VISUALIZING AND INSPECTING SENSITIVITIES IN LARGE MODELS USING DOT CHARTS AND STAR PLOTS, WITH APPLICATION TO COMPARTMENTAL MODELS

SUZAN GAZIOĞLU

Department of Mathematical Sciences
Montana Tech of the University of Montana
1300 W. Park Street
Butte, MT 59701
USA
e-mail: sgazioglu@mtech.edu

Abstract

Sensitivity analysis (SA) is a procedure for determining and quantifying the change in model behaviour as model input factors change. When working with models which consist of a large number of model inputs, identifying the most influential factors is of great interest. Even though some graphical tools are used to summarize results from SA, this approach is still rather new. Creating graphs that can help us visualize the results from the analysis can be very useful. In this study, two modern graphical techniques, namely, star plots and dot charts, which can be used for SA, are discussed. These visual summary tools help us to identify the most influential input factors at a glance. To demonstrate their use in SA, we utilize two different compartmental models of the global carbon cycle.

1. Introduction

Sensitivity analysis (SA) studies how the model response variables respond to changes in the model inputs. The use of systematic SA on large, complex computational models plays an important role in science. As complex numerical models are being increasingly applied for problem solving in many application areas, such as atmospheric science, combustion physics and engineering, biological systems, etc., the need for SA is becoming more and more apparent. Saltelli [19] considers SA as a prerequisite for model building in any discipline where models are used. As Tomović and Vukobratović [21] note, SA is typically applied to initial conditions, time-invariant or time-variant coefficients, sampling interval, sampling instant, characteristic frequencies, input frequency, temperature effect, delay, etc.. In this study, we use the terminology *input factor* to refer to all these characteristic elements (parameters, variables, coefficients, etc.).

SA methods based on sampling involve the generation and exploration of mappings from model input factors to model predictions. In these sampling-based SA methods, after the values for input factors are obtained by using a chosen sampling scheme, the model is run for a specified number of times N, and one or more output variables for each run are recorded. Following this step, appropriate statistical analyses, such as correlation measures, regression-based methods, and various nonparametric methods (such as correlation measures on ranks, Smirnov test, Cramér-von Mises test, and Mann-Whitney test) are performed on the output variables as dependent variables and on input factors as independent variables. These analyses are then used to assess input-output relationships and effect of uncertainties in inputs on the output variable(s).

Numerous graphical tools, such as scatterplots, histograms, radar charts, tornado charts, cobweb plots, and bar charts, have found applications in SA. According to a literature search by Cooke and van

Noortwijk [6], there is very little theoretical development for graphical methods in SA. As they point out, when graphical methods are used in SA, the focus is not on visualizing data as it is in the general sense, but rather on visualization to support the SA.

The objective of this study is to extend the use of two modern graphical techniques, namely, dot charts and star plots, to support SA. In the sections which follow, we first describe each test model we use to demonstrate the use of the two graphical tools in Section 2. The two test models are compartmental models of the global carbon cycle (GCC). In Section 3, we discuss the application of SA methods to GCC models. We then talk about graphical sensitivity techniques in Section 4. This is followed by the descriptions of dot charts and star plots, and their use in SA (see Sections 5 and 6). In Section 7, two different sensitivity measures are given and the use of dot charts and star plots to display the results from these measures on the two GCC models are presented. Finally, the discussions and conclusions are given in Section 8.

2. The Test Models

The two test models we adopted from the scientific literature and utilized in this study are compartmental models of global carbon cycle (GCC). They have been used to quantitatively describe the CO_2 distribution between atmosphere and oceans, atmosphere and terrestrial systems, and the responses of these reservoirs to the input resulting from fossil fuel burning and deforestation. Next, we briefly describe these two test models which, from here on, we will refer to as 'the 8-compartment model' and 'the 25-compartment model'. More detailed descriptions of both models along with their dynamic equations can be found in Gazioğlu [11].

2.1. The 8-compartment model

This compartmental model is adopted from Emanuel et al. [9]. It consists of eight well-mixed compartments and 15 transfer coefficients. The compartmental diagram of the model, consisting of 8 compartments and 18 transfer coefficients, is given in Figure 1.
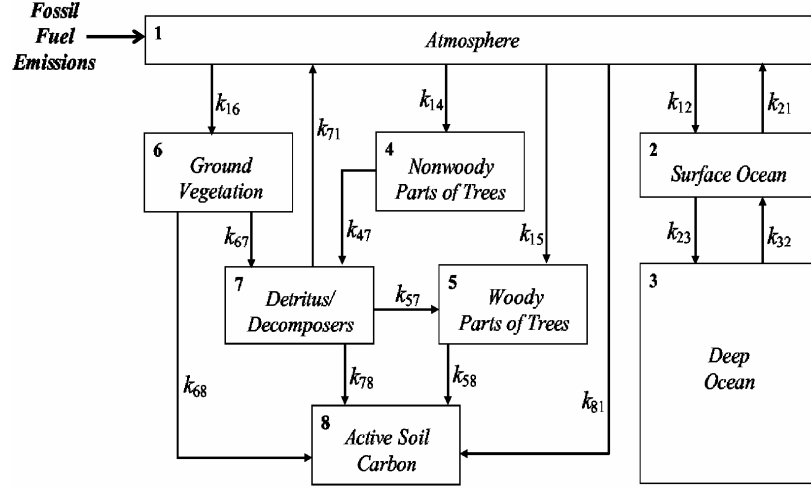

Figure 1. Diagram of the 8-compartment model.

The atmosphere is represented by a single compartment. Two compartments represent carbon in the 'surface ocean' and 'deep ocean'. Carbon in living plants is divided between 'tree' and 'ground vegetation' compartments. The 'tree' compartment is separated into two separate compartments, namely, 'nonwoody parts of trees' and 'woody parts of trees'. To represent carbon in dead parts of the terrestrial systems and their decomposers, two compartments are used. The 'detritus/decomposers' compartment corresponds to litter and its decomposers at the soil surface. Carbon input to this reservoir comes from death of above-ground parts of vegetation. The 'active soil carbon' compartment consists of carbon in soils and its decomposers. Carbon from death and initial decomposition of below-ground parts of vegetation and transport of decomposed material from the actively decaying litter layer is transferred into this compartment.

The CO_2 emissions due to fossil fuel burning and forest clearing enter the system through the atmosphere compartment. The flow of CO_2 between the compartments is described by a set of eight first-order, linear differential equations, which contain 23 uncertain model input factors   (15 transfer coefficients and 8 initial conditions). To keep this paper a manageable size, we only consider the initial conditions of the eight compartments (denoted by $x_1^\circ - x_8^\circ$) as model input factors, which are subject to uncertainty. Table 1 gives the nominal values of these initial conditions as well as the variability ranges.

Table 1. The 8-compartment model reference case initial compartment contents (in units of Gt C)

Description (Initial Conditions)	Input Factor	Nominal Value	Range
Atmosphere (1)	x_1°	622.40	497.92 – 746.88
Surface ocean (2)	x_2°	667.37	533.90 – 800.84
Deep ocean (3)	x_3°	37542.00	30033.60 – 45050.40
Nonwoody parts of trees (4)	x_4°	38.21	30.57 – 45.85
Woody parts of trees (5)	x_5°	634.47	507.58 – 761.36
Ground vegetation (6)	x_6°	59.32	47.46 – 71.18
Detritus/decomposers (7)	x_7°	108.22	86.58 – 129.86
Active soil carbon (8)	x_8°	1131.39	905.11 – 1357.67

2.2. The 25-compartment model

The model and its computer implementation are adopted from a technical report prepared for the US Department of Energy (see Emanuel et al. [10]). It consists of 25 well-mixed compartments, which represent three major components–atmosphere, oceans, and terrestrial systems–of

the cycle. The atmosphere is represented by one compartment, the oceans by 19 globally averaged depth layers, and the terrestrial systems by five compartments. Figure 2 shows the compartmental diagram of the model. In this model, CO_2 is released to the atmosphere by fossil fuel combustion. Deforestation also results in a direct transfer of carbon to the atmosphere from 'tree' compartments as well as a transfer to 'detritus/decomposers'. The relative magnitudes of transfers from the atmosphere to 'tree' and 'ground vegetation' are modified by land-use changes.
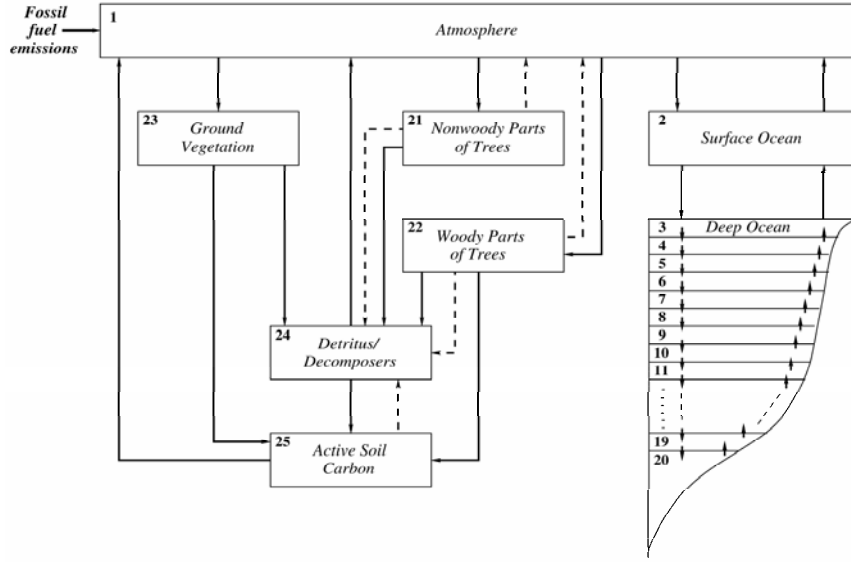

Figure 2. Diagram of the 25-compartment model.

For this model, there are 30 independent input factors that are subject to uncertainty. The description of the model input factors, their nominal values, variability ranges, and units are given in Table 2.

Table 2. The 25-compartment model input factors selected for sensitivity analysis

Description	Input Factor	Nominal Value	Range	Unit
Initial conditions:				
Atmosphere	CA0	548.80	510.7 – 596.0	Gt C
Nonwoody parts of trees	CF0	38.20	30.0 – 46.0	Gt C
Woody parts of trees	CW0	634.50	507.0 – 762.0	Gt C
Ground vegetation	CG0	59.30	47.0 – 72.0	Gt C
Detritus/decomposers	CD0	108.20	86.0 – 130.0	Gt C
Active soil carbon	CSL0	1131.00	905.0 – 1348.0	Gt C
Forest clearing:				
Fraction of forest clearing carbon transfered to atmosphere	PHIA	0.5	0.4 – 0.6	—
Fraction of forest clearing carbon transfered to detritus/decomposers	PHID	0.5	0.4 – 0.6	—
Ratio of soil to detrit./decomp. flux to forest clearing flux	PSIS	0.1	0.08 – 0.12	—
Fraction of forest clearing release that serves to decrease capacity for carbon storage in trees	SXIT	0.5	0.4 – 0.6	—
Reforestation:				
Rate of re-establishment of tree compartments	SIG	1.0E-6	0.8E-6 – 1.2E-6	$year^{-1}$
Rate coefficient controlling the time required for trees to dominate ground vegetation	SS	0.2	0.16 – 0.24	$year^{-1}$
Fraction of the change in capacity for carbon storage in trees that causes a change in capacity for storage in ground vegetation	EPS	0.5	0.4 – 0.6	—

Table 2 (continued)

Description	Input Factor	Nominal Value	Range	Unit
Physical & Chemical ocean:				
Depth of surface ocean	HM	75.0	$60.0 - 90.0$	m
Area of surface ocean	AREA	3.61E+14	$2.88E+14 - 4.33E+14$	m^2
Temperature change in surface ocean as a result of doubling atm. carbon content	DELTP	3.0	$1.5 - 4.5$	K
Total boron concentration in surface ocean	SIGB	4.1E-4	$3.27E-4 - 4.90E-4$	mol/L
Initial temperature of surface ocean	TEMP0	292.75	$290.75 - 294.75$	K
Chlorinity of surface water	CL	19.24	$15.0 - 23.0$	mL^{-1}
Relative humidity in atmosphere	RELHUM	0.75	$0.6 - 0.9$	—
Terrestrial turnover times:				
Nonwoody parts of trees	TF	1.75	$1.4 - 2.1$	year
Woody parts of trees	TW	25.00	$20.0 - 30.0$	year
Ground vegetation	TG	4.00	$3.2 - 4.8$	year
Detritus/decomposers	TD	2.00	$1.6 - 2.4$	year
Active soil carbon	TSL	100.00	$80.0 - 120.0$	year
Soil-forming fractions:				
Woody parts of trees	THW	0.1180	$0.094 - 0.14$	—
Ground vegetation	THG	0.3330	$0.26 - 0.40$	—
Detritus/decomposers	THD	0.0625	$0.05 - 0.075$	—
Intrinsic recovery times:				
Nonwoody parts of trees	TT2	20.0	$16.0 - 24.0$	year
Ground vegetation	TV2	4.0	$3.2 - 4.8$	year

3. Application of Sensitivity Analysis to
Global Carbon Cycle Models

In many works related to the SA of experiments, due to lack of knowledge about the distributions followed by the model input factors, the uniform/loguniform distributions are considered (Helton et al. [13]; Campolongo et al. [2]). For the ranges of variability, if no information exists, then different criteria may be adopted to obtain variability ranges for the model factors. For instance, Campolongo and Saltelli [3] use $\mp 20\%$ of the nominal value as a range of variability for some of the factors they consider. In the same article, they mention another criterion for calculating variability ranges: $(1\,/\,2K_0;\,2K_0)$, where K_0 is described as the nominal value of the input factor. In this study, for both GCC models under consideration, all input factors are assumed to follow uniform distributions and the uncertainty ranges (given in Tables 1 and 2) are set as $\mp 20\%$ of the nominal values.

After assigning a range and an appropriate probability distribution to each input factor, a sample input matrix is generated. There are several ways of producing input samples. In the SA framework, simple random sampling (SRS), Latin hypercube sampling, and importance sampling are commonly used. The purpose of all sampling techniques is the same: to obtain a good coverage of the sample space of the input factors. Comparison of these sampling methods and their effects on the results have been discussed by Helton and Davis [14, 15]. Because SRS is easy to implement, easy to explain, and because the models considered here are not computationally expensive to evaluate, SRS was used to generate the input matrix. Considering the large dimensions of the input spaces of the two GCC models, a sample size of $N = 5000$ was used. As a general rule suggested by Crosetto et al. [7], about 100 runs for each input factor are commonly performed. Based on this, it is believed that 5000 runs are sufficient to cover the whole input space in both models. However, to

assess the effect of different sample sizes on the analysis results and to show how to display these results in dot charts for easy comparison, we also considered $N = 100$.

In this study, SA methods based on Monte Carlo simulations are carried out by following three steps: (i) selection of random, independent sets of values for model input factors; (ii) initialization of the carbon cycle model; and (iii) simulation of GCC dynamics between years 1750 and 2100. Initialization of a carbon cycle model involves a calibration step in which the model is set up such that steady-state is maintained, i.e., the flux of CO_2 leaving compartment i is equal to the flux of CO_2 coming into that compartment, before the system in perturbed with any input. In other words, it is assumed that $dx_i / dt = 0$ in year 1750 when the model simulations are started. The 25-compartment model has a built-in calibration process in the model computer code but the computer code of the 8-compartment model does not. In order to enforce this important modelling assumption of initial steady-state, the computational procedures we introduced in a recent study can be used (see Gazioğlu and Scott [12]). One of these methods is based on treating a subset of model input factors as uncertain input factors while using others in the model calibration process. The other method, called windowing analysis, is based on adjusting the values of model input factors to achieve an acceptable match between observed and predicted model conditions. For detailed description of both methods, see the reference mentioned above.

After the initialization process puts the system in steady-state, the anthropogenic release of CO_2 from combustion of fossil fuels and changes in the land-use (mainly from deforestation) are introduced into each GCC model through the atmosphere compartment and the GCC dynamics are simulated. For each model, beginning from the preindustrial steady-state, the quantities of carbon stored in each compartment are calculated at annual intervals (from year 1750 to 2100). The 1751-2007 historic global annual CO_2 emissions from Boden et al.

[1] are used as input to the models. For predictions of CO_2 for the period 2008-2100, the scenario IS92a of the Intergovernmental Panel on Climate Change (IPCC) is used. The IS92a scenario (also known as the Business-as-Usual scenario) represents a middle-of-the range scenario due to modest and largely offsetting changes in the underlying assumptions. More information on this and the IPCC's other scenarios can be found in Houghton et al. [16].

4. Graphical Sensitivity Methods

A literature search done by Cooke and van Noortwijk [6] shows very little theoretical development for graphical methods in sensitivity analysis. There exist reference books in the literature, for example, Robbins [18]; Cleveland [4]; and du Toit et al. [8], which study visualization of univariate and multivariate data. As Cooke and van Noortwijk mention, when graphical methods are used in SA, the focus is not visualizing data as it is in the general sense, but rather on visualization to support this analysis.

Numerous graphical tools have found application in SA. Some of these are scatterplots, scatterplot matrix, overlay scatterplot, tornado charts, and cobweb plots. These graphs and their use in SA are described in detail by Cooke and van Noortwijk in Chapter 11 of the book "Sensitivity Analysis" (Edited by Saltelli et al. [20]). Here, we give a brief description of each and refer the interested reader to the aforementioned reference for more detail.

A useful non-quantitative screening technique is a sequence of scatterplots in which, each response variable (model prediction) appears on the vertical axis and each explanatory variable (model input factor) appears in turn on the horizontal axis. In SA, scatterplots of the input-output relationships are used as a guide to better understanding of the model behaviour. If the relationship is strong, this indicates that the considered model input has significant effect on the model output.

A scatterplot matrix (also called matrix plot) can also be produced. This type of plot displays the main features of the 2D relationships between each pair of variables without reference to the other variables. However, as the number of variables increases, it becomes harder to interpret the set of plots and obtain an overall sense of the data configuration. There is no doubt that generating scatterplots is the simplest SA method. One disadvantage of this technique is that, it requires drawing and inspecting a large number of plots, at least one plot for each model input factor. Considering that we may need to analyze several model outputs which may also be time dependent, then the number of plots needed becomes quite large.

Tornado charts are basically bar graphs of any global sensitivity measures (for instance, rank correlation coefficients) arranged in decreasing order of absolute value and arranged vertically.

Cobweb plots are used in identifying local probabilistically important factors. These plots give a picture of the joint distribution of the percentiles of up to 20 variables. Each parallel vertical line in a cobweb plot represents one variable, and the possible values of these variables are given on these lines. Then each set of values is marked on the vertical lines and connected by a jagged line.

Other graphical methods, such as histograms, radar charts, pie charts, and bar charts are also used in SA.

In the next two sections, we describe two modern graphical tools (namely, dot charts and star plots), which we suggest may also be used for SA. These plots can be drawn using the S language. Both S-PLUS and R, which are two implementations of S, have the 'dotchart' and 'stars' functions that can be utilized to create these two types of plots. S-PLUS is a commercial software sold by TIBCO Software Inc. R is an open source software and can be downloaded free of charge from http://www.r-project.org/.

5. Dot Charts

Dot charts are similar to bar charts with dots connected to the axis by lines instead of bars. According to research, dot charts are more easily understood than bar charts (see, for instance, Robbins [17]; Robbins [18]; Cleveland and Fisher [5]). They are less cluttered and easier to interpret.

The sensitivity measures calculated on a model output and input factors can be presented in a dot chart very effectively. Sensitivity measure is displayed on the horizontal axis and the input identifier goes up on the left side. The variable (model output) with the highest sensitivity measure is plotted furthest from the side and the variable with the lowest sensitivity measure is plotted closest to the side. See Subsection 7.1 for dot charts, we used to show results from an SA method applied to the 8-compartment model.

They are known to be more powerful especially to superpose additional data. Dot charts may be used not only to compare the influence of input factors on a single model output but also to compare the influence of input factors on a number of model outputs. For instance, in Subsection 7.1, we present dot charts, which are used to investigate if the degree of influence from an input factor changes when each compartment's CO_2 content in three different years are considered. On the same dot charts, we also display the results from two different numbers of model runs. This allows us to investigate the effect of sample size on the model predictions.

6. Star Plots

Star plots are used for representing multivariate data in two dimensions. They are made of sequences of equiangular rays called radii. Each ray extending from the center of a circle represents one variable. Then the measurements are linked with lines, which give the plot a star-like shape. The values of the measurements occupy a circle, and the fact that the starting points correspond to the end points facilitates comparison between cases.

Star plots can be used to visualize sensitivity measurements of the model response(s) to the uncertainties about the model inputs. Each ray in a star plot then represents one of the model input factors. The input factor furthest from the center of the circle indicates that this is the most influential input on the output. Conversely, an input plotted closest to the center is identified as the least influential input with the lowest sensitivity measure.

In Subsection 7.2, we use star plots to present results from an SA method applied to the 25-compartment model.

7. Results

The two types of graphs we suggest for use with SA can be used to display any quantitative parametric or nonparametric sensitivity measure. Here we present results from one parametric and one nonparametric approach. As the parametric approach, we considered standardized ranges as a measure of sensitivity for the 8-compartment model, and as the nonparametric approach, Cramér-von Mises test statistics as a measure of sensitivity for the 25-compartment model. First, we shall briefly describe these SA measures.

Standardized range: This statistic was calculated using predictions from the standard one-at-a-time (OAT) design in which each input factor is changed individually while others are kept constant and SA is performed to quantify the change in the model output.

First, the range of the CO_2 content of each compartment (say, compartment i) resulting from varying each input factor (say, initial condition j) over its range while keeping all other initial conditions at their nominal values, i.e., $Y_{ij}^{max} - Y_{ij}^{min}$ is calculated at year t. Then, this output range of compartment i is standardized by the nominal value of this compartment's initial content.

Because the compartmental outputs are time dependent functions of initial conditions of the compartments, to determine whether their importance changes through time, we examine the results from years 1900, 2000, and 2100. Each SR_{ij} is derived by using

$$SR_{ij} = [Y_{ij}^{max}(t) - Y_{ij}^{min}(t)] / x_i^{\circ},$$

where x_i° is the initial condition of compartment i at its nominal value. Also, to assess the effect of the number of model runs on the results, $N = 100$ and $N = 5000$ model simulations are considered.

Cramér-von Mises test: This is one of the nonparametric statistical tests that are often desirable in SA because of our limited knowledge of the input factors and their associated distributions.

The application of the Cramér-von Mises test to SA comes from the idea of partitioning the sample of an input factor under consideration into two sub-samples according to the quantiles of the output distribution. If the distributions of the input factor in the two sub-samples can be proven to be different, then the input is identified as an influential input. Using the test statistics as a sensitivity measure, the relative importance of the input factors for each output variable can be obtained (i.e., the higher the test statistic calculated between an input factor and an output variable, the more influential the input is on that output variable).

To express this in notation: The sample of a model input factor, say *X*, is partitioned into two sub-samples, say X_i and X_j, according to the quantiles of the distribution of the model output, say *Y*. Let X_i be size N_1 and X_j size N_2, and $S_1(x)$ and $S_2(x)$ be the empirical distribution functions of these two sub-samples, respectively. The total area between the two empirical distributions gives us the Cramér-von Mises test statistic, which is denoted by T_{CM} and is defined as

$$\mathbf{T}_{\mathrm{CM}} = \frac{N_1 N_2}{(N_1 + N_2)^2} \left(\sum_{i=1}^{N_1} [S_1(X_i) - S_2(X_i)]^2 \right.$$

$$\left. + \sum_{j=N_1+1}^{N_1+N_2} [S_1(X_j) - S_2(X_j)]^2 \right).$$

7.1. Dot charts of standardized ranges

The standardized ranges (SRs) of compartments due to varying each initial condition OAT are demonstrated as dot charts in Figures 3a and 3b. These figures allow us to judge how compartmental contents differ in terms of SRs within the same year and also between years. In each frame of these figures, we have superimposed the SRs computed on the predictions of the model outputs based on $N = 100$ and $N = 5000$ model runs, so that the influence of the number of model runs on the SRs can easily be visualized.
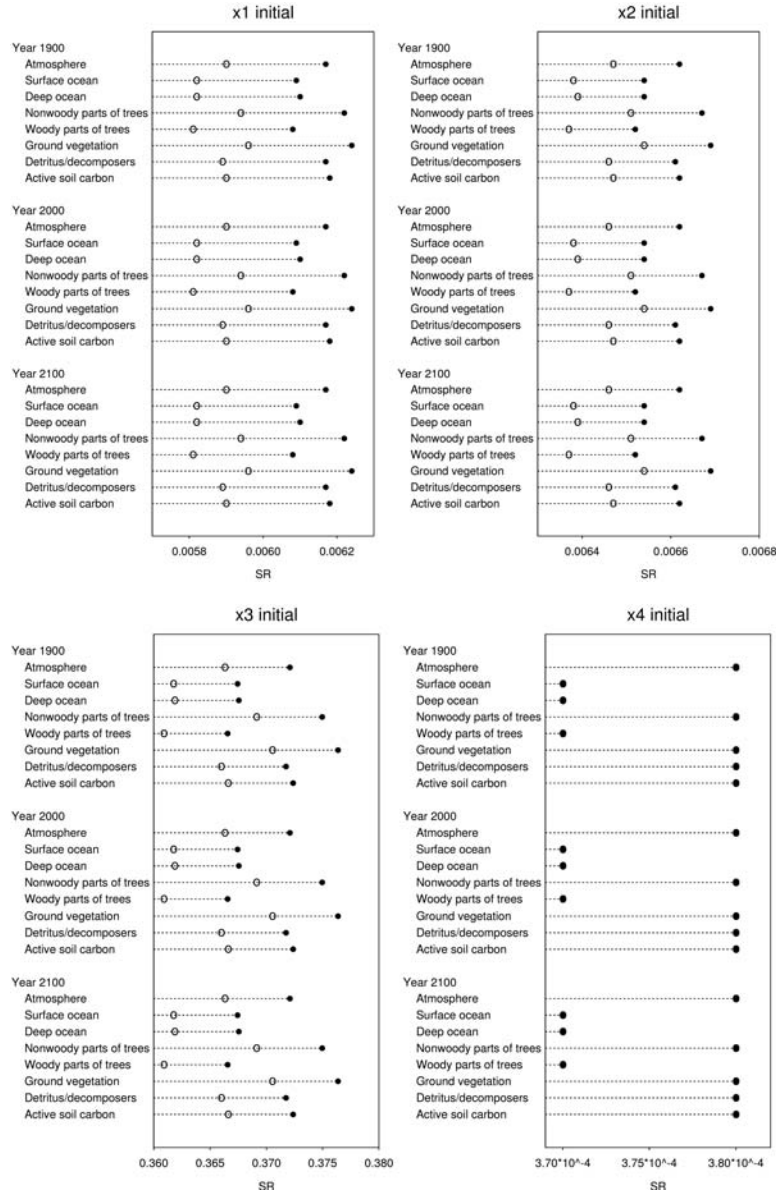
Figure 3a. Dot charts showing how each compartmental output of the 8-compartment model in years 1900, 2000, and 2100 is affected by the variation in the initial conditions x_1° through x_4° (above: x1 initial − x4 initial) in terms of standardized ranges. In each frame (∘) show results from $N = 100$ and (•) from $N = 5000$ model runs.
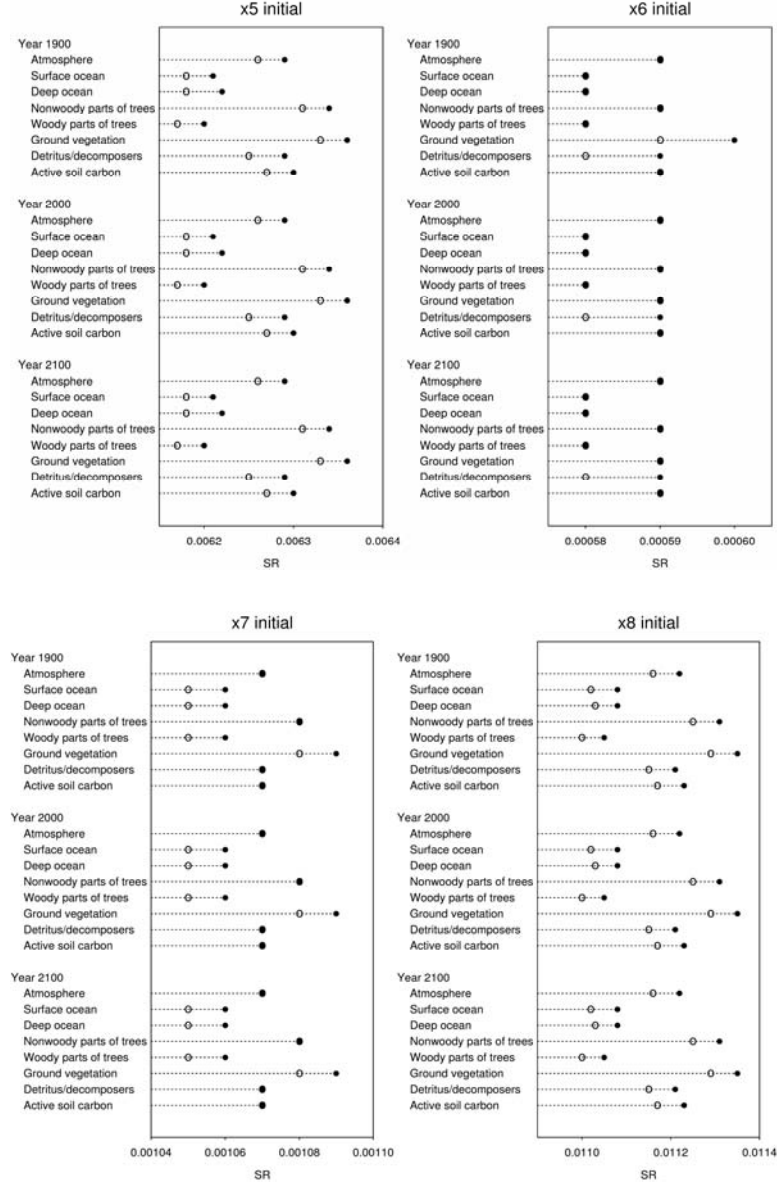
Figure 3b. Dot charts showing how each compartmental output of the 8-compartment model in years 1900, 2000, and 2100 is affected by the variation in the initial conditions x_5° through x_8° (above: x5 initial – x8 initial) in terms of standardized ranges. In each frame (∘) show results from $N = 100$ and (•) from $N = 5000$ model runs.

Even though the range of SR values are not the same, the degree of sensitivity of each compartment to the uncertainty in the initial conditions of atmosphere (x_1°) and both ocean compartments $(x_2^\circ$ and $x_3^\circ)$ does not appear to be changing, that is, the variation in all these three initial conditions influences the ground vegetation compartment the most, the nonwoody parts of trees compartment the second most, followed by the active soil carbon, atmosphere, and the detritus/decomposers compartments (see Figure 3a). The surface ocean, deep ocean, and the woody parts of trees compartments are the least influenced by these three initial conditions.

Compared to the SR results on x_3°, the SRs of all the compartments due to variation in the other seven initial conditions are much smaller (lower than 0.1), especially, the results due to uncertainty about x_4° and x_6° (see the lower-right frame in Figure 3a and the upper-right frame in Figure 3b). Considering how small the ranges of SRs are, it is difficult to say that changes in x_4° and x_6° affect any of the compartments at all. As seen in the lower-right frame of Figure 3a, the overlapping of the SRs from $N = 100$ (represented by open circles ∘) and from $N = 5000$ (represented by closed circles •) model runs indicates that there is no noticeable change in the results when different numbers of model iterations are considered in the analysis.

As Figures 3a and 3b show, the SRs hardly change from year to year no matter which N is used. The highest SR values result from the OAT design conducted on x_3°, which are in the range of 0.3611 and 0.3708 when $N = 100$, and between 0.3670 and 0.3765 when $N = 5000$ (see lower-left frame in Figure 3a).

In the dot chart showing the SRs obtained from the analysis where only x_7° was varied over its entire range, we see that with the results based on $N = 100$ model runs, the variation in x_7° is equally and most effective on nonwoody parts of trees and ground vegetation compartments followed closely by atmosphere, detritus/decomposers, and active soil carbon compartments, while the ocean and woody parts of trees compartments are the least influenced by x_7°. When $N = 5000$ model runs are considered, there is no change in the SR of four of the compartments, namely, nonwoody parts of trees, atmosphere, detritus/decomposers, and active soil carbon compartments. For the other three compartments, the SRs increase but only by 0.00001.

Again using the SRs of compartmental contents at three chosen years, we can also assess how the order of importance of the initial conditions change within the same compartment, i.e., which input factors are most influential. It is clear from the dot charts of the SRs given in Figures 3a and 3b that x_3° is the most influential initial condition on all compartmental outputs, followed in order by x_8°, x_2°, x_5°, x_1°, x_7°, x_6°, and x_4°. This order of relative importance based on the SR values does not change with the number of model runs and time.

As with x_1°, x_2°, and x_3°, the variation in the initial conditions of the woody parts of trees (x_5°) and active soil carbon (x_8°) compartments appear to influence the ground vegetation, nonwoody parts of trees, active soil carbon, atmosphere, and detritus/decomposers compartments more (in descending order of sensitivity in terms of SRs) (see Figure 3b). The effect of x_5° and x_6° on the two ocean and the woody parts of trees is relatively low.

7.2. Star plots of Cramér-von Mises test results

For the 25-compartment model, using simulation of $N = 5000$ runs, we calculated the Cramér-von Mises test statistics and presented the results (from year 2100) in Figures 4a – 4d using star plots. These plots allow us to compare the influences of the input factors on each compartmental output. The radius of the stars extending from the center of a circle represent a test statistic value computed on an input and an output variable. A long radius indicates that the corresponding input factor is important for the output variable under consideration. This graphical display proves to be a good way of picking out the most important input factors at a glance.

In Figures 4a and 4b, which show the influence of the 30 input factors on the outputs of the atmosphere and the three ocean compartments in 2100 in terms of the Cramér-von Mises test statistic based on both the 50th and the 90th quantile partitioning, respectively, it is clear that the input factors AREA and HM have quite a significant influence on these four compartments. We should note that since the examination of star plots for all 18 deep ocean compartments revealed that the results do not change significantly, to conserve space we present results only from two of the 18 deep ocean compartments, namely, 'deep ocean layer-5' and 'deep ocean layer-13'. Although the degree of importance of these two inputs is very close, the relative importance between these two most important inputs change with the two quantiles considered. For instance, for the surface ocean compartment with the 50th quantile partitioning AREA appears to be the most important and HM the second important input (see Figure 4a), but with the 90th quantile partitioning this order is reversed (see Figure 4b).
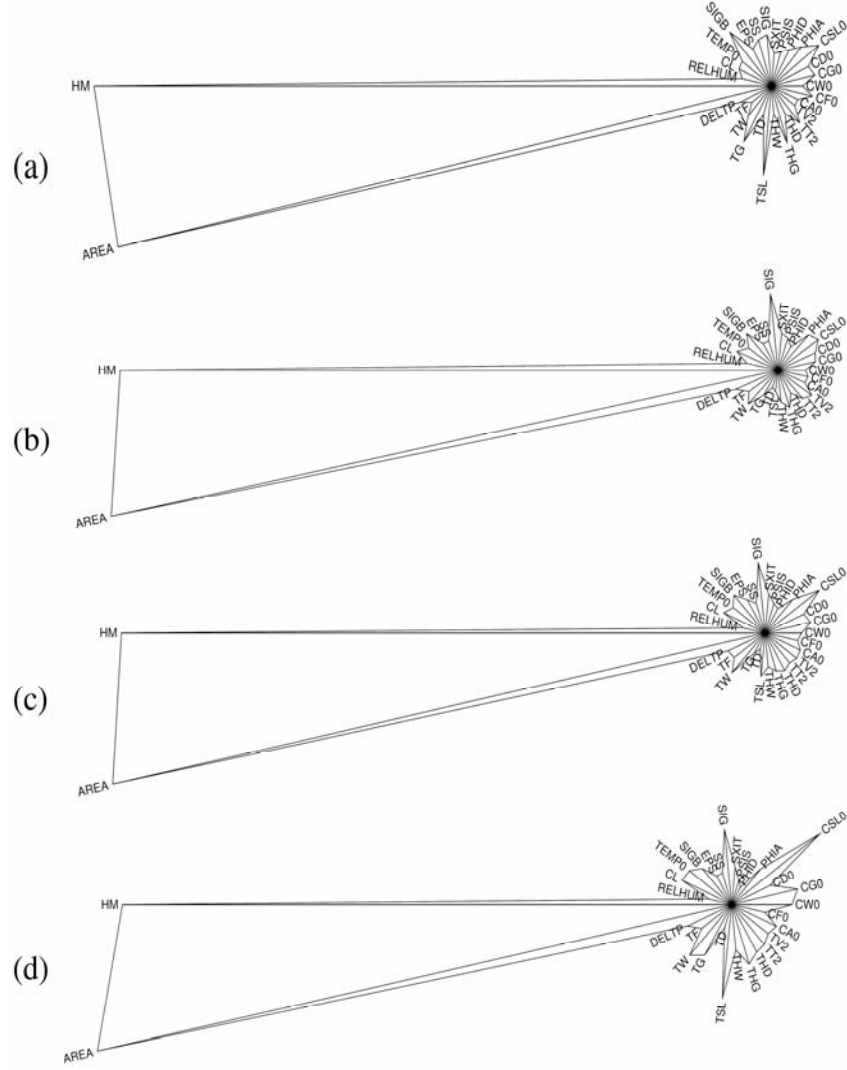
Figure 4a. Star plots of Cramér-von Mises test statistics based on the 50th quantile partitioning for the output of: (a) atmosphere; (b) surface ocean; (c) deep ocean-layer-5; and (d) deep ocean-layer-13 compartments of the 25-compartment model at $t = 2100$.
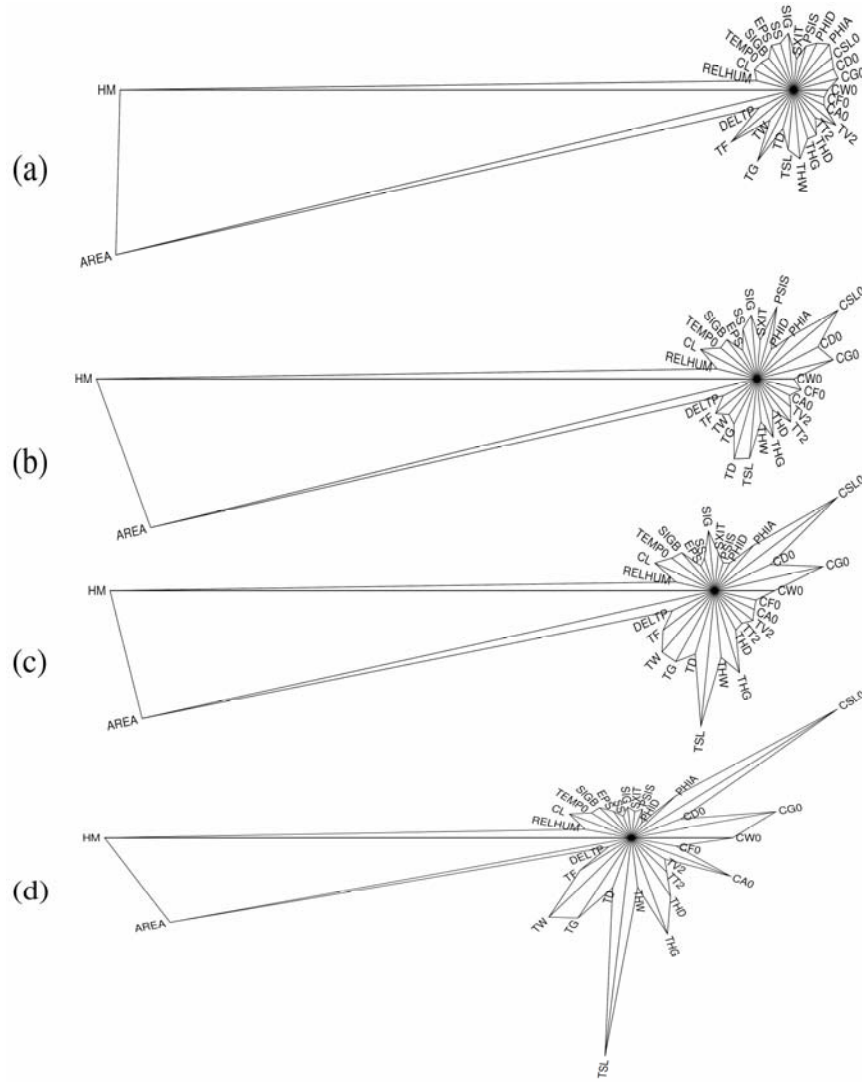
Figure 4b. Star plots of Cramér-von Mises test statistics based on the 90th quantile partitioning for the output of: (a) atmosphere; (b) surface ocean; (c) deep ocean-layer-5; and (d) deep ocean-layer-13 compartments of the 25-compartment model at $t = 2100$.
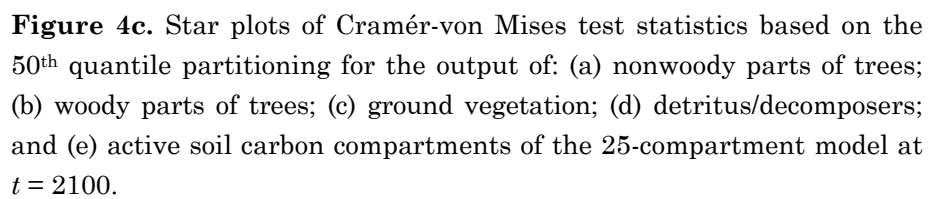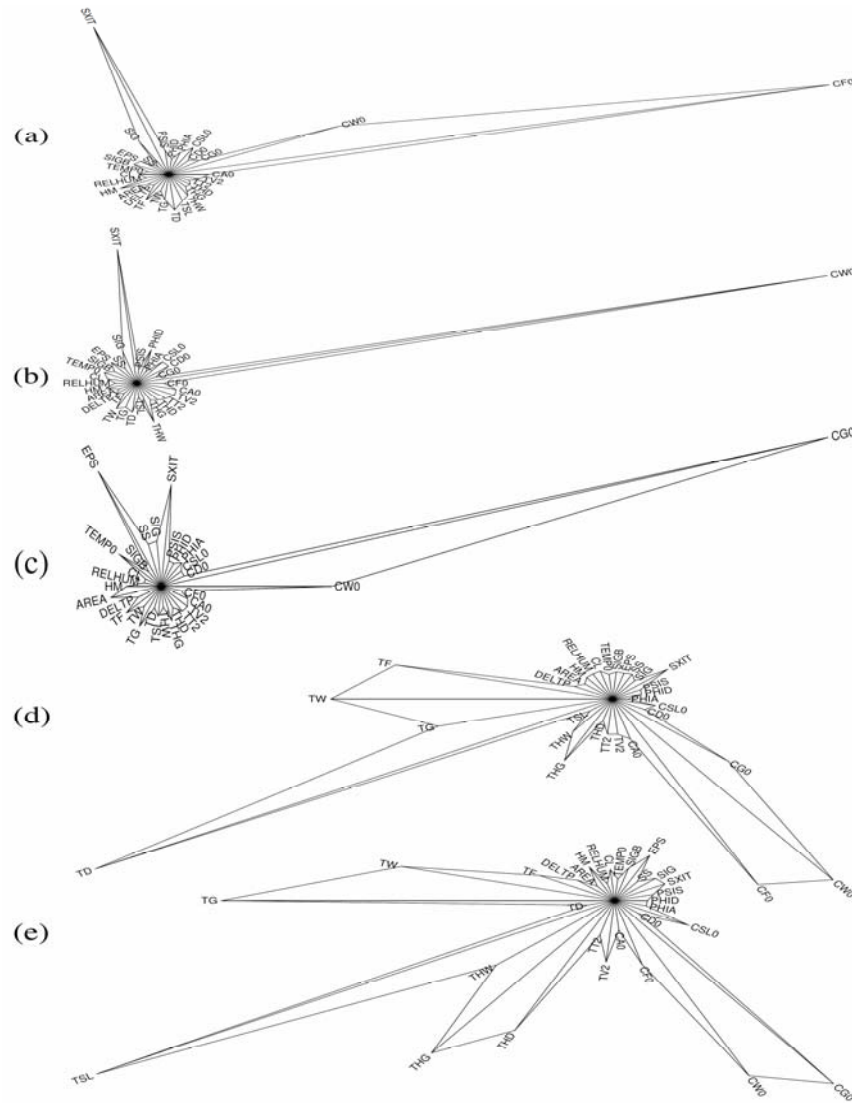
Figure 4c. Star plots of Cramér-von Mises test statistics based on the 50th quantile partitioning for the output of: (a) nonwoody parts of trees; (b) woody parts of trees; (c) ground vegetation; (d) detritus/decomposers; and (e) active soil carbon compartments of the 25-compartment model at $t = 2100$.

Figure 4d. Star plots of Cramér-von Mises test statistics based on the 90th quantile partitioning for the output of: (a) nonwoody parts of trees; (b) woody parts of trees; (c) ground vegetation; (d) detritus/decomposers; and (e) active soil carbon compartments of the 25-compartment model at $t = 2100$.

Compared to AREA and HM, the rest of the input factors do not appear to have much influence on the atmosphere and the three chosen ocean compartments (especially when the partitioning is based on the median). Figure 4a(a) shows that based on the 50th quantile partitioning the input factors TSL is the third and SIGB the fourth most influential inputs. Based on the 90th quantile partitioning (see Figure 4b(a)), TF is the third and TG the fourth most influential inputs, following AREA and HM. Similarly, an importance ranking of the input factors for the other output variables related to the other compartments can be obtained by examining their star plots.

8. Discussion and Conclusion

Models are needed to understand real world phenomena. However, any model is, at best, only an approximation of the system being modelled and is therefore inherently uncertain. This uncertainty can be reduced, and resulting model reliability increased, by carrying out sensitivity analysis (SA). This important modelling tool studies the effects of variations in model input parameters on the behaviour of model output(s). In this study, our focus is on the use of novel graphical tools with SA. While various commonly used graphical tools, such as scatterplots, histograms, radar charts, tornado charts, cobweb plots, bar charts, have found applications in SA, as Cooke and van Noortwijk [6] point out, there has been little theoretical development for the use graphical methods in SA.

We suggest the use of dot charts and star plots with SA. In this study, we have demonstrated their application to a pair of compartmental models, which have been used for global carbon cycle modelling. The sensitivity measures we computed and displayed in the dot charts and the star plots were the standardized ranges on a 8-compartment model and the Cramér-von Mises test statistics on a 25-compartment model.

The dot charts we present in this study allowed us to compare the SRs of all model outputs (the CO_2 contents of all eight compartments at time *t*) to the uncertainty in each model input factor. They are powerful tools, allowing us to superpose additional data and so judge not only how compartmental contents differ in terms of SRs within a given year but also between years (such as 1900, 2000, and 2100). Additionally, by superimposing the SRs computed on the model predictions based on two different numbers of model runs ($N = 100$ and $N = 5000$) we were able to easily see the influence of N on the SRs.

In these dot charts, we displayed the SRs on the horizontal axis and the compartment names on the vertical axis, making it easy to read the names of the compartments with no need to rotate or abbreviate them. Another advantage dot charts provide is that the scale along the horizontal axis does not require a zero baseline. This is important when sensitivity measures are similar to one another, as a zero baseline can hide detail. Dot charts with no zero base line make it easy to see even small differences.

Star plots can be used with SA to compare the relative behaviour of all model input factors on model output(s). The length of the ray for any given model input represents the extent to which model output is sensitive to that input, making it easy to identify which input factors dominate the sensitivities of model output. By comparing star plots, we were also able to identify similarities and differences in the sensitivities of two or more model outputs to the same model inputs.

We should note that these two types of graphs can be used to present not only the sensitivity measures we used in this study but also other sensitivity measures, such as sensitivity indices, the Pearson moment-correlation coefficients, the Spearman rank correlation coefficients, standardized regression coefficients, standardized rank regression coefficients, the Morris means, the Smirnov test statistics, the Mann-Whitney test statistics etc..

The use of dot charts and star plots to present SA results on our test models showed that both of these graphical tools convey information quite well. They allowed us to easily identify the most influential input factors at a glance. They both are great tools which display sensitivity results in a clear, easy to read, and easy to interpret fashion. As neither of these graphs relies on colour, there is no loss of clarity if they are used in black and white publications.

Both types of plots can be used in the preliminary stage of a study as screening tools, helping model users and/or developers identify which model input factors are influential, hence deserving close attention, and which ones are not so influential and can be left at their nominal values. This can significantly reduce the cost (both physical and computational) of an experimental study, especially when only a few of input factors have a significant effect on the model output(s).

Acknowledgement

References

[1]   T. A. Boden, G. Marland and R. J. Andres, Global CO_2 Emissions from Fossil-Fuel Burning, Cement Manufacture, and Gas Flaring: 1751-2008. NDP-030, CDIAC, Oak Ridge National Laboratory, (2011). Available at

http://cdiac.ornl.gov/ftp/ndp030/global.1751-2008.ems

[2]   F. Campolongo, S. Tarantola and A. Saltelli, Tackling quantitatively large dimensionality problems, Comput. Phys. Commun. 117 (1999), 75-85.

[3]   F. Campolongo and A. Saltelli, Sensitivity analysis of an environmental model: An application of different analysis methods, Reliability Engineering and System Safety 57 (1997), 49-69.

[4]   W. S. Cleveland, Visualizing Data, Hobart Press, Murray Hill, NJ, 1993.

[5]   W. S. Cleveland and N. I. Fisher, Good graphs for better business, The Quality Magazine 7(4) (1998), 64-68.

[6]   R. M. Cooke and J. M. Noortwijk, Graphical Methods, In Sensitivity Analysis, A. Saltelli, K. Chan and E. M. Scott, Eds., John Wiley & Sons, 245-264, West Sussex, 2000.

[7]   M. Crosetto, S. Tarantola and A. Saltelli, Sensitivity and uncertainty in spatial modelling based on GIS, Agric. Ecosyst. Environ. 81 (2000), 71-79.

[8]   S. H. C. du Toit, A. G. W. Steyn and R. H. Stumpf, Graphical Exploratory Data Analysis, Springer-Verlag, New York, 1986.

[9]   W. R. Emanuel, G. G. Killough, W. M. Post and H. H. Shugart, Modeling terrestrial ecosystems in the global carbon cycle with shifts in carbon storage capacity by land use change, Ecology 65 (1984), 970-983.

[10] W. R. Emanuel, G. G. Killough, W. M. Post, H. H. Shugart and M. P. Stevenson, Computer Implementation of a Globally Averaged Model of the World Carbon Cycle, TR010, DOE/NBB-0062, Oak Ridge National Laboratory, Oak Ridge, TN, 1984.

[11] S. Gazioğlu, Application of Sensitivity and Uncertainty Analyses to Linear, Time-Invariant Compartmental Models, Ph.D. Thesis, Department of Statistics, University of Glasgow, Glasgow, UK, 2002.

[12] S. Gazioğlu and E. M. Scott, Sensitivity analysis of linear time-invariant compartmental models with steady-state constraint, J. Appl. Stat. 38 (2011), 2485-2509.

[13] J. C. Helton, J. D. Johnson, M. D. McKay, A. W. Shiver and J. L. Sprung, Robustness of an uncertainty and sensitivity analysis of early exposure results with the MACCS reactor accident consequence model, Reliability Engineering and System Safety 48 (1995), 129-148.

[14] J. C. Helton and F. J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering and System Safety 81 (2003), 23-69.

[15] J. C. Helton and F. J. Davis, Sampling-Based Methods for Uncertainty and Sensitivity Analysis, Sandia National Laboratories, SAND99-2240, New Mexico, 2000.

[16] J. T. Houghton, B. A. Callander and S. K. Varney, (Editors), Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment, WMO/UNEP, International Panel on Climate Change, Bracknell, 1992.

[17] N. B. Robbins, Dot Plots: A Useful Alternative to Bar Charts, (2006). Available at

http://cdiac.esd.ornl.gov/ndps/ndp030.html

[18] N. B. Robbins, Creating More Effective Graphs, John Wiley & Sons, Hoboken, NJ, 2005.

[19] A. Saltelli, What is Sensitivity Analysis, In Sensitivity Analysis, A. Saltelli, K. Chan and E. M. Scott, Eds., John Wiley & Sons, 3-13, West Sussex, 2000.

[20] A. Saltelli, K. Chan and E. M. Scott, (Editors), Sensitivity Analysis, John Wiley & Sons, West Sussex, 2000.

[21] R. Tomović and M. Vukobratović, General Sensitivity Theory, American Elsevier Publishing Company, Inc., New York, 1972.

∎