2012 IEEE International Conference on
Automation Science and Engineering (CASE)
Seoul, Korea. August 20-24, 2012

Fitting a Functional Structural Plant Model based on Global Sensitivity Analysis

Yubin LIN, Mengzhen KANG *, Jing HUA

Abstract— **Calibration of a functional structural plant model is a challenging task because of the complexity of model structure. Parameter estimation through gradient-based optimization technique was highly dependent on initial parameter values. This motivated the use of global sensitivity analysis technique to choose parameter subset in fitting the data sequence. Global sensitivity indices were computed using the source sink ratio as the output of interest, which regulates all organ growth. By fitting on chrysanthemum data from nine sampling dates, it is shown that sensitivity analysis method helps to identify the influential parameters for a given sampling date. As a result, fitting process is less dependent on the initial parameter values. Current work provides a new method of calibrating a plant growth model with multiple outputs.**

I. INTRODUCTION

Mathematical model is a powerful tool which is used in various domains, in order to simulate, explain and predict the corresponding phenomena. In agriculture, dynamic plant models have been developed with different levels of complexity. Among them, Processed-Based Models (PBM) aim at describing the yield formation by modeling the photosynthesis process and biomass partitioning regulated by sink-source balance [1]. Moreover recently, structural information of plant are taken into classical PBM, leading to Functional-Structural Plant Models [2], which bring visual plants with complex behavior. Among them, GreenLab model has a generic model frame and it has been used to retrieve the traces of plant development and growth, for crops [3][4] and trees [5].

Parameterizing dynamic plant growth model is a critical issue in model application. In general, there are three approaches to decide the value of parameters [6]: finding them from published literature, measuring every parameter when the system is small, or using some parameter estimate techniques to fit the model output to measurement. For the last case, when there are lots of parameters to be estimated, it is not a trivial work. There are risks of (1) over-parameterization that can limit the model from being generalized; (2) data redundancy which costs too much in data acquisition; (3) sensitivity to initial parameter values when using gradient-based optimization techniques.

Sensitivity Analysis (SA) refers to techniques identifying how the uncertainty of system output can be decomposed into the different perturbation of the input in the model [7]. SA methods have been widely used in research of economy, environment and ecology [8][9][10], which can help to get in-depth understanding of the model structure, simplify and calibrate the model [7]. The aim of current work is use SA to aid parameterizing of dynamic plant growth model with multiple outputs. GreenLab is chosen for case study, because it is one of few FSPMs that has experienced some fitting exercise. In this discrete system, part of parameter values can be obtained directly by observation, while some parameters controlling biomass production and allocation need to estimated by model inversion.

Previous studies [3][11][12] have concentrated on fitting the GreenLab model using Generalized Least Squares Method (GLSM). The fitting targets are biomass of different type of organs, at individual organ level and/or plant level [4]. As measuring individual organ biomass is very tedious, attempt has been made by testing fitting process using more sparse data [12]. But, still, problems exist in two aspects: (a) satisfactory solution is not guaranteed since the gradient descent algorithm usually converges to a local minimum; (b) updating all parameters during the fitting process often leads to a fitting failure. On the other hand, the plant growth is generally sequential in its life circle, e.g., from vegetative to generative stage, thus not all parameters are involved at a given plant age. Inspired by this observation, in this paper, we use the Global Sensitivity Analysis method to identify a subset of parameters for fitting. The paper is organized as follows: Sec.II presents briefly the GreenLab model and parameters to be estimated; Sec.III-A presents the Global Sensibility Analysis method and how it is used in current context; Sec.III-B describes the fitting process used in experiment; Sec.IV shows the results, followed by discussion and conclusion.

II. GREENLAB MODEL

GreenLab [13] is a generic plant model simulating two basic process of plant: development (organogenesis) and growth (organ expansion). The organogenesis is simulated with a dual-scale automaton, which gives the number of organs that participate biomass production and allocation. At each time interval, called Growth Cycle (GC), plant structure is updated according to the organogenesis model. Combined with organ size computed from the functional model (see

This work was supported by 863 Program of China (#2012AA101906-2) and the National Natural Science Foundation of China (#31170670).

Yubin LIN is with National Key Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.

Mengzhen KANG and Jing HUA are with the State Key Laboratory of Management and Control for Complex Systems, Beijing Engineering Research Center of Intelligent Systems and Technology, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China; Dongguan Research Institute of CASIA, Cloud Computing Center, Chinese Academy of Sciences, Songshan Lake, Dongguan, 523808, China. mengzhen.kang@ia.ac.cn

below) and geometrical parameters (e.g., branching angel), three dimensional virtual plant can be simualted.

Biomass production of a plant is calculated using the following equation [3]:

$$Q(i) = \frac{E(i)S_{\mathrm{P}}}{R \cdot k}\left(1 - \exp\Big(-k\frac{S(i)}{S_{\mathrm{P}}}\Big)\right), \tag{1}$$

where $E(i)$ is the photosynthetic production potential at GC i, S_{P} is the ground projection area of plant, R is the photosynthesis coefficient, k is analogous to the extinction coefficient of Beer-Lambert's Law [14], $S(i)$ is the total leaf area at GC i. The initial biomass $Q(0)$ is from the seed or cutting.

Biomass are distributed among growing organs according to their sink strength. For an organ o of age j, its sink strength is defined by:

$$p_o(j) = P_o \cdot f_o(j), \tag{2}$$

where o represent the organ type (b: leaf blade, p: leaf petiole, i: internode, f: female, m: male); $f_o(j)$ is the sink variation function [13], P_o is the relative sink strength of organ o, with the sink strength of leaf blade set to 1 as a reference. $f_o(j)$ is an empirical function describing change of organ sink strength according to organ age.

Summing up the sink strength of all organ, we get the total demand of plant at GC i:

$$D(i) = \sum_o P_o\Big(\sum_{j=1}^{i} N_o(i,j) \cdot f_o(j)\Big), \tag{3}$$

where $N_o(i,j)$ is the number of organs, given by the organogenesis model.

According to the sink strength, biomass acquired by organ of age j at GC i is:

$$\Delta q_o(i,j) = \frac{P_o \cdot f_o(j)}{D(i)} \cdot Q(i). \tag{4}$$

In Eqn.4, the ratio $Q(i)/D(i)$ represents the biomass availability for each organ, called source sink ratio.

The biomass of an organ o is the accumulated biomass since it appeared in plant:

$$q_o(i,j) = \sum_{k=1}^{j} \Delta q_o(i-j+k,k) \qquad (i \geq j). \tag{5}$$

Summing up the biomass of all individual organs of the same property gives the total biomass of organs, which are measured in reality.

Above is a brief description of GreenLab kernel equations describing organ expansion. In case of Dicotyledons plants with secondary growth, there are other parameters regarding to the thickening process of stem. A parameter S_c is used in quantifying the sink strength of secondary growth in stems [5].

III. METHOD

A. Global Sensitivity Analysis

Approaches of SA can be divided into two classes: local and global SA. In local methods, the sensitivity measure of a single factor (parameter) is calculated around a fixed point, while other parameters are set to constant; the result is applicable only to linear models. Because of its simplicity, this method is widely used. Global methods estimate the effect on the output from a parameter using point sample from the whole input space. The global methods is model free and more accurate compared to the local one, but it needs more model evaluations and thus is computational costly.

For GreenLab model, as the linearity of GreenLab model is unknown, a variance-based global sensitivity analysis (Sobol's method) is used to explore the model properties through Monte Carlo simulations [15]. The method is presented briefly below, considering a model whose output Y being a function of k parameters $(X_1, X_2, \ldots, X_k)$:

$$Y = f(X_1, X_2, \ldots, X_k). \tag{6}$$

The Sobol's method expands the function f into terms of increasing dimensions:

$$f = f_0 + \sum_i f_i + \sum_i \sum_{j>i} f_{ij} + \cdots + f_{12\ldots k}, \tag{7}$$

in which each term relies only on the parameters in its index. By using Monte Carlo simulations, we consider the variance of each term $V(f_{i_1\ldots i_s})$ as a measure of sensitivity. In particular, $V(f_i)$ is $V[E(Y|X_i)]$. Divide them by the unconditional variance $V(Y)$, we get the first-order sensitivity index (SI) [16]:

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)}. \tag{8}$$

For additive models, which include linear models, the sum of all first-order sensitivity indices equal to 1.

B. Fitting Process

Previous fitting exercise showed that when all parameters are estimated simultaneously by fitting all the data, the risk of fitting failure is high since the parameters seldom converge to satisfactory result and rely on the initial value for iteration. As plant growth process is sequential, it is possible to choose a subset of parameters, while other parameters being fixed, according to the sensitivity index of a given output to parameters. There are five output sequence of interest. Sensitivity can be computed for different parameters, but global information is missing that can tell the sensitivity of the model system to a certain parameter. For such case of multiple correlated outputs, one solution is to decompose the data upon a complete orthogonal basis and then to compute sensitivity indices on each component of the decomposition [18]. In another word, this can make sure that all outputs are independent. However, it is worthy to notice that, in GreenLab model, the different outputs are correlated, as they are all dependent on the source sink ratio (Eqn.4, 5). This

TABLE I

DESCRIPTION OF NOTATIONS

Notation	Description
K	number of parameters
N	number of target files
p_k	the kth parameter
S_k	the sensitivity index of Q/D respect to p_k
η	threshold of SI

TABLE II

FITTING PROCESS

```
for n = 1 to N
    for k = 1 to K
        estimate S_k
        if S_k > η
            select p_k
    fitting the selected parameters to the first n
    target files using GLSM.
fitting all parameters to all target data.
```

led to the idea of using the source sink ratio, rather than the organ biomass, to analysis and select the sensitive parameters for fitting.

The policy of selecting subset of parameters is as follows: (1) making global SA of the dynamic Q/D, to obtain the sensitivity indices for all parameters; (2) reading a target data file from the initial to final sampling dates; if the SI of a parameter is above a threshold for a certain sampling date, this parameter is included in fitting current data, otherwise it remains fixed; (3) continuing step 2 until all data were fit; fitting all parameters for all target data. For the sake of clarity we introduce the notations in Table I. The fitting process is presented in Table II.

IV. RESULT

A. Sensitivity Analysis on Q/D

Sensitivity index of Q/D to different parameters were estimated using Monte Carlo simulation. Fig.1 gives the sensitivity indices of Q/D with respect to all parameters. It can be seen that the SI of parameter R is always high during plant growth, because it effects linearly on biomass acquisition of plant; the SIs of parameter $Q(0)$ decreases with plant age, which means the effect of initial biomass reserve becomes less important when plant grows up; the SI of P_i^1 disappeared around cycle 45, because of the start of flowering; parameters S_c, P_i^2 and P_f take the relay at later stage. According to this figure, it is expected that $Q(0)$ and P_i^1 are estimated and fixed early.

B. Fitting results

Data of chrysanthemum were collected at nine sample dates, from initial planting to harvest time (opening of flower) [17]. For each sampled plant, measured data includes the total dry weight (DW) of leaves in main stem (Q_b^1), main stem (Q_i^1), leaves in branches (Q_b^2), stem in branches (Q_i^2), and flowers (Q_f). Parameters to estimate include the initial

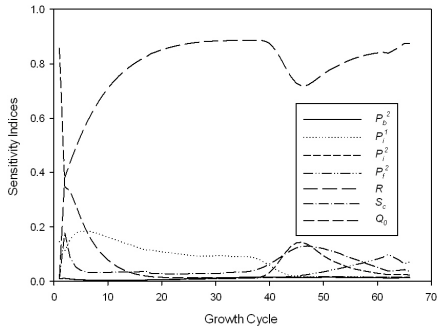

Fig. 1. Sensitivity index of source sink ratio (Q/D)

TABLE III

MEASURED DATA

GC	Q_b^1	Q_i^1	Q_b^2	Q_i^2	Q_f
6	0.1555	0.0675	-	-	-
17	0.7810	0.3854	-	-	-
22	0.9688	0.6304	0.1762	0.0196	-
34	1.6564	1.3290	0.3246	0.0470	-
39	2.0766	2.1888	0.5090	0.2252	-
44	1.9492	2.6276	0.6176	0.3724	0.1368
49	2.2860	2.4868	0.5454	0.8656	0.4008
54	2.3440	2.5092	0.6136	1.0626	0.6058
66	2.2816	2.4762	0.7724	1.7546	1.8778

biomass from cutting ($Q(0)$), photosynthesis coefficient (R) (Eqn.1), sink strength of main stem internode (P_i^1), sink strength of branch leaf (P_b^2), sink strength of branch internode (P_i^2), sink strength of flower (P_f, Eqn.2) (main stem flower was picked off), and coefficient for secondary growth of main stem (S_c). Measured data at 9 sampling dates are showed in Table III.

The threshold of SI, η, is set to 0.1; the number of target data files is 9 (N=9); the number of parameters is 7 (K=7). Initial values of all parameters were set to 0.1. By applying the fitting procedure as in Sec.III-B, the subset of parameters chosen at each sampling date is as follows (see Table IV).

Fitting result converged although the final parameter values are far from their initial values, as shown in Table V.

The target data and fitting curves are shown in Fig.2.

TABLE IV

PARAMETERS CHOSEN AT EACH FITTING STEP

Step	Parameters chosen
1	P_i^1, R, $Q(0)$
2	P_i^1, R
3	P_i^1, R
4	P_i^1, R
5	P_i^1, R, S_c
6	P_i^2, R, S_c
7	P_i^2, R, S_c
8	P_i^2, P_f^2, R, S_c
9	P_f^2, R
10	P_b^2, P_i^1, P_i^2, P_f^2, R, S_c, $Q(0)$

TABLE V

FITTING RESULT WITH INITIAL VALUE BEING 0.1. DIRECT FITTING MEANS ALL PARAMETERS ARE ESTIMATED SIMULTANEOUSLY BY FITTING ALL OF TARGET DATA.

Parameter	Initial value	Final value	Direct fitting
P_b^2	0.1	0.0395	0.0433
P_i^1	0.1	0.4672	0.1000
P_i^2	0.1	0.2596	0.3836
P_f^2	0.1	2.4028	3.6851
R	0.1	0.000081	0.000076
S_c	0.1	5.5331	10.7197
$Q(0)$	0.1	0.0726	0.1000
RMS		0.0284	0.0294

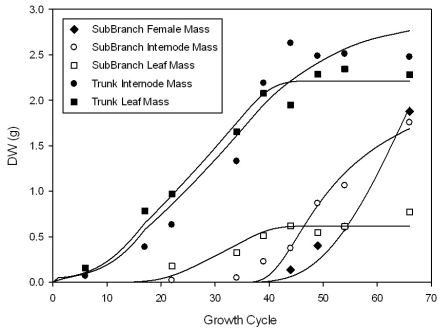

Fig. 2. Fitting results on chrysanthemum data from 9 sampling dates, including the total dry weight of leaves in main stem, main stem, leaves in branches, stem in branches, and flowers.

Different initial parameter values were tried for the fitting iteration (see Table VI). The experiment showed that the fitting converged regardless of the initial values when aided by sensitivity analysis. Instead, if all parameters are fit simultaneously, the fitting failed, as can be seen from the parameter values and error with direct fitting.

V. CONCLUSIONS AND DISCUSSION

In this work, we have presented the application of a SA method for improving the data fitting robustness, with a case study on GreenLab model. Fitting on chrysanthemum data demonstrates the value of sensitivity analysis for selecting parameter set for the fitting: using different initial parameter values, fitting can be achieved with similar final values.

Calibration of a functional structural plant model is a challenging task because of the complexity of model structure. Like in real plant, the yield and other organ biomass are joint results of the genetic background and environmental conditions, both influencing the development and growth process of plant. Using a model free analysis technique helps greatly in understanding the behavior of model and plant. For example, analysis has been made on maize plant using the biomass production Q as the output [19]. Current work proved the advantage of using global sensitivity analysis in choosing parameters in dynamic fitting. Actually, it makes use of the fact that the plant growth is sequential: for the case of chrysanthemum, the plant starts by vegetative growth of leaves and stem extension; when flowers buds start to growth, the main stem organs finish extension, while branch stem and flowers become the dominant sink. So it is reasonable to estimate dominant parameters for the active process. As a result, the fitting process is less dependent on priori knowledge of possible parameter values.

Compared to other applications of SA on plant model, current work has made use of a key variable that adjust the growth of organs, source sink ratio (Q/D), instead of computing the principle components for analysis. According to prevailing hypothesis that plant growth is regulated by source sink ratio [1], this choice is not only reasonable but also more mechanistic compared to statistical approach. The sensitivity indices can also aid to decide the sampling date and content to same labor and time in measurement. This is useful when the online calibration of plant model is needed. However, in line with the no free lunch rule, the fitting relies on the sensitivity analysis process, which is costly and dependent also on sampling space. Moreover, the number of parameters to be chosen once rely on the threshold η, another parameter to be tuned.

TABLE VI

FITTING RESULT WITH INITIAL VALUE BEING 5. DIRECT FITTING MEANS ALL PARAMETERS ARE ESTIMATED SIMULTANEOUSLY BY FITTING ALL OF TARGET DATA.

Parameter	Initial value	Final value	Direct fitting
P_b^2	5	0.0395	3.5171
P_i^1	5	0.4674	5.0000
P_i^2	5	0.2596	4.2202
P_f^2	5	2.4023	4.9921
R	5	0.000081	0.000019
S_c	5	5.5299	6.1047
$Q(0)$	5	0.0727	5.0000
RMS		0.0284	0.2350

REFERENCES

[1] L. Marcelisa, E. Heuvelinkb, and J. Goudriaan, "Modelling biomass production and yield of horticultural crops: a review," *Scientia Horticulturae*, vol. 74, pp. 83–111, 1998.

[2] J. Vos, J. Evers, G. Buck-Sorlin, B. Andrieu, M. Chelle, and P. de Visser, "Functional-structural plant modelling: a new versatile tool in crop science," *Journal of Experimental Botany*, vol. 61, pp. 2101–2115, 2009.

[3] Y. Guo, Y. Ma, Z. Zhan, B. Li, M. Dingkuhn, D. Luquet, and P. de Reffye, "Parameter optimization and field validation of the functional-structural model greenlab for maize," *Annals of Botany*, vol. 97, pp. 217–230, 2006.

[4] M. Kang, J. Evers, J. Vos, and P. de Reffye, "The derivation of sink functions of wheat organs using the greenlab model," *Annals of Botany*, vol. 101, pp. 1099–1108, 2008.

[5] V. Letort, P.-H. Cournède, A. Mathieu, P. de Reffye, and T. Constant, "Parametric identification of a functional-structural tree growth model and application to beech trees (fagus sylvatica)," *Functional Plant Biology*, vol. 35, pp. 951–963, 2008.

[6] Y. Jin, H. Yue, M. Brown, Y. Liang, and D. Kell, "Improving data fitting of a signal transduction model by global sensitivity analysis," in *American Control Conference, 2007. ACC '07*, New York, USA, 2007, pp. 2708–2713.

[7] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis, the primer ed.* England: John Wiley & Sons, 2008.

[8] P. Mandal, A. Srivastava, and M. Negnevitsky, "Improving performance of nn based electricity price forecasting using sensitivity analysis," in *Power Systems Conference and Exposition, 2009. PSCE '09*, Seattle, WA, 2009, pp. 1–8.
[9] D. Makowski, C. Naud, M.-H. Jeuffroy, A. Barbottin, and H. Monod, "Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction," *Reliability Engineering System Safety*, vol. 91, pp. 1142–1147, 2006.
[10] J. Caribonia, D. Gatelli, R. Liska, and A. Saltelli, "The role of sensitivity analysis in ecological modelling," *Ecological Modelling*, vol. 203, pp. 167–182, 2007.
[11] Y. Ma, M. Wen, Y. Guo, B. Li, P.-H. Cournède, and P. de Reffye, "Parameter optimization and field validation of the functional-structural model greenlab for maize at different population densities," *Annals of Botany*, vol. 101, pp. 1185–1194, 2007.
[12] Y. Ma, M. Wen, B. Li, Y. Guo, P.-H. Cournède, and P. de Reffye, "Calibration of greenlab model for maize with sparse experimental data," in *Second International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA)*, Beijing, China, 2006, pp. 188–193.
[13] H. Yan, M. Kang, P. de Reffye, and M. Dingkuhn, "A dynamic, architectural plant model simulating resource-dependent growth," *Annals of Botany*, vol. 93, pp. 591–602, 2004.
[14] J. Vose, N. Sullivan, B. Clinton, and P. Bolstad, "Vertical leaf area distribution, light transmittance, and application of the beer-lambert law in four mature hardwood stands in the southern appalachians," *Canadian Journal of Forest Research*, vol. 25, pp. 1036–1043, 1995.
[15] A. Saltelli, "Making best use of model evaluations to compute sensitivity indices," *Computer Physics Communications*, vol. 145, pp. 280–297, 2002.
[16] I. Sobol, "Sensitivity analysis for non-linear mathematical models," *Mathematical Modeling & computational Experiment*, vol. 1, pp. 407–414, 1993.
[17] M. Kang, E. Heuvelink, S. Carvalho, and P. de Reffye, "A virtual plant that responds to the environment like a real one: the case for chrysanthemum," *Submitted to New Phytologist*.
[18] K. Campbell, M. McKay, and B. Williams, "Sensitivity analysis when outputs are functions," *Reliability Engineering & System Safety*, vol. 35, pp. 951–963, 2008.
[19] Q. Wu and P.-H. Cournède, "Sensitivity analysis of greenlab model for maize," in *Third International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA)*, Beijing, China, 2009, pp. 311–318.
[20] Z. Zhan, P. de Reffye, F. Houllier, and B. Hu, "Fitting a functional-structural growth model with plant architectural data," in *International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA)*, Beijing, China, 2003, pp. 236–249.

VI. APPENDIX

A. Estimation of Sensitivity Index

Given the model $y = f(x_1, x_2, \ldots, x_k)$ with independent random variables, the joint probability density function of input can be expressed as:

$$P(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} p_i(x_i). \tag{9}$$

Variance of model out y is:

$$\begin{aligned}\mathbb{V}(y) &= \iint\cdots\int \big(f(x_1, x_2, \ldots, x_k) - \mathbb{E}(y)\big)^2 \prod_{i=1}^{k} p_i(x_i)\mathrm{d}x_i \\ &= \iint\cdots\int f^2(x_1, x_2, \ldots, x_k) \prod_{i=1}^{k} p_i(x_i)\mathrm{d}x_i - \mathbb{E}^2(y)\end{aligned} \tag{10}$$

When input factor x_j is fixed to $\tilde{x}_j$, the conditional variance of y is:

$$\begin{aligned}&\mathbb{V}(y|x_j = \tilde{x}_j) \\ &= \iint\cdots\int \big(f(x_1, x_2, \ldots, \tilde{x}_j, \ldots, x_k) - \mathbb{E}(y|x_j = \tilde{x}_j)\big)^2 \\ &\quad \cdot \prod_{\substack{i=1 \\ i\neq j}}^{k} p_i(x_i)\mathrm{d}x_i \\ &= \iint\cdots\int f^2(x_1, x_2, \ldots, \tilde{x}_j, \ldots, x_k) \prod_{\substack{i=1 \\ i\neq j}}^{k} p_i(x_i)\mathrm{d}x_i \\ &\quad - \mathbb{E}^2(y|x_j = \tilde{x}_j)\end{aligned} \tag{11}$$

In order to remove the dependence on $\tilde{x}_j$, integrating $\mathbb{V}(y|x_j = \tilde{x}_j)$ over the probability density function of $\tilde{x}_j$ one obtains:

$$\begin{aligned}&\mathbb{E}\big(\mathbb{V}(y|x_j)\big) \\ &= \iint\cdots\int f^2(x_1, x_2, \ldots, x_j, \ldots, x_k) \prod_{i=1}^{k} p_i(x_i)\mathrm{d}x_i \\ &\quad - \int \mathrm{E}^2(y|x_j = \tilde{x}_j) p_j(\tilde{x}_j)\mathrm{d}\tilde{x}_j\end{aligned} \tag{12}$$

Subtract Eqn.12 from Eqn.10, obtaining

$$\mathbb{V}(y) - \mathbb{E}\big(\mathbb{V}(y|x_j)\big) = \int \mathbb{E}^2(y|x_j = \tilde{x}_j) p_j(\tilde{x}_j)\mathrm{d}\tilde{x}_j - \mathbb{E}^2(y) \tag{13}$$

To avoid a double loop in computing Eqn.13 using Monte Carlo simulation, it can be rewritten as [15]:

$$\begin{aligned}&\int \mathbb{E}^2(y|x_j = \tilde{x}_j) p_j(\tilde{x}_j)\mathrm{d}\tilde{x}_j \\ &= \int \Big\{ \iint\cdots\int f(x_1, x_2, \ldots, \tilde{x}_j, \ldots, x_k) \prod_{\substack{i=1 \\ i\neq j}}^{k} p_i(x_i)\mathrm{d}x_i \Big\}^2 \\ &\quad \cdot p_j(\tilde{x}_j)\mathrm{d}\tilde{x}_j \\ &= \iint\cdots\int f(x_1, x_2, \ldots, \tilde{x}_j, \ldots, x_k) f(x'_1, x'_2, \ldots, \tilde{x}_j, \ldots, x'_k) \\ &\quad \cdot \prod_{\substack{i=1 \\ i\neq j}}^{k} (p_i(x_i)\mathrm{d}x_i) \prod_{\substack{i=1 \\ i\neq j}}^{k} (p_i(x'_i)\mathrm{d}x'_i)\mathrm{d}\tilde{x}_j \\ &= \iint\cdots\int f(x_1, x_2, \ldots, x_j, \ldots, x_k) f(x'_1, x'_2, \ldots, x_j, \ldots, x'_k) \\ &\quad \cdot \prod_{i=1}^{k} (p_i(x_i)\mathrm{d}x_i) \prod_{\substack{i=1 \\ i\neq j}}^{k} (p_i(x'_i)\mathrm{d}x'_i)\end{aligned} \tag{14}$$

The integral in Eqn.14 is the expectation value of the function F of a set of $(2k-1)$ factors:

$$\begin{aligned}&F(x_1, x_2, \ldots, x_j, \ldots, x_k, x'_1, x'_2, \ldots, x'_{j-1}, x'_{j+1}, \ldots, x'_k) \\ &= f(x_1, x_2, \ldots, x_k) f(x'_1, x'_2, \ldots, x'_{j-1}, x_j, x'_{j+1}, \ldots, x'_k)\end{aligned} \tag{15}$$

Therefore, the sensitivity measure for a generic factor x_j is:

$$S_j = \frac{\mathbb{V}\big(\mathbb{E}(y|x_j)\big)}{\mathbb{V}(y)} = \frac{U_j - \mathbb{E}^2(y)}{\mathbb{V}(y)} \tag{16}$$

where

$$U_j = \int \mathbb{E}^2(y|x_j = \tilde{x}_j) p_j(\tilde{x}_j) \mathrm{d}\tilde{x}_j \tag{17}$$

can be computed using a single Monte Carlo loop.

B. Generalized Least Squares Method

Considering a generic model

$$[y_1, y_2, \ldots, y_n] = F(\theta_1, \ldots, \theta_p) \tag{18}$$

which consists of p parameters and n outputs, the objective is to adjust the parameters to best fit a observed data set $Y = [Y_1, \ldots, Y_n]$. Generalized least squares method estimates parameters by minimizing the squared Mahalanobis length of the residual vector [20]:

$$L_\Omega(\theta) = (Y - F(\theta))^{\mathrm{T}} \Omega (Y - F(\theta)), \tag{19}$$

where Ω is a $n \times n$ matrix which weights the different observations.

Let J be the Jacobian matrix

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial \theta_1} & \cdots & \frac{\partial y_1}{\partial \theta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial \theta_1} & \cdots & \frac{\partial y_m}{\partial \theta_n} \end{bmatrix} \tag{20}$$

in which the element can be estimated by

$$\frac{\partial y_v(\theta)}{\partial \theta_k} \approx \frac{\partial y_v(\theta + h_k e_k) - y_v(\theta)}{h_k} \tag{21}$$

where e_k is the kth column vector of the $p \times p$ identity matrix, and h_k the step length. Consequently, the model can be approximated at $\theta^{(q)}$ by

$$F(\theta) = F(\theta^{(q)}) + J(\theta^{(q)})(\theta - \theta^{(q)}) + o(||\theta - \theta^{(q)}||). \tag{22}$$

Let $\Delta Y = Y - F(\theta^{(q)})$, $\Delta\theta = \theta - \theta^{(q)}$ and omit the higher-order infinitesimal, $L_\Omega(\theta)$ becomes

$$L_\Omega(\theta) = \Big(\Delta Y - J(\theta^{(q)})\Delta\theta\Big)^{\mathrm{T}} \Omega \Big(\Delta Y - J(\theta^{(q)})\Delta\theta\Big) \tag{23}$$

Setting the gradient of $L_\Omega(\theta)$ with respect to $\Delta\theta$ to zero, we get

$$\Delta\theta = \Big(J(\theta^{(q)})^{\mathrm{T}} \Omega J(\theta^{(q)})\Big)^{-1} J(\theta^{(q)})^{\mathrm{T}} \Omega \Delta Y \tag{24}$$

Then the parameters are updated iteratively

$$\theta^{(q+1)} = \theta^{(q)} + \Delta\theta \tag{25}$$

until converged.