quantitative method is considered (GU, 2002), providing a basis for comparison since its sensitivity measures are more reliable and robust than the screening methods. Moreover, as quantitative methods usually require high numbers of model evaluations, an 1D model of centrifugal compressors (AUNGIER, 2000; MONJE *et al.*, 2014) was utilized, as his computational processing cost is low (seconds per model run), allowing a great number of model evaluations. If this novel screening strategy, combining EE and FF methods, proves to be reliable, it can decrease computational costs for identifying interactions between factors.

2. METHODOLOGY

This section presents the combination strategy of EE and FF screening methods, the SS-ANOVA quantitative method and the 1D model of centrifugal compressor used herein. As presented in Figure 1, the 1D model was parameterized considering three variables: number of blades (Z_{fb}), axial height (ΔZ) and blade thickness (t). The isentropic efficiency of the centrifugal compressor is chosen as interest output, which provides insight into variables that influence the equipment performance. Moreover, both screening sensitivity analyses (EE and FF) are computed in parallel. The EE method seeks to properly rank the input variables importance and the FF method focus on identifying the two-factor interaction influence. In order to evaluate this combination strategy, both method results are taken to a comparison with a more reliable and variance-based method: the SS-ANOVA.

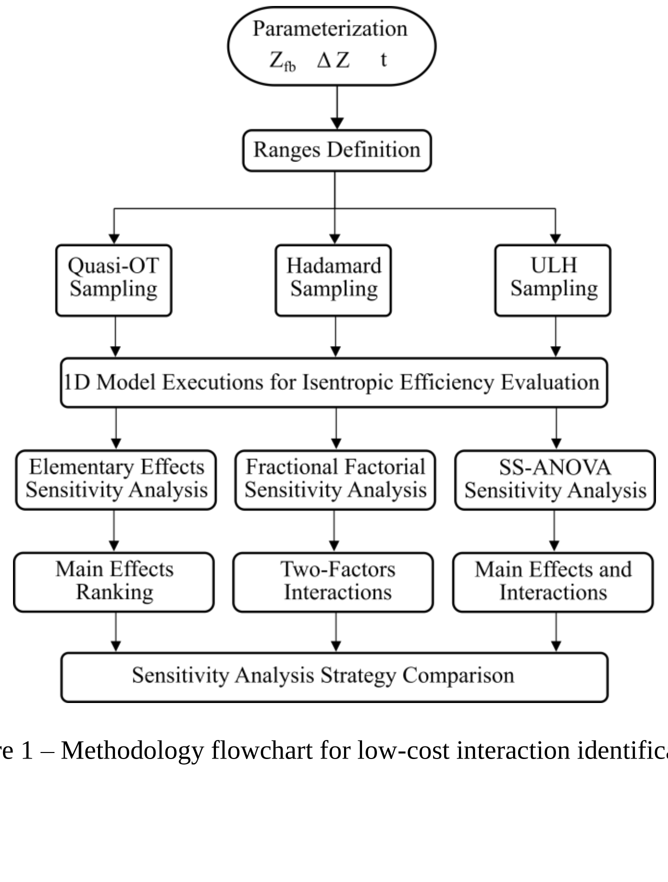

Figure 1 – Methodology flowchart for low-cost interaction identification.

2.1 1D Model for S-CO_2 Centrifugal Compressors

The centrifugal compressor's 1D model used in this study was based on Aungier (2000) and Monje *et al.* (2014). The entire model consists on the evaluation of continuity, energy conservation, velocity triangle and Equation of State (EOS) (R.SPAN; W.WAGNER, 1994) until convergence. The loop of these governing equations is performed on impeller leading and trailing edges, vaneless diffuser and volute, providing a full preliminary assessment of fluid flow behavior in the equipment and avoiding high Mach numbers, condensation and gas-like behavior for supercritical CO_2.

The preliminary geometry obtained by the 1D model is the first step in the centrifugal compressor design, which is followed by a Computational Fluid Dynamics (CFD) model, as an already established strategy for achieving high machine performances (XIA *et al.*, 2021). Herein, the 1D model is chosen due to its fast and cheap executability, necessary for SS-ANOVA to be assessed since it is an expensive SA method. The inlet thermodynamic and operational conditions are summarized on Table 1.

Table 1 – 1D model inlet thermodynamic and operation conditions.

Variable	Values
Inlet Pressure (P_{in})	7.372 MPa
Inlet Temperature (T_{in})	313.15 K
Pressure Ratio (PR)	3.4
Mass Flux ($\dot{m}$)	55.56 kg/s
Flow Coefficient (φ)	0.012
Head Coefficient (μ)	0.549

The input parameters chosen are elucidated in Figure 2. The axial height (ΔZ), number of blades (Z_{fb}) and blade thickness (t) were selected to evaluate their impact on isentropic efficiency. In order to properly assess the effectiveness of screening methods combination, each variable range of variation was fixed as follows: $20 < \Delta Z < 150$ mm; $8 < Z_{fb} < 14$ blades and $1 < t < 4$ mm.

Figure 2 – Parameters chosen for Sensitivity Analysis.

The 1D model validation was performed with the test 'CBC_081202_1003' recorded in the Sandia report (WRIGHT *et al.*, 2010). Basically, the 1D model geometry and operational inputs were settled as the same as Sandia's tests parameters and the total pressure (P_{t6}) was compared, as presented in Table 2. It can be noticed that all comparisons presented errors below 2%, providing good agreement between model and experimental data.

Table 2. Comparison between experimental test and the 1D model.

n [rpm]	$\dot{m}$ [kg/s]	T_{in} [K]	P_{in} [MPa]	PR	P_{t6} Exp.	P_{t6} 1D	Error (%)
55000	3.969	304.4	7.722	1.18	9.101	9.273	1.89%
55000	3.719	304.4	8.067	1.26	10.135	10.283	1.46%
55000	3.402	304.4	8.136	1.31	10.687	10.794	1.00%
55000	2.540	304.4	8.170	1.42	11.583	11.683	0.86%
55000	1.670	304.4	8.205	1.47	12.066	12.104	0.31%

2.2 Combining Elementary Effects and Fractional Factorial Methods

The EE Method (MORRIS, 1991) is recognized as a cheap and effective method for screening a model, recognizing the few important input variables among the many that can be contained in a model. If a model has k independent inputs, then X_i, (i=1,...,k), the Elementary Effect is defined as

$$EE_i = \frac{Y(X_1, X_2, \ldots, X_i + \Delta, \ldots, X_k) - Y(X_1, X_2, \ldots, X_k)}{\Delta} \tag{1}$$

in which Δ is the step for the discretized input space. The sampling strategy suggests that r elementary effects are computed for each input variable. Each one of the sample space trajectories contains k+1 points. If r = 10 is chosen and three variables are considered, the total cost of the method is: r(k+1) = 40 model executions, which is relatively cheap in comparison with quantitative methods. Moreover, as the number of variables in the model increase, the more the difference in costs becomes evident.

In order to properly rank the input variable importance, a main effect sensitivity measure is proposed as the average of each input EE distribution on absolute values (μ*), equation (2), which assesses the overall influence of the variable (CAMPOLONGO; CARIBONI, 2007), being sufficient to provide a reliable ranking of variables.

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r}\left|EE_i^j\right| \tag{2}$$

The sampling strategy maximizes the r trajectories dispersion in the sample space. First, it generates a high number of trajectories (M = 500), and then it selects r trajectories with the highest 'spread', based on the definition of 'distance' (d_{ml}) between trajectories:

$$d_{ml} = \begin{cases} \sum_{i=1}^{k+1}\sum_{j=1}^{k+1}\sqrt{\sum_{z=1}^{k+1}\left[X_i^m(z) - X_j^l(z)\right]^2}, & m \neq l \\ 0, & otherwise \end{cases} \tag{3}$$

in which $X_i^m(z)$ indicates the z^{th} coordinate of the i^{th} point of the m^{th} Morris' trajectory. Summing up, it selects the combination of r trajectories with the highest value of distance. This sampling methodology is called 'optimal sampling' (CAMPOLONGO *et al.*, 2007) as it considers all possible combinations of trajectories. However, this combinatorial optimization approach turns the sampling process unfeasible, as the total number of combinations for an optimal approach is M!/[r!(M-r)!]. To overcome this problem, instead of picking r optimal trajectories (OT) directly from the original set of M; (M-1) trajectories are first picked on the first step; Moreover, in the second step, (M-2) trajectories are chosen. This procedure is repeated until only r trajectories are left. This set of r trajectories is the so-called quasi-Optimal Trajectories (quasi-OT) (GE *et al.*, 2014). The total number of combinations considered in this approach decreases to (M+n)(M-n+1)/2.

The FF method is a screening method based on the orthogonality of Hadamard matrices, which produces a useful DoE for two-factor interaction identification (SALTELLI *et al.*, 2008). The Hadamard matrices are composed by -1 and +1, indicating the lower and upper bounds of the parameter range of variation, respectively. Its order (n) must be higher than the number of input variables (k) present in the model and it is always a power of 2. In order to further improve the properties of the FF DoE, its construction uses the proper Hadamard matrix and its opposite, as elucidated in Equation (4), resulting in a cost of 2n model realizations.

$$M_n = \left[\frac{H_n}{-H_n}\right] \tag{4}$$

In such a M_n design, an equal number of simulations are allocated to each corner of the sample space, which is helpful in determining which combinations of parameters are working together to influence the value of Y (SALTELLI *et al.*, 2008). Moreover, the sensitivity measure used to account for the effect of a parameter or two-factor combination in FF analysis, $ME_r(Y)$, is presented in Equation (5).

$$ME_r(Y) = \frac{1}{2n}\sum_{j=1}^{2n} x_{jr}\, y_{jr} \tag{5}$$

in which x_{jr} is the Hadamard matrix column of the specific input variable and y_{jr} is the response vector after the DoE is assessed by the model in the analysis.

Finally, the combination of both screening methods can be useful in determining the importance of two-factors interaction in comparison with main effects. The EE method is responsible for providing the overall variable influence on the model and the FF would indicate if the variables influence is due to interactions with other variables. Therefore, the total cost of such SA combination strategy would be the simple sum of both method costs: r*(k+1) + 2n.

2.3 Smoothing Spline ANOVA Method

The SS-ANOVA implemented in the software ModeFRONTIER (GU, 2002) was chosen as the quantitative SA tool. Usually, the recommended number of simulations for smoothing spline analysis is k^3, representing the method's computational cost. This classical analysis is based on variance (ANOVA), working through the decomposition of main effects and two-factor interactions, measuring the percentage of its contribution to the global variance (RATTO; PAGANO, 2010), defined as

$$f(X) = f_0 + \sum_{j=1}^{n} f_j(x_j) + \sum_{j<i}^{n} f_{j,i}(x_j, x_i) \tag{6}$$

The SS-ANOVA can be built as the minimization of the usual least square functional (estimates the fitness of the model f) subjected to the constraint J (controlling the smoothness of the model and avoiding overfitting), using the Lagrange method (KIM; GU, 2004).

3. RESULTS AND DISCUSSION

This section is divided in two main parts: the cost evaluation between screening method's combination and variance-based method, and the effectiveness assessment of screening methods to identify two-factor interactions and variable importance.

3.1 Screening Methods Cost Evaluation

The computational cost of the combined screening methods and the SS-ANOVA were compared for a different number of input variables, as shown in Figure 3. As already demonstrated, the cost for the combined EE and FF method is r(k+1) and for SS-ANOVA is k^3. Therefore, the cost for the combined screening methods increases linearly, whilst the SS-ANOVA cost increases exponentially, resulting in almost the same cost for 3 and 4 input variables, and in major differences for high-dimensional problems.

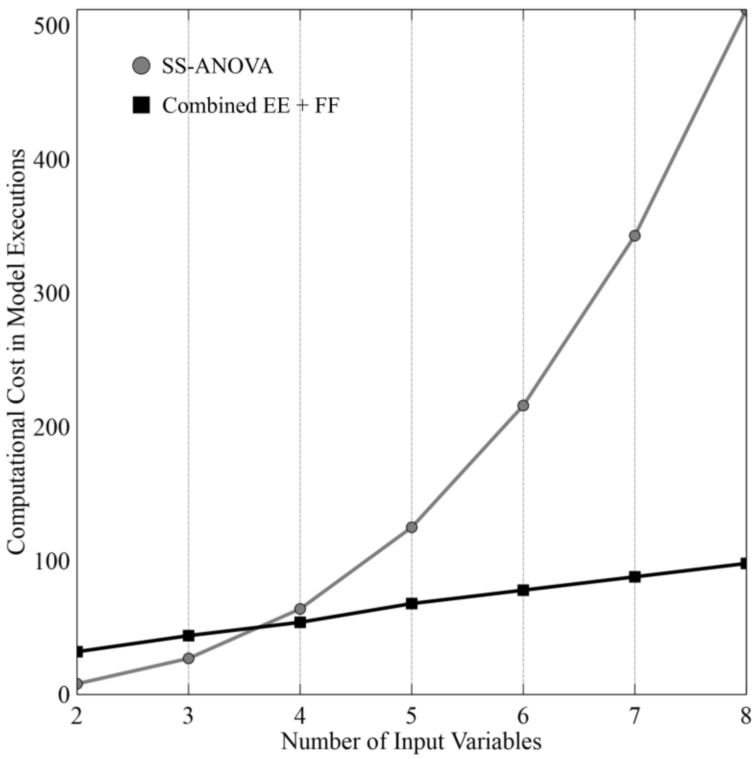

Figure 3 – Cost predictions for combined screening methods and SS-ANOVA in a number of model executions.

As the number of input variables in a model increase, the more the combined screening methods become helpful. For instance, in a model containing 8 input variables, the combined screening methods would require only 17.6% of the SS-ANOVA cost. If this novel screening methodology is effective in identifying both important and interacting variables, it would reduce the computational costs for SA of high-dimensional models, turning this preliminary assessment feasible. In order to inexpensively evaluate the effectiveness of the combined screening method, only 3 input variables were considered herein utilizing the low-cost 1D model of centrifugal compressors.

3.2 Two-factor Interaction and Influent Variables Identification Through Screening Methods

In order to properly evaluate if the combined screening methods can identify the two-factor interactions, the 1D model was executed through all three DoE. First, the Quasi-OT sampling used 40 model executions, allowing the computation of Morris' EE for all variables, presented in Figure 4. The number of blades (Z_{fb}) and axial height (ΔZ) were ranked as more impacting on isentropic efficiency than the blade thickness (t).

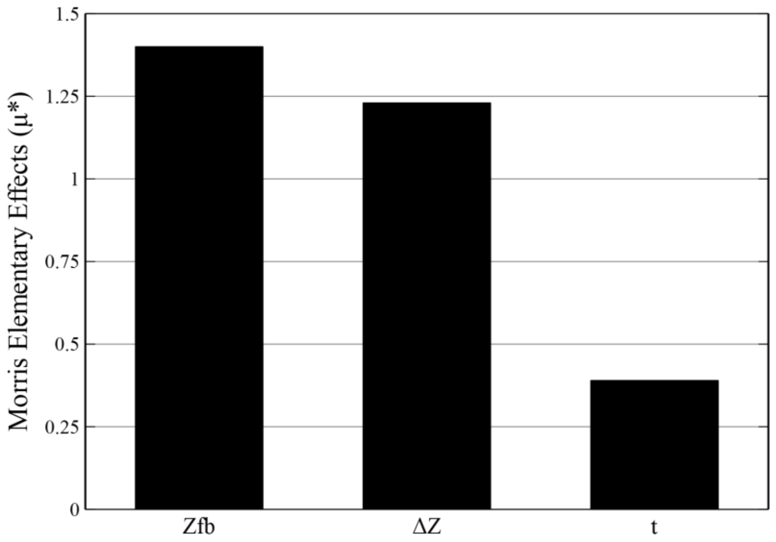

Figure 4 – Histogram of an absolute average of Morris' EE method.

Moreover, the Hadamard matrices DoE was executed, using 8 model runs, and the FF single and two-factor effects (ME_i) were computed, as presented in Figure 5. The FF methodology identified a strong interaction between number of blades (Z_{fb}) and axial height (ΔZ), presenting the thickness (t) as the second more impacting variable towards isentropic efficiency.

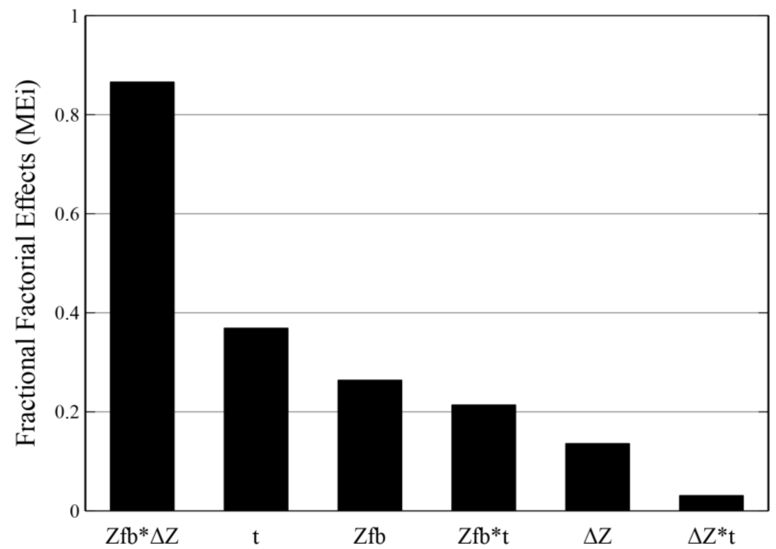

Figure 5 – Histogram of single and two-factor effects of FF method.

Therefore, for purposes of factor fixing (i.e., treating non-impacting variables as constants) the combination of screening methods indicates that all variables should stay in the analysis since they all have a significant impact on isentropic efficiency. Moreover, it states that the interactions between variables are not negligible. The total cost for this combined screening method analysis was 48 model executions.

Finally, a ULHS DoE of 27 cases was the initial estimated size for the SS-ANOVA sampling. However, a 60 cases DoE was used for ensuring good properties for the analysis, determining the total cost for the variance-based method. After running the 1D model, the main and interaction effects for SS-ANOVA method were also computed, as presented in Figure 6. Obviously, the axial height (ΔZ) and the interaction between number of blades (Z_{fb}) and axial height (ΔZ) were considered the most impacting effects, with the blade thickness (t) on intermediary ranking position.

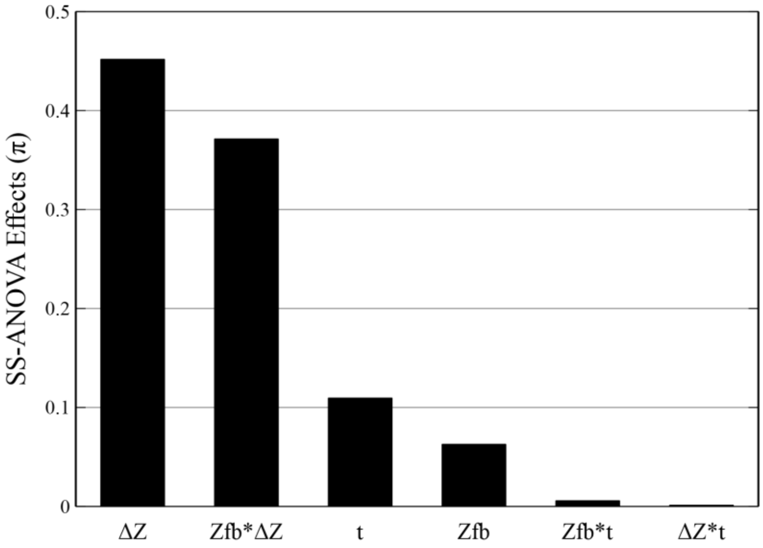

Figure 6 – Histogram of main and interacting effects of SS-ANOVA method.

Therefore, the SS-ANOVA method corroborates the results presented by the combined screening SA. Thus, the most influential variables are: Number of blades and axial height; and the strength of interaction effects between these variables. In practical terms, running the combined screening SA instead of the expensive SS-ANOVA method would classify the same variables as influential.

4. CONCLUSIONS

The 1D model of a S-CO_2 centrifugal compressor was studied herein due to its low computational cost, in order to evaluate the combination of two SA screening methods towards its capacity to identify two-factor interactions. To properly assess the effectiveness of the combination, a variance-based method, SS-ANOVA, was used for corroboration. The study's highlights are:

- As the number of variables of a model increase, the more the combined SA screening methodology is advantageous in relation to variance-based methods such as the SS-ANOVA;
- The SA combined screening method has proven to be an effective and low-cost methodology for two-factor interaction identification;
- The interaction between the number of blades and axial height needs to be investigated in phenomenological aspects;

Furthermore, the combined SA screening methodology should be investigated on models with a higher number of input variables to be better scrutinized by the scientific community. Moreover, the application of this technique on more expensive models would also evidence the computational costs advantageous of the method.

5. ACKNOWLEDGEMENTS

We gratefully acknowledge the support of the RCGI – Research Centre for Gas Innovation, hosted by the University of São Paulo (USP) and sponsored by FAPESP – São Paulo Research Foundation (2014/50279-4) and Shell Brasil.

6. REFERENCES

AUNGIER, R. H. **Centrifugal Compressors - A Strategy for Aerodynamic Design and Analysis**. New York: ASME PRESS, 2000.v. 53. Disponível em: file:///C:/Users/User/Downloads/fvm939e.pdf.

CAMPOLONGO, F.; CARIBONI, J. Sensitivity analysis: how to detect important factors in large models. **European Commission: Directorate General Joint Research Centre. EUR**, [*s. l.*], v. 22798, p. 25, 2007. Disponível em: http://www.jrc.cec.eu.int.

CAMPOLONGO, F.; CARIBONI, J.; SALTELLI, A. An effective screening design for sensitivity analysis of large models. **Environmental Modelling and Software**, [*s. l.*], v. 22, n. 10, p. 1509–1518, 2007.

CAMPOLONGO, F.; SALTELLI, A.; CARIBONI, J. From screening to quantitative sensitivity analysis. A unified approach. **Computer Physics Communications**, Italyv. 182, n. 4, p. 978–988, 2011. Disponível em: http://dx.doi.org/10.1016/j.cpc.2010.12.039.

FENG, K.; LU, Z.; YANG, C. Enhanced Morris method for global sensitivity analysis : good proxy of Sobol ' index. **Structural and Multidisciplinary Optimization**, [*s. l.*], v. 59, p. 373–387, 2019.

GE, Q.; CIUFFO, B.; MENENDEZ, M. Combining screening and metamodel-based methods: An efficient sequential approach for the sensitivity analysis of model outputs. **Reliability Engineering and System Safety**, [*s. l.*], v. 134, p. 334–344, 2014. Disponível em: http://dx.doi.org/10.1016/j.ress.2014.08.009.

GU, C. **Smoothing Spline ANOVA Models**. New York: Springer-Verlag, 2002. 2002.

KIM, Y. J.; GU, C. Smoothing spline Gaussian regression: more scalable computation via efficient approximation. **Journal of the Royal Statistical Society**, [*s. l.*], v. 66, n. 2, p. 337–356, 2004.

MONJE, B.; SÁNCHEZ, D.; SAVILL, M.; PILIDIS, P.; SÁNCHEZ, T. A Design Strategy for Supercritical CO2 Compressors. *In*: ASME TURBO EXPO 2014, 2014, Dusseldorf, Germany. **Anais** [...]. Dusseldorf, Germany: ASME, 2014.

MORRIS, M. D. Factorial Sampling Plans for Preliminary Computational Experiments. **Technometrics**, USAv. 33, p. 161–174, 1991.

NAGEEB EL-HELALY, S.; HABIB, B. A.; ABD EL-RAHMAN, M. K. Resolution V fractional factorial design for screening of factors affecting weakly basic drugs liposomal systems. **European Journal of Pharmaceutical Sciences**, [*s. l.*], v. 119, n. February, p. 249–258, 2018. Disponível em: https://doi.org/10.1016/j.ejps.2018.04.028.

R.SPAN; W.WAGNER. A New Equation of State for Carbon Dioxide Covering the Fluid Region from the Triple-Point Temperature to 1100 K Pressures up to 800 MPa. **Jornal of Chemical**, [*s. l.*], v. 25, n. 6, 1994.

RATTO, M.; PAGANO, A. Using recursive algorithms for the efficient identification of smoothing spline ANOVA models. **Advances in Statistical Analysis**, [*s. l.*], v. 94, n. 4, p. 367–388, 2010.

SALTELLI, A. (Andrea). **Global sensitivity analysis : the primer**. [*S. l.*]: John Wiley, 2008. 2008.

SALTELLI, A.; ALEKSANKINA, K.; BECKER, W.; FENNELL, P.; FERRETTI, F.; HOLST, N.; LI, S.; WU, Q. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. **Environmental Modelling and Software**, [*s. l.*], v. 114, n. March 2018, p. 29–39, 2019. Disponível em: https://doi.org/10.1016/j.envsoft.2019.01.012.

SALTELLI, A.; RATTO, M.; ANDRES, T.; CAMPOLONGO, F.; CARIBONI, J.; GATELLI, D.; SAISANA, M.; TARANTOLA, S. **Global Sensitivity Analysis. The Primer**. Chichester, England: John Wiley & Sons Ltd, 2008. 2008.

WRIGHT, S. A.; RADEL, R. F.; VERNON, M. E.; ROCHAU, G. E.; PICKARD, P. S. **Operation and Analysis of a Supercritial CO2 Brayton CycleSandia National Laboratories**. [*S. l.: s. n.*], 2010. Disponível em: https://linkinghub.elsevier.com/retrieve/pii/S0898589816301735. .

XIA, W.; ZHANG, Y.; YU, H.; HAN, Z.; DAI, Y. Aerodynamic design and multi-dimensional performance optimization of supercritical CO2 centrifugal compressor. **Energy Conversion and Management**, [*s. l.*], v. 248, n. 28, p. 114810, 2021. Disponível em: https://doi.org/10.1016/j.enconman.2021.114810.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.