Sensitivity Analysis and Dimension Reduction of a Steam Generator Model for Clogging Diagnosis

Sylvain Girard[a,b,c,*], Thomas Romary[b], Jean-Melaine Favennec[a], Pascal Stabat[c], Hans Wackernagel[b]

[a] *EDF R&D, STEP P1C, 6 Quai Watier, 78400, Chatou, France*
[b] *Mines ParisTech, Centre de Géosciences/Équipe de Géostatistique, 35 avenue Saint Honoré, 77305, Fontainebleau, France*
[c] *Mines ParisTech, Centre Énergétique et Procédés, 60 Boulevard Saint Michel, 75006, Paris, France*

Abstract

Nuclear steam generators are subject to clogging of their internal parts which causes safety issues. Diagnosis methodologies are needed to optimize maintenance operations. Clogging alters the dynamic behaviour of steam generators and particularly the response of the wide range level (WRL - a pressure measurement) to power transients. A numerical model of this phenomenon has previously been developed. Its input variables describe the spatial distribution of clogging and its output is a discretization of the WRL dynamic response.

The objective of the present study is to characterize the information about the clogging state of a steam generator that can be inferred from the observation of its WRL response. A methodology based on several statistical techniques is implemented to answer that question. Principal component analysis reveals that clogging alters the WRL response mainly in two distinct ways. Accordingly, the output can be summarized into a vector of dimension 2. A sensitivity analysis is carried out to rank the input variables by magnitude of influence. It has shown that they can be divided into two

*Corresponding author. Tel.: +33130878050.
Email addresses: `sylvain-2.girard@edf.fr` (Sylvain Girard), `thomas.romary@mines-paristech.fr` (Thomas Romary), `jean-melaine.favennec@edf.fr` (Jean-Melaine Favennec), `pascal.stabat@mines-paristech.fr` (Pascal Stabat), `hans.wackernagel@mines-paristech.fr` (Hans Wackernagel)

groups corresponding to the two sides of the steam generator. Finally, sliced inverse regression is used to reduce the input dimension from 16 to 2. A sampling issue that arises when the input dimension is high is addressed.

The simplification of the original problem yields a diagnosis methodology based on response surfaces techniques.

Keywords: sensitivity analysis, principal component analysis (PCA), sliced inverse regression (SIR), bootstrap, steam generators, clogging.

1. Introduction

Pressurized light water nuclear power plants mainly consist of two separated water loops that exchange heat. The water from the primary loop goes first through the reactor where it is heated by the nuclear reaction and then through heat exchangers called steam generators (SGs) where it transfers heat to the water of the secondary loop. Steam exits the SGs by their upper opening and then flows through the turbines. A SG consists of a cylindrical tank (approx. 20 m high and 3m wide) that contains the secondary steam-liquid mixture. The primary water enters the SG at its bottom and goes through a bundle of U shaped tubes. Eight circular plates called tube support plates (TSPs) maintain the tube bundle. The tubes fit in circular holes drilled in the TSPs. These holes are surrounded by additional quatrefoil holes to let the secondary steam-liquid mixture flow through. A SG diagram can be found in figure 1.

SGs internal elements foul with iron oxides carried by the secondary feedwater. This causes clogging of the quatrefoil holes that induces safety issues. Means to estimate TSP clogging are needed to optimize maintenance operations. The pressure difference measured between the steam dome and the bottom of the SG is called the wide range level (WRL). Previous studies [1, 2] has shown that the shape of the WRL response curve to a power transient is altered by the clogging state of the TSPs and derived a diagnosis method that utilizing this link. The principle of the method is to compare measured response curves with simulations using with a mono-dimensional SG model. To assess the method's potential and make it reliable, it is necessary to characterise *how much information about the clogging state can be inferred from the WRL response.* This issue breaks down into three closely related questions:

- how does TSP clogging affect the shape of the WRL response?

- Are these effects different in nature and magnitude depending on the location inside the SG?

- What is the simplest formulation of input and output variables that captures these effects ?

The methodology presented here to answer these questions relies on computer intensive statistical methods. As the CPU time for a transient simulation with the 1D SG model is around 5 min, large samples of response curves corresponding to different clogging configurations can be generated.

Sensitivity analysis [3] and principal component analysis (PCA) [4] have been carried out to address the first two questions and the simplification of the output. The results suggested the use of a dimension reduction technique called sliced inverse regression (SIR) [5] to simplify the input. Along the process, bootstrap techniques were used to assess the robustness of the results and help with the interpretation. The SG numerical model and the statistical method that have been used are described in section 2. The results are presented and discussed in section 3.

2. Model and methods

2.1. Mono-dimensional steam generator model

The SG type examined 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 in figure 1.

The SG model has been developed with the Modelica language using Dymola software.

Its main elements are:

- primary fluid flow inside the U-tubes (single-phase flow);

- secondary fluid flow 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.

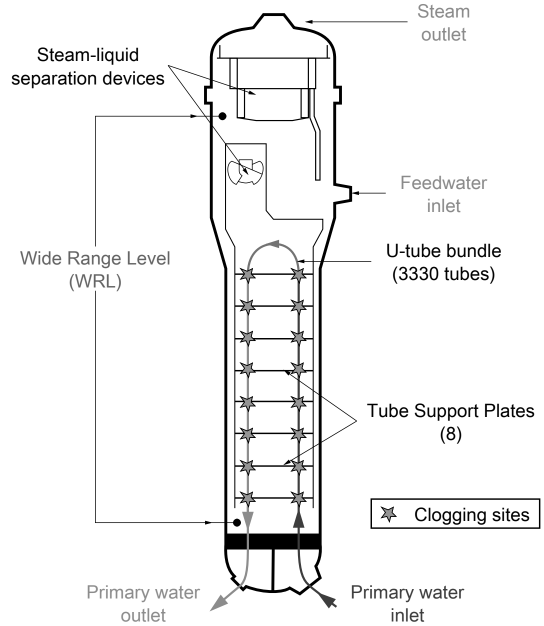

Figure 1: Westinghouse type 51 steam generator.

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 the primary fluid enters the SG) and one for the *cold leg* (*i.e* countercurrent exchanging side, where the primary fluid exits the SG). The exchanging channels are composed of 20 evenly spaced meshes. 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 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 computer intensive methods. Additional details about the model can be found in [2].

2.1.1. Model output definition

A power transient is simulated by varying the model boundary conditions. The transient used in the clogging diagnosis method is a roughly linear power decrease from nominal power to 40% of nominal power in an average time of 1148 s. It is modelled by a linear variation of primary inlet enthalpy and secondary outlet steam flow rate. The feed water flow rate is being determined by the control system. The model output is a vector, $\mathbf{w}$, of dimension 1148. Its coordinates are the values of the WRL at each 1 s time step.

2.1.2. Model input definition

There are 8 TSPs in the SGs under study and two 1D channels so the vector describing the clogging state, $\mathbf{x}$, is of dimension 16. Each of its coordinates is a *clogging ratio* associated to a half-TSP. Clogging ratios are defined as the ratio of the blocked area to the total area of the holes without clogging:

$$x_i = \frac{(\text{clogged area of half-TSP })_i}{(\text{total holes area of half-TSP })_i} \quad . \tag{1}$$

Clogging affects the WRL response by increasing the singular pressure drop at TSP crossings. In the model, the corresponding pressure drop coefficients depends on the clogging ratios through a function derived from experiments conducted on a 1:4 scale mock-up of TSPs and tubes [6].

2.1.3. Preliminary analysis

The singular pressure drop at a TSP crossing increases with the clogging ratio and steam fraction and decreases with the pressure of the steam-liquid mixture. The pressure is nearly the same in the two legs and it decreases as the secondary mixture rises inside the SG. The steam fraction equals zero at the bottom of the SG (liquid alone) and increases as the fluid rises and gets heated by the tubes. Its increase is sharper on the hot leg. From this, clogging is expected to have a greater impact in the hot leg than in the cold leg and in the higher parts of the SG than in the lower.

2.2. Sensitivity analysis of a functional output model

Sensitivity analysis studies how perturbations of the model input variables generate perturbations on its output variables. Here, general information about how does TSP clogging affects the WRL response is sought without any particular clogging configuration in mind. Hence, a *global* sensitivity analysis method [3] has been used. It consists in estimating sensitivity indices called Sobol' indices through a Monte Carlo computation scheme. They are presented in section 2.2.1. A preprocessing issue is addressed in section 2.2.2. Sensitivity analysis is usually applied to univariate or small dimensional output models, hence reduction of the output dimension was needed. Section 2.2.3 details how a convenient projection basis can be constructed using PCA. Eventually, section 2.2.4 describes how the validity of the results can be assessed with bootstrap confidence intervals.

2.2.1. Sobol' indices

Let us first derive Sobol' indices for a univariate output model. Let f be a function that represents the model, $\mathbf{x}$ the input vector of size n and y the scalar output.

$$\begin{aligned} f\ :\ & \mathbb{I}^n \to \mathbb{R} \\ & \mathbf{x} \mapsto y = f(\mathbf{x}) \end{aligned} \tag{2}$$

The input can be scaled to take values in $[0,1]$ so $\mathbb{I}^n$ denotes the n-dimensional unit hypercube.

ANOVA-representation. Assuming f is an integrable function, consider the following decomposition,

$$f(\mathbf{x}) = f_0 + \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} f_{i_1\ldots i_s}(x_{i_1},\ldots,x_{i_s}) \quad , \tag{3}$$

where f_0 is a constant and the f_{ij} are functions of subsets of (x_i). The double sum means that there is a function $f_{i_1\ldots i_s}(x_{i_1},\ldots,x_{i_s})$ for each 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' [7] has shown that under the following condition on the summands of (3),

$$\int_0^1 f_{i\ldots j}(x_i,\ldots,x_j)\,\mathrm{d}x_k = 0 \quad \text{for} \quad k = i_1,\ldots,i_s \quad , \tag{4}$$

the decomposition exists and is unique. It is then called the ANOVA-representation of f. It follows from condition (4) that the summands in (3) are orthogonal and can be expressed as integrals of f.

Order 1 Sobol' indices. If f is square integrable, then the $f_{i_1\ldots i_s}$ are also square integrable. Squaring and integrating (3) raises

$$\int_0^1 f^2(\mathbf{x})\,\mathrm{d}\mathbf{x} - f_0^2 = \sum_{s=1}^{n} \sum_{i_1<\cdots<i_s}^{n} \int_0^1 f^2_{i_1\ldots i_s}\,\mathrm{d}x_{i_1}\ldots\mathrm{d}x_{i_s} \quad . \tag{5}$$

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

$$D = \int_0^1 f^2 \, \mathrm{d}\mathbf{x} - {f_0}^2 \tag{6}$$

and

$$D_{i_1 \dots i_s} = \int_0^1 f^2_{i_1 \dots i_s} \, \mathrm{d}x_{i_1} \dots \mathrm{d}x_{i_s} \quad , \tag{7}$$

and the following equality holds:

$$D = \sum_{s=1}^{n} \sum_{i_1 < \dots < i_s}^{n} D_{i_1 \dots i_s} \ . \tag{8}$$

In other words, D measures the variability due to variations of all the input variables while $D_{i_1 \dots i_s}$ represents the variability caused by variations of the variables from the subset $(x_{i_1}, \dots, x_{i_s})$. Equation (8) 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 Sobol' index of a subset of variables $(x_{i_1}, \dots, x_{i_s})$ by the following ratio,

$$S_{i_1 \dots i_s} = \frac{D_{i_1 \dots i_s}}{D} \quad , \tag{9}$$

where s is called the order of the index. Order 1 Sobol' indices, $S_i = {}^{D_i}/_{D}$, measure the influence of each half TSP clogging ratio *alone* while higher order indices measure the interactions. With 16 input variables there are already 120 order 2 indices. Estimating Sobol' indices requires numerous model evaluation so higher order indices were not computed.

Total Sobol' indices. the input variables are strongly physically linked so completely ignoring interactions could be misguiding. As a palliative, consider the sum D_i^{tot} of the variances caused by all subsets that include a given variable x_i. Dividing this quantity by D, the overall variance, one defines total Sobol' indices, S_i^{tot} :

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

The difference between the total index and the order 1 index of a given variable represents its interactions with other variables.

A simple calculation [8] shows that D_i^{tot} and the variance $D_{\neg i}$ caused by the subset of all variables except x_i, sum up to D :

$$D_i^{tot} = D - D_{\neg i} \quad . \tag{11}$$

Hence, each total Sobol' index can be deduced from the estimation of the variance of one subset of variables.

Computation Scheme. Sobol' [7] has demonstrated that the variances corresponding to subsets of variables can be expressed as integrals. Monte Carlo estimates of these integrals are provided in [9]. Estimating order 1 and total indices of n input variables with a Monte Carlo sample size of N requires $(2n + 1) \times N$ model evaluations.

In this context, standard Monte Carlo relying on pseudo-random numbers is only moderately effective. This is due to the tendency of pseudo-random sequences to aggregate into clusters which is detrimental especially in high dimension. Substantial improvement is achieved by using a quasi-Monte Carlo procedure based on low-discrepancy uniform sequences such as Sobol' sequences [10]. Here a Sobol' sequence has been used to generate the samples, following the procedure prescribed by Sobol' [9]. The input vector coordinates vary from 0 to 0.65 which covers most of practical clogging cases.

2.2.2. Preprocessing of the output

The increased pressure drop due to clogging alters both the full power 'static' values of the WRL and its dynamic behaviour. The 'static' value is presently used for cursory diagnosis of clogging. Examining the dynamic response is meant to retrieve more detailed information and to sidestep the issue of sensor bias. As the range of variation of the WRL 'static' value over the years of plant operation is large compared to the dynamic variations of the WRL during a power transient, it has been necessary to pre-process the data by removing the 'static' value trend. Indeed, Sobol' indices computed on unprocessed output reflect only the variance due to differences in WRL initial value.

A straightforward corrective action would be to subtract from each curve its initial value. However, this would arbitrarily eliminate the variance of the first sequential variables. Subtracting a constant is a crude correction and choosing this constant to be equal to the value taken by the initial variable concentrates all available accuracy on the beginning of the curves.

Let $w : t \mapsto w(t)$ be the WRL response function. For a given time t_0, one can write a Taylor expansion of w of the form

$$w(t) = w(t_0) \, + \, w'(t_0)(t - t_0) \, + \, w''(t_0)\frac{(t - t_0)^2}{2} \, + \cdots \quad . \tag{12}$$

Averaging (12) for t from 0 to 1148 makes the temporal mean appear as the first term of the 'average' expansion. Subtracting the temporal mean instead of the initial value is a means to distribute the error along the time interval. In this way no assumption is made *a priori* about the most informative part of the response curves.

Five sample WRL response curves are displayed in the left panel of figure 2. The difference in 'static' value can be appreciated by the difference in initial value. However, differences in the shape of the curves are difficult to distinguish. In the right panel of figure 2, the same curves are presented with their temporal mean subtracted. Their shapes appear more contrasted. For instance, the circle-marked and x-marked curves are approximately equidistant from the square-marked one in the left graph but the right graph shows that the circle-marked curve has a much more similar shape.

The validity of the subtraction of a constant has been investigated using the PCA results in section 3.1.2.

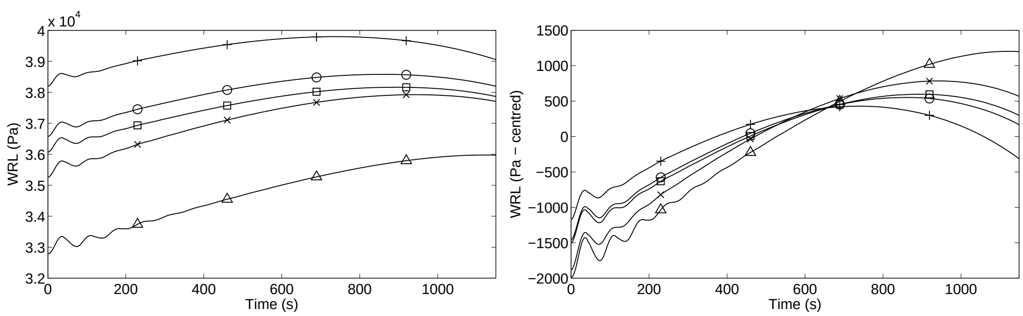

Figure 2: Simulated WRL response curves before and after subtraction of temporal means.

2.2.3. Reduction of output dimension

The most straightforward implementation of sensitivity analysis consists in considering the value of the WRL at each time step as distinct output variables. However this yields a large number of indices which makes the ranking of the input variables and the analysis of their interactions cumbersome. Moreover, this approach does not take into account the functional nature of the output because it is blind to the high correlation of the sequential variables. One way to tackle this issue is to expand the time series onto an appropriate orthogonal basis [11–13].

PCA is a simple method to obtain such a basis directly from the data. The optimization criterion used in PCA is the maximization of variance along the directions of the basis. As our aim is to attribute shares of overall variances to input variables, using a variance based method for dimension reduction makes sense.

Principal components. The principle of PCA is to progressively build an orthogonal projection basis by adding directions so that the spanned space fits the data the most adequately. Considering the output vector, $\mathbf{w} = (w(t))$, as a random vector leads to define principal components (PCs) as ordered linear combinations of the original variables, $(w(t))$, that are orthogonal and have maximum variance [4]. For a set of p variables, up to p PCs can be found. Coordinates along each PC are called *scores*; they constitute a new set of variables, each representing a smaller share of total variance than the previous one. The sample used for the sensitivity analysis can be used to estimate the PCs and their scores: the eigenvectors of its empirical covariance matrix are the PCs and their variances are the corresponding eigenvalues. The scores are then easily obtained by projection.

A common practice in PCA is to use centred reduced variables to avoid scale problems. For instance, if a variable lies in the interval $[10^3, 10^4]$ while the others vary from 0 to 10, it could cause most of the variance of the dataset while varying, relatively, as much as the others. Both types of PCA (with 'raw' variables and centred reduced variables) have been used in this study, leading to different types of interpretation.

Sensitivity analysis on scores. Eigenvalues usually drop quite quickly in magnitude. PCs of low eigenvalue describe small fluctuations in the dataset and one can neglect them without losing information. Keeping only the r most prominent PCs allows to sum up the effect of clogging on the shape of the WRL response in a manageable number of variables.

The WRL response curves lie in a p-dimensional space. Excluding the $p-r$ PCs of lowest variances comes down to selecting the r-dimensional subspace that most nearly encloses the data, the curves being very 'flat' along the directions left aside. Then, Sobol' indices can be computed on PC scores in exactly the same way as sequential indices. In addition of being less numerous, these new Sobol' indices present the advantage of being linked to PCs whose shapes can be interpreted. A more sophisticated approach using the notion of generalized sensitivity indices has been proposed by [13]. Here, the small number of PCs with a substantial eigenvalue made it unnecessary.

2.2.4. Assessing indices validity

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. In addition, the chosen computation scheme sometimes induces aberrations, such as slightly negative indices or sums of indices that exceed one, due to slow convergence of the Monte Carlo estimates; confidence intervals allow to decide if these irregularities can be overlooked or if larger samples should be used.

For each sensitivity index S, an estimator $\hat{S}$ has been computed. Building a confidence interval consists in finding $\hat{S}_{lo}$ and $\hat{S}_{up}$ so that the two events $S < \hat{S}_{lo}$ and $S > \hat{S}_{up}$ have both a given small probability. As little is known about the distributions involved, bootstrap methods are particularly indicated as they are robust and distribution free.

Bootstrap confidence intervals. The general idea behind bootstrap is to draw conclusions about a given estimator by using the empirical distribution upon which the estimator is based. The estimator $\hat{S}$ is linked to the sample used to compute it, $\boldsymbol{\zeta}$, by a function ϕ: $\hat{S} = \phi(\boldsymbol{\zeta})$. A *bootstrap sample* $\boldsymbol{\zeta}^{(b)}$ is obtained by drawing uniformly with replacement from $\boldsymbol{\zeta}$. For each bootstrap sample, a *bootstrap replication* $\hat{S}^{(b)} = \phi(\boldsymbol{\zeta}^{(b)})$ is computed in the same way as the estimator. It is 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 percentile intervals but include a correction of bias and an *acceleration* that compensates for variation of the standard error of S with the value of S. These two corrections consist in shifts of the percentiles finally chosen from the empirical distribution.

Details about bootstrap confidence intervals and their derivation can be found in the book by Efron and Tibshirani [14].

2.3. Sliced inverse regression

The dimensionality of the model input was chosen on a physical basis: it is the most detailed description of clogging that can be reasonably described by a 1D model. However, as the results of the sensitivity analysis will show, variations in shape of the response can be satisfactorily accounted by a smaller number of variables. Discarding irrelevant variables is necessary to ensure that the diagnosis method is only used in its applicability domain. It also reduces the size of the space to be sampled and allows for the construction of meaningful graphic representations.

Among dimension reduction techniques, sliced inverse regression (SIR) [5] bears several practical advantages. It is quite robust, easy to implement and based on a very generic model. A generic formulation of the SIR method is given in section 2.3.1. For SIR to be really useful, it is necessary to determine the dimension the input space can be reduced down to without loss of information. A bootstrap technique intended to address this issue is presented in section 2.3.2. Finally, section 2.3.3 details what adjustments are needed in the multivariate output case.

2.3.1. SIR principle

The basic idea behind SIR is to find a limited number of linear combinations of the predictors that are sufficient to retrieve the information from the regression.

The following model is assumed,

$$y = g(\boldsymbol{\beta}_1'\mathbf{x}, \dots, \boldsymbol{\beta}_q'\mathbf{x}, \epsilon) \quad , \tag{13}$$

where $(\boldsymbol{\beta}_k)$ is a family of unknown vectors, g is an unknown function taking value in $\mathbb{R}^{q+1}$ and ϵ is independent from $\mathbf{x}$.

The space $\mathrm{Span}[(\boldsymbol{\beta}_k)]$ is called the *efficient dimension reduction* (e.d.r.) subspace and its elements e.d.r. *directions*. This terminology emphasizes the fact that g is arbitrary and that the $\boldsymbol{\beta}_k$ themselves are not identifiable.

The inverse regression curve $\mathbb{E}(\mathbf{x}|y)$ lies in $\mathbb{R}^n$. If model (13) holds, it stays always close to a q-dimensional subspace. Appropriate conditions on the distribution of $\mathbf{x}$ will ensure that it falls into the e.d.r. subspace.

Confining the inverse regression curve to the e.d.r. subspace. Consider the following condition,

Condition 1 (Linearity)**.** *For any* $\mathbf{b}$ *in* $\mathbb{R}^n$*, the conditional expectation* $\mathbb{E}(\mathbf{b}'\mathbf{x}|\boldsymbol{\beta}'_1\mathbf{x}, \ldots, \boldsymbol{\beta}'_q\mathbf{x})$ *is linear in* $\boldsymbol{\beta}'_1\mathbf{x}, \ldots, \boldsymbol{\beta}'_q\mathbf{x}$*.*

Such a condition is difficult to check because it involves the unknown $(\boldsymbol{\beta}_k)$. It is however satisfied if $\mathbf{x}$ has an elliptically symmetric distribution, such as the Gaussian distribution [5].

Theorem 1. *Under model (13) and condition 1, the centred inverse regression curve* $\mathbb{E}(\mathbf{x}|y) - \mathbb{E}(\mathbf{x})$ *lies in the subspace spanned by* $(\Sigma\boldsymbol{\beta}_\mathbf{k})$*, where* Σ *is the covariance matrix of* $\mathbf{x}$*.*

Hence, substituting $\mathbf{x}$ by its standardized version implies under condition 1 that the inverse regression curve is contained in the e.d.r. subspace.

SIR algorithm. The model can produce a sample of WRL responses for an arbitrary distribution of $\mathbf{x}$. This sample can then be used to estimate, first the inverse regression curve and then, in the same manner as in section 2.2.3, the q-dimensional subspace, that most adequately contains it.

The following algorithm given by Li [5] has been used:

1. Standardize $\mathbf{x}$ using its empirical covariance matrix $\widehat{\Sigma}$:
   $\tilde{\mathbf{x}}_i = \widehat{\Sigma}^{-1/2}(\mathbf{x}_i - \bar{\mathbf{x}})$.

2. Divide the range of variation of y into H slices, $I_1, \ldots, I_H$, each containing a proportion p_h of the N observations.

3. Compute the slice averages, $(\hat{\mathbf{m}}_h)$, of the input individuals:
   $\forall h \in \{1, \ldots, H\},\ \hat{\mathbf{m}}_h = p_h \sum_{\{i|y \in I_h\}} \tilde{\mathbf{x}}_i$.

4. Compute $\widehat{V} = \sum_{h=1}^{H} p_h\, \hat{\mathbf{m}}_h \hat{\mathbf{m}}'_h$.

5. Find $(\hat{\boldsymbol{\eta}}_k)$, the family of eigenvectors of $\widehat{V}$ sorted by decreasing eigenvalues.

6. Output $(\hat{\boldsymbol{\beta}}_k) = (\widehat{\Sigma}^{-1/2}\hat{\boldsymbol{\eta}}_k)_{k \in \{1,\ldots,q\}}$.

In the last step of the algorithm, the $n - q$ eigenvectors with the smallest eigenvalues are left aside. It is necessary to determine the dimension of the e.d.r. subspace in order to avoid missing information or including spurious directions.

2.3.2. e.d.r. subspace dimension determination

Li [5] proposed a statistical test to determine the dimension q. Unfortunately, it relies on an assumption of Gaussian distribution for $\mathbf{x}$. As a non Gaussian distribution has been investigated here, the bootstrap approach devised by Liquet and Saracco [15] has been preferred.

Let B_K and $\widehat{B}_K$ be the matrices whose columns are respectively the vectors $(\boldsymbol{\beta}_k)$ and their estimators $(\hat{\boldsymbol{\beta}}_k)$ with k in $\{1, \dots, K\}$. Let P_K and hP_K be the Σ-orthogonal and $\widehat{\Sigma}$-orthogonal projectors onto the spaces spanned by these same vectors,

$$P_K = B_K(B'_K \Sigma B_K)^{-1} B'_K \Sigma \quad ; \quad \widehat{P}_K = \widehat{B}_K(\widehat{B}'_K \widehat{\Sigma} \widehat{B}_K)^{-1} \widehat{B}'_K \widehat{\Sigma} \quad . \tag{14}$$

The following risk function,

$$R_k = \frac{1}{k} \mathbb{E}\Big[\text{Trace}(P_k \widehat{P}_k)\Big] \quad , \tag{15}$$

expresses the closeness of the two subspaces: a value close to 1 indicates a good match while a value close to 0 reveals important differences.

A bootstrap estimate $\hat{R}_k$ of R_k can be formed as follows [14]: for a given bootstrap replication of the sample used to conduct the SIR, the plug-in estimator of R_k is

$$\widehat{R}_k^{(b)} = \frac{1}{k} \mathbb{E}\Big[\text{Trace}\Big(\widehat{P}_k \widehat{P}_k^{(b)}\Big)\Big] \quad . \tag{16}$$

Then ,for $\mathcal{B}$ bootstrap replications, the bootstrap estimate is

$$\widehat{R}_k = \frac{1}{\mathcal{B}} \sum_{b=1}^{\mathcal{B}} \widehat{R}_k^{(b)} \quad . \tag{17}$$

2.3.3. Multivariate output SIR

Several approaches have been proposed to adapt SIR to a multivariate context [16]. In this paper, the following adaptation of model (13) is adopted:

$$\mathbf{y} = g(\boldsymbol{\beta}'_1 \mathbf{x}, \dots, \boldsymbol{\beta}'_q \mathbf{x}, \epsilon) \quad , \tag{18}$$

where $\mathbf{y}$ stands for the multivariate output.

Then, building on what has been done for the sensitivity analysis in section 2.2.3, SIR has been carried out with the scores of the r selected PCs as output variables. Then, a method called Pooled Marginal Slicing (PMS) has been applied to the r-dimensional output [16].

Pooled Marginal Slicing principle. Applying the SIR algorithm up to step 4 to each of the r components of the multivariate output yields a set of weighted covariance matrices $(\widehat{V}_i)_{i\in\{1,\ldots,r\}}$. A convex combination with a set of weights (w_i) can be formed,

$$\widehat{V}_{pool} = \sum_{i=1}^{r} w_i \widehat{V}_i \quad . \tag{19}$$

The e.d.r. directions are finally estimated by executing the second half of the SIR algorithm with $\widehat{V}_{pool}$.

3. Results and discussion

3.1. Sensitivity analysis of the SG model

First, sequential indices are presented in section 3.1.1. Then the projection basis obtained by PCA is presented in section 3.1.2. It is compared to PCs obtained with plant data. Finally, section 3.1.3 details the 'compact' sensitivity indices computed with the reduced dimension output.

3.1.1. Sequential Sobol' indices

The size of the Monte Carlo samples has been fixed to 1000 so a total of 33000 transient simulations have been run for the sensitivity analysis.

Sequential order 1 and total indices are represented in figures 3 and 4. Indices are grouped in graphics by hot and cold leg variables. The shade of the curves corresponds to the height of the TSPs: light curves are associated to the lower TSP and dark ones to the higher. The error bars represent the bounds of the BC_a confidence intervals. Only a few of them are presented for readability but no discrepancies have been observed on the whole set. There are a few negative order 1 indices which is caused by lack of convergence of the Monte Carlo estimates. It seems legitimate to consider them as null because their error bars are roughly centred on the baseline and the corresponding total indices are all close to zero and have very short error bars.

Both sequential order 1 and total sets of indices display two sharp contrasting behaviours for each leg. The ranking 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 the preliminary analysis conducted in section 2.1.3. As stated in section 2.2.1, the difference between total indices and order 1 indices measures the amount of interaction. Comparison of figure 3 and 4 shows that there are only limited interactions and that they involve only the highest TSPs.

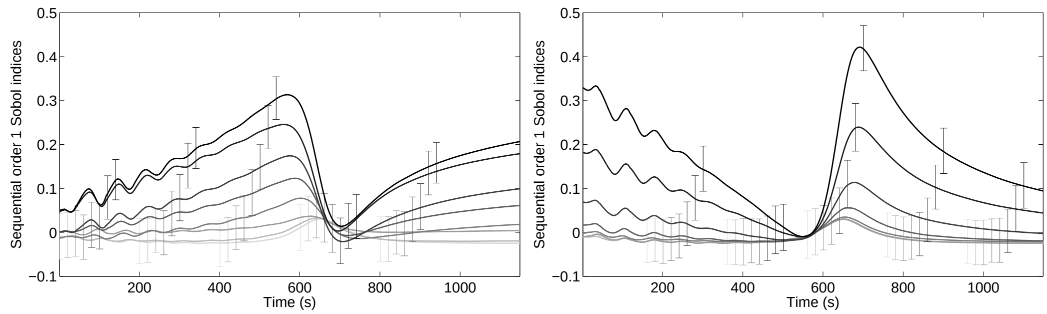

Figure 3: Sequential order 1 Sobol' indices (l. hot leg; r. cold leg).

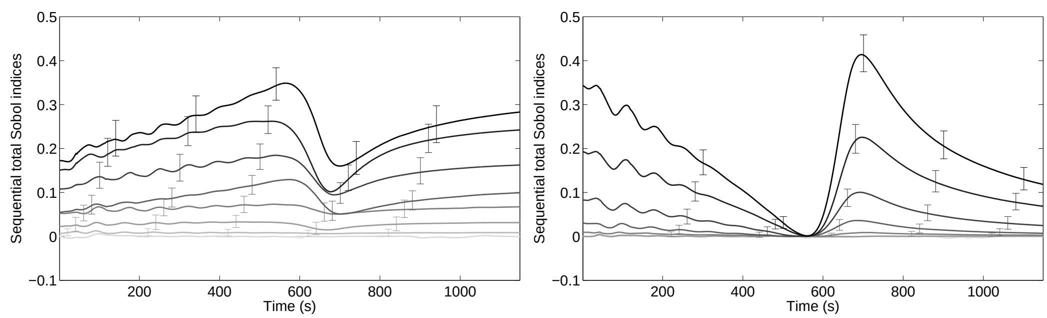

Figure 4: Sequential total Sobol' indices (l. hot leg; r. cold leg).

3.1.2. Dimension reduction of the model output

Sequential indices revealed that the impact of clogging changes qualitatively with the SG leg and quantitatively with the level of the TSPs. However, the large number of sequential indices makes it difficult to estimate precisely the impact of each TSP and the interactions. In order to reduce the output dimension, a 'raw' and a normalized PCA have been carried out. The resulting PCs have been compared to those obtained with plant data.

The first 10 PCs obtained with a uniform sample with 'raw' and normalized variables are displayed in figure 5. The normalized PCs in the right panel of figure 5 have been multiplied by the square root of their eigenvalue. Hence, it is actually the sequential correlation coefficients between the time steps variables and the PCs that are represented. In both cases, the first 2 PCs account for more than 99.9 % of the overall variance. It shows in the normalized variables graphic: the correlation coefficients of the next PCs almost do not departs from the baseline meaning that these PCs are only marginally correlated with the original variables. The low variance PCs can be seen in more details in the left panel because they all have an $\mathcal{L}_2$ norm equal to one. They are rather disorderly and do not look like any general feature of the curves except from the oscillations in the beginning that have been identified as numerical artefacts.

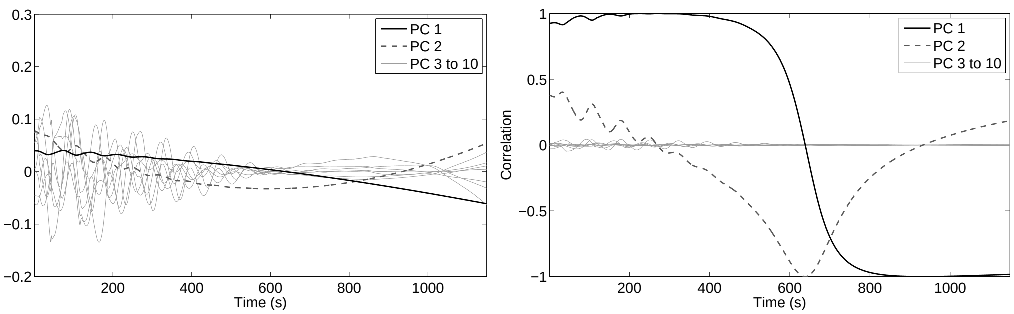

Figure 5: First 10 PCs obtained with the uniform sample (l. 'raw' variables; r. normalized variables)

The first 2 'raw' PCs are polynomials of degree 1 and 2. The first PC increases the global slope of the curves by spinning it round a fixed point around time 650 s. The second PC increases the curvature by 'bending' the curves with two fixed points at times 250 s and 900 s. The PC 1 correlation coefficients curve is S-shaped with 2 plateaux at $+1$ and -1 from approximately 0 s to 400 s and 800 s to 1148 s. In between there are two sharp inflexions. This means that the original variables of the beginning of the time interval are highly correlated with PC 1 while those at the end are highly anti-correlated. The PC 1 correlation coefficients curve is V-shaped and points towards 1 around 650 s. Only the time steps variables of the middle of the interval are substantially correlated with the PC 2.

PCA on measured data. A PCA has been carried out on 291 measured response curves from 5 EDF units. The 97 processed transients (there are 3 SGs per unit) spread over a period of 10 years and each unit has undergone a chemical cleaning at some points. Hence, the data include a wide array of clogging configurations, from very low clogging just after the chemical cleaning, to very high clogging just before.

The first 3 PCs obtained without preprocessing and the first 2 PCs obtained with the preprocessing described in section 2.2.2 are displayed in figure 6. On the left panel, the first PC is nearly a constant and its scores are proportional to the temporal mean of the curves. The PCs are orthogonal by construction so PC 2 and 3 are very similar to PC 1 and 2 from the right panel. This validates the chosen preprocessing. The PCs obtained with measurements are similar to those found with the simulations. This shows that the main effects of clogging on WRL are correctly represented by the model and that PCA is an appropriate tool to represent them.

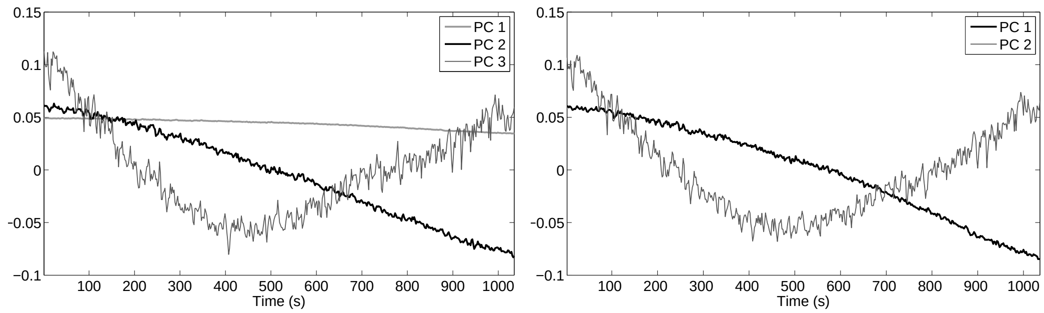

Figure 6: First PCs obtained on 291 measured response curves (l. before subtraction of the temporal mean; r. after subtraction of the temporal mean.

3.1.3. Sobol' indices of reduced dimension output

The first 2 PCs obtained in the previous section account for almost all of the variance of the sample. Comparison with plant data showed that they satisfactorily represent the effects of clogging on the WRL response. It is straightforward to select those 2 PCs to build a projection basis for the reduced dimension output sensitivity analysis. Sobol' indices computed with 'raw' and normalized PC scores did not differ fundamentally and only the former are presented here.

The results of the sensitivity analysis conducted on the first 2 sets of standardized PC scores are displayed in figure 7. Each couple of bars corresponds to a TSP. They are lined up from bottom to top in ascending order, hot leg first. The light bars represent total indices and the dark bars represent order 1 indices. The length difference of the two bars of a couple represents the interaction in which the input variable is involved. The error bars indicate the bounds of the BC_a confidence intervals.

As for the sequential indices, total indices have shorter 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 3.1.1 leads to

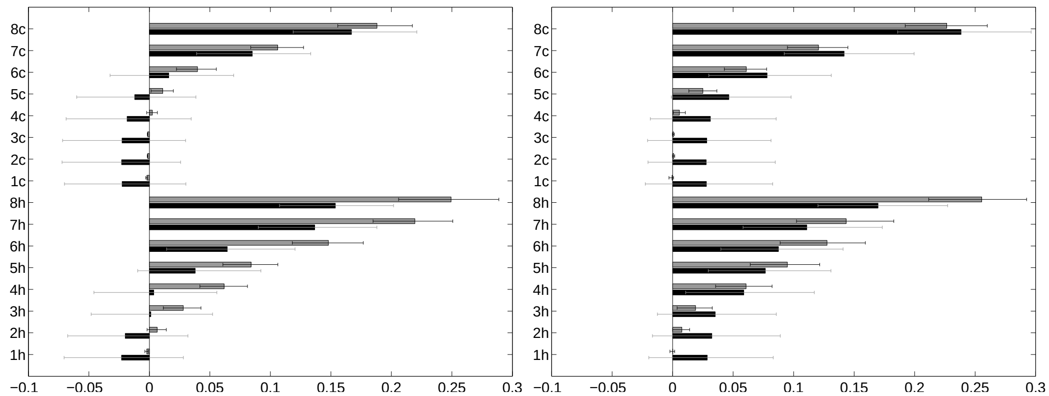

Figure 7: Order 1 and total Sobol' indices computed with PC scores (l. PC 1; r. PC 2).

consider them as null. There is another aberration: the PC 2 cold leg total indices are lower than the order 1 indices. However, the values of the total indices are always within the confidence intervals of order 1 indices and the confidence intervals of the total indices are small. Thus, it seems sensible to assume that these order 1 indices actually equal the total indices and that no interaction is involved here. The ranking of the indices is again in agreement with the preliminary analysis. On the whole, interactions are rather limited. Taking the confidence intervals into account, only TSP 4 to 8 on the hot leg and to a lesser extent TSP 7 and 8 on the cold leg seem to be involved in substantial interactions.

3.2. Dimension reduction of the SG model input

The sequential Sobol' index curves in figures 3 and 4 are almost proportional. This suggests that clogging of TSPs of a same leg affect the WRL response in a similar manner. In addition, sensitivity analysis on the PC scores showed that there are little input variable interactions. These observations give credibility to model (13) so SIR is well indicated to simplify the model input. The high dimension of the input raises a sampling issue. It is addressed in section 3.2.1. Then, results of PC-wise univariate SIR and multivariate SIR are detailed in section 3.2.2 and 3.2.3.

3.2.1. Note on sampling scheme

Gaussian sampling. A simple means to satisfy condition 1 is to choose a Gaussian distribution for $\mathbf{x}$. A sample of 10^4 response curves for clogging ratios following a multivariate Gaussian distribution of mean 0.65/2 and standard error 0.65/6 has first been simulated. A few individuals with negative or very high clogging ratios have been trimmed without affecting too much the elliptic symmetry of the distribution. The first two PCs obtained with this sample are similar in shape to those found with a uniform sample displayed in figure 5. However, the shares of explained variance are different: the ratio of the first eigenvalue to the second is much higher in the Gaussian sample case. The standard deviation of the sequential variables is also globally lower in the Gaussian case, especially in the middle of the time interval. This is due to the fact that the Gaussian sample covers a volume much smaller than the uniform sample. At best, the Gaussian sample can efficiently cover the hypersphere inscribed into the hypercube $[0, 0.65]^{16}$. This would not cause much trouble in low dimension, but here the hypercube looks more like a sea urchin than a cube: it has $2^{16} = 65536$ 'corners' having each a volume approximately 4.25 times higher than the volume of the inscribed hypersphere. The previous observations tend to show that the Gaussian sample is unable to capture what happens inside the 'corners' of the hypercube. Yet, sampling extensively the hypercube while preserving an elliptic contour for $\mathbf{x}$ is rendered difficult by its shape. Indeed, trimming and re-weighting a uniform sample, following for instance the guidelines of Cook and Nachtsheim [17], is unlikely to succeed because the probability that at least one individual out of a 10^4 size sample falls into the inscribed hypersphere is lower than 0.04!

Flexibility of the linearity condition. Condition 1 is actually weaker than elliptic symmetry and SIR can yield sensible results in cases that does not exactly comply with it. It has been shown by Diaconis and Freedman [18] that most low-dimensional projections from a high-dimensional data set are approximately Gaussian. Hall and Li [19] extended this result showing that low-dimensional projections of high-dimensional data are almost linear. As an illustration, a simulation example of e.d.r. directions correctly identified by SIR with a uniform sample in dimension 10 is given in the rejoinder of [5]. Here the dimension is higher and the data are relatively smooth because they are produced by a model so it can be expected that SIR would work in spite of a violation of condition 1. The bootstrap dimension determination method has been successfully tested with a strongly non elliptically distributed input [15]. A 10^4 size uniform sample of WRL response curves has been simulated in order to investigate the model's behaviour inside the 'corners'. The results obtained with this uniform sample are presented below.

3.2.2. Marginal slicing

Marginal slicing has been applied to the first two PCs of the data set. Using 'raw' or normalized PCA made but little difference so only the results with normalized PCA are presented here. The number of slices had also a very limited influence. Here, 33 slices of cardinal 303 have been used.

Bootstrap estimates of the risk function for the e.d.r. space dimension have been computed using 500 bootstrap replications. Corresponding box plots for the uniform samples are given in figure 8. In both plots, the mean of $\widehat{R}_k$ is first close to 1, then decreases steeply down to around 0.7 and eventually climbs up until it reaches 1 for $k = 16$. The variance of $\widehat{R}_k$ is close to 0 on the initial plateau, then it soars at the beginning of the drop in mean and eventually decreases regularly down to 0 as k increases up to 16. The increase in mean in the third part of the plots is a consequence of the growth of the basis. Additional directions progressively restrict the angular domain where the directions found with the bootstrap replications may differ from those found with the original sample. This shows through the progressive reduction of the variance of $\widehat{R}_k$ as k increases. When k equals 16, $\widehat{P}_k$ and the $\widehat{P}_k^{(b)}$ are proportional to the identity.

The dimension of the e.d.r. space is given by the highest value of k for which the mean of $\widehat{R}_k$ is nearly 1 and its variance nearly 0 [15]. Here, it is equal to 2 for both sets of PC scores. The first two directions found with the uniform sample are displayed in figure 9. The Gaussian sample yielded similar results but the directions were a little less monotonous which goes against physical reasoning. It was not able to retrieve the second e.d.r. direction with the PC 1 scores.

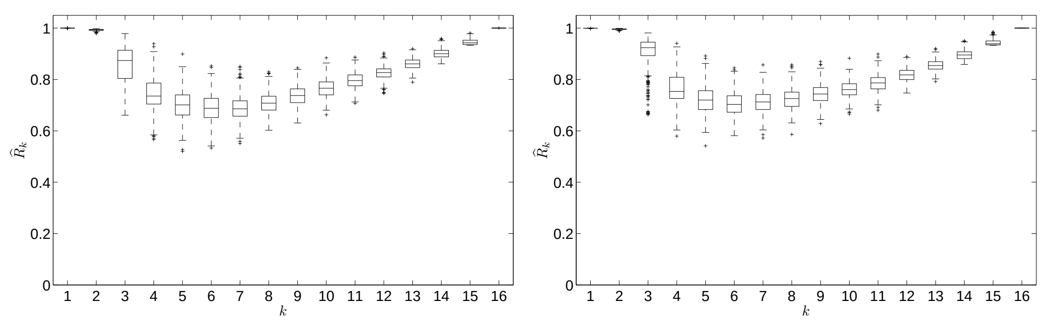

Figure 8: Box plots of $\widehat{R}_k$ values with uniform sample for the PC 1 (l.) and PC 2 (r.) scores.

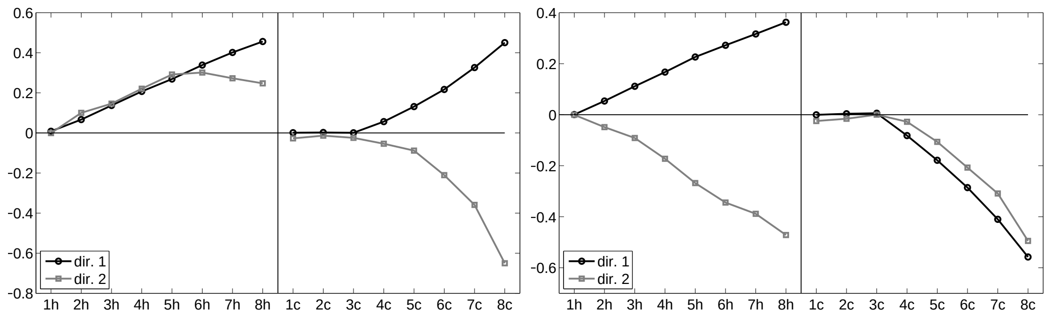

Figure 9: e.d.r. direction obtained with uniform sample for the PC 1 (l.) and PC 2 (r.) scores.

3.2.3. Pooled Marginal Slicing

The dimension of the e.d.r. space yielded by PMS (see section 2.3.3) with the two sets of PC scores is equal to 2 as can be seen on the left panel of figure 10. The right panel of figure 10 displays an orthogonal basis of the plane spanned by the first 2 directions found with the uniform sample. The vectors have been combined so that hot and cold legs are as separated as possible between. They are normalized to have a $\mathcal{L}_1$-norm equal to 1 so that the coordinates vary in the same range as clogging ratios. Using PMS made SIR more robust to changes in the sampling scheme. Indeed, the basis obtained with the Gaussian sample was nearly the same as the one displayed in figure 10. The $\widehat{R}_k$ values also indicated that the e.d.r. space is a plane but in a less obvious manner.

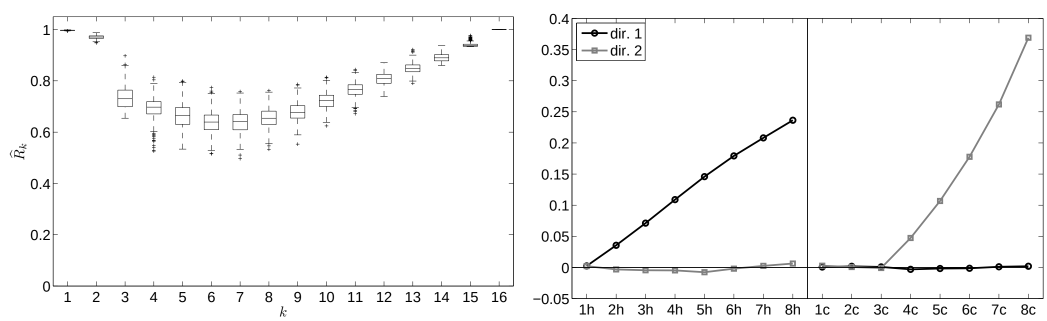

Figure 10: Box plots of $\widehat{R}_k$ values (l.) and e.d.r. space basis (r.) obtained by PMS with the uniform sample.

The two e.d.r. directions found correspond to weighted averages of the clogging ratios of each leg. This means that a clogging diagnosis based on WRL response curve analysis will consist of hot and cold average clogging ratios.

The results illustrate the fact that SIR can provide interesting results even when the linearity condition is not fully satisfied. When the input dimension is large, elliptically contoured distributions are unlikely to be able to efficiently cover the domain of interest. In such cases, uniform sampling is a straightforward alternative to Gaussian sampling and the bootstrap method proposed by [15] can be used to determine the e.d.r. subspace dimension.

4. Conclusion

A methodology combining several statistical techniques has been carried out with a 1D SG model. It allowed to characterize the information about the clogging state of a SG that can be inferred from its WRL response to a power transient. The study has shown that :

- clogging affects the WRL response in two distinct ways. It alters its global slope and its curvature.

- These effects depend on the leg of the SG and the elevation of the clogging sites. Clogging of the hot leg and cold leg have a different impact and the former is predominant. The higher is the clogging site in the SG, the greater is the magnitude of the alteration.

- The WRL response curves can be resumed by vectors of size 2, each coordinate describing respectively the global slope and the curvature of the curves. The clogging state of individual half-TSPs cannot be identified by analysing the WRL response. The diagnosis actually consists in average clogging ratios of each legs.

The low dimensions of the simplified input and output provide a convenient framework for future development of a diagnosis methodology. Two response surfaces, one for each direction of the e.d.r. subspace basis, can be built by swapping the input and output. Then, any measured WRL response can be projected on the PC basis yielding two coordinates. The average clogging of each leg is then indicated by the heights of the two response surfaces associated to the couple of coordinates.

The methodology can be easily adapted to other diagnosis contexts. Here are three remarks to serve that purpose. The derivation of the diagnosis method is mainly based on the dimension reduction achieved with PCA and SIR. However, the basis of the e.d.r. subspace that SIR outputs may not be the most pertinent for the diagnosis. The sensitivity analysis provides valuable insights on the role of each input variable and suggests meaningful combinations of the e.d.r. directions found with SIR.

When the input dimension is high and it is suspected that important features may appear only for extreme values of the input variables, the SIR should be carried out with both a Gaussian sample and a uniform sample. Possible inconsistencies in the results can be caused by a too strong violation of the linearity hypothesis in the uniform sample case. In such a situation, the sensitivity analysis can be used to remove the least influential input variables prior to the SIR. Keeping only the 3 to 5 most prominent variables allows to build a Gaussian sample that covers a more reasonable portion of the hypercube domain.

Finally, in situations where the output is more complex, that is, when there are more PCs with non-null eigenvalue, the reduction of the output dimension can be carried out in a more sophisticated way. One drawback of using the PC scores as the new output variables is that this choice is independent of the input. Using Hotelling's theory of most predictable variates, Li et al. [20] have proposed an extension of SIR that relies on the data to find the output projection basis.

References

[1] M. Midou, J. Ninet, A. Girard, J.-M. Favennec, Estimation Of SG TSP Blockage: Innovative Monitoring Through Dynamic Behavior Analysis, in: 18$^{\text{th}}$ International Conference on Nuclear Engineering ICONE18, 2010.

[2] J. Ninet, J.-M. Favennec, Determination of Applicability of EDF Steam Generator Monitoring Algorithm to Pressurized Water Reactors Worldwide, Tech. Rep. 1021079, EPRI, 2010.

[3] A. Saltelli, Global sensitivity analysis: an introduction, in: Proc. 4th International Conference on Sensitivity Analysis of Model Output (SAMO'04), 27–43, 2004.

[4] I. T. Jolliffe, Principal Component Analysis, 2$^{\text{nd}}$ edition, Springer, 2002.

[5] K. C. Li, Sliced Inverse Regression For Dimension Reduction (with discussion), Journal of the American Statistical Association 86 (414) (1991) 316–327.

[6] J. Pillet, N. Nimambeg, C. Pinto, Résultats des Essais Monophasiques sur la Maquette P2C pour la Détermination des Pertes de Charge induites par le Colmatage, Tech. Rep. H-I84-2009-02838, EDF, 2010.

[7] I. M. Sobol', Sensitivity Estimates for Nonlinear Mathematical Models, in Matem. Modelirovanie, 2 (1)(1990) 112 – 118, English Transl.: MMCE 1 (4).

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

[9] I. M. Sobol', Global Sensitivity Indices for Nonlinear Mathematical Models and their Monte Carlo Estimates, Mathematics and Computers in Simulation 55 (2001) 271–280.

[10] G. E. B. Archer, A. Saltelli, I. M. Sobol', Sensitivity Measures, ANOVA-like Techniques and the Use of Bootstrap, Journal of Statistical Computation and Simulation 58 (1997) 99–120.

[11] K. Campbell, M. McKay, B. Williams, Sensitivity Analysis when Model Outputs are Functions, Reliability Engineering & System Safety 91 (10-11) (2006) 1468–1472.

[12] M. Lamboni, D. Makowski, S. Lehuger, B. Gabrielle, H. Monod, Multivariate Global Sensitivity Analysis for Dynamic Crop Models, Field Crops Research 113 (2009) 312–320.

[13] M. Lamboni, H. Monod, D. Makowski, Multivariate Sensitivity Analysis to Measure Global Contribution of Input Factors in Dynamic Models, Reliability Engineering & System Safety 96 (2010) 450–459.

[14] B. Efron, R. Tibshirani, An Introduction to the Bootstrap, Chapman & Hall, 1993.

[15] B. Liquet, J. Saracco, Application of the Bootstrap Approach to the Choice of Dimension and the α Parameter in the SIR α Method, Communications in Statistics Simulation and Computation 3 (6) (2008) 1198–1218.

[16] L. Barreda, A. Gannoun, J. Saracco, Some Extensions of Multivariate Sliced Inverse Regression, Journal of Statistical Computation and Simulation 77 (1) (2007) 1–17.

[17] R. D. Cook, C. J. Nachtsheim, Reweighting to Achieve Elliptically Contoured Covariates in Regression, Journal of the American Statistical Association 89 (426) (1994) 592–599.

[18] P. Diaconis, D. Freedman, Asymptotics of graphical projection pursuit, The Annals of Statistics 12 (3) (1984) 793–815.

[19] P. Hall, K. C. Li, On Almost Linearity of Low Dimensional Projections From High Dimensional Data, The Annals of Statistics 21 (2) (1993) 867–889.

[20] K. C. Li, Y. Aragon, K. Shedden, C. Thomas Agnan, Dimension Reduction for Multivariate Response Data, Journal of the American Statistical Association 98 (461) (2003) 99–109.