Available online at www.sciencedirect.com

ScienceDirect

Procedia Social and Behavioral Sciences 2 (2010) 7658–7659

Procedia
Social and Behavioral Sciences

Sixth International Conference on Sensitivity Analysis of Model Output

Optimal Sensitivity Analysis under Constraints: Application to fisheries

Jean-Pierre Gauchi[a]*, Sigrid Lehuta[b], Stéphanie Mahévas[b]

[a]*INRA, MIA(UR341), Domaine de Vilvert, Jouy-en-Josas, 78352 France*
[b]*IFREMER, Dpt Ecologie et Modèles pour l'Halieutique, Nantes, 44311 France*

Abstract

We propose a new optimal (in some sense) methodology to conduct a global sensitivity analysis (GSA) of numerical model outputs relative to model inputs when three specific constraints exist. The three constraints considered here are: (i) A computation time that is too long to perform all simulations required when using the usual methods of GSA, typically those based on LHS, Sobol sequences, etc., and typically when only hundreds of simulations are possible; (ii) The inputs are not independent because some structural correlations or functional relationships exist between them (or part of them), or bounded combinations of inputs exist; (iii) Qualitative inputs are present in addition to quantitative inputs.

Keywords: Sensitivity Analysis; D-optimality; PLS regression; Halieutics.

1. Introduction

The two main innovative aspects of the proposed methodology are based on the construction of a *D*-optimal simulation design and the use of the PLS regression (Tenenhaus, 1998). This new methodology allows us to compute specific Sensitivity Indexes, referred to as SI-VIP, already defined in Schwob et al. (2009) and Ellouze et al. (2009), but in the presence of the three specific constraints given in the abstract. We show the effectiveness of this methodology by addressing a real problem for fisheries of the IFREMER Research Centre: the anchovy fishery, located in the Gulf of Gascogne (France). The numerical model used is the IFREMER-ISIS model.

2. Methodology steps

Step 1: Construction of a correlated network of candidate simulations. We refer to this network as X_{0c} (N_0 x p) where N_0 is large (we typically start from an LHS and correlate it afterwards) and p is the input number.

* Corresponding author.
E-mail address: Jean-Pierre.Gauchi@jouy.inra.fr.

doi:10.1016/j.sbspro.2010.05.165

Step 2: Postulation of a second-degree polynomial model *P* and construction of its matrix: X_{0cm} (N_0 x *s*) where *s* is the number *p* plus the number of the chosen interaction terms (if some qualitative inputs exist, they are transformed into their indicator variables) or quadratic terms.

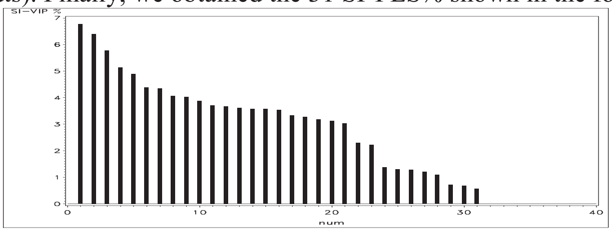
Step 3: Calculation of the maximum determinant of the information matrix of *P* (the ideas for Steps 3 and 4 are taken from the theory of optimal design of experiments for linear models (see Atkinson & Donev, 1986, for a teaching statement)).

Step 4: Calculation of the *D*-optimal simulation design. The optimal simulation design X^*_n is a matrix (*n* x *p*), where $s <= n <= N_0$, a subset extracted from X_{0c} which is optimal under the criterion of *D*-optimality. In other words, among all *n*-designs, this will be the one presenting the greatest determinant of its associated model information matrix. Using D_{max}, we calculate the *D*-efficiency as defined in the reference given above. To find X^*_n, we used a very well known algorithm by means of the Optex procedure of SAS software.

Step 5: Analysis of results and calculation of sensitivity indices SI-VIP. Since many terms in the model are not significant, we first use the PLS-BQ method (Gauchi & Chagnon, 2001) to select the significant terms relative to the Q2G change (as defined in Lazraq et al. (2003)). Consequently, the polynomial *P* becomes the simplified polynomial *P*' with very few terms. We then compute a PLS model with these terms, and if the R^2x100 > 75%, the SI-VIP are computed.

3. Results

We now give some of the results obtained for the halieutic problem mentioned in the introduction after applying our methodology, where 21 inputs were considered. Step 1 led to a candidate network X_{0c} composed of 6854 lines and 21 columns; Step 2 led to X_{0cm} (6854 x 296); Step 3 led to $D_{max} = 1.881 \times 10^{-239}$; Step 4 led to the *D*-optimal simulation design X^*_n (790 x 21); Step 5, after having computed **only 790 simulation runs**, led, for one of the biological responses, to a PLS model (with a 98% explanation rate) based on 31 terms (main effects of inputs and some interactions between inputs). Finally, we obtained the 31 SI-PLS% shown in the following figure.



There is not enough room here to give formulas and more details on the proposed new methodology, but they are necessary for a good understanding of it. It is fully described in a technical report (Gauchi et al., 2010), as well as in a paper in progress.

4. References

Atkinson, A. C. and Donev, A. N., 1992 : “Optimum Experimental Designs ”, Clarendon Press, Oxford, UK.

Ellouze, M., Augustin, J.-C., Gauchi, J.-P., 2010: “Sensitivity Analyses applied to an exposure model of Listeria monocytogenes in smoked cold salmon”, In Press in *Risk Analysis.*

Gauchi, J.-P. and Chagnon, P., 2001: “Comparison of selection methods of explanatory variables in PLS regression with application to manufacturing process data”, *Chemometrics and Intelligent Laboratory Systems*, 58, 171-193.

Gauchi, J.-P., Lehuta, S., Mahévas, S., 2010 : “ Analyse de sensibilité optimale sous contraintes : application à un problème halieutique”. Rapport Technique de l’Unité INRA-MIA (UR341) n°2010-1.

Lazraq, A, Cléroux, R., Gauchi, J.-P., 2003: “Selecting both latent and explanatory variables in the PLS1 regression model”, *Chemometrics and Intelligent Laboratory Systems*, 66, 117-126.

Schwob, C., Gauchi, J.-P., Huiji, S., Chambon, L., 2009: “Two efficient sensitivity analysis methods to address factor prioritization in the context of structural computations”, Submitted to *RESS*.

Tenenhaus, M., 1998 : “La Régression PLS, théorie et pratique ”, Technip Publisher, Paris.