Assessing Prediction Uncertainty in the BIGMOD Model: A Shuffled Complex Evolution Metropolis Algorithm Approach

Fernando, T.M.K.G.[1], H.R. Maier[1], G.C. Dandy[1] and B.F.W. Croke[2]

[1]School of Civil and Environmental Engineering, The University of Adelaide, SA 5005, Australia
[2]Mathematical Sciences Inst./Integrated Catchment Assessment and Management Centre, Australian National University, ACT 0200, Australia
Email: gfernand@civeng.adelaide.edu.au

Keywords: *Parameter uncertainty, prediction, salinity, Markov Chain Monte Carlo, shuffled complex evolution*

EXTENDED ABSTRACT

There are a number of large scale computer simulation models being used in Australia to support the development and implementation of salinity management strategies in the River Murray system. BIGMOD is one such model used by the Murray Darling Basin Commission (MDBC) to model the major processes of flow, salinity, extractions and diversions.

The BIGMOD model was originally calibrated against historical flow and salinity data from Dartmouth Dam to upstream of Lake Alexandrina in South Australia using traditional parameter estimation methods. The calibrated model has been used extensively to provide salinity predictions to Murray Darling basin managers who have been making important salinity management decisions based on the predictions. The accuracy and reliability of the model predictions play a vital part in the success of the salinity management strategy. However, like predictions obtained with any other complex and large scale hydrological and water quality model, the BIGMOD model predictions contain uncertainty. Traditional uncertainty analysis approaches can be used successfully to quantify parameter uncertainty and prediction uncertainty in simple hydrological models. However, these methods cannot be used with large integrated models, such as BIGMOD, as computational requirements are too high.

This study focuses on estimating the confidence intervals for BIGMOD salinity predictions by quantifying the uncertainty associated with significant parameters in the BIGMOD model using the recently developed Shuffled Complex Evolution Metropolis (SCEM-UA) algorithm. The SCEM –UA algorithm is an effective and efficient evolutionary Markov Chain Monte Carlo (MCMC) sampler, which combines the search capabilities of the Shuffled Complex Evolution (SCE-UA) algorithm with the Metropolis algorithm.

Part of the South Australian reach of the River Murray system, from Lock 5 to Lock 1, was modelled in this study. Prior knowledge of the modelling processes under investigation was used to select the significant parameters whose uncertainty was assessed using the SCEM-UA algorithm. The BIGMOD model calibration was carried out in 2 stages. First, the flow was calibrated against flow data upstream of Lock 1. Then salinity was calibrated against salinity data at Morgan. Uncertainty associated with 2 significant parameters, travel time and dead storage, was estimated in each reach during the SCEM-UA BIGMOD model calibration. Confidence intervals for salinity predictions at Morgan and flow upstream of Lock 1 were estimated.

The results indicate a high level of uncertainty for each of the 16 parameters modeled. However, this variation is likely to be due to the fact that measured flow and salinity data are only available at one location each, and that different combinations of parameter values of the model can achieve the same output values. The high degree of correlation between parameters was confirmed by the fact that the confidence bounds on the predictions obtained were quite narrow, despite the high degree of uncertainty associated with individual parameters.

In addition to providing insight into the model parameters and confidence intervals for the model predictions, the SCEM-UA method produced slightly better results when compared with the predictions obtained with traditional model calibration. This highlights the ability of automated calibration methods to determine good model parameter vectors without any *a priori* knowledge of the system being modelled.

1. INTRODUCTION

Over the last 2 decades, large scale computer simulation models have become extremely useful tools for managing large river basins. BIGMOD (Close 1996) is one such model used by the Murray Darling Basin Commission (MDBC) to model and predict major processes such as flow, salinity, extractions and diversions in the River Murray system. The model has been used extensively to provide salinity predictions to Murray Darling basin managers who have been making important salinity management decisions based on the predictions. The accuracy and reliability of the model predictions play a vital part in the success of the development and implementation of salinity management strategies in the River Murray system. However, like predictions obtained with any other complex and large scale hydrological and water quality model, the BIGMOD model predictions contain uncertainty. This may arise from various sources. For example, the BIGMOD model itself can contain uncertainties due to the approximations in the system representations; the available data defining the input conditions and calibration events likely contain measurement errors and calibrated parameters can have uncertainties. Traditional uncertainty analysis approaches can be used successfully to quantify parameter uncertainty and prediction uncertainty in simple hydrological models. However, these methods cannot be used with large integrated model such as BIGMOD, as computational requirements are too high.

Unlike simple hydrological models, the BIGMOD model simulates a number of complex processes. Consequently, the number of parameters involved in representing these processes is considerably higher when compared with a simple hydrological model. Moreover, BIGMOD simulates the River Murray by dividing the river system into a number of river reaches. Each of these reaches can contain up to 50 parameters, which have to be estimated during model calibration. The BIGMOD model contains more than 200 river reaches, making traditional calibration approaches extremely labour intensive. Furthermore, the success of traditional parameter estimation techniques is also strongly dependent on the experience of the modeller.

Automatic methods for model calibration have been developed to overcome the major issues associated with manual model calibration. Such methods are easier to implement and take advantage of the speed and power of computers to estimate parameters and have therefore become increasingly popular. However, most of these methods encounter difficulties when trying to find global parameter estimates (Duen *et al.* 1992, Sorooshian *et al.* 1993). The shuffled complex evolution (SCE-UA) global optimisation algorithm proposed by Duen *et al.* (1992) is generally considered to be the most efficient and effective global parameter optimisation algorithm used in hydrological model calibration. However, automatic model calibration methods are aimed at finding a single best set of parameter values for a given system and do not take account of parameter uncertainty.

Uncertainty analysis techniques, such as evaluation of the likelihood ratio (Beven and Binley 1992), Markov Chain Monte Carlo (MCMC) algorithms (Kuczera and Parent 1998) and bootstrap techniques have been developed and successfully applied to hydrological models to estimate parameter uncertainties. Among them, MCMC algorithms have become increasingly popular for estimating the posterior probability distribution of parameters in hydrological models, as they are able to successfully manage nonlinear, complex models. However, MCMC methods require *a priori* definition of a sampling distribution. The choice of the sampling distribution determines the explorative capabilities of the Markov Chain sampler and its rate of convergence. A poor choice of the proposal distribution will result in slow convergence of the Markov Chain and an inability to recognise when convergence to a limiting distribution has been achieved (Vrugt *et al.* 2003).

An adaptive MCMC sampler entitled the Shuffled Complex Evolution Metropolis algorithm (SCEM-UA) was proposed by Vrugt et al. (2003) to improve the search efficiency of the MCMC sampler. It operates by merging the strengths of the Metropolis (Metropolis *et al.* 1953) and SCE-UA algorithms (Duan *et al.* 1992). The SCEM-UA algorithm is a MCMC sampler that provides an estimate of the most likely set of parameter values and underlying posterior distribution within a single optimisation run. Previous studies suggested that the adaptive capabilities of the SCEM-UA algorithm can significantly reduce the number of model simulations required to estimate the posterior distribution of the parameters when compared with traditional Metropolis-Hastings samplers (Vrugt *et al.* 2006, Vrugt *et al.* 2003).

This study focuses on estimating the confidence intervals for BIGMOD salinity predictions by quantifying the uncertainty associated with significant parameters in the BIGMOD model using the SCEM-UA algorithm. The SCEM-UA algorithm estimated parameter values are compared with the parameter values of the MDBC calibrated model. Modelled salinity predictions at Morgan, as well as flow predictions upstream of Lock 1, are also compared with observed salinity values and flow values at Morgan and upstream of Lock 1, respectively. Furthermore, the predictions are compared with MBDC calibrated model predictions.

2. METHODOLOGY

In this study, the BIGMOD modelling suite was combined with SCEM-UA to quantify the uncertainties associated with significant model parameters and to assess the confidence intervals for model predictions.

2.1. BIGMOD Model Calibration

The BIGMOD model simulates the River Murray system by dividing the river into a number of reaches. In each river reach, major processes, such as flow, salinity, losses, extractions and diversions, are modelled using different techniques. For example, flow is modelled by using hydrologic flow routing techniques and salinity is routed by tracking parcels of water called 'markers'. The model was originally calibrated against the historical data from Dartmouth Dam to upstream of Lake Alexandrina in South Australia in 2 stages (MBDC 2002). First, the flow routing and transmission loss calibration was carried out by setting the flow at an upstream gauging station in the model to the recorded data and routing the flow down to the next flow gauging station. The parameters estimated during this calibration process include the relationship between flow and travel time and the flow versus loss relationship. The salinity routing was calibrated next using the estimated parameter values from the flow calibration. The salinity calibration was accomplished by following a similar approach to the flow calibration. As part of the salinity calibration, salinity at an upstream gauging station was set to measured data and then routed to the next available salinity gauging station. The dead storage in each reach was calibrated during this process.

2.2. Shuffled Complex Evolution Metropolis Algorithm

The SCEM-UA algorithm is an adaptive MCMC sampler, which generates multiple sequences of set of parameter values which converge to the stationary posterior distribution for a sufficiently large number of simulations. The basic steps in the SCEM-UA algorithm can be summarised as follows. First, an initial population of points that is distributed randomly throughout the feasible parameter space is generated. For each point, the posterior density is computed. Next, the population of parameter values is partitioned into a number of complexes. In each complex, a parallel sequence is launched from the point that exhibits the highest posterior density. A new candidate point in each sequence is generated using a multivariate normal distribution. The new candidate point is added to the current sequence by testing the Metropolis-annealing (Metropolis et al, 1953) criterion. Finally, the new candidate point is shuffled into the original population of complexes. The evolution and shuffling procedures are repeated until each of the parameters converges to a stationary posterior target distribution. A detailed description of the SCEM-UA method can be found at Vrugt et al. (2003).

3. CASE STUDY

Part of the South Australian reach of the River Murray system, from Lock 5 to Lock 1, was modelled in this study (Figure 1). This river section was divided into 9 reaches in the BIGMOD model, details of which are given in Table 1. The total number of reaches modelled was 8, as reach number 162 was not modelled in this study.

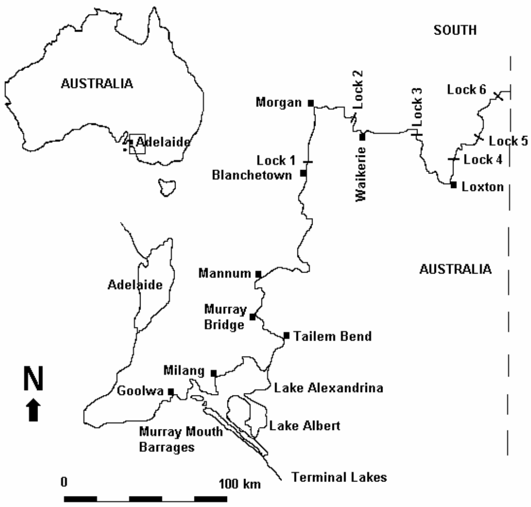

Figure 1. Location map of the modelled part of the River Murray system

The significant parameters to be estimated during the SCEM-UA calibration process were identified as the flow travel times and dead storages in each reach. The flow travel time is modelled within BIGMOD as a relationship with flow and was also used to estimate the salinity travel time. For each reach, MDBC estimated 10 travel time values corresponding to the different flow levels, as shown in Table 1. Instead of modelling 10 travel times separately for each reach, the travel times

Table 1. Reach descriptions and MBDC estimated parameter values for dead storages and travel times

Reach No	Description	Dead storage in reach (ML)	Travel time (days)									
			0	5,000	10,000	20,000	30,000	50,000	70,000	100,000	140,000	300,000
112	Lock 5 - Berri	22415	0.22	0.76	0.76	0.90	1.12	2.28	1.94	2.38	1.94	1.94
162	Berri - Lock 4	6085	0.06	0.19	0.19	0.22	0.28	0.57	0.48	0.60	0.48	0.48
116	Lake Bonney	0	0	0	0	0	0	0	0	0	0	0
117	Lock 4 - Lock 3	45000	0.25	0.80	0.80	0.90	1.60	3.95	6.00	4.60	2.50	2.50
118	Lock 3 - Woolpunda	7359	0.05	0.12	0.12	0.12	0.12	0.38	0.74	0.44	0.32	0.32
163	Woolpunda - Waikerie	13101	0.06	0.17	0.17	0.17	0.16	0.53	1.01	0.61	0.44	0.44
164	Waikerie - Lock 2	7540	0.05	0.13	0.13	0.13	0.12	0.39	0.76	0.45	0.33	0.33
119	Lock 2 - Morgan	28075	0.17	0.55	0.55	0.54	0.70	0.75	1.55	1.55	1.55	1.55
165	Morgan - Lock 1	28925	0.15	0.48	0.48	0.46	0.60	0.65	1.35	1.35	1.35	1.35

were calibrated in this study by multiplying the MDBC calibrated values by a factor.

Consequently, a single multiplication factor was optimised for each reach by assuming that the MBDC calibrated flow-travel time relationship shapes were accurate. Dead storage corresponds to the water in deep holes and in weir pools which is present even at very low flows, and is the most significant parameter estimated during the salinity calibration.

Observed data from May 1983 to May 1990 were used in this study. The first 2 years of data were used to warm up the model and data from May 1985 to May 1990 were used for model calibration and uncertainty estimation.

In the SCEM-UA model, the sample size was set to 1,000 and the number of complexes was set to 10. These values were obtained by trial and error and are similar in order of magnitude to values suggested in previous studies. In the preliminary studies, the BIGMOD model parameters converged to a minimum after about 15,000 model evaluations. Therefore, the maximum number of simulations was set to 30,000, as this enabled the final 10,000 samples to be used to obtain the uncertainty estimates.

The BIGMOD model calibration was conducted in 2 stages. First, the travel times for each reach were calibrated against the flow upstream of Lock 1. The average travel time values for each reach were computed by using the 1,000 travel time values obtained from the final sample. These calibrated values were then used to calibrate against the salinity data. The dead storages were estimated by calibrating the modelled salinity values against measured salinity at Morgan. Dead storage values were also estimated by using a multiplication factor to increase the computational efficiency of the calibration process. The multiplication factors were allowed to vary between 0.1 to 10 to give enough search space to find the optimum values.

4. RESULTS AND DISCUSSION

This section discussed the preliminary results obtained by analysing the final parameter sample, which contained 1,000 probable values for each parameter.

The posterior distributions of the travel time multiplication factors were estimated using the final 1,000 samples and are shown in Figure 2 for 4 reaches. For some reaches, most of the travel time multiplication factors were less than 1.0, suggesting that the travel time values were overestimated as part of the MDBC calibration process. For example, for reach 112, around 75% of the estimated values were less than 1.0. As a result, approximately 75% of the travel times computed by the SCEM-UA algorithm were smaller than those estimated with the traditional calibration approach. The mean and median travel time values estimated using SCEM-UA were less than the MDBC estimated values, as shown in Figure 3.

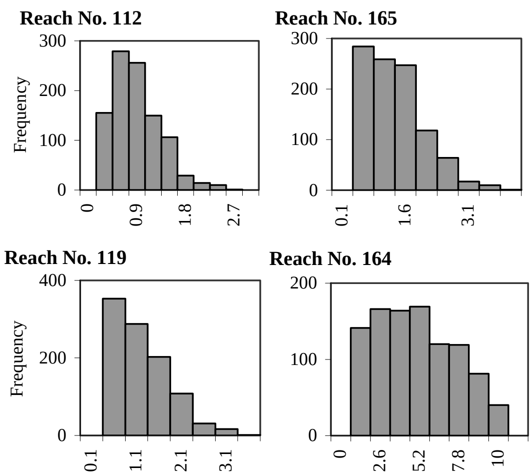

Figure 2. Posterior distribution of travel time multiplication factors for reach No. 112, 165, 119 and 164

Table 2. Comparison of SCEM-UA calibrated dead storages and MDBC calibrated dead storages (in ML)

	Reach No.							
	112	162	117	118	163	164	119	165
SCEM-UA (mean)	68595	6963	9574	4775	5403	4577	6025	152247
SECM-UA (median)	66791	6366	8748	4142	4906	4055	5818	151528
MDBC	22415	6085	45000	7359	13101	7540	28075	28925

On the other hand, for some reaches, such as reach No. 165, most of the estimated values were scattered in the vicinity of 1.0. Consequently, the mean and median travel time values were similar to the MDBC estimated values (Figure 3).

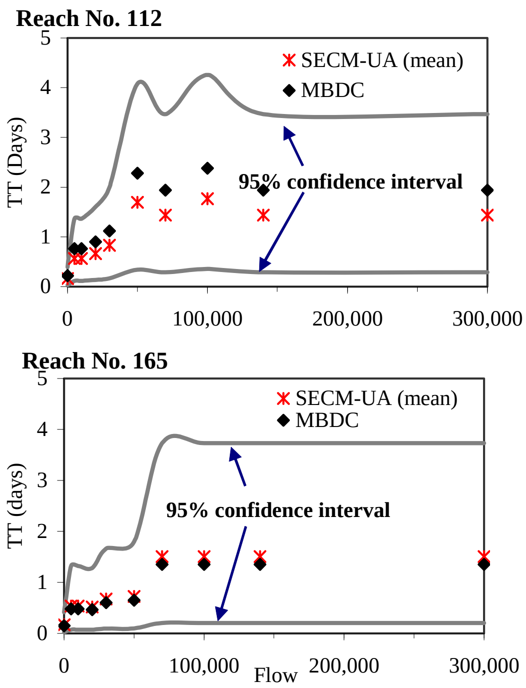

Figure 3. Comparison of MDBC estimated travel times with SECM-UA calibrated travel times for reach no. 165 and 112.

The SCEM-UA parameter optimisation approach resulted in significantly higher travel times for some reaches. For instance, for reach No. 164, only about 10% of the travel time multiplication factors were less than 1.0 and more than 50% were greater than 4. In such cases, the SCEM-UA estimated travel times were significantly greater than the MDBC estimated travel times.

The SCEM-UA estimated dead storage multiplication factors followed a similar variation to the travel time multiplication factor estimates (Figure 2 and Figure 4).

For some reaches, the SCEM-UA estimated dead storages were strongly related to the MDBC estimated values. In particular, for reach number 162, the mean and median dead storages estimated were of the same order of magnitude as the MDBC estimated dead storages, as shown in Table 2. For some reaches, the SCEM-UA estimated values were significantly lower than the MDBC estimated values. For instance, for reach number 119, the MBDC estimated a dead storage of 28,075 ML. However, most of the SCEM-UA calibrated values were in the range of 6,000 ML (Figure 4), even after allowing dead storages to vary from 2,807.5 ML to 280,750 ML during optimisation. Reach no. 117 also indicated a similar variation (Table 2). On the other hand, for some reaches, the SCEM-UA approach resulted in significantly higher dead storages (eg. reach number 112 and 165).

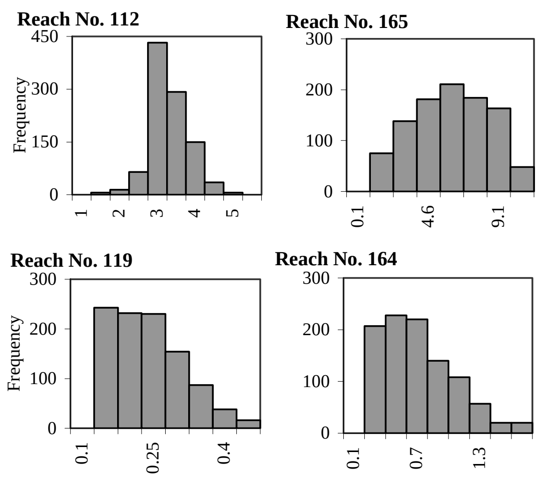

Figure 4. Posterior distribution of dead storage multiplication factors for reach No. 112,165, 119 and 164

Uncertainties in model predictions were assessed by conducting 1,000 BIGMOD simulations with the parameter values from the final sample. The salinity at Morgan, as well as flow upstream of Lock 1, were extracted from the BIGMOD model outputs and further analysed and compared with the MDBC calibrated model outputs.

The results obtained indicate that the salinity values computed with the SCEM-UA algorithm estimated parameters were similar to the MDBC predictions. The 95% confidence intervals for the salinity time series (Figure 5) were relatively narrow, even though they were obtained using 16 parameters, each of which had a great deal of uncertainty associated with it. This suggests that the parameters obtained using the SCEM-UA algorithm are highly correlated. This makes sense from a physical perspective and provides further insight into the wide range of parameter variations obtained as part of the calibration process. As flow and salinity data are only available at one location each, and different parameters for travel time and dead storage need to be found for a number of reaches, there is a large number of combinations of parameters (e.g. small travel time in reach one followed by a large travel time in reach two and vice versa etc.) that are able to match the measured flow and salinity data. This highlights the importance of incorporating as much *a priori* system knowledge into the automated calibration process as possible, so that the most physically plausible parameter vectors can be selected from those obtained as part of the SCEM-UA analysis. In other words, even though there are many parameter combinations that are able to reproduce the measured data, not all of them would make physical sense. As the algorithm only considers the impact of the combination of parameters on the predictions at a single measuring point, the relative values of parameters in individual reaches are not necessarily physically plausible.

The plots in Figure 5 also indicate that the measured data do not fall within the confidence bounds of the predictions. This suggests that not all sources of uncertainty have been included in the analysis. Additional sources of uncertainty could include the uncertainty of other model parameters or the structure of the model itself (e.g. not all of the processes affecting salinity are described adequately in the model).

The prediction accuracy of the models calibrated by MDBC and using the SCEM-UA algorithm was compared by calculating the root mean square error and correlation coefficient between observed and modelled values. For the SCEM-UA approach, the mean of the 1,000-modelled values was used as the predicted value. The results indicate that the SCEM-UA model calibration method is slightly better than the conventional BIGMOD model calibration method, for both salinity and flow predictions (Table 3).

Table 3. Error indicators for MDBC calibrated salinity and flow values and SCEM-UA calibrated salinity and flow values. (MOS= Salinity at Morgan, L1UF= Flow at up stream of Lock 1, RMSEs in EC units for salinity and in ML for flow)

		RMSE	r
MOS	MDBC	125	0.86
	SCEM-UA	120	0.87
LIUF	MDBC	8314	0.92
	SCEM-UA	7388	0.95

5. CONCLUSION

In this study, the recently developed SCEM-UA algorithm was applied to the large scale simulation model BIGMOD, which models flow and salinity in the River Murray, Australia, to quantify parameter and output uncertainty. The results indicate a high level of uncertainty for each of the 16 parameters modeled. However, this variation is likely to be due to the fact that measured flow and salinity data are only available at one location each, and that different combinations of model parameters can achieve the same model output. The high degree of correlation between parameters was confirmed by the fact that the confidence bounds on the predictions obtained were quite narrow, despite the high degree of uncertainty associated with individual parameters. This highlights the importance of incorporating as much *a priori* knowledge into the calibration process as possible, as not all parameter combinations that result in a good match between modeled and predicted data make physical sense.

The results obtained also indicate that not all measured salinity data fall between the modeled confidence bounds. This suggests that not all sources of uncertainty were included in the analysis.

In addition to providing insight into the model parameters and confidence intervals for the model predictions, the SCEM-UA method produced slightly better results when compared with the predictions obtained with traditional model calibration. This highlights the ability of automated calibration methods to determine good set of model parameter values without any *a priori* knowledge of the system being modeled. However, as pointed out above, such knowledge is required to choose the physically most plausible set of parameter values generated.

6. ACKNOWLEDGMENTS

Funding for this project was provided by the Centre for Natural Resource Management, Department of Water, Land and Biodiversity Conservation, South Australia (Project 054125). The authors would also like to thank Pradeep Sharma and Andrew Close from the MDBC for their valuable input and advice.

7. REFERENCES

Beven, K.J. and A.M. Binley, (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, 6, 279-298.

Close, A. (1996), A new daily model of flow and solute transport in the River Murray, In *proceedings of the 23rd Hydrology and water resources symposium*, Hobart, Australia, 21-24 May, 173-178.

Duan, Q., V.K. Gupta and S. Sorooshian (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, 28, 1015-1031.

Kuczera, G. and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *Journal of Hydrology*, 211, 69-85.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth and A.H. Teller (1953), Equations of state calculations by fast computing machines, *Journal of Chem. Phys*. 21, 1087-1091.

Murray Darling Basin Commission (2002), Setting up of MSM-BGIMOD modelling suite for the River Murray system, Technical report No 2002/5.

Sorooshian, S., Q. Duan and V.K. Gupta (1993), Calibration of rainfall-runoff models: Application of global optimisation to the Sacramento soil moisture accounting model, *Water Resources Research*, 29, 1185-1194.

Vrugt, J.A., H.V. Gupta, S.C. Dekker, S. Sorooshian, T. Wagener and W. Bouten (2006), Application of stochastic parameter optimization to the Sacramento soil moisture accounting model, *Journal of hydrology*, 325, 288-307.

Vrugt, J.A., H.V. Gupta, W. Bouten and S. Sorooshian (2003), A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resources Research*, 39(8), 1201. doi: 10.1029/2002WR001642.

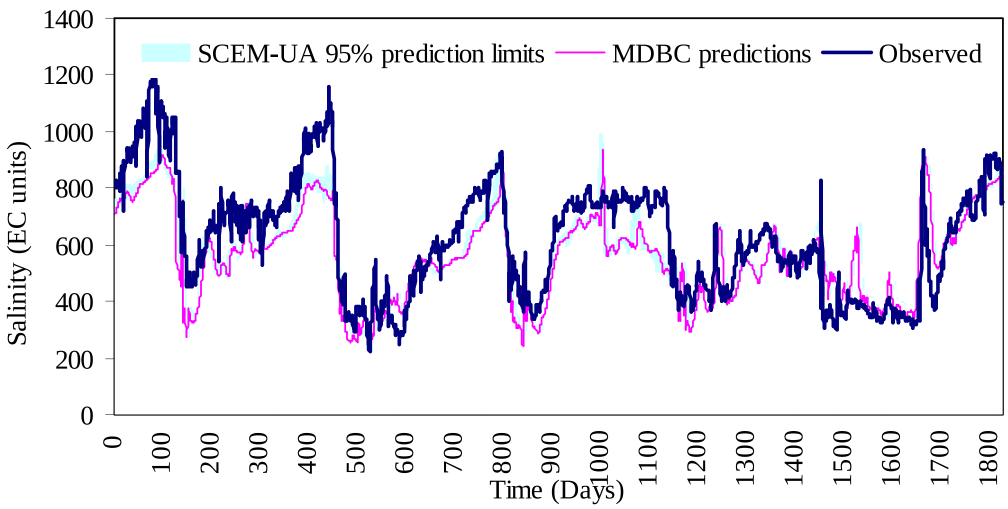

Figure 5. The 95% prediction limits for SCEM-UA BIGMOD model predictions and BIGMOD predictions obtained with MBDC calibrated parameters