MORE Sensitivity Analysis of the MSM-BIGMOD River Murray Flow and Salinity Model

Ravalico, J.K. [1], G.C. Dandy [1] and H.R. Maier [1]

[1] School of Civil and Environmental Engineering, The University of Adelaide, South Australia
Email: jravalic@civeng.adelaide.edu.au

Keywords: *Sensitivity analysis, River Murray, environmental modelling, decision-making, natural resources management*

EXTENDED ABSTRACT

The MSM-BIGMOD model of the River Murray in South Eastern Australia is a comprehensive flow and salinity routing model, used to assess the impacts of potential changes in river management on river flow and salinity levels. The modelling suite consists of a combination of five models that have been developed over a period of several years. Sensitivity analysis of the model is particularly important, given that decisions are made about management of the River Murray based on outputs from the model. The large number of model inputs and parameters arising from the inclusion of the many tributaries, storages, drains, and diversions pose a challenge for traditional sensitivity analysis methods, such as one-at-a-time parameter perturbation methods.

The Management Option Rank Equivalence (MORE) method of sensitivity analysis is an innovative method of sensitivity analysis developed especially for use with complex models used for decision-making. The MORE method assesses the sensitivity of management decisions based on model output, to changes in the model inputs, in order to provide a sensitivity analysis in the decision context. The MORE method is based on the premise that potential management options are ranked based on model output. The analysis then attempts to locate the Rank-Equivalence Boundary (REB), which is the surface in parameter space where the management options become equal in rank. In this implementation the MORE method uses a genetic algorithm to search for the minimum and maximum normalised Euclidean distances between the calibrated model parameter vector and the REB in parameter space.

This research applies the MORE method to MSM-BIGMOD in order to assess the sensitivity of the choice between two different salt interception schemes (SIS), in different reaches of the river, to variations in a multiplication factor of the travel time and dead storage at different flows, for the 9 reaches of the river, shown in Figure 1.

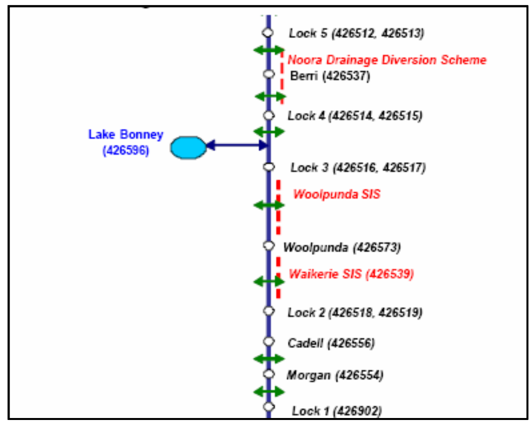

Figure 1. Section of the River Murray analysed for sensitivity to travel time and dead storage parameters (MDBC 2002)

The sensitivity results obtained found a minimum normalised Euclidean distance to the REB equal to 0.783, which, given a uniform variation in the multiplication factor across the 18 parameters, was equivalent of varying the travel time and storage for each reach by ±18.5% of their original values. Within this variation there would be no change in the management decision, indicating that the decision is not highly variable. The results also found a maximum Euclidean distance to the REB of 3.453, which was 90% of the maximum possible distance from the calibrated model parameter vector. This indicates that there are very few solutions which are guaranteed to change the management decision. Over 99% of the parameter space lies in a region where it is uncertain whether the management decision will change, indicating that there is considerable variation in sensitivity in different directions in parameter space, and that further sensitivity analysis may be useful.

By providing information on possible variations in management decisions based on variations in parameters, the MORE method is shown to be an effective method of sensitivity analysis for use in decision-making.

1. INTRODUCTION

Sensitivity analysis is recognised as a key component of model development (Saltelli et al. 2000). Not only can sensitivity analysis form a part of model calibration, informing the modeller of which parameters will have considerable impact on the model when altered, but it can also allow identification of ways that the model can be simplified, thus helping to avoid over-parameterisation (Tarantola et al. 2002). Further, it can help modellers to identify any unexpectedly strong dependencies on parameters which should not be highly influential, allowing correction of the model (Saltelli and Scott 1997). In integrated and environmental modelling, it is often the case that uncertainties are poorly understood, while taking large samples or repeating experiments is often not possible (Norton 1996). In this case, where there may also be limited resources, sensitivity analysis provides a way for modellers and decision-makers to identify key parameters, and allocate resources towards further data collection appropriately.

As the size of a model increases, simultaneously with the complexity of its interactions, sensitivity analysis becomes increasingly important. In the case of large and integrated models, non-linearities and non-monotonicity are common place, and model outputs may not always be intuitive. Thus, reliable and thorough sensitivity analysis is the key to informing modellers of the internal workings of the model. Well established techniques for sensitivity analysis, such as Fourier Amplitude Sensitivity Testing (FAST) (Cukier et al. 1978; Saltelli and Bolado 1998), and the method of Sobol' (Sobol' 1993; Sobol' 2001) use analysis of variance to assess the contribution of the variance in a single parameter to the variance in the output, while methods such as Morris one-at-a-time factor screening (Morris 1991), attempt to identify unimportant parameters such that further analysis can be simplified.

In the case where a model is being used for decision making, it is important to determine whether the model outputs from alternative management options are significantly different from one another, given the possible variation in the model parameters. In this instance, a new type of sensitivity analysis is required. The Management Option Rank Equivalence (MORE) method (Ravalico et al. 2006) is a new method of sensitivity analysis, which endeavours to provide an assessment of sensitivity of the management decision to variation in the model inputs and parameters. The method uses parameter bounding techniques and optimisation to determine the minimum change in model inputs that may cause a change in the preference ranking of potential management options. This research is an application of the MORE method to the MSM-BIGMOD modelling suite for flow and salinity in the River Murray.

The MSM-BIGMOD modelling suite is a comprehensive flow and salinity model of the River Murray in South-Eastern Australia. Beginning with the inflows from Dartmouth Dam, the model incorporates tributaries, storages, weirs, irrigation and urban diversions, salt interception schemes, drainage diversions, wetlands and flood runners. The model operates through a process of hydrological routing, which involves dividing the river into reaches, each with different flow parameters, and variations due to the different inputs (MDBC 2002).

The modelling suite is a combination of five models: MSM, a monthly simulation model that computes irrigation demands, resources assessment and water accounting, MODFLW, which converts monthly values computed in MSM into daily input files for use in BIGMOD, GETDVM, which creates monthly inputs from MSM for BIGMOD, BIGMOD, which is a daily flow and salinity routing model from the Hume Dam to Lake Alexandrina, and BIGARKW, which is used to analyse the results of MSM and BIGMOD. Of these, BIGMOD and MSM are the key calculation models, and can be run separately or sequentially using outputs from MSM as inputs to BIGMOD (MDBC 2002).

The large number of inputs and parameters in the modelling suite, and their potential interactions, prohibits standard single parameter variation as a method of sensitivity analysis. Further, use of the model in decision-making and the importance of the decisions made, make MORE sensitivity analysis an ideal option. The BIGMOD model is selected in this research as a starting point for sensitivity analysis of the entire modelling suite.

2. MORE METHOD OF SENSIVITY ANALYSIS

The Management Option Rank Equivalence (MORE) method of sensitivity analysis was developed specifically for use with models used in decision-making. It provides an assessment of the sensitivity of a management decision to changes in model inputs. The aim of the method is to assist decision-makers to assess the suitability of a particular model in making specific management decisions, given known or approximated parameter uncertainties, as well as providing assistance in selecting between different management options.

In a situation where a decision-maker is selecting between two or more potential management options, choices are often made through ranking of management choices, with the option that is ranked highest selected to be put into practice, and the options with lower rankings discarded. These rankings may occur on the basis of one or several different model outputs.

The situation that is of interest to the decision-maker is that where changes in model inputs cause a change in the ranking of potential management options, such that the decision previously made would be altered. In the case where a decision-maker is considering two options, the question of changes in the rank of management options can be considered by investigating different regions of the parameter space. The parameter space can be separated into three regions; one where the rank of the management options will alter, one where they will not, and a boundary region, separating the two previous regions, where the ranks of the two options are equal. It is this Rank-Equivalence Boundary (REB) that is of interest to decision-makers, as it represents the boundary of a change in management option. Given a large number of parameters, and a model which may be non-linear and non-monotonic, as is the case with many integrated models, locating the rank equivalence boundary is problematic. The MORE method of sensitivity analysis overcomes this problem by searching the REB for the parameter vector that is the minimum distance (D_{min}) from the calibrated model parameter vector (**x**), and that which is the maximum distance (D_{max}) from the calibrated model parameter vector, with parameter ranges normalised to [0,1]. D_{min} and D_{max} are then used to characterise the REB by separating the parameter space into three distinct regions.

The first of these, S, is identified as a hypersphere with radius D_{min} and centre **x**. This represents the region of parameter space over which the parameter vector can vary, without altering the ranking of management options. The second of these, C, is identified as the region of parameter space outside of the hypersphere of radius D_{max}, and represents that region of parameter space where it is certain that there will be a change in the rank of management options. The final region, U, which is represented by the volume within the hypersphere of radius D_{max} and centre **x**, but not including S, is an uncertain region – where we are unsure whether the ranking of management options will change or not. Thus, based on potential uncertainties in the model inputs, a decision-maker can assess whether their choice of management option is sufficiently robust, or whether there needs to be further data collection, or review of the model being used before a decision can be made.

The MORE method also provides a decision-maker with information regarding changes in sensitivity in different model directions in parameter space, or different combinations of parameter changes. If the parameter space is viewed as the entire feasible region of model parameters, then the feasible space has unitary volume, since the parameter ranges are normalised to [0,1]. This space is divided into three subsets that each have their own volume. The volume of the inner sphere, S, gives us the fraction of the total parameter space where the management decision will not change, the volume of the region C gives the fraction of the feasible space where the decision is certain to change, and the volume of region U gives the fraction of the feasible space that is uncertain. A large volume of the region U indicates a significant change between the minimum and maximum distances from the calibrated model parameter vector to the REB. This in itself indicates that rather than being uniformly sensitive, the sensitivity of the management decision is variable in different directions in parameter space. In other words, while some combinations of individual parameter changes may not affect the decision significantly, other combinations, which may have a similar combined value, are likely to have a substantial effect. The volume of the set U can be used as an indicator to determine whether more information on the location of the REB is required. Where U is small, the sensitivity can be considered to be reasonably uniform, and resultantly further analysis is not required. However, where U is large it is desirable to undertake further analysis on the model.

In the situation that the hyperspheres bounding set S and set C do not fall entirely within the parameter space, the volume is too complex to calculate exactly. Due to the high dimensionality of the parameter space, and the potential inaccuracy of Monte-Carlo approximation techniques over the considerable search space, a hyper-cube with the same volume and centre as the hypersphere in question is used to approximate the volume which lies within parameter space.

Further information gained from the method can assist a decision-maker in determining the required tolerances for the model parameters. The minimum distance from the calibrated model parameters to the REB indicates the smallest combined change in parameter values before the management decision made becomes incorrect. Thus if a decision-maker can ensure that the combined parameter uncertainties would not lead to a combined change in model parameters equal to or greater in magnitude than D_{min}, then they can be assured that their decision is robust.

3. MSM-BIGMOD FLOW AND SALINITY MODEL

The MSM-BIGMOD flow and salinity model is a comprehensive model of the River Murray in South Eastern Australia. The model includes irrigation and urban diversions from the river, salt interception schemes and drainage diversions, storages and weirs on the river, wetlands and flood runners, as well as the tributaries. Historical rainfall and evaporation data are used to calculate gains and losses from river reaches and storages or weirs.

Flow routing in the river is based on hydrological routing, with the downstream flow of each reach calculated as:

$$q_{out} = q_{in} + S_{start} - S_{end} - d - \frac{E \times A}{100} - L_{hf} - L_{cm} \qquad (1)$$

where q_{out} is flow out of the reach, q_{in} is flow into the reach, S_{start} is the reach storage at the start of the day, S_{end} is the reach storage at the end of the day, d is the distributed diversion, E is the Net evaporation rate, A is the surface area of the reach, L_{hf} are the high flow losses from each reach and L_{cm} are the continuous monthly losses. The storage in each reach is considered to be a function of upstream flows, and the corresponding storage values for each flow are contained within a lookup table as inputs into the model. For each reach, 10 different flow values are stored with corresponding travel time values, particular to that reach. Using the travel times and dead storage for each reach, the reach storage can then be derived (MDBC 2002).

Salinity within the model is treated as being contained within parcels of water. The flows upstream and downstream of the reach or sub-reach under consideration are determined prior to salinity calculations for each time step. A water balance is used and the movement of the parcels of water with particular salinity concentrations is tracked, providing the overall salinity of the water within a particular reach. Lakes are assumed to be fully mixed, and weirs may be modelled in the same way as a normal reach, or may be fully mixed. Water lost or gained due to evaporation or rainfall will concentrate or dilute the salinity as appropriate, however, water lost due to high flow or constant monthly losses does not affect salinity (MDBC 2002).

Changes to the Murray River are assessed by running the model over a benchmark period from 1891 to 2000 for flow modelling and 1975-2000 for salinity modelling, under current conditions, as well as under the proposed conditions. The flow and salinity outputs from the two different runs are compared to assess the impact that the proposed changes would have on the current condition of the river over a considerable period of time.

4. ANALYSES

The BIGMOD modelling suite is currently used to assess the impact of changes in flow and salt interception into the river on the salinity levels in electrical conductivity units (EC) at Morgan, in South Australia. The model is run twice over the benchmark period, once with the current data, and once with the altered data, and the change in EC at Morgan assessed. In this study BIGMOD was used to assess the impact of two management options, involving the implementation of a salt interception scheme (SIS) in the river reach between Lock 3 and Lock 2. One option considered locating the SIS upstream of Woolpunda, while the other option considered an SIS downstream of Woolpunda.

4.1. Distance metric

In order to determine the distance between the calibrated model parameters and the REB, a number of distance metrics can be selected. These include, but are not limited to, the Manhattan (or City Block) distance, the Infinity Norm distance or the Euclidean Distance. While each of these measures have their advantages and disadvantages, the normalised Euclidean distance is selected in this instance, which is given by:

$$d_2 = \sqrt{\sum_{i=1}^{k} \left(\frac{x_i - x_i^{'}}{x_{i\max} - x_{i\min}} \right)^2}$$

Where d_2 is the normalized Euclidean distance, x_i is the original value of the ith parameter, x_i' is the new value of the ith parameter, and k is the total number of parameters under investigation.

Using the normalised Euclidean distance transforms parameter space to represent a unit hypercube. This also allows the sensitivity results to be assessed using the different properties of Euclidean geometry.

4.2. Search Algorithm

The computational efficiency and accuracy of the MORE method is reliant upon selection of an appropriate search method to locate the parameter vectors on the REB that are the minimum and maximum distance from the calibrated model parameter vector, **x**. For problems of high dimensionality, gradient methods tend to converge to local optima and can be highly dependant on the starting point (Elbeltagi et al. 2005). Evolutionary algorithms (EAs) are search algorithms that mimic natural biological processes, such as evolution, in order to optimize an objective function. EAs have been found to outperform traditional mathematical optimisation techniques in comparative studies (Elbeltagi et al. 2005).

Given the likely complexity of the REB, a genetic algorithm (GA) is used for this implementation of the MORE method. The GA is an evolutionary algorithm, based on Darwinian principles of survival of the fittest (Goldberg 1989). Each parameter vector is considered a chromosome, with each individual parameter considered to be a gene on the chromosome. The fitness of each chromosome as the problem solution is a combination of two measures, the distance from the calibrated model parameter vector to the solution represented by the chromosome and the constraint of location on the REB. The fitness is based first on the amount that the constraint is violated, with those chromosomes with the smallest violation considered the fittest. For chromosomes without constraint violation, the distance is evaluated, and those with either the minimum or maximum distance (depending on which search is being performed) from the calibrated model parameter vector are considered the fittest

The GA used in this instance is real coded. Each gene on the chromosome contains a real parameter value. In order to prevent repetition of results through parameter inheritance, the child parameter value which would be inherited from the parent is randomly selected from a normal distribution with a mean corresponding to the parent value, and a standard deviation of one sixth the distance between the two parent values (Gibbs et al. 2005). The standard deviation is selected such that there will be only minor overlap (less than 0.5%) between the distributions generated from each parent chromosome.

Elitism is incorporated within the GA, such that the fittest chromosome from each generation is preserved and included in the tournament of the next generation, replacing the least fit chromosome of the tournament winners. A mutation operator is also included in the GA, to increase diversity of the solutions. A chromosome is selected for mutation with a probability of 0.5. Once a chromosome is selected for mutation, one of the parameters of the chromosome is randomly selected to be replaced with a value randomly generated from the parameter distribution.

4.3. Model Analysis

For illustration purposes of the implementation of the MORE method, the model was run for a shortened period from April 1983 to May 1992.

To further assist with the computational efficiency of running a sensitivity analysis on such a large model, the model has been sectioned, such that the analysis can be performed on the reaches of the model downstream of lock 5. As well as reducing model runtime, this reduces the number of parameters under investigation, making it feasible to run a thorough sensitivity analysis.

The parameters selected for sensitivity analysis are the travel time and dead storage in the reaches of the river from lock 5 to lock 1, shown in Figure 1. There are 9 reaches in this section of the river, giving 18 sets of values for travel time and dead storage. The values of travel time and dead storage for each reach of the river vary based on the flow in the river at that particular time-step. For the analysis, it was assumed that the general shape of the relationship between flow and storage, or flow and travel time, is fairly well known. Hence the analysis was performed using a multiplier factor to vary the model parameters. Rather than altering each value individually, the values for storage or travel time at each flow level are multiplied by the same factor, which ranges between 0.1 and 10. It is the sensitivity of the management decision to these multiplication factors that is assessed.

The management decision that is the subject of this sensitivity analysis is the selection between two potential salt interception schemes, one between lock 3 and Woolpunda and the other between Woolpunda and Lock 2. There will be different levels of salt removal from the two schemes and the management choice is which scheme should be implemented. The decision is based entirely on changes to the 95th percentile of salinity in EC units at Morgan over the time period considered. The management option that has the lower 95th percentile EC value will be selected. Management option 1 for the analysis is implementing a scheme between Lock 3 and Woolpunda, and management option 2 is implementing a scheme between Woolpunda and Lock 2. In both cases, there was assumed to be no other SISs operating in either reach. Using the calibrated model parameters, management option 1 reduces the 95th percentile by 21 EC and is ranked first, while management option 2 is ranked second, with a reduction of 16 EC from the base case.

The MORE method was run, varying the travel time and dead storage for the data nodes corresponding to the starting and finishing points of the six reaches from Lock 5 to Lock 1. For each of the data nodes, there are two parameters; the scaling factor for the travel times (t) and the scaling factor for the dead storage (s). All parameters have a calibrated value of 1, a minimum of 0.1 and maximum of 10. As a GA is used, all analyses are repeated ten times with different random number seeds, in order to minimise the impact of any random effects resulting from the stochastic nature of the algorithm. The three best solutions are shown in Section 5.

5. SENSITIVITY RESULTS

The results for the best three runs of the MORE method are shown in Figure 2 and 3 and Table 1.

Based on the smallest value found for D_{min} and calculating the maximum equal change in all parameters, the results indicate that provided all parameters stay within ±18.5% of their original values, the implementation of the SIS upstream of Woolpunda will remain the preferred management option, base on the EC outputs at Morgan. This indicates that the decision is quite robust in relation to the travel time and dead storage parameters.

Despite this, the volume of set S, shown in Figure 3 appears very small. This is due to the considerable range that the parameters are tested over. In this analysis parameters ranges vary between -90% and +1000% of the original parameter value. The considerable parameter range in combination with the high dimensionality of the parameter space, results in smaller volumes being representative of large changes in individual parameters.

The values found for D_{max} are 90% of the maximum possible Euclidean distance from the calibrated model parameter vector to the extremities of parameter space. This can be seen in Figure 2, which also shows a comparison with the values of D_{min}. This, in combination with the skewed position of the calibrated model parameter vector in parameter space, indicates that there is likely to be very little area within the parameter space where a change in management options is certain to occur. This is supported by the cube approximation of the volume of C, which is shown in Table 1 to be zero. The small volume where there would be a certain change of management option is missed by the necessary volume approximation.

As shown in Table 1 and Figure 3, the volume of U is almost 1, indicating that almost the entire volume of the selected parameter space lies within the uncertain region. The large size of U also indicates that sensitivity is highly varied in different parameter directions. This can also be seen in Figure 4, which shows the individual normalised parameter changes for the best of the MORE runs. The large variation in parameter changes between the minimum and maximum variation is evident. It should be noted that the individual parameter changes in Figure 4 do not give general information about the sensitivities of the decision to the individual parameters. The changes are indicative of sensitivity to the parameters in only two specific directions, which are the direction of the minimum and maximum Euclidean distances, and may vary considerably in other different directions.

In order to gain a better understanding of the different contributions of each parameter at different locations in parameter space it would be beneficial to perform further sensitivity analysis on the model.

The results show that the decision of SIS selection is robust in relation to the travel time and dead storage factors in the reaches of the River Murray between Lock 5 and Lock 1.

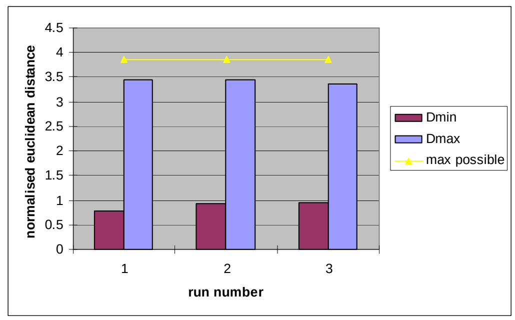

Figure 2. Minimum and maximum distance to REB for three different implementaions of the MORE method

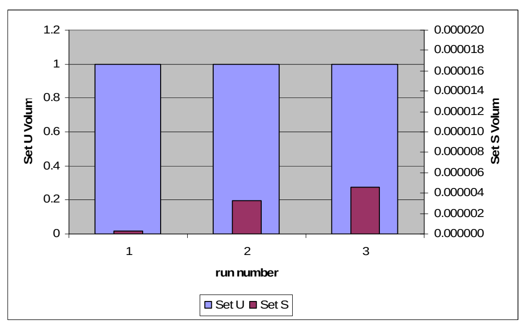

Figure 3. volumes of set S and set U

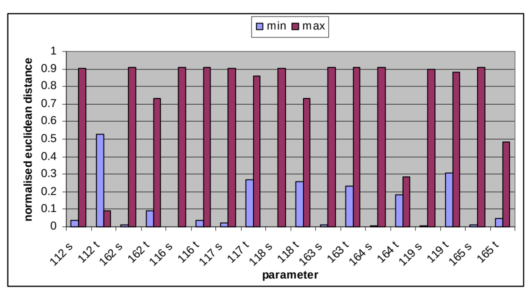

Figure 4. Individual normalised parameter changes for the best min and max runs achieved

Table 1: Numerical sensitivity results from five repetitions of the MORE method

run	D_{min}	D_{max}	S	U	C
1	0.783	3.453	2.73E-07	0.999	0
2	0.931	3.453	3.32E-06	0.999	0
3	0.951	3.365	4.55E-06	0.999	0

6. CONCLUSIONS

The MORE method has been used effectively to analyse the sensitivity of the selection of an SIS, using the BIGMOD model of the River Murray. The implementation of the MORE method represents a new direction in sensitivity analysis which investigates the sensitivity of decisions made based on model output, which may involve more than one model output, rather than simply investigating the sensitivity of the outputs to the parameters individually. In this instance the decision of whether to implement an SIS between Lock 3 and Woolpunda, or Woolpunda and Lock 2 in order to reduce salinity at Morgan was investigated. It was found that a reasonable amount of variation in the model parameters of travel time and storage was needed across the 9 reaches under investigation, before it was preferable introduce the SIS downstream of Woolpunda rather than upstream, as indicated by the original model parameters.

7. ACKNOWLEDGEMENTS

The authors would like to acknowledge Barry Croke and Pradeep Sharma for their assistance in preparation of the BIGMOD model for analysis. The authors would also like to acknowledge funding support for this research from the Australian Research Council, Murray Darling Basin Commission and the Department of Water Land and Biodiversity Conservation of the Government of South Australia.

8. REFERENCES

Cukier, R. I., H. B. Levine and K. E. Shuler (1978), Nonlinear Sensitivity Analysis of Multiparameter Model Systems. *Journal of Computational Physics*, 26: 1-42.

Elbeltagi, E., T. Hegazy and D. Grierson (2005), Comparison among five evolutionary-based optimimzation algorithms. *Advanced Engineering Informatics*, 19: 45-53.

Gibbs, M. S., G. C. Dandy, H. R. Maier and J. B. Nixon (2005), Selection of Genetic Algorithm Parameters For Water Distribution System Optimization, World Water & Environmental Resource Congress, ASCE, Anchorage, AK, USA.

Goldberg, D. E. (1989), Genetic Algorithms in Search, Optimization, and Machine Learning, Addison- Wesley Publishing Company, Inc.

MDBC (2002), Setting up of MSM-BIGMOD modelling Suite for the River Murray System. Canberra, Murray Darling Basin Commission.

Morris, M. D. (1991), Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2): 161-174.

Norton, J. P. (1996), Roles for deterministic bounding in environmental modelling. *Ecological Modelling*, 86: 157-161.

Ravalico, J. K., G. C. Dandy and H. R. Maier (2006), Rank-Equivalence method for Sensitivity Analysis of an Integrated Model of a River Catchment., Proceedings of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software", Burlington, USA, International Environmental Modelling and Software Society.

Saltelli, A. and R. Bolado (1998), An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics & Data Analysis*, 26(4): 445-460.

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

Saltelli, A. and M. Scott (1997), Guest editorial: The role of sensitivity analysis in the corroboration of models and its link to model structural and parametric uncertainty. *Reliability Engineering & System Safety*, 57(1): 1-4.

Sobol', I. M. (1993), Sensitivity Estivmates for Nonlinear Mathematical Models. *Mathematical Modelling & Computational Experiment*, 1: 407-414.

Sobol', I. M. (2001), Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1-3): 271-280.

Tarantola, S., N. Giglioli, J. Jesinghaus and A. Saltelli (2002), Can global sensitivity analysis steer the implementation of models for environmental assessments and decision-making?. *Stochastic Environmental Research and Risk Assessment (SERRA)*, 16(1): 63-76.