Proceedings of ICONE19
19th International Conference on Nuclear Engineering
May 16-19, 2011, Chiba, Japan

ICONE19-43254

TOWARDS A BETTER UNDERSTANDING OF CLOGGED STEAM GENERATORS: A SENSITIVITY ANALYSIS OF DYNAMIC THERMOHYDRAULIC MODEL OUTPUT

Sylvain Girard
EDF
Phone: +33130878050
sylvain-2.girard@edf.fr

Thomas Romary
Mines ParisTech
Phone: +33164694773
thomas.romary@mines-paritech.fr

Jean-Melaine Favennec
EDF
Phone: +33130878536
jean-melaine.favennec@edf.fr

Pascal Stabat
Mines ParisTech
Phone: +33140519152
pascal.stabat@mines-paristech.fr

Hans Wackernagel
Mines ParisTech
Phone: +33164694760
hans.wackernagel@mines-paristech.fr

1 Introduction

Internal parts of Steam Generators (SGs) foul up with iron oxides which causes Tube Support Plate (TSP) clogging which brings about concerns about safety and performance. Means to estimate TSP clogging are needed to optimize the costly maintenance operations. Previous work showed that the dynamic response to a power transient of the wide range level measurement contains informations about the clogging state of steam generators. A diagnosis method based on the comparison of plant WRL response with simulated responses is being developed by EDF. In order to achieve better understanding of the effect of clogging on SGs dynamic behaviour and to assess the potential of a diagnosis method based on the analysis of this behaviour through simulations, it is necessary to determine *i)* what features of the WRL response curves are characteristic of clogging and *ii)* the relative impact of each half TSP on these features.

2 Materials and methods

The objective of the present study is to analyse the sensitivity of the model output (WRL dynamic response) to its input parameters (TSP clogging ratios). A method based on the ANOVA-decomposition and a Monte Carlo computation scheme has been used to compute order 1 and total sensitivity indices for each half-TSP. As the model output is functional, a principal component analysis (PCA) on the samples used to compute sensitivity indices had to be carried out to reduce the dimensionality of the model output and compute 'compact' order 1 and total sensitivity indices for each major principal component. Finally, estimation variability was assessed by construction of BC_a bootstrap confidence intervals.

3 Results

A sensitivity analysis of the dynamic output of a SG model provided a better understanding of the effect of TSP clogging and valuable insights for the development of diagnosis methods: *i)* sequential Sobol indices revealed different behaviours for each SG leg; *ii)* PCA highlighted two main features of the responses curves that characterize clogging; *iii)* Sobol indices on the PCA reduced output allowed to quantify the importance of each half TSP; *iv)* the convergence of the computations was checked with bootstrap confidence intervals.

References

G.E.B. Archer, A. Saltelli, and I.M. Sobol'. *Journal of Statistical Computation and Simulation*, 58:99–120, 1997.

K. Campbell, M.D. McKay, and B.J. Williams. *SAMO 2004*. 2005.

J.G. Collier and J.R. Thome. Oxford University Press, 1996.

B. Efron and R. Tibshirani. Chapman & Hall, 1993.

T. Homma and A. Saltelli. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

I.T. Jolliffe. Springer, 2002.

Lamboni et al. *Field Crops Research*, 113:312–320, 2009.

Midou et al. *ICONE18*. 2010.

I.M. Sobol'. *English Transl.: MMCE*, 1(4), 1993.

I.M. Sobol'. *Mathematics and Computers in Simulation*, 55:271–280, 2001.

G. Yadigaroglu. *27^{th} Short Courses - MCMF*. 2010.

Proceedings of ICONE19
19th International Conference on Nuclear Engineering
May 16-19, 2011, Chiba, Japan

ICONE19-43254

TOWARDS A BETTER UNDERSTANDING OF CLOGGED STEAM GENERATORS: A SENSITIVITY ANALYSIS OF DYNAMIC THERMOHYDRAULIC MODEL OUTPUT

Sylvain Girard
EDF
Phone: +33130878050
sylvain-2.girard@edf.fr

Thomas Romary
Mines ParisTech
Phone: +33164694773
thomas.romary@mines-paritech.fr

Jean-Melaine Favennec
EDF
Phone: +33130878536
jean-melaine.favennec@edf.fr

Pascal Stabat
Mines ParisTech
Phone: +33140519152
pascal.stabat@mines-paristech.fr

Hans Wackernagel
Mines ParisTech
Phone: +33164694760
hans.wackernagel@mines-paristech.fr

Keywords: steam generator, clogging, thermohydraulic model, sensitivity analysis, principal component analysis.

Abstract

Tube support plate clogging of steam generators affects their operating and requires frequent maintenance operations. A diagnosis method based on dynamic behaviour analysis is under development at EDF to provide means of optimisation of maintenance strategies. Previous work showed that the dynamic response to a power transient of the wide range level measurement contains informations about the clogging state of steam generators. The diagnosis method consists of comparisons of the measured dynamic response with simulations on a mono-dimensional dynamic steam generator model for various input clogging configurations. In order to assess the potential of this method, a sensitivity analysis has been conducted through a quasi-Monte Carlo scheme to compute sensitivity indices for each half tube support plate's clogging ratio. Sensitivity indices are usually defined for scalar model outputs. Principal component analysis has been used to determine a small subset of variables that condense the information about the shape of the response curves. Finally, estimation variability was assessed by construction of bootstrap confidence intervals. The results showed that half of the preselected input variables have negligible influence and allowed to rank the most important ones. Interactions of input variables have been estimated to exert only a small influence on the output. The effects of clogging on the steam generator dynamics has been characterised qualitatively and quantitatively.

1 Introduction

Steam Generators (SG) are affected by fouling of their internal elements. This causes clogging of the quatrefoil holes of the Tube Support Plates (TSP) that can induce safety issues. Means to estimate TSP clogging are needed to optimize maintenance operations.

Clogging reduces the open cross sectional area of the TSPs and thus induces higher pressure drop. This alters the operating point of the SG as well as its dynamic behaviour. Previous studies (Midou et al., 2010) demonstrated that the shape of the Wide Range Level (WRL – the pressure difference measured between the steam dome and the bottom of the downcomer) response curve to a power transient is determined by the clogging state of the TSPs. A diagnosis method based on comparison between measured response curves and simulated ones for varying clogging states is being developed by EDF. A mono-dimensional SG model has been created with Modelica and the Dymola software for that purpose. It is able to simulate the SG behaviour during power transient phases. A 60 % power decrease with a 3 $\%.min^{-1}$ rate is performed on French nuclear reactors every three months which allows for frequent diagnosis. The input variables of the model are the clogging ratios of each half TSP, one for the hot and cold legs of the U-tube bundle. Clogging ratios are defined as the ratio of the closed area to the total area of the holes without any clogging. The output of the model consists of the 1200 values of WRL (1 per second) given the clogging ratios of the 16 half TSP.

In order to assess the performance of the method, it is necessary to determine *i)* what features of the WRL response curves are characteristic of clogging and *ii)* the relative impact of each half TSP on these features.

If it happens, for instance, that the clogging ratio of a given half TSP has a negligible influence, it is irrelevant to keep it as an input variable because little or no information about this ratio is contained into the WRL response curve. Discarding irrelevant variables reduces the dimension of the clogging state space that has to be sampled to produce the diagnosis.

The objective of the present study is to analyse the sensitivity of the model output (WRL dynamic response) to its input parameters (TSP clogging ratios). A sensitivity analysis has been conducted through a quasi-Monte Carlo scheme to compute sensitivity indices for each half TSP clogging ratio. As the model output is functional, a Principal Component Analysis (PCA) on the samples used to compute sensitivity indices had to be carried out to reduce the dimensionality of the model output and compute 'compact' order 1 and total sensitivity indices for each major principal component. Sequential WRL temporal indices have also proved to be meaningful. Finally, estimation variability was assessed by construction of bootstrap confidence intervals.

2 Materials and methods

2.1 Mono-dimensional steam generator model

The steam generator (SG) model has been developed with the Modelica language using the Dymola software. Modelica is an object-oriented language especially designed for modelling physical systems. It relies on a third party compiler and solver for simulation. These roles are being assumed here by the Dymola software. More specifically, we use a solver named DASSL, provided by Dymola, which is capable of solving differential algebraic equations.

The SG type studied here is the Westinghouse 51. EDF currently operates 48 of these; most of them being about 30 years old. A diagram representing the principal elements of a SG is given on figure 1.

The main elements of the model are :

- primary fluid flowing inside the U-tubes (single-phase flow) ;
- secondary fluid flowing outside the U-tubes (two-phase flow) ;
- thermal transfer between the two fluids and through tube interfaces ;
- two-phase singular pressure drops *e.g.* at the TSP quatrefoil holes ;
- steam-liquid separation devices ;
- feed water flow rate control system.

All these elements are mono-dimensional but the exchanger part is modelled as two channels: one for the *hot leg* (*i.e* concurrent exchanging side, where primary fluid enters the SG) and one for the *cold leg* (*i.e* countercurrent exchanging side, where primary fluid exits the SG). The exchanging channels are composed of 20 meshes evenly spaced.

The choice of mono-dimensionality and of the number of meshes is driven by the applications for which the model has been developed. On the one hand, it must be able to simulate the dynamic response of a SG precisely enough so that information about clogging ratios spatial distribution is not lost by averaging processes. On the other hand, computation time for simulation must not exceed five minutes so that it can be used in Monte Carlo methods.

Variations of the boundary conditions of the model allows to simulate a power transient of the plant. The power decrease used in the clogging diagnosis method is modelled by a linear variation of primary and secondary inlet enthalpies and secondary outlet steam flow rate. The feed water flow rate is being determined by the control system.

Heat transfer coefficients are computed using correlations available in the literature. Pressure drops at the quatrefoil holes are computed using a specific correlation derived from experiments conducted at EDF R&D on a 1:4 scale mock-up of TSP and tubes.

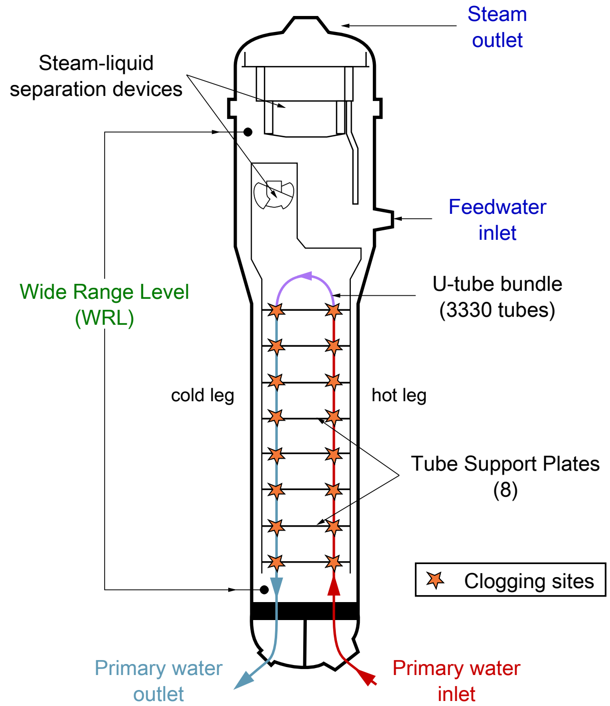

Fig. 1: *Westinghouse type 51 SG principal elements*

The two phase flow (secondary fluid) in the riser has been modelled using a simple mixture formulation known as the *homogeneous-equilibrium model* (Yadigaroglu, 2010) in which equal phase velocities and thermal equilibrium between the phases is assumed. This choice has been driven by the will to limit the number of unknown parameters to be inferred or calculated by correlations. Correlations available in the literature are usually designed for flows in tubes which can be quite different from a flow in a tube that contains a tube bundle and TSPs. Very little is known about the regime of the two-phase flow in a SG which makes the modelling of velocity slip tricky.

Another argument in favour of the homogeneous formulation is the high level of turbulence that the flow exhibits and the presence of the TSPs. Indeed, going through the quatrefoil holes tend to deconstruct and mix the flow while it slows down the steam more than the liquid. Visual observations made on the experiments mentioned above suggest that we are in presence of a slug-bubbly flow in the upper part of each interval between TSPs (under the TSP) while the flow is bubbly with very small steam structures from the top of the TSP up to approximately the middle of the interval. Flow regimes are described in the book by Collier and Thome (1996, Chap. 1). In the lower part of the intervals, the ratio of the volume of the bubbles to their surface is low and so is the ratio of the 'floating' force to frictions that is responsible of the higher velocity of steam. These rough observations were made on a freon-water mixture without heating for a few mass fractions. When the results of these experiments are available, it will be possible to examine this question in much more details.

Implementing velocity slip in the SG model is being considered for future work. The impact of this amelioration on the results shown here are expected to be negligible.

2.2 Global sensitivity analysis

Sensitivity analysis in the broad sense studies how perturbations of model input variables generate perturbations on the output variables. Local sensitivity analysis is concerned by the effect on the output of small perturbations of the input variables around a given point, whereas global sensitivity analysis looks at the variability of the output on the whole variation domain. Here we focus exclusively on global sensitivity analysis because we want to obtain general information about how TSP clogging affects the WRL response without any particular clogging distribution in mind.

In this study, the components of the input vector are the 16 half TSP clogging ratios so $n = 16$ because Westinghouse 51 SGs have 8 TSP.

Clogging ratios were assumed to range from 0% to 65% as these values cover most real cases.

Let f be a function that represents the model, $\mathbf{x}$ the input variables and u a scalar output variable of the model.

$$\begin{aligned} f\ :\ & \mathbb{I}^n \to \mathbb{R} \\ & \mathbf{x} \mapsto u = f(\mathbf{x})\ . \end{aligned} \quad (1)$$

For the sake of clarity, in the following general derivation of sensitivity indices the input variables are assumed to take their values in $[0, 1]$ so $\mathbb{I}^n$ denotes the n-dimensional unit hypercube. In (1), the model output consists of a single scalar. We will actually consider several of them and compute a set of sensitivity indices for each one independently.

Quite simple sensitivity indices can be defined when the model is linear or monotonous but this is seldom the case for complex thermohydraulic models. Variance based methods require a large number of simulations but they are model free. Among them is a method introduced by Sobol' (1993) based on the ANOVA-representation of the function f (ANOVA stands for ANalysis Of VAriance) and a Monte Carlo computation scheme. The next section relates the principal theoretical aspects of this method. The following derivation can be found in greater detail in (Sobol', 2001).

2.2.1 ANOVA-representation

Assuming f is an integrable function, let us consider the following decomposition

$$f(\mathbf{x}) = f_0 + \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} f_{i_1\ldots i_s}(x_{i_1}, \ldots, x_{i_s})\ , \quad (2)$$

where f_0 is a constant and the double sum means that there is a function $f_{i_1\ldots i_s}(x_{i_1}, \ldots, x_{i_s})$ for every possible family of input variables: from $f_1(x_1)$ to $f_n(x_n)$ then all the $f_{ij}(x_i, x_j)$ with $1 \leq \cdots < i < j \leq n$ and so on up to $f_{1\ldots n}(x_1 \ldots x_n)$. The number of terms in this decomposition is 2^n.

Sobol' (1993) has proved that if we impose to the summands of (2) the following condition

$$\int f_{i\ldots j}(x_i, \ldots, x_j)\, \mathrm{d}x_k = 0 \quad \text{for} \quad k = i_1, \ldots, i_s\ , \quad (3)$$

the decomposition exists and is unique. It is then called the ANOVA-representation of f. It follows that the summands in (2) are orthogonal and can be expressed as integrals of f.

2.2.2 Sensitivity indices

If we assume f to be square integrable, the $f_{i_1\ldots i_s}$ are also square integrable. Squaring and integrating (2) raises

$$\int f^2(\mathbf{x})\, \mathrm{d}\mathbf{x} - f_0^2 = \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} \int f^2_{i_1\ldots i_s}\, \mathrm{d}x_{i_1} \ldots \mathrm{d}x_{i_s}\ . \quad (4)$$

Now if $\mathbf{x}$ is a random vector uniformly distributed in $\mathbb{I}^n$ then $f(\mathbf{x})$ and $f_{i_1\ldots i_s}(x_{i_1}, \ldots, x_{i_s})$ are random variables whose variances are respectively

$$D = \int f^2\, \mathrm{d}\mathbf{x} - {f_0}^2 \quad (5)$$

and

$$D_{i1} \ldots i_s = \int f^2_{i_1\ldots i_s}\, \mathrm{d}x_{i_1} \ldots \mathrm{d}x_{i_s}\ , \quad (6)$$

and we have the following equality:

$$D = \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} D_{i_1\ldots i_s}\ . \quad (7)$$

In other words, D measures the variability caused by variations of all the input variables while the $D_{i_1\ldots i_s}$ represent the variability caused by variations of variables in subsets $(x_{i_1}, \ldots, x_{i_s})$. Equation (7) states, as expected, that the overall variability is the sum of the variabilities caused by all the possible subsets of input variables.

This leads to define the global sensitivity index of a subset of variables $(x_{i_1}, \ldots, x_{i_s})$ by the following ratio

$$S_{i_1\ldots i_s} = \frac{D_{i_1\ldots i_s}}{D} \quad (8)$$

where s is called the order of the indice.

Order 1 sensitivity indices, $S_i = {}^{D_i}/_{D}$, measure the influence of each half TSP clogging ratio *alone*, that is without considering interactions. Reliable computation of sensitivity indices usually requires numerous model evaluations so it might be off-putting to try to estimate indices of order higher than 1 or 2. In our study we restrict ourselves to order 1 indices due to the high number of input variables. However, as our input variables are highly physically linked, we expect combined effects to play a significant role and completely ignoring them could be misguiding.

As a palliative let us consider the variance of the subset of variables constituted by all the variables but one. Indeed, its complementary to the overall variance is the variance of all the subsets that comprise the one left aside (Homma and Saltelli, 1996). For an input variable x_i, we have the following equality

$$D_i^{tot} = D - D_{\bar{i}} \tag{9}$$

where D_i^{tot} is the variance of all the subsets that comprises x_i and $D_{\bar{i}}$ is the variance of the subset constituted by all variables but x_i. Dividing equation (9) by the overall variance D leads to define *total sensitivity indices* for each input variable as follows

$$S_i^{tot} = \frac{D_i^{tot}}{D} \,. \tag{10}$$

The difference between total sensitivity indices and order 1 sensitivity indices for a given input variable measures the influence of all the combinations that involve the variable. Thus we can estimate how much of the overall variance is not accounted for by order 1 indices and determine what variables are responsible for it.

2.2.3 Computation scheme

Sobol' (1993) has demonstrated the following theorem

Theorem 1 (Variance of a subset of variables)**.** *Let* $\mathbf{y}$ *be a subset of variables of the complete family of input variables* $\mathbf{x}$*. Let* $\mathbf{z}$ *be its complementary i.e.* $\mathbf{x} = (\mathbf{y}, \mathbf{z})$*. The variance* $D_\mathbf{y}$ *on the output of* f *induced by variations of the variables of* $\mathbf{y}$ *is equal to*

$$D_y = \int f(\mathbf{x}) f(\mathbf{y}, \mathbf{z}') \, d\mathbf{x} d\mathbf{z}' - {f_0}^2 \,. \tag{11}$$

The prime symbol indicates here that $\mathbf{z}'$ comes from a different realisation of $\mathbf{x}$ seen as a random variable.

By applying theorem 1 to the case of $\mathbf{y}$ being a single variable x_i and $\mathbf{z}$ being all the input variables but x_i, order 1 and total indices can be computed by estimating integrals : one for f_0, one for D and two for each input variable.

This can be achieved through a Monte Carlo approach. Let $\boldsymbol{\xi} = (\boldsymbol{\xi}_j)$ and $\boldsymbol{\xi}' = ({\boldsymbol{\xi}'}_j)$ be two independent samples of size N uniformly distributed in $\mathbb{I}^n$. The following approximations hold for high enough N

$$\frac{1}{N}\sum_{j=1}^{N} f(\boldsymbol{\xi}_j) \to f_0 \,, \qquad \frac{1}{N}\sum_{j=1}^{N} f^2(\boldsymbol{\xi}_j) \to D + {f_0}^2 \,, \tag{12}$$

$$\frac{1}{N}\sum_{j=1}^{N} f(\boldsymbol{\xi}_j) f({\xi_j}^{(1)}, \ldots, {\xi_j}^{(i-1)}, {\xi'_j}^{(i)}, {\xi_j}^{(i+1)}, \ldots, {\xi_j}^{(n)}) \to D_i + f_0^2 \,, \tag{13}$$

$$\frac{1}{N}\sum_{j=1}^{N} f(\boldsymbol{\xi}_j) f({\xi'_j}^{(1)}, \ldots, {\xi'_j}^{(i-1)}, {\xi_j}^{(i)}, {\xi'_j}^{(i+1)}, \ldots, {\xi'_j}^{(n)}) \to (D - D_i)^{tot} + f_0^2 \,. \tag{14}$$

In this study, random samples of size 1000 have been used. First computations made with samples of size 400 only displayed aberrations that indicated poor convergence.

For each TSP, two sets of 1000 clogging configurations have been built to generate the second members in each term of the sums in equations (13) and (14). The first set is obtained from the first sample by replacing the clogging ratio of the TSP by the corresponding ratio from the second sample. The second set is obtained from the first sample by replacing all the clogging ratios but the one of the TSP by the corresponding ratio from the second sample. Simulations have also been run with the untouched first sample in order to compute f_0 and D and the first members of the terms of the sums in equations (13) and (14). Eventually, running additional 1000 simulations with the second sample allowed a cross-check by switching the role of the sets of clogging configurations. Thus, a total of 33000 simulations have been run.

LP_τ sequences (also known as Sobol sequences) were used to generate the random samples as they are stated to make the sum in equations (12), (13) and (14) converge faster (Archer et al., 1997).

The mean computation time for a simulation with the SG model is 3 minutes on a regular workstation. Hence the total computation time is around 2 month and a half. However, it is possible to launch several parallel threads without any further development so the computations had been carried through in about two weeks on a single workstation.

2.3 Monte Carlo algorithms with Dymola

2.3.1 Initialization of states

A dynamic model is a set of equations that describe the temporal evolution of states. Before running a simulation with Dymola, the states need to be initialized. In the case of this study, the sought after states are those of stable conditions for a given clogging configuration.

For ordinary differential equation models, imposing all the derivatives to equal 0 and setting the values of the boundary conditions and parameters (such as clogging ratios, referred to as *input variables* in section 2.2) is usually sufficient. For more complex models, it is needed to supply values of certain variables called *iteration variables.* When differential algebraic equations are involved, the initialization is likely to fail if the supplied values are too far from 'real' ones, that is values for which the system has a solution. There are about 500 iteration variables for the SG model and most of them need to be estimated with significant precision.

Once a set of appropriate values of the iteration variables has been found for a set of parameters, a convenient means to find values for another set of parameters is to use dynamic simulations. A first set of values for the zero clogging case has been found during the development of the model through a stepwise procedure: going from a cold state with no thermal exchanges and small flow rates and slowly increasing toward the nominal operating point of the SG. This state provides a starting point for the determination of others by dynamic simulation.

For instance, if the stable state corresponding to a homogeneous clogging of 5 % is needed, the procedure would be:

i) load the zero clogging stable state values of iteration variables ;
ii) set the clogging ratios to 5 %;
iii) run a dynamic simulation with Dymola ;
iv) monitor the simulation for stability (the WRL and circulation ratio are convenient indicators). Go back to step iii) if the states are not stabilized;
v) save the final states values. They can now be used to initialize the model with all derivatives set to 0.

The new stable state can then be used for dynamic simulation of a power transient.

Changing the clogging configuration amounts to introducing a perturbation into the system. This results in a damped transient state until the new sought for stable state is reached. The more important is the perturbation, the longer the resulting oscillations will take to dampen. If the perturbation is too important, even the dynamic initialization can fail.

2.3.2 Optimization of the stable states research

The computation scheme described in section 2.2.3 implies that numerous stable state researches have to be performed. Therefore this procedure has to be optimized in order to minimize the computation cost of model evaluations and allow for samples of significant size. Optimizing the research of stable states consists in using as a departure point in step i a state that corresponds to a clogging configuration similar to the one sought for: this minimizes the perturbation so that the damping to stable state is as smooth and swift as possible.

Defining the distance between two input configurations of parameters is simple when there are only a few of them and when they affect the model linearly. However, the reasons that motivated the choice of the method presented in section 2.2 precisely make this task non trivial.

The question is to find a function τ that estimates the computation time needed to go from a clogging configuration $\boldsymbol{\theta^1} = (\theta^1_{h_1}, \dots, \theta^1_{h_8}, \theta^1_{c_1}, \dots, \theta^1_{c_8})$ to another, $\boldsymbol{\theta^2}$ where the half-TSPs are numbered from bottom to top, the first 8 being on the hot leg and the last 8 on the cold leg.

A heuristic can be derived from simple physical considerations on how clogging affects the SG. Three empirical rules have been used here.

i) The perturbation, and thus computation time, will increase with differences of individual clogging ratios.
ii) As the secondary mixture goes through the SG its steam mass fraction increases and singular pressure losses are substantially more important on richer mixtures. Thus, differences of clogging ratios will have a stronger impact on computation time for transition when they are in the hot leg and the higher parts of the SG.
iii) The TSPs singular pressure drops follow an approximately exponential relation with clogging ratios (this has been observed with the experiments referred at in section 2.1). Thus, computation time will increase, probably at an exponential pace, with the mean values of each couples of clogging ratios $(\theta^1_{x_i}, \theta^2_{x_i})$.

Let us assume that the effects of all clogging ratio differences add up linearly (which is a crude hypothesis) and that effects i) and iii) are multiplicative. The sought for function has the following form

$$\tau(\boldsymbol{\theta^1}, \boldsymbol{\theta^2}) = \sum_1^8 w_{hi}\, \varphi\left(|\theta^1_{h_i} - \theta^2_{h_i}|\right)\, \gamma\left(\frac{\theta^1_{h_i} + \theta^2_{h_i}}{2}\right) + w_{ci}\, \varphi\left(|\theta^1_{c_i} - \theta^2_{c_i}|\right)\, \gamma\left(\frac{\theta^1_{c_i} + \theta^2_{c_i}}{2}\right), \quad (15)$$

where the (w_{hi}) and (w_{ci}) are unknown weights that account for rule ii) and φ and γ are two unknown monotonically increasing functions that account respectively for rules i) and iii).

Finding valid values for all the 16 weights seems somewhat illusory considering the roughness of the approach. Using the linear approximation made before, a new set of weights can be defined by averaging groups of consecutive terms in equation (15).

The computation time for 50 transitions from a stable state to another were measured and the weights in (15) were estimated by multiple linear regression using these measurements as second members for various trial φ and γ functions. The validity of these functions was assessed by the maximum error of the computation times for transition estimated with the resulting τ function.

This way, a set of 6 weights could be found with satisfactory goodness of fit: three for the lower, middle and upper part of the SG for both the hot and cold legs. The φ and γ functions eventually selected are

$$\varphi\ :\ [0,1] \to \mathbb{R},\quad x \mapsto e^{(7.0\,x)} - 1 \qquad \text{and } \gamma\ :\ [0,1] \to \mathbb{R},\quad x \mapsto e^{(8.4\,x)} - 1\ , \quad (16)$$

and the weights values are

$$\begin{aligned} &w_{h\ low} = 0.09 \quad w_{h\ mid} = 0.29 \quad w_{h\ up} = 0.61 \\ &w_{c\ low} = 0.05 \quad w_{c\ mid} = 0.08 \quad w_{c\ up} = 0.11\ . \end{aligned} \quad (17)$$

The maximum error is 120 s which is sufficient to ensure computation time of a few minutes. All the weights are positive and respect rule ii). The weights can be interpreted as some crude sensitivity indices for the 6 groups of TSPs.

2.3.3 Implementation of the Monte Carlo computing scheme

The τ function enables to estimate the computation time needed to go from one state to another. An initial set of some 100 stable states has been obtained and was used in the following algorithm:

i) draw a clogging configuration from the samples constructed for the sensitivity analysis ;
ii) using the τ function, find the 'nearest' stable state available in the stable states database;
iii) if estimated computation time is above a limit t_{lim1}, go back to i);
iv) do the transition and dynamic simulation. Store the simulated WRL response curve;

v) if estimated computation time is above a limit t_{lim2}, add the stable state used for simulation to the database;
vi) go back to i until all the complete sample has been treated.

Multiple threads can be easily run at the same time. The two limits t_{lim1} and t_{lim2} have to be guessed at first and can be adapted as the algorithm goes by. For the algorithm to eventually deal with all the clogging configurations, it has been necessary to add a loosening condition regarding t_{lim1}.

2.4 Preprocessing model output

The aim of this study is to characterize the influence of each half TSP clogging ratios on the shape of the WRL response curve. It is thus necessary to select a set of output variables that correctly represent the shape of the curves and allows to compute sensitivity indices that are interpretable.

2.4.1 Sequential indices

The simplest possible set of output variables is constituted of the values taken by the WRL at each time step. This way $16 \times 1200 = 19200$ order 1 indices and 19200 total indices are obtained. The families of 1200 indices corresponding to a given half TSP will be then referred as the 16 *sequential indices* whose temporal variations during the transient we study.

The starting point of the response curves are quite far from each other because clogging alters the operating point of the SG. Sensors uncertainties and model representativeness issues make it unreliable to use this information in the model based diagnosis method (Midou et al., 2010). Differences in the shape of the curves are of small magnitude compared to distances between individual variation domains: the *static* effect (starting point) hides the *dynamic* effect (shape of the curves). Thus, sequential indices computed without any pre-treatment do not display any appreciable variation because most of the variance of the curves samples is actually due to differences of starting points. The alternative used both for the diagnosis method and the sensitivity analysis presented here is to subtract from each curve its temporal mean. This can be seen as ignoring the first term of the Taylor expansion of the underlying functions and studying the subsequent ones.

Subtracting also from each variable its mean allows to reduce the imprecision due to the fact that the value of f_0 is important compared to the summands of equations (13) and (14). This effect described in (Sobol', 2001) has been quantified when computing confidence intervals as described in section 2.5.

Sequential indices allow to follow the evolution of the relative importance of each input variable throughout the transient. However, considering the 1200 points of the response curves as independent variables is somehow missing the functional nature of the output which is precisely what we are trying to catch.

2.4.2 Principal component analysis

A classical approach to the problem of reducing the dimension of functional data is to use projection on carefully chosen subspaces. In this study, principal component analysis (PCA) has been chosen both because it does not require any inference to be made to construct the subspaces and, as our aim is to attribute shares of overall variances to input variables, using a variance based method for dimension reduction makes good sense. Campbell et al. (2005) and Lamboni et al. (2009) have made similar uses of PCA in a context of sensitivity analysis.

Let $\mathbf{WRL} = (WRL(t))$ be the vector of output variables considered in previous subsection. Principal components (PCs) of $\mathbf{WRL}$ are ordered linear combinations of the $(WRL(t))$ that have maximum variance and are orthogonal to all the preceding PCs. For a set of p variables, up to p PCs can be found, each having less variance than the preceding. Here the number of variables equals the number of time steps so $p = 1200$. Let $(\mathbf{pc}_i)$ be the family of principal components of $\mathbf{WRL}$. The $pc_i(t)$ are the coefficients in the linear combinations and are named the *loadings* of the i-th PC. Thus, $\mathbf{WRL}$ can be written

$$WRL(t) = \sum_{i=1}^{p} pc_i(t) \, \langle \mathbf{WRL}, \mathbf{pc}_i \rangle \, . \tag{18}$$

By construction, the variance of the PCs are the eigenvalues of the empirical covariance matrix of the sample. They represent the share of the overall variance that is accounted for by each PC. In order to achieve significant dimension reduction, one hopes that the first PCs will account for most of the variance of the variables. In this latter case, projection on the subspace generated by the first few PCs provide a restricted number of variables (the corresponding loadings) that contain most of the information about variance of the whole. The construction of the PCs ensures that such a projection is optimal in terms of sum of variances. The book by Jolliffe (2002) provides PCs derivation, properties and applications in much details and examples.

The remark about subtraction of temporal means made in section 2.4.1 still holds for sensitivity indices computed on PCs loadings: without this preprocessing, the first PC represents more than 99 % of the overall variance because the differences of variation domains are of greater magnitude than variations in scale and shape of the curves.

Once again, subtraction of the means of each variables is helpful to achieve better precision.

A common practice when doing a PCA is to use centred reduced variables. This presents the advantage of normalizing the variables and avoiding the problem of discrepancies in units or range of variation. For instance, if a variable is in the range $[10^3, 10^4]$ while the others vary from 0 to 10, it will cause most of the variance of the set while varying, relatively, as much as the others.

A drawback of working with correlation instead of variances is that the PCs cannot be interpreted as easily. Indeed, when using variances, the PCs can be represented as perturbation of the mean of the sample which allows for simple interpretation of the loadings.

In this study, both approaches will be used and compared.

2.5 Bootstrap confidence intervals

It is important to estimate the accuracy of the computed sensitivity indices. One wants to know for instance, if the ranking of the indices can be trusted as it is or if groups of input variables should be considered. Also, the chosen computation scheme often induces aberrations, such as slightly negative indices or sums of indices that exceed one, due to slow convergence of the sums in equation (13) and (14); confidence intervals allow to decide if these irregularities can be overlooked or if larger sample sizes should be sought.

For each computed sensitivity index $\hat{S}$, we seek $\hat{S}_{lo}$ and $\hat{S}_{up}$ so that the two events $S < \hat{S}_{lo}$ and $S > \hat{S}_{up}$, where S is the real value of the sensitivity index, have both a probability of α. In this study, α is set to 0.05. As little is known about the distributions involved, bootstrap methods are particularly indicated as they are robust and distribution free.

The general idea behind bootstrap is to draw conclusions about a given estimator by using the empirical distribution upon which the estimator is based. Consider $\hat{S} = \phi(\boldsymbol{\zeta})$ the estimator of S computed using a sample $\boldsymbol{\zeta} = (\boldsymbol{\xi}, \boldsymbol{\xi'})$ of size $2N$. A *bootstrap sample* $\boldsymbol{\zeta}^*$ is obtained by drawing uniformly with replacement from $\boldsymbol{\zeta}$. For each bootstrap sample, a *bootstrap replication* $\hat{S}^* = \phi(\boldsymbol{\zeta}^*)$ is computed in the same way as the estimator was from the original sample. It is now possible to draw inferences on the underlying distribution followed by $\hat{S}$ by analysing the empirical distribution of the bootstrap replications.

Bootstrap *percentile* confidence intervals are constructed by taking the α and $1 - \alpha$ percentiles of the empirical distribution obtained after re-sampling. The *bias-corrected and accelerated* (shortened BC_a) intervals used in this study are derived from the percentiles intervals but include a correction of bias and an *acceleration* that compensates for variation of standard error of S with the value of S. These two corrections occur by shifting the percentiles chosen from the empirical distribution.

Details about bootstrap confidence intervals and their derivation can be found in the book by Efron and Tibshirani (1993). A rule of thumb for the number of necessary bootstrap samples to be used for the construction of confidence intervals is that it should be above 1000. However, it does not require any additional evaluation of the numerical model. Thus, BC_a confidence intervals have been computed for all sequential order 1 and total indices with 2000 bootstrap samples in less than one hour with a regular workstation.

3 Results

3.1 Sequential indices

As an example, 7 WRL response curves of sample $\boldsymbol{\xi}$ after subtraction of their temporal means are displayed in figure 2. The curves are of various shapes and are crossing each other. This is characteristic of a non-linear behaviour and is auspicious for the resolution of a diagnosis method based on curves comparison.

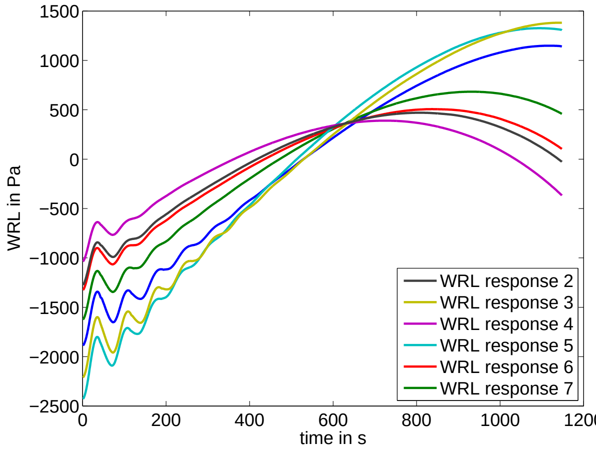

Fig. 2: *7 WRL response curves with their temporal means subtracted*

The results obtained for both order 1 and total sensitivity indices are displayed in figures 3 and 4. Red lines correspond to the hot leg and blue ones to the cold leg and the colours darken as the TSPs are higher. The error bars represent the bounds of the BC_a confidence intervals. There are only a few of them for clarity but no discrepancies have been observed on the whole set.

On all presented graphics with a time abscissa, the end time is 1148 s instead of 1200 s. This reflects the fact that actual 60 % power transients performed on plants are usually a little steeper than 3 $\%.s^{-1}$; 1148 s is the mean of the observed stopping times in (Midou et al., 2010).

Both sequential order 1 and total sets of indices display two sharp contrasting behaviours for each leg. The hierarchy of the indices is the same in all cases: the higher the TSP is positioned in the SG, the higher are the corresponding sensitivity indices. This is in agreement with physical intuition as stated in section 2.3.2.

As stated in section 2.2.2, the difference between total indices and order 1 indices measures the importance of combined effects.

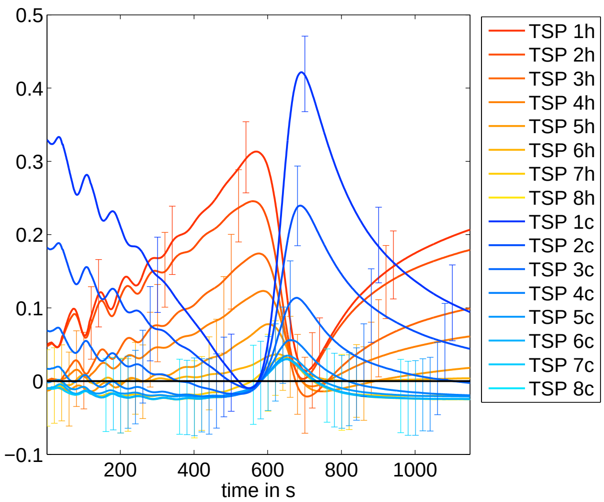

Fig. 3: *Sequential order 1 sensitivity indices with BC_a error bars*

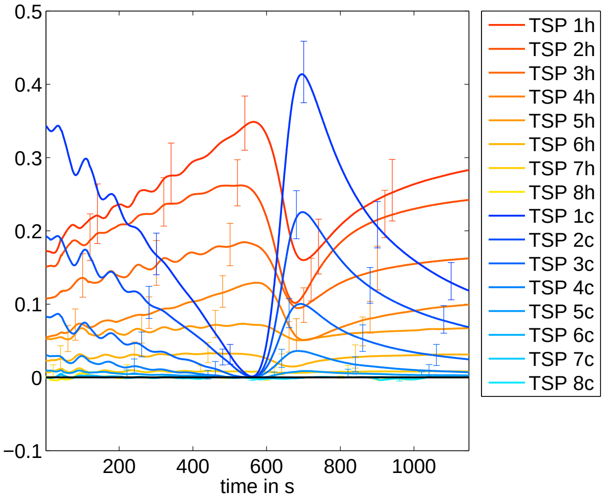

Fig. 4: *Sequential total sensitivity indices with BC_a error bars*

Four phases are discernible during the transient: *i)* from 0 to 250 s, top hot and cold TSPs prevail with some combined effects for hot leg; *ii)* from 250 s to 650 s, cold TSP have negligible impact and hot combined effects decrease; *iii)* from 650 s to 900 s, the three highest cold TSPs predominate while most of the hot impact is on the form of combined effects; *iv)* from 900 s until end time, hot leg order 1 effects catch up with combined effects while cold leg indices decline again.

There are a few negative order 1 indices but their error bars are roughly centred on the baseline. The corresponding total indices are estimated much more precisely and have near zero values so it is legitimate to consider these indices to be null. On the whole, BC_a intervals are shorter for total indices than for order 1 indices. Their lengths are approximately proportional to the value of the total indices while even very small indices are estimated with substantial uncertainties for order 1 indices.

The different behaviour of the hot and cold legs and the different phases observed during the transient will allow for refinement of the curves comparison methodology. An important result is that only the 4-5 highest TSP of the hot leg and the 3 highest of the cold leg have substantial effect during a significant time interval. From the SG safety analysis point of view, this would imply that lower TSP clogging has little or no effect on SG dynamic behaviour as far as the pressure field is concerned. They might however alter the mass repartition in the SG by modifying the mass fraction of the mixture. For the dynamic behaviour based diagnosis method, this means that no accurate clogging diagnosis can be made about the lower TSP if the WRL is chosen as informative variable.

3.2 Principal Components

A PCA has been conducted on the curves from the $\boldsymbol{\xi}$ sample and the two samples constructed by 'resampling' with $\boldsymbol{\xi'}$ all together after subtracting from each curves its temporal mean and from each variable $WRL(t)$ the mean of the WRL values at time t. The first ten PCs obtained with 'raw' variables are represented on figure 5 and the first ten PCs obtained after normalizing the variables are represented on 8. The first two PCs account for more than 99.9 % of the overall variance: the first (in dark blue) accounts for 95.54 % and the second (in green) for 4.40 %. Thus the choice of the first two loadings as output variables is straightforward. A resemblance with the order 1 and 2 Legendre polynomials can be noted. Legendre polynomials is a base of function commonly used for projection of functional data (Campbell et al., 2005). Subsequent PCs (in light blue) a rather disorderly and do not look like any general feature of the curves bundle apart from the oscillations in the beginning that have been identified as numerical artefacts. Discarding these PCs is a means to increase the signal-to-noise ratio in addition to radically reduce the dimension of the model output.

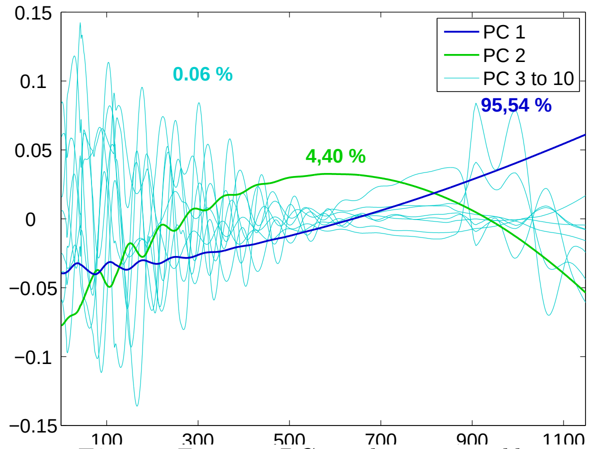

Fig. 5: *First 10 PCs with raw variables*

A convenient way to interpret the physical meaning of the PCs is to represent them as perturbations of the mean of the curves used for the PCA. On figure 6 the solid black line is the mean of the curves bundle while the '-' and '+' lines are respectively the mean minus or plus k times the first PC ; the value of k is arbitrarily chosen but identical for the two PCs. Figure 7, represent the effect of the second PC on the shape of the response curve with the same conventions. The first PC increases the global slope of the curve by swivelling it round a fixed point around time 650 s. The second PC increases the curvature by 'bending' the curve with two fixed points at times 250 s and 900 s. The fixed points of the perturbation of the mean by the first two PCs roughly corresponds to the time limits of the four phases described in section 3.1.

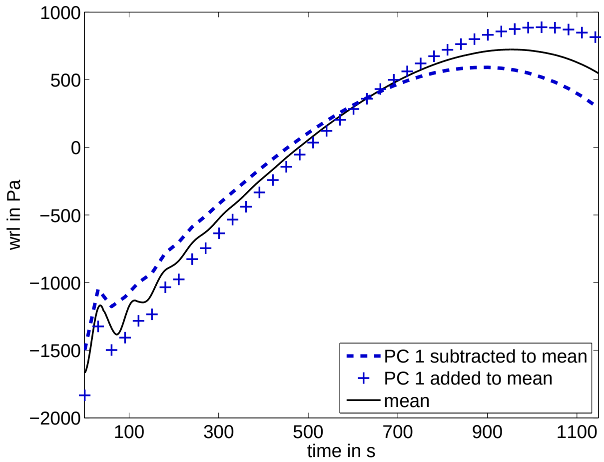

Fig. 6: *Effect PC 1 with raw variables represented as perturbations of the mean*

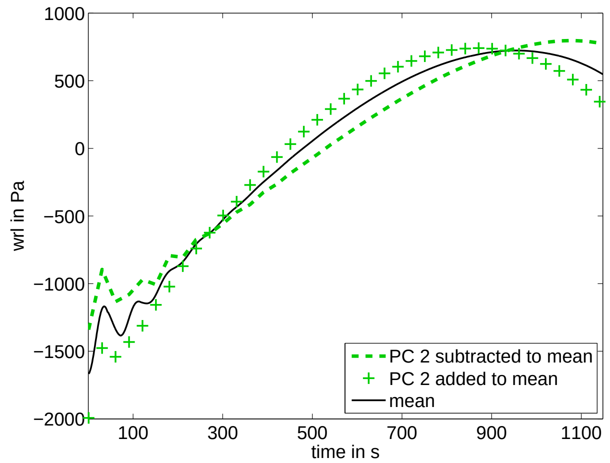

Fig. 7: *Effect PC 2 with raw variables represented as perturbations of the mean*

The result of the PCA conducted with centred reduced variables is presented on figure 8 with the same colour convention as in figure 5. Once again, the first two loadings are designated output variables as the first PC accounts for 83.83 % and the second for 16.11 %. Using standardized PCA does not enable for simple interpretation of the PC but the first two PCs still have distinct shapes (PC 1 and PC 2 are respectively antisymmetric and symmetric with respect to the axis $t = 650s$) while the subsequent ones are chaotic.

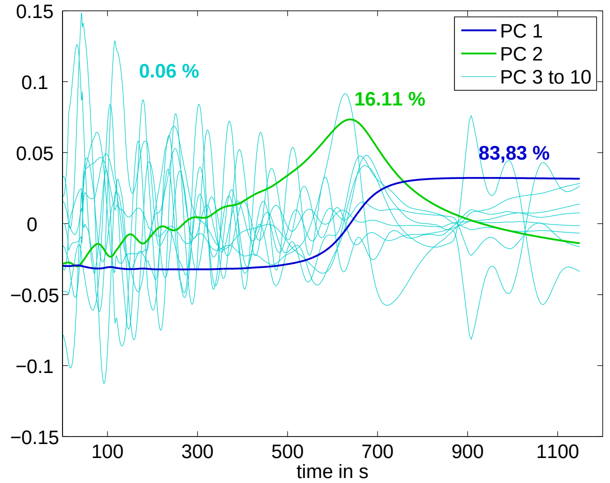

Fig. 8: *First 10 PCs with centred reduced variables*

For the sensitivity analysis, standardized PCA (figure 8) has been preferred to plain PCA (figure 5 in this study. The first reason for this choice, is that non-normalized PCs are subjected to a scale effect referred to in section 2.4.2. The first PC is the one that tunes the maximum reached in the terminal part of the curves (figure 6). Thus, its important eigenvalue (95.54 %) mostly reflects the variations of amplitude in the response curves which biases the estimation of the relative importance of the PCs. The second reason is that the relatively small eigenvalue of the second variance-PC (4.40 %) means that the corresponding effects to be estimated through sensitivity analysis would be also small and thus uncertainties would be more important: the approximations stated by equations (13) and (14) are erratic when the seconds members of the terms in the sums are too small.

3.3 Sensitivity indices of the reduced dimension model output

The results of the sensitivity analysis conducted on the first two loadings obtained through standardized PCA are displayed on figures 9 and 10 respectively. The TSPs are lined up on the ordinates in ascending order (TSP 1 is the lowest in the SG) with the hot leg TSP at the bottom and the cold leg ones above. For each TSP, two bars represent the corresponding order 1 and total indices: the bar above in lighter hue represents the total index while the other represents the order 1 index. The difference in length between the two bars indicates combined effects. The error bars indicate the bounds of the BC_a confidence intervals.

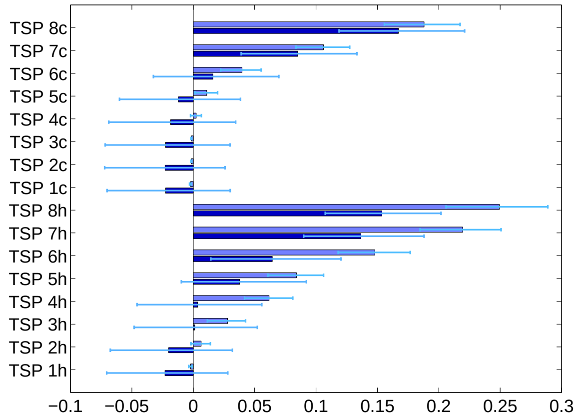

Fig. 9: *PC 1 total and order 1 indices with BC_a error bars*

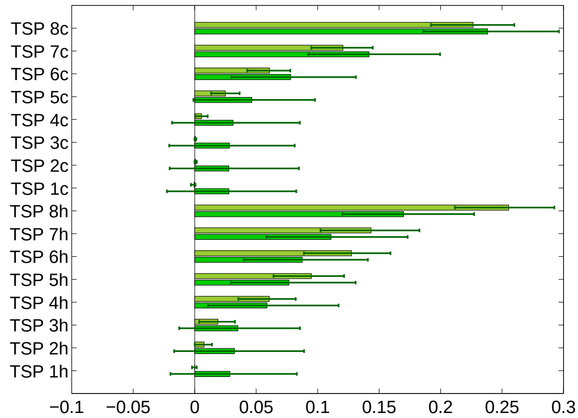

Fig. 10: *PC 2 total and order 1 indices with BC_a error bars*

As for the sequential indices, total indices have short confidence intervals and their length is proportional to the value of the indices while order 1 indices have longer confidence intervals of constant size. A few order 1 indices for PC 1 are negative. The same reasoning as in section 3.1 leads to consider them as null.

All the PC 2 couples of indices of cold leg TSPs and the first of hot leg TSPs are in wrong order: the order 1 indices exceed the total indices. However, the values of the total indices are always comprised in the confidence intervals of order 1 indices and the confidence intervals of the total indices are small. Thus, it seems legitimate to assume that interactions can be neglected for these variables and that order 1 indices actually equals total indices.

Total sensitivity indices for PC 1 are very low for TSPs 1 to 3 of the hot leg and 1 to 5 of the cold leg. Order 1 indices of the cold leg nearly equals total indices; as confidence intervals are quite large for the order 1 indices, combined effects are considered negligible. There are substantial combined effect TSP 6 to 8 of the hot leg and and probably for TSP 4 and 5 but for these last two the important uncertainties on order 1 indices prevent from being positive. As for sequential indices, the higher in the SG a TSP is, the higher are its sensitivity indices.

The ratio of the two PC's eigenvalues is 5, so the most important effects are those related to PC 1 but those related to PC 2 are not negligible. Sensitivity indices of PC 2 display the same features as those of PC 1: same 'important' TSPs, same ranking of the indices and combined effects present only on the hot leg.

Eventually, the result of the sensitivity analysis suggests a reduction of the model input dimension from 16 to 9 because the lower TSPs total indices are very low. With the choice of the WRL measurement as informative variable, clogging ratios of the lower TSPs cannot be estimated by dynamic behaviour analysis. However, clogging mostly affects the SG by increasing the pressure losses and the WRL is an 'integration' of all the pressure losses along the height of the SG. Thus, it can be concluded that the impact safety of clogging of the lower TSP is very limited.

4 Conclusion

In order to achieve better understanding of the effect of clogging on SGs dynamic behaviour and to assess the potential of diagnosis method based on analysis of this behaviour through simulations, a sensitivity analysis has been conducted on a mono-dimensional SG model.

A method based on the ANOVA-decomposition and a Monte Carlo computation scheme was used to compute order 1 and total sensitivity indices for each half-TSP. Sequential indices were first computed and exhibited different behaviour for the hot and cold leg as well as four distinct phases during the transient. More compact indices could be computed by conducting a PCA beforehand: the first two PCs contain most of the available information and represent two modes of alteration of the SG behaviour by clogging. The model output could then be reduced to the first two sets of loadings and two sets of sensitivity indices have been computed.

The sensitivity indices indicated that lower TSPs clogging had little or no influence on the WRL response of the SG. This means that no diagnosis can be made about it through the dynamic behaviour analysis method with the WRL as informative variable. However, this also suggests that the influence of the lower TSP clogging on the overall dynamic behaviour of the SG is of second order as the WRL is an accurate indicator of it.

BC_a confidence intervals were used to check the validity of the conclusions drawn from the computed sensitivity indices. They allowed to validate the methodology presented in this study.

There are substantial combined effects for the upper hot leg TSPs only. Thus, although being not exhaustive, the present study can be considered as a thorough analysis of the effect of TSP clogging on SG dynamic.

A new version of the SG model is currently under development at EDF. Indeed, an important hypothesis is being made when modelling the SG exchanging part by two disconnected channels: this amounts to neglecting all transport phenomena between the two legs. In the new version, transport phenomena between these two channels are taken into account by connections whose pressure drop coefficients are being tuned so that corresponding flow rates at nominal power are equal to those computed with the certified EDF 3D CFD code THYC. These transversal connections are expected to impact the sensitivity indices, probably by increasing the amount of combined effects between the hot and cold legs.

References

G.E.B. Archer, A. Saltelli, and I.M. Sobol'. Sensitivity measures, ANOVA-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58:99–120, 1997.

K. Campbell, M.D. McKay, and B.J. Williams. Sensitivity analysis when model outputs are functions. In K.M. Hanson and F.M. Hemez, editors, *Sensitivity Analysis of Model Output (SAMO 2004)*, pages 81–89. Los Alamos National Laboratory, March 2005.

J.G. Collier and J.R. Thome. *Convective Boiling and Condensation*. Oxford University Press, 1996.

B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, 1993.

T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

I.T. Jolliffe. *Principal Component Analysis, 2^{nd} edition*. Springer, 2002.

M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, and H. Monod. Multivariate global sensitivity analysis for dynamic crop models. *Field Crops Research*, 113:312–320, 2009.

M. Midou, J. Ninet, A. Girard, and J.-M. Favennec. Estimation of SG TSP blockage: Innovative monitoring through dynamic behavior analysis. In *18^{th} International Conference on Nuclear Engineering ICONE18*, may 2010.

I.M. Sobol'. Sensitivity estimates for nonlinear mathematical models, in Matem. Modelirovanie, 2 (1)(1990) 112 – 118. *English Transl.: MMCE*, 1(4), 1993.

I.M. Sobol'. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55:271–280, 2001.

G. Yadigaroglu. Basic models for two-phase flows. In *27^{th} Short Courses - Modelling and Computation of Multiphase Flows*. Swiss Federal Institute of Technologie (ETH), February 2010.