

**HYDROLOGY AND FISH POPULATION DYNAMICS IN THE OKAVANGO BASIN: MANAGING FOR
UNCERTAINTY IN A DATA POOR ENVIRONMENT**

BY

ANNA McDANIEL CATHEY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

Chair: Rafael Muñoz-Carpena

Co-chair: Greg Kiker

Committee:

Mike Allen

Mark Brown

Mathew Cohen

Piotr Wolski

Contents

Chapter 1. Introduction	5
Motivation	11
Objectives	13
References	13
Chapter 2. Uncertainty, sensitivity, and impacts for scenario modeling using the Pitman hydrologic model in the Okavango Basin	19
Abstract.....	19
1. Introduction	20
2. Methods.....	26
2.1 The Pitman Model.....	26
2.2 Input Factor Selection	28
2.3 Morris Global Sensitivity Analysis	34
2.4 FAST Variance Based Global Sensitivity Analysis.....	35
2.5 Climate Change Uncertainty	36
3. Results.....	37
3.1 Morris Method	37
3.2 FAST Variance Based Global Sensitivity Analysis.....	38
3.3 Climate Change Uncertainty	41
4. Conclusions	42
Acknowledgements	45
References	45
Tables.....	54
Figure Captions	59

Chapter 3. Objectively defining uncertainty in the face of data scarcity using a reservoir model of the Okavango Delta, Botswana	73
Abstract.....	73
1. Introduction	74
2. Methods.....	78
2.1 The Study Site: Okavango Delta, Botswana	78
2.2 The ORI model.....	79
2.3 Global Sensitivity and Uncertainty Analysis	82
2.4 Monte Carlo Filtering	86
2.5 Selection of Prior Input Probability Distributions	87
3. Results and Discussion	90
3.1 Modified Morris Global Sensitivity Analysis.....	90
3.2 FAST Variance Based Global Sensitivity	91
3.3 Monte Carlo Filtering	92
3.4 FAST Variance Based Global Uncertainty Analysis	93
4. Conclusions	95
Acknowledgements	97
References	98
Tables.....	104
Figure Captions	109
Chapter 4. A fish population model for the Okavango Delta, Botswana: How the flood pulse drive population dynamics	119
Abstract.....	119
1. Introduction	120
2. Methods.....	126
2.1 Fish data.....	126
2.2 Model structure	128
2.3 Global Sensitivity and Uncertainty Analysis	132
2.4 Monte Carlo filtering.....	133
2.5 Model Optimization	134
2.6 Probability density functions	135

3. Results.....	140
3.1 Global Sensitivity Analysis.....	141
3.2 Monte Carlo filtering.....	142
3.3 Global Uncertainty Analysis.....	143
4. Conclusions.....	143
Acknowledgments.....	145
References.....	146
Tables.....	156
Figure Captions.....	158
Chapter 5. Conclusions.....	171
Limitations.....	174
Appendix A. ORI Reservoir Model High Performance Computing GSA/UA Script and Instructions.....	176
Appendix B. Fish model data and code.....	178

Chapter 1. Introduction

The Okavango Basin, located in southern Africa, is characterized as a large (530,000 km²), endorheic, and physiologically diverse basin that delivers a unique and ecologically important flood pulse to the Okavango Delta. The Basin's headwaters and the majority of the contributing area begin in the more humid and steep Angola. The Basin is formed by two major Rivers the Cubango and the Cuito River. The Cubango River forms part of the boundary between Namibia and Angola before converging with the Cuito River to form the Okavango River. The Okavango River then flows through the Caprivi Strip in Namibia where the topography is flatter and the climate is drier than in the headwaters. After flowing through Namibia, the Okavango River empties to the Okavango Delta, entirely contained within Botswana, which like Namibia is also dry and extremely flat. Here, on the edge of the Kalahari Desert, is the Basin's terminal end. The water is evapotranspired in the Delta before ever reaching the sea.

Average rainfall in the upstream Angolan portion of the Basin is as high at 1,300 mm yr⁻¹, while the two downstream basin nations, Namibia and Botswana, are much more arid with rainfall averaging approximately 560 mm yr⁻¹ in Rundu, Namibia (Mendelsohn and Obeid, 2004).

Angola also has more topography, with elevations in this portion of the Basin ranging between 1,200 and 1,800m, while the Delta is exceptionally flat, with elevations ranging between 900 and 1,000m (Mendelsohn and Obeid, 2004). The eastern headwaters also have different hydrologic characteristics from the western headwaters with the western headwaters demonstrating more consistent base flows. The northwestern portion of the Basin is underlain

by volcanic and metamorphic rock while the northeastern portion is underlain by Kalahari sands. Thus, the western headwaters are hydrologically flashier while the eastern headwaters have higher base flows and less seasonal variability (Hughes et al., 2006).

Rainfall in the Angolan headwaters occurs between December and March. It takes two months for the flood pulse to reach the top of the Delta and another four to five months for it to attenuate through the Delta, finally reaching the distal ends in August. Flooding extents in the Delta range from 6,000 – 12,000 km² (Wolski et al., 2006) and the entire alluvial fan is approximately 40,000 km² (Gumbrecht et al., 2005). The Delta is extremely smooth and flat with maximum local relief generally less than 2 or 3 m (McCarthy et al., 2003; Gumbrecht et al., 2001).

Over 70 fish species (Kolding, 1996) live in the Okavango Delta and there are three main fishing industries: subsistence, small scale commercial, and recreational (Merron 1991, Mosepele and Kolding 2003). According to Mosepele (2001), 65% of the people in northern Ngamiland depend on fish as a source of livelihood. Monitoring and preserving the health of this resource is important for both economic and ecological reasons. In the Okavango Delta no studies have been done to research how fish populations respond to the flood pulse. However, there have been studies that show how the annual flood pulse produces other ecologic responses, such as through nutrient enrichment, bursts in primary productivity, and hatching zooplankton (Hoberg et al., 2002; Merron, 1991).

The Okavango Integrated River Basin Management Project (IRBM, 2009) is a transboundary association between Angola, Namibia, and Botswana that promotes an adaptive approach to

management in the Okavango Basin. Adaptive Management (AM) can be loosely defined as managing in an uncertain world with a built in plan for learning by doing (Walters & Holling, 1990). AM explicitly calls for the use of models and the acknowledgement of uncertainty (Johnson et al., 1997; Walters, 1986; Walters & Holling, 1990). Models can't be validated because of the ubiquitous nature of uncertainty in observations and measurements (Saltelli et al., 2008). However, models can be given pedigrees that assess their 'fitness for purpose'. Global sensitivity and uncertainty analysis (GSA/UA) and Monte Carlo (MC) filtering are three tools that bring transparency to the modeling process and judge model quality. Uncertainty analysis quantifies model reliability. Sensitivity analysis identifies the key factors that contribute to the model's uncertainty. MC filtering reduces uncertainty and may also improve the description of uncertainty. Together these three methods add to transparency in modeling and apply value to model results for their intended purposes. (Saltelli et al., 2008)

There are several approaches for conducting a GSA/UA on mathematical models ranging from simple one-at-a time (OAT) variational methods to more sophisticated global techniques. With OAT techniques the variation of the model output is explored through changing a single input at a time. This traditional sensitivity analysis method is limited since it explores a prescribed (and usually small) parametric range, can only consider a few inputs, and cannot account for interactions between inputs (Saltelli et al., 2008).

When model responses are non-linear and non-additive, which is the case with most complex models, OAT techniques are not sufficient and global techniques, which evaluate the input factors concurrently throughout the entire parametric space, should be used (Saltelli et al.,

2008). Different types of global sensitivity methods may be selected based on the objective of the analysis and the model structure (Cacuci, 2003; Saltelli et al., 2008). A model evaluation framework around two global techniques, a screening method (Morris, 1991) and a quantitative variance-based method (Saltelli, 1999), can be quite useful when faced with a large number of inputs. This two-step process has been used in several recent model investigations (Muñoz-Carpena et al., 2007; Muñoz-Carpena et al., 2010; Fox et al., 2010; Jawitz et al., 2008; Chu-Agor et al., 2011). The screening method culls the inputs so that only the most important ones go on for further analysis in the quantitative variance based FAST sensitivity and uncertainty analyses. The least important inputs are set to constants and ignored for further analysis. This reduces model complexity and focuses effort on the most important aspects of the model.

The Method of Morris (Morris, 1991) is an OAT sensitivity method, but because it explores the entire parametric space, it is also considered a global method. The Morris method is qualitative and is less computationally intensive than other quantitative methods, thus its use as a screening tool to identify the most sensitive inputs. Morris proposes two sensitivity indices: μ represents the mean or general magnitude of effect of the uncertainty of an input on the model output and σ estimates higher-order relationships, such as nonlinear and interaction effects. σ gives the standard deviation of the results and represents the spread of the importance that each input has based on the values chosen for the other inputs. The Modified Method of Morris (Campolongo et al., 2007) has a number of improvements over the original Method of Morris. It allows an analysis of models with multiple outputs, allows factors to be grouped, and

has a more effective sampling strategy at no additional computation cost. Furthermore, the enhanced sensitivity index μ^* is approximately as good as indices based on variance methods.

Once the most sensitive factors are identified using the Modified Method of Morris, a quantitative sensitivity analysis may be conducted on those most sensitive factors using the Fourier Amplitude Sensitivity Test (FAST) (Cukier et al., 1978; Koda et al., 1979). FAST uses Fourier analysis to decompose the variance of model outputs into first order variances for each input. The Extended FAST technique (Saltelli et al., 1999) allows for the additional quantification of higher levels of variance. These higher levels of variance describe the interactions between inputs. Extended FAST defines S_i as a measure of global sensitivity. S_i is the ratio between the variance attributed to a single input divided by the total model variance. In a model where there are no interactions the sum of all S_i 's across all inputs is equal to one. In models where there are interactions this sum is greater than one. One limitation of the FAST technique is that it cannot be used on factors that are interdependent (Crosetto and Tarantola, 2001).

Monte Carlo (MC) Filtering can be used to calibrate a model as well as decrease uncertainty (Saltelli et al, 2008; Rose et al., 1991). In MC filtering, model outputs are classified as either behavioral (B) or non-behavioral (\bar{B}) based on whether or not they exceed some predefined threshold. Two subsets of the input set X_i are defined as X_i/B and X_i/\bar{B} . The two-sided Smirnov test is used on each input factor independently to determine if X_i/B is statistically different from X_i/\bar{B} . For this test, the null hypothesis is that the distribution of the inputs producing B is equal

to the distribution of inputs producing \bar{B} . If X_i/B is significantly different from X_i/\bar{B} then a new PDF can be assigned to the input based on X_i/B .

The remote location of the basin and the recent civil war in Angola (1975-2002) has resulted in sparse data sets and a high level of uncertainty regarding the basin's diverse physiology, climate, and hydrology. Most of the stream gauge data is intermittent and the basic physiologic data are on rough scales and/or in discrete sampling locations. Therefore, issues of data scarcity and resulting uncertainty are compounded in this area. When data for the inputs show no apparent distribution, such as normal or triangular, its PDF can be set to uniform (Muñoz-Carpena et al., 2007). The uniform distribution allows for equal probability of selection across the defined range. Because of the data scarcity in this location this is generally the approach to setting PDF's for inputs in this work. This conservative approach inherently introduces additional uncertainty. In response to this, this work will show that MC Filtering can be used to objectively redefine and truncate prior PDFs within their predefined ranges and decrease model uncertainty. This is especially useful because one of the largest critiques of the GSA/UA process is the subjectivity through which the PDFs are defined. In data poor areas such as the Okavango Basin this problem is exacerbated and the objective truncation of the PDFs is particularly useful.

Equifinality is the principle that given a model with multiple inputs it is possible to get exactly the same output with a variety of input sets (Beven, 2001, Beven and Freer, 2001). This is sometimes referred to as 'getting the right answer for the wrong reason.' Equifinality is an important concept in hydrologic modeling due to the large number of inputs. It is especially

crucial in reference to predictive modeling. Predictive modeling often involves changes in state, such as development and climate change scenarios. When the state of the system changes, internal processes and relationships change. Following this, if a model is getting the right answer for the wrong reason under one state, it is possible that it will no longer get the right answer when internal processes change. Traditionally, the goal of inverse optimization techniques is to find the optimal set of inputs for a given model to match some measured data (Mertens et al., 2006). This is especially problematic when considering the issue of equifinality. Though this work does employ inverse optimization style techniques, it does not seek an optimal parameter set. Instead, it looks for evidence of these models as useful tools in the presence of equifinality. This is done in two ways (1) through identifying best fit model input sets and running these through non-stationary scenarios and (2) through identifying key inputs which do not display equifinality. Through these two methods, insights into the confidence in model behavior can be gained even in the presence of equifinality.

Motivation

As a transboundary Basin, there are conflicting management agendas in the Okavango watershed. The IRBM transboundary project (2009) advocates for an adaptive approach to management in the Basin. Adaptive Management explicitly calls for the acknowledgement of uncertainty and the use of models. Uncertainty regarding physiographic, ecologic, and hydrologic data in the Basin is especially high because of its large size, remote location, and recent civil war in Angola. Several models have been calibrated in the basin and are being used to assess climate change and development scenarios. Hughes et al., (2006) calibrated the Pitman rainfall runoff model in the Okavango Basin. Uncertainty analysis has been conducted

on the Pitman model in several basins in South Africa (Hughes et al., 2010) however; no GSA/UA has been conducted for this model in the Okavango Basin. The ORI (Okavango Research Institute) reservoir model has been calibrated in the Okavango Delta (Wolski et al., 2006) however; no GSA/UA has been conducted for this model in the Okavango Delta. Climate change and development scenario modeling has been conducted using both the Pitman and ORI models (Andersson et al., 2006; Wolski, 2009). Conducting a GSA/UA on these models will provide a transparent qualification of their use for scenario modeling, enable model simplification through identifying unimportant inputs, and provide a deeper understanding of model structure and interactions.

Additionally, quantitative studies show how the annual flood pulse produces an ecologic response in nutrients and zooplankton (Hoberg et al., 2002) and qualitative observations and conceptual models have been proposed for how the flood impacts fish population dynamics (Hoberg et al., 2002; Merron, 1991). A major recommendation in the *Environmental Flow Module Special Report for Fish* (Mosepele, 2009) is for the development of a quantitative relationship between the flood pulse and fish population dynamics in the Delta. A fish population model has yet to be developed for the Delta. Conducting GSA/UA and MC filtering during the development of a fish model will prevent over-parameterization, reduce uncertainty, and allow for an investigation of evidence of the flood pulse as a driver despite issues of equifinality.

Objectives

The objective of this work is to provide a transparent audit of the usefulness of hydrologic and fish population models of Okavango Basin. Two hydrologic models will be investigated: one of the Basin and one of the Delta, both of which have previously been developed and calibrated in the area and are currently being used to test development and climate change scenarios. For this research these two models will be analyzed for sensitivity, uncertainty, and 'fitness for purpose.' A fish population model for the Okavango Delta based on the annual flood pulse will also be developed, calibrated, tested, and analyzed for sensitivity, uncertainty, and 'fitness for purpose.' This analysis will add depth of meaning to the climate change and development scenarios that are being simulated using the existing hydrologic models. It will also corroborate the existing conceptual models and qualitative observations that describe fish population dynamics as being driven by the food pulsed hydrology of the Delta.

References

- Andersson, L., Wilk, J., Todd, M., Hughes, D.A., Earle, A., Kniveton, D., Layberry, R., Savenije, H.H.G., 2006. Impact of climate change and development scenarios on flow patterns in the Okavango River. *Journal of Hydrology*, 331, 43-57.
- Beven, K., 2001. How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5, 1–12.
- Beven, K.J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *Journal of Hydrology*, 249, 11–29.

- Cacuci, D.G., Mihaela, I.B. , Ionel, M.N., 2003. Sensitivity and uncertainty analysis. Chapman & Hall/CRC Press, Boca Raton, Florida.
- Campolongo, F., Cariboni, J., Saltelli, A., Schoutens, W., 2005. Enhancing the Morris Method. in K.M. Hanson and F.M. Hemez (Eds). Sensitivity Analysis of Model Output. Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004). Los Alamos National Laboratory, Los Alamos, USA, pp.369–379.
- Chu-Agor, M.L., Munoz-Carpena, R., Kiker, G., Emanuelsson, A., Linkov, I., 2011. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling & Software*, 26(5), 593-604.
- Crosetto, M., Tarantola, S., 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. *International Journal of Geographical Information Science*, 15(5), 415-437.
- Cukier, R.I., Levine, H.B. and Schuler, K.E., 1978. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26, 1–42.
- Fox, G.A., Muñoz-Carpena, R., Sabbagh, G.J., 2010. Influence of flow concentration on parameter importance and prediction uncertainty of pesticide trapping by vegetative filter strips *Journal of Hydrology*, 384, 164–173.
- Gumbrecht, T., McCarthy, T.S., Merry, C.L., 2001. The topography of the Okavango Delta, Botswana, and its tectonic and sedimentological implications. *South African Journal of Geology*. 104, 243-264.

- Gumbrecht, T., McCarthy, T.S., Bauer, P., 2005. The micro-topography of the wetlands of the Okavango Delta, Botswana. *Earth Surface Processes and Landforms*, 30, 27-39.
- Hoberg, P., Lindholm, M., Ramberg, L., Hessen., D.O., 2002. Aquatic food web dynamics on a floodplain in the Okavango delta, Botswana. *Hydrobiologia*, 470, 23–30.
- Hughes, D. A., Andersson, L., Wilk, J., Savenije, H.H.G., 2006. Regional calibration of the pitman model for the Okavango River. *Journal of Hydrology (Amsterdam)*, 331(1-2), 30-42.
- Hughes, D.A., Kapangaziwiri, E., Sawunyama, T., 2010. Hydrological model uncertainty assessments in southern Africa. *Journal of Hydrology*, 387(3-4), 221-232.
- Jawitz, J.W., Muñoz-Carpena, R., Muller, S., Grace, K.A., James A.I., 2008. Development, Testing, and Sensitivity and Uncertainty Analyses of a Transport and Reaction Simulation Engine (TaRSE) for Spatially Distributed Modeling of Phosphorus in South Florida Peat Marsh Wetlands: U.S. Geological Survey Scientific Investigations Report 2008-5029. 109 p.
- Johnson, F.A., Moore, C.T., Kendall, W.T., Dubovsky, J.A., Caithamer, D.F., Kelley, J. R., Williams, B.K., 1997. Uncertainty and the management of mallard harvests. *Journal of Wildlife Management*. 61, 202-216.
- Koda, M., McRae, G.J., Seinfeld, J.H., 1979. Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics*, 11, 427–444.

- McCarthy, J.M., Gumbrecht, T., McCarthy, T., Frost, P., Wessels, K., Seidel, F., 2003. Flooding patterns of the Okavango wetland in Botswana between 1972 and 2000. *Ambio*, 32, 453-457.
- Mendelsohn J, el Obeid S. 2004. *Okavango River: Flow of a Lifeline*. Capetown(South Africa): Struik.
- Merron, G.S., 1991. The ecology and management of the fishes of the Okavango Delta, Botswana, with particular reference to the role of the seasonal flood. PhD dissertation. Rhodes University, Grahamstown, South Africa.
- Mertens, J., Stenger, R., Barkle, G.F., 2006. Multiobjective inverse modeling for soil parameter estimation and model verification. *Vadose Zone Journal*, 5, 917–933.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometric*, 33, 161–174.
- Mosepele, K., 2001. Preliminary description of the Okavango Delta fishery. Unpublished technical report. Fisheries Section, Ministry of Agriculture, Gaborone, Botswana.
- Mosepele, K., Kolding, J., 2003. Fish stock assessment in the Okavango Delta: preliminary results from a length based analysis. In: Bernard, T., Mosepele, K., Ramberg, L., (Eds.), *Proceedings of the Environmental Monitoring of Tropical and Subtropical Wetlands Conference*, 4–8 December, 2002, Harry Oppenheimer Okavango Research Centre, University of Botswana, Maun, Botswana, pp. 363–390.

- Mosepele K., 2009. Okavango River Basin Technical Diagnostic Analysis: Environmental Flow Module Specialist Report, Country: Botswana, Discipline: Fish. OKACOM 2010 Project Management Unit, Luanda, Angola. Sponsored by GEF.
- Muñoz-Carpena, R., Zajac, Z., Kuo, Y.M., 2007. Evaluation of water quality models through global sensitivity and uncertainty analyses techniques: application to the vegetative filter strip model VFSSMOD-W. Transactions of ASAE, 50(5), 1719-1732.
- Muñoz-Carpena, R., Fox, G.A., Sabbagh, G.J., 2010. Parameter importance and uncertainty in predicting runoff pesticide reduction with filter strips. Journal of Environmental Quality, 39(1) 1–12.
- Okavango Integrated River Basin Management Project (IRBM). 2009. Okavango Integrated River Basin Management Project Annual Report 2009, USAID Contract Number: LAG-I-811-99-00018-00, Burlington, Vermont.
- Rose, K. A., Smith, E. P., Gardner, R. H., Brenkert, A. L., Bartell, S. M., 1991. Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. Journal of Forecasting, 10, 117-133.
- Saltelli, A., 2004. Sensitivity analysis in practice: A guide to assessing scientific models. Wiley, Hoboken, NJ, 2004.
- Saltelli, A., Tarantola, S. Chan, K.P., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics, Vol. 41, No. 1, pp.39–56.

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S.,
2008. *Global Sensitivity Analysis: The Primer*. Wiley and Sons, Sussex, England.

Walters, C. J., 1986. *Adaptive management of renewable resources*. MacMillan, New York, New
York, USA.

Walters, C., Holling, C.S., 1990. Large-scale management experiments and learning by doing.
Ecology, 71, 2060-2068.

Wolski, P., H.H.G. Savenije, M. Murray-Hudson, T. Gumbricht. 2006. Modelling of the flooding in
the Okavango Delta, Botswana, using a hybrid reservoir-GIS model, *Journal of
Hydrology*, 331, 58-72.

Wolski, P., 2009. *Assessment of hydrological effects of climate change in the Okavango Basin*.
Okavango River Basin Transboundary Diagnostic Analysis Technical Report. EPSMO.
Capetown, South Africa.

Chapter 2. Uncertainty, sensitivity, and impacts for scenario modeling using the Pitman hydrologic model in the Okavango Basin

Abstract

Understanding and improving the reliability of models in watersheds in the presence of uncertainty is crucial for applying hydrologic models in a responsible management context. The Okavango Basin is a large and remote watershed (530,000 km²) that delivers an annual flood pulse to the Okavango Delta, an internationally renowned wetland. Sparse data sets and large gaps in knowledge regarding the diverse basin physiography and hydrology are a major concern in the Basin. An important step toward sustainably developing the water resources of the Okavango Basin is through an exploration of the potential hydrologic scenarios and inherent uncertainty of their modeled outcomes. This work presents two-tiered regionalized examination of the Pitman hydrologic model for the Okavango Basin in southern Africa through a state-of-the-art global sensitivity and uncertainty analysis (GSA/UA). The Pitman model is a monthly rainfall/runoff model that is widely used in southern Africa especially in ungauged basins. In this work, the 24 sub-basins are grouped into two regional arrangements to compare the significance of the choice of arrangement. The Morris GSA method is first used to qualitatively rank the importance of the inputs and then the variance-based Fourier Amplitude Sensitivity Test (FAST) GSA/UA method is used to quantitatively identify the most sensitive inputs in terms of contributed variance to simulated outputs. Model uncertainty is investigated under a climate change scenario, which predicts increasing temperature and rainfall in the next thirty years. Results show that inputs related to infiltration, temporal rainfall distribution,

groundwater runoff, and the evaporation coefficient are the most important in the model. Additionally the western headwater is the most important region. Results also show that the choice of regional arrangement does not make a significant difference in the model outcome. Introducing climate change conditions to the watershed increases uncertainty by approximately 18%. These results and conclusions are especially useful within the context of Adaptive Management when interpreting model uncertainty, determining the usefulness of model predictions in a non-steady-state condition, and focusing monitoring and management efforts.

Keywords: Sensitivity analysis, uncertainty analysis, Okavango, Pitman, scenario modeling, hydrologic modeling, regionalization, Morris Method, FAST method, Adaptive Management

1. Introduction

Hydrologic modeling is widely used in management programs for predicting the impacts of development and climate change on water supply for both human and environmental needs (Beckers et al., 2009; Hughes, 2002; Jayakrishnan et al., 2005; Koster et al., 2000; Stiegliz et al., 1997, Wolksi 2009). These models can't be truly validated because of the ubiquitous nature of uncertainty in observations and measurements (Beven and Binley, 1992). However, they can be given pedigrees that judge their quality as a function of their 'fitness for purpose' (Saltelli et al., 2008). Global sensitivity and uncertainty analysis (GSA/UA) are tools used during model analysis to judge model fitness under different applications. Uncertainty analysis quantifies the overall uncertainty of a model and sensitivity analysis identifies the key factors that contribute to that uncertainty. Using these tools, a modeler can assign the level of confidence to a model for its intended purpose. (Saltelli et al., 2008)

The Okavango Basin is a large and remote watershed (530,000 km²) that delivers an annual flood pulse through the Okavango River from Angola, through Namibia, and then into Botswana, which contains the Okavango Delta. The Okavango Basin has a distinct and heterogeneous physiological and climatic organization. Rainfall rates, geology, and topography vary considerably throughout. Rainfall in the Angolan headwaters is 1,300 mm yr⁻¹ but only 560 mm yr⁻¹ in downstream Namibia and Botswana (Mendelsohn and Obeid, 2004). The Angolan headwaters are also mountainous, with elevations ranging between 1,200 and 1,800m, while the Botswana portion of the Basin is extremely flat, with elevations ranging between 900 and 1,000m (Mendelsohn and Obeid, 2004). The eastern headwaters also behave differently from the western headwaters with the western headwaters having more consistent base flows. Rainfall is evenly distributed within the Basin from the east to the west (Wilk et al., 2006). However, the geology is different and can explain this difference in the base flow behavior (Hughes et al., 2006). The western portion of the Basin is underlain by volcanic and metamorphic rock while the eastern portion is underlain by Kalahari sands. Thus, the western headwaters are hydrologically flashier while the eastern headwaters have higher base flows and less seasonal variability.

As a transboundary basin, each of the three countries that contain portions of the Okavango River values the resource in a different way. Angola is upstream basin state and the major land holder. Namibia owns a sliver of land in the middle. Botswana contains the terminal end and the famed Okavango Delta, which is an international Ramsar site (Ramsar, 2011). The different values that the three countries place on the Okavango Basin and their different physiographic

environments create challenges for management of the resource. As the three countries develop and face climate change they must find ways to equitably distribute the river flows. Hydrologic models can be useful in this circumstance because they can simulate hydrologic response to proposed development and climate change. However, without an uncertainty analysis managers cannot contextualize the uncertainty of these models and they may provide unjustified confidence regarding the predictability of the system. Additionally, the remoteness of the basin and the recent civil war in upstream Angola (1975-2002) has resulted in sparse data sets and a high level of uncertainty regarding the diverse basin physiography and hydrology. Most of the gauge data is intermittent and the basic physiographic data are on rough scales and/or discrete sampling locations (Hughes et al, 2006). Therefore, issues of data scarcity and uncertainty are compounded in this location.

The Pitman model (Pitman, 1973; Hughes et al., 2006) (Figure 1) was developed in southern Africa specifically for ungauged catchments (Hughes, 1995). It is widely used in the area and has been calibrated and validated in the Okavango Basin (Hughes et al., 2006). An uncertainty and sensitivity analysis for the Pitman model has been conducted for four basins in South Africa (Hughes et al., 2010). The results of this analysis show that uncertainty and sensitivity in the Pitman model varies from basin to basin depending on physiography and climate. This current research builds upon these results and involves a GSA/UA of the Pitman rainfall runoff hydrologic model specifically for the Okavango Basin emphasizing the issue of data scarcity regarding basin physiography.

This GSA/UA is specifically applied with management purposes in mind such as interpreting model uncertainty under stationary conditions and climate change scenarios, identifying gaps in knowledge, and focusing monitoring efforts. The Okavango Integrated River Basin Management Project (IRBM, 2009) promotes an adaptive approach to management in the Okavango Basin. Adaptive Management (AM) can be loosely defined as managing in an uncertain world with a built in plan for learning by doing (Walters & Holling, 1990). AM explicitly calls for the use of models and the acknowledgement of uncertainty (Johnson et al., 1997; Walters, 1986; Walters & Holling, 1990).

The publication *AM for Water Resources Project Planning* (NRC, 2004) offers six elements of successful AM programs: (1) management objectives are routinely revised (2) a model of the system is developed and revised (3) a range of management choices are investigated (4) outcomes are monitored and evaluated (5) mechanisms for learning are incorporated into future decisions (6) a collaborative structure for stakeholder participation and learning is constructed and revised. GSA/UA is useful in three of these steps. In step (2) models can be revised through inputs identified in sensitivity analysis. The unimportant inputs can be set to constants to avoid over-parameterization and simplify the model structure (Saltelli et al., 2008). For this work, inputs will be defined as any variable that is not a time series. In step (3) uncertainty analysis can be used to evaluate management outcomes and assign value to the model. In step (5) mechanisms for learning can be based on gaps in knowledge that are identified through sensitivity analysis. (Saltelli et al., 2008)

There are various approaches for conducting sensitivity analyses on mathematical models ranging from simple one-at-a time (OAT) variational methods to more sophisticated global techniques. With OAT variational derivative techniques the variation of the model output is examined through the variation of one model input at a time. Sometimes, as an alternative, a crude variational approach is selected in which, incremental ratios are taken by moving factors one at a time from the base line by a fixed amount (for example, 5%) without prior knowledge of the factor uncertainty range. These traditional sensitivity analysis methods are limited since they explore a fixed parametric range, can only consider a few inputs, and do not take into account interaction between inputs (Saltelli et al., 2005).

When model responses are non-linear and non-additive, as with most complex models, these simple derivative techniques can be misleading and global techniques that evaluate the input factors of the model concurrently over the entire parametric space (described by probability density functions) should be used. Different types of GSA/UA methods can be selected based on the objective of the analysis (Cacuci et al., 2003; Saltelli et al., 2000; Saltelli 2004). This study proposes a model evaluation framework (Muñoz-Carpena et al., 2007; Muñoz-Carpena et al., 2010; Fox et al., 2010; Jawitz et al., 2008; Chu-Agor et al., 2011) around two such modern global techniques, a screening method (Morris, 1991) and a quantitative variance-based method (Cukier et al., 1978; Saltelli et al., 1999). This two-step method only considers inputs and does not investigate structural uncertainty or sensitivity. The screening method allows for the identification of the most important inputs and a resulting reduction of inputs, based on this importance, for further investigation the quantitative variance-based Extended Fourier

Amplitude Sensitivity Test (FAST). This two- step process is useful because of the computational time that is required in using the FAST method. Both the Morris and FAST methods are pre and post processed in the software Simlab (2011). The proper use of global sensitivity methods can yield three main products for the Okavango application: (1) an assurance of the model's behavior (confidence in theoretical behavior), (2) a ranking of importance of the parameters for different outputs, and (3) the identification of the type of influence of the important parameters (first or higher order interactions) (Saltelli, 2008). These products can aid managers through assigning confidence in model results and strategically allocating monitoring efforts.

Through uncertainty analysis, alternative model scenarios can also be investigated. Equifinality is the principle that given a model with multiple inputs it is possible to get exactly the same output with a variety of input sets (Beven, 2001, 2006; Beven and Freer, 2001). This is sometimes referred to as 'getting the right answer for the wrong reason.' Equifinality is an important concept in hydrologic modeling due to the large number of inputs. It is especially crucial in reference to predictive modeling. Predictive modeling often involves changes in state such as development and climate change scenarios. When the state of the system changes, i.e. non-stationary, internal processes and relationships change. Following this, if a model is getting the right answer for the wrong reason under one state, it is possible that it will no longer get the right answer when internal processes change. Therefore, any predictive model that investigates a change in state should incorporate an analysis that quantifies the change in uncertainty that accompanied that change in state. This assigns confidence in model results for management purposes.

In this work, a global sensitivity and uncertainty analysis (GSA/UA) of the Pitman model in the Okavango Basin will be conducted to explore the uncertainty of the model under stationary and non-stationary conditions as well as identify the inputs that are most responsible for that uncertainty. The specific objectives of this work are threefold. The results of these analyses will be used to (1) quantify the uncertainty of the model results, (2) identify the most important or most sensitive inputs and regions, and (3) investigate the degree of equifinality in the model and assess the impacts of this equifinality on the use of the model in a practical predictive capacity. Two objective functions are analyzed in this process: the mean annual flow and the Nash-Sutcliffe coefficient of efficiency (ceff) (Nash and Sutcliffe 1970) which evaluates the model's goodness of fit by comparing the modeled and measured monthly flow. These results will yield important products for managers. (1) Confidence in model results is assigned so that managers can use the model along with additional lines of evidence in an informed manner. (2) Gaps in knowledge are identified allowing for strategic monitoring to close those knowledge gaps. (3) The uncertainty of the predictive capacity of the model under a climate change scenario is assessed so that confidence in climate change scenarios can be assigned.

2. Methods

2.1 The Pitman Model

The Pitman model (Pitman, 1973) is a semi-distributed monthly time step rainfall runoff watershed model developed in southern Africa (Figure 2). Since its inception the model has undergone a number of revisions, the most recent being the addition of more explicit surface water/groundwater interactions (Hughes, 2004). The model has also been integrated into SPATSIM (Spatial and Time Series Information Modeling), a GIS platform to facilitate spatial

analyses (Hughes et al., 2002; Hughes and Forsyth, 2006). The Pitman model is one of the most commonly used rainfall runoff models in southern Africa.

Data requirements for running the Pitman model (Figure 2) include monthly rainfall and evaporation time series, basin and sub-basin delineations, and physical characteristics for each sub-basin such as soil transmissivity and storativity, slope, interception and soil absorption rates. Additionally, optional anthropogenic inputs and lags include abstractions by irrigation, dams, and reservoirs. Storage processes such as interception, soil moisture holding capacity, and groundwater systems are simulated as simple water storage tanks within each sub-basin. Kapangaziwiri (2008) provides a more detailed description of the general model.

The Pitman model has been set up and calibrated in the Okavango River (Hughes et al., 2006). This calibration will be used as a baseline for the following sensitivity and uncertainty analysis. In the model, the Okavango River Basin, above the Delta, was divided into 24 sub-basins. Seventeen of these sub-basins have stream gauges at their outlet. However, most of these records were discontinued prior to 1975 and only two of the downstream-most gauges, which are located in Namibia and Botswana, have continuous contemporary data. The model was calibrated over the period of 1960 and 1972 and validated between 1991 and 1997. Rain gauge measurements were used in calibration (Hughes et al., 2006). Because of the civil war in Angola (1975-2002), which contains the majority of the watershed, rain gauge readings were discontinued. TRMM (Tropical Rainfall Measuring Mission) SSM/I (Special Sensor Microwave Imager) remotely sensed rainfall measurements are available for the entire watershed from 1991 to 1997 (Wilk et. al, 2006) and were used for model validation (Hughes et al., 2006). The

Hargreaves equation (Hargreaves and Allen, 2003) was used to calculate actual evapotranspiration from the water equivalent of extraterrestrial radiation, temperature, and the difference between mean monthly maximum and minimum temperatures. Soil parameters for the Basin were obtained from FAO (Food and Agricultural Organization of the United Nations) data. Geologic and topographic inputs are derived from USGS data (Persits et al., 2002). Water abstractions are based on population density and are assumed to be negligible. The Pitman model was manually calibrated by Hughes et al. (2006) to surface flows in the Basin where gauges are. Calibration parameters included surface runoff, soil moisture storage, groundwater recharge, and soil moisture evaporation. The three objective functions that were used to minimize the difference between observed and simulated follows during calibration include the coefficient of determination (ceff) and the mean monthly percentage of error (Hughes et al., 2006). The calibrated Pitman model simulated outflow at the downstream Mukwe sub-basin with a ceff of 0.851 and a mean monthly error of +1.7% (Hughes et al., 2006).

2.2 Input Factor Selection

Ideally, different probability distributions would be produced for each of the model inputs in each of the 24 sub-basins based on measured physical variability. However, due to the lack of data in the area there is little basis for defining different PDF's for each of the sub-basins. On the other hand, lumping the Basin as a whole and producing one PDF for each of the inputs would disregard its physical heterogeneity. Therefore, a regional approach was taken. The Basin was split into regions based on geology, topography, and rainfall and PDF's for each of the inputs are produced in each of the regions. Two regional approaches were taken in order to compare the importance of the regionalization concept: a) one with three regions and b) one

with five regions (Figure 3). In the three region approach the basin was separated into eastern headwaters, western headwaters, and southern receiving waters. The eastern and western headwaters are geologically different from each other and both receive higher rainfall rates than the southern receiving waters. This regional approach mimics the same regionalization that is referred to by Hughes et al. (2006) in the calibration of the model. In the five region approach the southern region was subdivided into the middle basins, Omatako basins, and southern basins. These subdivisions reflect the difference in rainfall and topography in the area. The subdivision of the southern region in the five regional approach also reflects the fact that the Omatako River is hydrologically unique and though it is hydrologically connected, it is now considered a fossil river and there is no record of it contributing flow to the system (Crerar, 1997).

Probability density functions (PDF's) were developed for nineteen inputs in the Pitman model (Table 1). The basin is a fairly pristine system, so inputs regarding anthropogenic disturbances such as dams and reservoirs were not considered. Uniform distributions were assigned to all of the input PDF's. This is considered to be the most conservative approach when there is a lack of data available for describing specific types of distributions (Muñoz-Carpena et al., 2010). In an ideal situation PDF's are based on experiments, literature values, physical bounds, or expert opinion. For example, Kapangaziwiri and Hughes (2008) developed an estimation method for inputs into the Pitman model based on physiological basin characteristics that are largely defined by South Africa's Agricultural Geo-referenced Information System (AGIS, 2007).

However, compared to South Africa, the Okavango Basin is located in a remote area and there

is significantly less data available and fewer experts in the area. Additionally, many of the inputs are empirical (though physically relevant) and are used as scaling factors or describe the shape of relationships among system components. As a result of the lack of data and the empirical nature of the inputs, the true physical ranges of these inputs are largely unknown. When information is lacking, PDF's can be bound by some standard percentage of the calibrated value. For this reason the inputs were categorized as having low, medium, and high uncertainty. The inputs that were identified as having low uncertainties were bounded by $\pm 15\%$ of the calibrated value, medium $\pm 30\%$, and high $\pm 45\%$ of the calibrated values. Each input for each sub-basin within a region is varied independently from its calibrated value within its assigned uncertain range based on the regional value assigned by the matrix. In an area where more data is available autocorrelation tests could be done to show dependence between data in each sub-basin. However, here no such data exists and inputs are assumed independent. The results of the GSA/UA are largely dependent upon the input PDF's and setting these PDF's is extremely important. The following paragraphs define the input factors and the rationale behind the PDF's that were developed from them.

Monthly rainfall in the Pitman model is disaggregated into four sub-monthly inputs that use the rainfall distribution input factor (*RDF*) to describe the temporal heterogeneity of the input distribution. Lower *RDF*'s create even temporal rainfall distributions and higher values result in flashier and heavier rainfall events within each month. In the calibrated model, *RDF* was set to 0.7 for each sub-basin. Hughes et al. (2003) used values in the Kafue Basin that ranged between 0.6 and 1.28. The Kafue Basin is located approximately 800 km northeast of the

Okavango Basin in Zambia. The Okavango Basin covers a large area going from mountainous temperate headwaters to flat semi-arid receiving waters. It is unlikely that monthly rainfall is distributed the same way in the headwaters as it is in the southern portions of the basin. The uncertainty for this input was set to medium.

In the model, interception (Figure 2) is defined for the two different vegetation types (*PI1* and *PI2*). Vegetation type 1 represents non-forested land cover and vegetation type 2 is forested. Pitman (1973) asserts that interception in southern Africa can range between 0 and 8 mmday^{-1} . De Groen (2002) considers a range of 2–5 mmday^{-1} but cites that established thresholds for South Africa are 1 to 2 mmday^{-1} and as much as 7 mmday^{-1} when litter interception is included. In the calibrated model the interception rates are set to 1.5 for vegetation type 1.0 and 4.0 for vegetation type 2. This is a physically based input but is only allowed to range across two vegetation types. The uncertainty for these two inputs was set to medium.

Evapotranspiration is based on the ratio between potential and actual evaporation at different levels of soil moisture (R), the area of the sub-basin covered by type 2 vegetation ($AFOR$), a factor that scales the evapotranspiration for vegetation type 2 (FF), and the riparian strip factor (RSF). R determines the shape of a linear relationship between actual and potential evaporation loss at different moisture storage levels. The value of R is bounded between 0 and 1. Actual evaporation is calculated according to Eq. 1 where PE is potential evaporation, S is current soil moisture storage, and ST is the maximum soil moisture storage. R is a physical parameter, but there is little data for the area and the uncertainty level was set to medium.

$$E = PE \left[1 - \left\{ 1 - R \left(1 - \frac{PE}{PE_{MAX}} \right) \right\}^{-1} \left(1 - \frac{S}{ST} \right) \right] \quad (1)$$

AFOR represents the percent of the basin covered by type 2 vegetation. The GLC2000 land cover map from the Global Vegetation Modeling Unit (GLC, 2000) was used to derive these parameters. Because *AFOR* is based on remotely sensed data, its uncertainty was set to low. *FF* allows type 2 vegetation to have greater *ET* than type 1 and is set to 1.3 in all of the sub-basins. This categorizes the entire Okavango Basin by two types of vegetation. Its uncertainty is set to medium. *RSF* determines the water loss due to evapotranspiration in areas adjacent to the channel. There is a great deal uncertainty associated with *RSF* as it is a fairly empirical input with very little data available for verification. Its uncertainty level was set to high.

In the Pitman model, infiltration is governed by the amount of monthly rainfall and the inputs *ZMIN* and *ZMAX*. *ZMIN* and *ZMAX* provide boundary conditions for a triangular distribution of absorption rates. If rainfall is greater than the absorption rate defined by that month's rainfall then overland runoff occurs. *ZMIN* and *ZMAX* are manually fitted to approximate the Kostiakov (1932) equation to calculate infiltration. The Kostiakov equation is an empirical model that assumes infiltration decreases in time during a rainfall event according to a power function. The uncertainty for these inputs was set to high because of the empirical nature of the equation, the differences in scale between rainfall events and the model's monthly time step, and the lack of data describing the soils in the area.

Soil moisture storage is represented by the maximum depth of water storage (*ST*). If *ST* (mm) is filled, then the additional rainfall becomes runoff. The uncertainty for *ST* was set to medium

because there is some physical evidence for relative differences in these values based on the geology of the region but there is still very little data regarding actual values. Hortonian overland flow is calculated according to two inputs: the runoff generated at the maximum soil moisture (FT) and a power function that allows a nonlinear relationship between runoff and soil moisture (POW). Kapangaziwiri (2008) give values for FT (mm month^{-1}) ranging between 0.4 and 43.4 for basins with low drainage density and 0.4 to 72.3 in basins with high drainage density. The model was calibrated with values ranging from 0 to 38 with a great deal of variation between regions. Because this range is so large, the uncertainty for FT was set to high. POW simulates the curve that represents how runoff slows as moisture in the soil decreases. Values for POW in the calibrated model vary between 2.5 and 4.0. There is little data describing accurate values for this input and its uncertainty was set to high.

Groundwater recharge in the Pitman model is one-dimensional and governed by the rate of recharge at the maximum soil moisture storage (GW) and a power function that describes the non-linear relationship between recharge and soil moisture ($GPOW$). GW is the groundwater equivalent to FT and $GPOW$ is the groundwater equivalent to POW . Xu and Beekman (2003) compiled literature values on ranges of groundwater recharge in southern Africa. They found maximum rates ranging from 4.2 to 420 mm month^{-1} . Because this range is so large, the uncertainty for $GPOW$ and GW was set to high.

There are a number of additional groundwater accounting parameters in the model: drainage density ($DDENS$), transmissivity (T), storativity (S), the depth of the aquifer below the channel at which groundwater ceases to flow (RWL), the groundwater slope (GWS), and the riparian strip

factor (*RSF*). Kapangaziwiri (2008) states that estimates of drainage density from 1:250,000 maps are three times greater than those from 1:50,000 maps. T (m^2d^{-1}) is equal to the soil permeability time the aquifer thickness. Milzow et al., (2009) cites that the aquifer thickness in the Okavango Delta varies between 70 and 400m. S (m^3) is equal to the product of specific storage and aquifer thickness. According to Singhal and Gupta (1999) storativity in unconfined aquifers generally ranges between 0.05 and 0.30. Milzow et al., (2009) use a specific yield (or drainable porosity) of 0.05 in a MODFLOW model of the Okavango Delta. The calibrated values of S ranged between 0.001 and 0.05 throughout the basin. The uncertainty for *DDENS*, T , S , *RWL*, and *GWS* are set to medium.

2.3 Morris Global Sensitivity Analysis

The GUA/GSA of the Pitman model begins with a qualitative screening assessment using the Morris Method (Morris, 1991). The PDF for each model input (X_i , $i = 1, 2, \dots, k$) is divided into p discrete levels. A matrix of model runs is then formed by selecting values OAT from the PDF's in trajectories that efficiently sample throughout input space for each PDF. Though the sensitivity calculation is one-at-a-time derivative, it may actually be considered global because it samples throughout the multivariate parametric space. The elementary effects (F_i) are obtained according to the equation below, where Δ represents the step size across the levels.

$$F_i = \frac{y(X_1, X_2, \dots, X_{i-1}, X_{i+\Delta}, \dots, X_k) - y(X_1, X_2, \dots, X_k)}{\Delta} \quad (2)$$

Morris proposes two sensitivity indices: μ_i and σ_i for each input. The index μ_i represents the average magnitude of change in the model output resulting from varying each input within its PDF. This is essentially the direct importance or significance of the input. Because the model is

run in batch with varying inputs sets, each μ_i depends on all of the other inputs and can vary in each model run. Thus, σ_i is the standard deviation of μ_i , and estimates higher-order relationships such as nonlinear and interaction effects. The Modified Method of Morris (Campolongo et al., 2007) has a number of improvements over the original Method of Morris. It allows for an analysis of models with multiple outputs, allows factors to be grouped, and has a more effective sampling strategy at no additional computation cost. Furthermore, the enhanced sensitivity index μ^* is approximately as good as indices which are based on variance methods (Campolongo et al., 2007).

2.4 FAST Variance Based Global Sensitivity Analysis

Once the most sensitive factors are identified using the Modified Method of Morris, a more quantitative sensitivity analysis was conducted on only the most sensitive factors using the variance-based Fourier Amplitude Sensitivity Test (FAST) (Cukier et al., 1978; Koda et al., 1979). FAST uses Fourier analysis to decompose the variance of a model output into variances for each input. The Extended FAST technique (Saltelli et al., 1999) allows for the quantification of higher levels of variance. These higher levels of variance describe the interactions of inputs. $V(Y)$ is the summation of the first order variance in each input and also the residual which is the variance attributed to all interactions. Thus, $V(Y)$ describes the total variance of a single input including higher levels of variance (Eq. 3).

$$V(Y) = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < l} V_{ijl} + \dots + V_{123\dots k} \quad (3)$$

FAST also defines S_i as a measure of global sensitivity. S_i is the ratio of the variance that is attributed to a single input divided by the total model variance. In a model in which there are

no interactions the sum of all S_i 's across all inputs is equal to one. In models where there are interactions this sum will be greater than one. For the extended FAST GSA/UA, a model is run for $C = Mk$ iterations, where k is the number of inputs, an M is a value that ranges between 100's and 1000's (Saltelli et al., 1999). Since the extended FAST method uses a randomized sampling procedure, it provides an extensive set of outputs that can be used in the global, Monte-Carlo type uncertainty analysis, of the model. Thus, PDF's, cumulative probability functions (CDFs), and percentile statistics can be derived for each output of interest.

2.5 Climate Change Uncertainty

The previous methods are used to show how the model behaves under stationary conditions. However, models are often used to model non-stationary scenarios such as climate change where the baseline conditions change. Under these conditions, the original validation may no longer be valid. This section presents a method to investigate the change in model uncertainty under a non-stationary climate change scenario. The best fit model simulations ($ceff > 0.70$) are inspected to identify the variety of input scenarios that provide a good model fit in the stationary simulation. A threshold $ceff$ of 0.70 is chosen which allows 0.31% of the runs (466) to be included as best fit simulations. This number was determined to be appropriate for both optimizing the definition of 'optimal' model outputs while at the same time maximizing the number of passable simulations. These simulations are chosen from the FAST three region model and are selected only from those runs that have a $ceff$ of greater than 0 in each of the three regions. The three region version is chosen over the five region version because the model was originally calibrated using the three region approach. Wolski's (2009) climate change projections predict that temperature in the Okavango Basin will increase between 2.3

and 3°C and rainfall will increase between 0 and 20% within the next 30 years. For this work, a scenario that involves an increase in temperature of 2.6°C and an increase in rainfall of 15% will be considered. The uncertainty is compared between the original scenario and the climate change scenario. The goal of this exercise is not to calculate the uncertainty of the climate change scenario, but rather the difference between the uncertainty of the model under stationary versus a possible non-stationary changing scenario.

3. Results

The Morris and FAST GSA/UA methods were run using a three region approach that includes the western headwaters, eastern headwaters, and southern basins as well as a five region approach in which the southern basin is subdivided into middle, Omatako, and southern regions (Figure 3). The objective functions are the mean annual flow and the coeff of the monthly flows.

3.1 Morris Method

In the Morris Method three region approach the importance of both direct effects (x axis) and indirect effects or interactions (y axis) is evident (Figure 4). The western region contributes the most important inputs. Also of interest are the inputs *ZMAX* and *RDF* which are important in all three regions. These inputs relate to rainfall and interception (Table 1). The inputs that are less important are identified and set as constants in the next quantitative and more computationally intensive variance-based sensitivity analysis. A cutoff of 0.2 on the direct effect axis was used to separate the important inputs from the less important. This cutoff excludes the inputs that are clustered around the origin and for clarity are not labeled. Though this cut-off is based on visual inspection, the low level of importance shown for some of the input factors that are still

included later in the following variance-based analysis demonstrates that the cutoff is satisfactory.

In the Morris Method five region approach, as with the three region approach, both the direct and indirect effects are important as evidenced by the even spread on both the x and y axes (Figure 5). Similarly, *ZMAX* and *RDF* continue to be important in most of the five regions. A cutoff at 0.1 on the direct effects axis is used to distinguish important inputs for further analysis in the FAST GSA/UA (Table 2). Again, as with the three region approach, the unimportant inputs are set to constants in future analyses. The low level of importance shown for some of the input factors that are still included in the following analysis demonstrates that the cutoff is satisfactory.

The regional approaches are compared to identify how regionalization affects input importance (Table 2). The addition of the Omatako and middle regions does have an effect on the GSA results; however this effect is not as overwhelming as might be expected given the proportional size of the new regions. Only two inputs from the Omatako region and one from the middle region fall above the cutoff for important inputs in the five regional approach.

3.2 FAST Variance Based Global Sensitivity Analysis

After the Morris Methods narrowed down the most important inputs, the variance-based FAST analysis was then used to quantitatively compute the uncertainty and sensitivity in the three and five regional approaches. The same two objective functions were considered: the mean annual flow and the ceff. Again the ceff is the preferred objective function because it is able to take into account the monthly variation in flow. However, when conducting the uncertainty

analysis the use of the ceff becomes problematic because the uncertainty is measured between regions. This comparison is not practical for three reasons. The first reason is that measured data is required to compute the ceff and in the eastern headwaters and Omatako region there is a limited amount of measured flow data. During calibration, 12 to 69 monthly measured flow values were available for calibration in the middle reservoirs and none were available in the Lower Omatako (Hughes et al., 2006) . The second problem is that the sub-basins must confluence in order to compute a ceff for an entire region. In the eastern headwaters and the middle region the sub-basins do not confluence. Finally, given the calibrated inputs (Hughes, 2006), the eastern headwater region underestimates monthly flows and results in a poor model fit and a low ceff. In the equation for the ceff the numerator and denominators are both squared, and as a result the poorer fit runs have exaggerated values for the ceff. Because of these issues, the uncertainty analysis is conducted using the mean annual flow as the objective function when comparing regions. But when the uncertainties of the regional approaches are compared at the southern terminus of the River, the ceff is used as the objective function. For the GSA these issues are not a problem and the ceff is the objective function.

For the GUA the mean annual flow is presented as regular and also normalized by the regional catchment areas. This is done because the contributing area for the southern region is much larger and actually contains the water from the eastern and western Regions and so it has larger flows and as a result larger uncertainties associated with those flows (Figure 6). Through normalization the differences in the flow variability due solely to the sizes of the regions are negated.

Normalizing the three regions by their areas shows the southern region becoming relatively more certain and the eastern region relatively less certain (Figure 6). Normalizing the five regions by their areas shows the southern region becoming more certain and the eastern and western region becoming less certain (Figure 6). Omatoko shows very little normalized uncertainty, which is likely due to the small quantity of water that it contributes due to its arid climate. The three and five region *ceff*'s of the monthly flow at the southern outlet are also compared (Figure 7). This analysis shows that the addition of two regions does not add to the uncertainty of the model results and are therefore somewhat superfluous from a management perspective. The statistics of the uncertainty analysis confirms the similarities (Table 3). This table can also be used to compare the results of this study to those of Hughes et al. (2010). In Hughes study the *ceff* that the Pitman model produced for the uncertainty analysis ranged between 0.21 and 0.83 in various South African Basins. However, here the mean is 0.23 to 0.25. The comparatively poor fit that is achieved in the Okavango Basin is likely due to selection of probability distributions that were defined in this study and the catchment size. This study assigned uniform distributions to the inputs while Hughes et al. (2010) used triangular distributions. Additionally, this is a much larger watershed with 24 sub-basins while the Hughes study focused on smaller watersheds (174-898 km²) with a maximum of three sub-basins.

The quantitative FAST GSA was also conducted for both of the regional simulations (Figures 8 and 9). In each of these simulations *ZMAX* in the western region is the most important input. Additionally, the western RDF, southern RDF, and southern *ZMAX* show a high level of importance in each regional simulation. Both of the regional simulations display a relatively

large degree of interactions and higher order indexes which is in agreement with the Morris results.

3.3 Climate Change Uncertainty

Hydrologic models are frequently used to predict conditions under non-stationary scenarios. In these situations the original validation performed under the stationary conditions may not be valid. This issue is possibly exacerbated when there are equifinality concerns as there may be many equally probably input sets that are able to simulate equally good outputs and no way of distinguishing between them. When the state of the system then changes and internal relationships are modified, what was once an equifinal result, may no longer be so. Investigating the degree of equifinality and the change in uncertainty between a stationary and non-stationary scenario (such as climate change) gives insights into the usefulness of the Pitman model as a predictive tool.

For this analysis, the best fit model simulations from the FAST three region GSA/UA were isolated to identify the variety of input scenarios that provide a good model fit. Similar to Gupta et al.'s (1998) depiction and Hamby and Tarantola's (1999) cobweb plots, input trajectories are shown that map the top 30 model simulations (Figure 10). The lines in this figure serve to demonstrate the various trajectories that best fit input sets may take as well as the values themselves. The lack of pattern or trend shows that there are many varying input sets that are able to create a good model fit and thus a high level of equifinality. This is expected given the data scarcity, the resulting widely varying output PDF's, and the large number of inputs. The best fit GUA model simulations ($ceff > 0.70$) were then run through a

climate change scenario based on Wolski's (2009) projections and the uncertainty was compared between the original scenario and the climate change scenario. In the climate change scenario rainfall was increased by 15% and temperature was increased by 2.6 degrees throughout the basin. The goal of this exercise was not to calculate the uncertainty of the climate change scenario, but rather how the uncertainty of the model changes under a possible non-stationary condition. Results show an average increase in flow between the original simulations and the climate change simulations by 28%. The uncertainty is also increased with an 18% wider 95% confidence interval. This demonstrates that the uncertainty of the model increases by 18% under this non-stationary condition. It is unclear this increased uncertainty is related to the added input of rainfall versus the changing relationships between storages and flows within the model due to changing hydrologic patterns.

4. Conclusions

Evaluating model uncertainty and identifying gaps in knowledge are important for the Adaptive Management process. Doing so gives managers a measure of confidence in model results and identifies areas for strategic monitoring whereby model results can be improved. This two-step and multi-tiered regionalization GSA/UA of the Pitman model highlights the most important and least important inputs, quantifies model uncertainty, compares regional groupings, and explores the uncertainty of a non-stationary model scenario. The conclusions are useful when interpreting model uncertainty, determining the usefulness of model predictions, identifying gaps in knowledge, and focusing monitoring efforts.

Through the sensitivity analysis, the most important and least important inputs were identified within each region. The inputs are qualitatively ranked free of regional specification based on the ceff (Table 4). The ranking is based on the summation of the Morris μ^* and σ sensitivity indexes in both regional simulations for each input. The most sensitive inputs are *GPOW*, *RDF*, *ZMAX*, *FF*, and *GW*. The least sensitive inputs are *T*, *S*, *RWL*, *DDENS*, and *GWS*. Input importance is also mapped on the model structure to identify the most important processes (Figure 11). This map shows that rainfall distribution, catchment absorption, evapotranspiration, and groundwater recharge are the most important processes and groundwater runoff is the least important process. This map is especially useful in the NRC (2004) Adaptive Management processes step 2: developing and revising the model. Using the GSA results, the most sensitive inputs (i.e. *GPOW*, *RDF*, *ZMAX*, *FF*, and *GW*) should be focused on first through additional research and monitoring. This will close the data gap, narrow the PDF's, and decrease the number of input combinations that will create equally good model fits.

Another point of interest is the large number of important inputs from the western region in both the three and five region simulations (Figure 12). This is likely the result of the distribution of rainfall in the northern portion of the basin and the larger size of the western headwater region relative to the eastern headwater region. The majority of the rain falls in the headwaters and this initial input drives much of the system. In the calibrated model, the eastern and western headwaters receive on average between 1000 and 1280 mm of rainfall yr^{-1} while the southern region receives between 540 and 640 mm yr^{-1} . Additionally, the western region is much larger than the eastern region and occupies 22% of the total basin the 11% that

is occupied by the eastern region. Because of the importance of the headwater regions, monitoring efforts that seek to improve model results should focus on these areas.

The uncertainty analysis that is normalized by the regional areas shows that the flows at the downstream outlet are more certain than those in the headwater regions. This is not necessarily intuitive as one might initially think that uncertainty would be compounded in the downstream direction. This effect may be due to the more arid climate in the southern portion of the watershed. In this region there is less input of water and more evapotranspiration. The southern reaches of the rivers are losing reaches with the sum of the calibrated mean annual flow in the eastern and western headwaters summing to $10,128 \text{ m}^3 \times 10^6$ and the mean annual flow at the southern outlet at Mohembo only $8,465 \text{ m}^3 \times 10^6$ (Hughes et al., 2006). Thus, since there are fewer important hydrologic inputs in the southern portion of the basin, the certainty of the mean annual flow, when normalized by the area of the contributing basin, is increased in the downstream direction. This reinforces how focusing monitoring efforts in the headwaters will strategically decrease model uncertainty.

Given the data scarcity, and resulting wide PDF's that were set for the inputs, there are multiple input sets that can simulate the same goodness of model fit. This issue of equifinality was further analyzed to evaluate the effects that it may have on using the model in a predictive capacity for climate change projections. It was shown that, after isolating the best fit model simulations, the 95% confidence interval of the average annual flow increased by 18% between the original simulations and the climate change simulations. Thus, the uncertainty in the model is compounded under a changing state such as climate change. It is unclear whether this

increase in uncertainty is due to the increased rainfall, the changing hydrologic relationships, or both. A manager may interpret this as an 18% increase in uncertainty between predictions under stationary and this climate change scenario. This level of uncertainty is significant in the context of management decisions.

Data scarcity presents many challenges for modelers and policy makers when calibrating, interpreting model results, and applying models in predicative and dynamic capacities. However, there are ways of managing for these challenges. Adaptive Management is specifically designed for determining a course of action within the context of uncertainty. The tools presented here are couched within the context of Adaptive Management and quantitatively analyze model uncertainty in the light of practical management issues. This work identifies which inputs should be focused on to reduce model uncertainty (*GPOW*, *RDF*, *ZMAX*, *FF*, and *GW*) and also shows there is an 18% increase in uncertainty when applying the model to a non-stationary climate change scenario. Thus, the model is shown to have significance and applicability even in the presence of high levels of uncertainty and equifinality.

Acknowledgements

The authors would like to thank the National Science Foundation's AM-W3 IGERT at the University of Florida and Mark Brown for their assistance with funding.

References

AIGS., 2007. Agricultural Geo-Referenced Information System. Retrieved March, 2011 from <http://www.agis.agric.za>

- Beckers, J., Smerdon, B., and Wilson M., 2009. Review of Hydrologic Models for Forest Management and Climate Change Applications in British Columbia and Alberta. Forrex Forum for Research and Extension in Natural Resources Society, Kamloops, British Columbia, Canada
- Beven, K. J., and Binley, A.M., 1992. The future of distributed models: Model calibration and uncertainty prediction, *Hydrological Processes*, 6, 279-298.
- Beven, K., 2001. How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5, 1–12.
- Beven, K.J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *Journal of Hydrology*, 249, 11–29.
- Beven K., 2006. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320: 18–36.
- Cacuci, D.G., Mihaela, I.B. , Ionel, M.N., 2003. Sensitivity and uncertainty analysis. Chapman & Hall/CRC Press, Boca Raton, Florida.
- Campolongo, F., Cariboni, J., Saltelli, A., Schoutens, W., 2005. Enhancing the Morris Method. in K.M. Hanson and F.M. Hemez (Eds). *Sensitivity Analysis of Model Output. Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output (SAMO 2004)*. Los Alamos National Laboratory, Los Alamos, USA, pp.369–379.

- Chu-Agor, M.L., Munoz-Carpena, R., Kiker, G., Emanuelsson, A., Linkov, I., 2011. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling & Software*, 26(5), 593-604.
- Crerar, S., 1997. Hydrology of the Okavango River system upstream of the Okavango Delta. Appendix A – Initial Environment Evaluation. In: Feasibility study on the Okavango River to Grootfontein Link of the eastern National Water Carrier, vol. 4, part 3: Appendices, 31 pp.
- Cukier, R.I., Levine, H.B. and Schuler, K.E., 1978. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26, 1–42.
- De Groen, M.M., 2002. Modelling interception and transpiration at monthly time steps; introducing daily variability through Markov chains. PhD thesis, IHE-Delft. Swets & Zeitlinger, Lisse, The Netherlands.
- Fox, G.A., Muñoz-Carpena, R., Sabbagh, G.J., 2010. Influence of flow concentration on parameter importance and prediction uncertainty of pesticide trapping by vegetative filter strips *Journal of Hydrology*, 384, 164–173.
- GLC. 2000. Global Land Cover 2000. European Commission Joint Research Centre, Global Environment Monitoring Unit.
- Gupta, H.V., Sorooshian S., Yapo, P.O., 1998. Towards improved calibration of hydrologic models: multiple and incommensurable measures of information, *Water Resources Res.* 34, 751–763.

- Hamby, D.M. and S. Tarantola, 1999. Exploring Sensitivity Analysis Techniques for the Assessment of an Environmental Transport Model. In: Proceedings of ESREL '99 - The Tenth European Conference on Safety and Reliability. Garching, Germany. September 13-17, 1999.
- Hargreaves G.H., and R.G. Allen, 2003. History and evaluation of Hargreaves evapotranspiration equation, *Journal of Irrigation and Drainage Engineering*, 1, 53–63.
- Hughes, D.A. (1995). Monthly rainfall-runoff models applied to arid and semiarid catchments for water resource estimation purposes. *Hydrological Sciences Journal*, 40(6), 751-769.
- Hughes, D.A., 2002. The development of an information modelling system for the regional water resource assessments. In: H.J. van Lanen and S Demuth (Eds), FRIEND 2002 – Regional Hydrology: Bridging the gap between research and practice. Proceedings of the Fourth International Conference., Cape Town, March 2002. IAHS Publ. 274, pp. 43–49.
- Hughes, D., Görgens, A., Middleton, B., and Hollingworth, B., 2002. Regional water resource assessment in the SADC region. In: Proceedings of the Fourth International FRIEND Conference, Cape Town, South Africa, March 2002. IAHS Publ. No. 274. 2002, pp. 11–18.
- Hughes D.A., Mwelwa E., Andersson L., Wilks J., 2003. southern Africa FRIEND Phase II – Regional Water Resources and River Flow Modelling. Institute for Water Research, Rhodes University, South Africa.

- Hughes, D.A., 2004. Incorporating ground water recharge and discharge functions into an existing monthly rainfall-runoff model. *Hydrological Sciences Journal*, 42(2) 297-311.
- Hughes, D.A., Forsyth, D., 2006. A generic database and spatial interface for the application of hydrological and water resource models. *Computers and Geosciences*, 32, 1389-1402.
- Hughes, D. A., Andersson, L., Wilk, J., Savenije, H.H.G., 2006. Regional calibration of the pitman model for the Okavango River. *Journal of Hydrology (Amsterdam)*, 331(1-2), 30-42.
- Hughes, D.A., Kapangaziwiri, E., Sawunyama, T., 2010. Hydrological model uncertainty assessments in southern Africa. *Journal of Hydrology*, 387(3-4), 221-232.
- Jayakrishnan, R., Srinivasan, R., Santhi, C., Arnold, J.G., 2005. Advances in the application of the SWAT model for water resources management. *Hydrological Processes*, 19 (3), 749-762.
- Jawitz, J.W., Muñoz-Carpena, R., Muller, S., Grace, K.A., James A.I., 2008. Development, Testing, and Sensitivity and Uncertainty Analyses of a Transport and Reaction Simulation Engine (TaRSE) for Spatially Distributed Modeling of Phosphorus in South Florida Peat Marsh Wetlands: U.S. Geological Survey Scientific Investigations Report 2008-5029. 109 p.
- Johnson, F.A., Moore, C.T., Kendall, W.T., Dubovsky, J.A., Caithamer, D.F., Kelley, J. R., Williams, B.K., 1997. Uncertainty and the management of mallard harvests. *Journal of Wildlife Management*. 61, 202-216.
- Kapangaziwiri, E., 2008. Revised parameter estimation methods for the Pitman monthly rainfall-runoff model. Master's thesis, Rhodes University, South Africa.

- Kapangaziwiri, E., Hughes, D., 2008. Towards revised physically-based parameter estimation methods for the Pitman monthly rainfall-runoff model. *Water South Africa*, 32(2), 183-191.
- Koda, M., McRae, G.J., Seinfeld, J.H., 1979. Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics*, 11, 427–444.
- Koster R.D., Suarez M.J., Ducharne A., Stieglitz M., Kumar P., 2000. A catchment-based approach to modeling land surface processes in a general circulation model, 1. Model structure. *Journal of Geophysical Research*, 105(D20) 24,809–24,822.
- Kostiakov, A.N., 1932. On the dynamics of the coefficient of water percolation in soils and on the necessity for studying it from a dynamic point of view for purposes of amelioration. *Transactions of 6th Congress of International Soil Science Society. Moscow*, pp. 17–21.
- Mendelsohn, J., el Obeid, S., 2004. *Okavango River: The Flow of a Lifeline*. Struik Publishers, Cape Town, South Africa.
- Milzow C., Kgotlhang L., Kinzelbach W., Meier P., Bauer-Gottwein P., 2009. The role of remote sensing in hydrological modelling of the Okavango Delta, Botswana. *Journal of Environmental Management*, 90(7) 2252-2260.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometric*, 33, 161–174.

- Muñoz-Carpena, R., Zajac, Z., Kuo, Y.M., 2007. Evaluation of water quality models through global sensitivity and uncertainty analyses techniques: application to the vegetative filter strip model VFSSMOD-W. *Transactions of ASAE*, 50(5), 1719-1732.
- Muñoz-Carpena, R., Fox, G.A., Sabbagh, G.J., 2010. Parameter importance and uncertainty in predicting runoff pesticide reduction with filter strips. *Journal of Environmental Quality*, 39(1) 1–12.
- Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.
- National Research Council (NRC). 2004. Adaptive Management for Water Resources Planning, Panel on Adaptive Management for Resource Stewardship, Committee to Assess the U.S. Army Corps of Engineers Methods of Analysis and Peer Review for Water Resources Project Planning, Water Science and Technology Board, Ocean Studies Board, Division on Earth and Life Studies, Washington: National Academies Press.
- Okavango Integrated River Basin Management Project (IRBM). 2009. Okavango Integrated River Basin Management Project Annual Report 2009, USAID Contract Number: LAG-I-811-99-00018-00, Burlington, Vermont.
- Persits, F., Ahlbrandt, T.S., Tuttle, M., Charpentier, R.R., Brownfield, M.E., Takahashi, K., 2002. Maps showing geology, oil and gas fields and geological provinces of Africa Ver. 2.0 (29A, Economic geology, geology of energy sources No. OF 97-0470-A). United States (USA): U. S. Geological Survey, Reston, VA, United States (USA).

- Pitman, W.V., 1973. A mathematical model for generating monthly river flows from meteorological data in South Africa. Hydrological Research Unit, Univ. of the Witwatersrand, Report No. 2/73.
- Ramsar, 2011. The Ramsar Convention on Wetlands. Retrieved March, 2011 from <http://www.ramsar.org>
- Saltelli, A., Tarantola, S. Chan, K.P., 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, Vol. 41, No. 1, pp.39–56.
- Saltelli, A., Tarantola, S. and Campolongo, F., 2000. Sensitivity analysis as an ingredient of modeling. *Statistical Science*, 15, 377–395.
- Saltelli, A., 2004. *Sensitivity analysis in practice: A guide to assessing scientific models*. Wiley, Hoboken, NJ, 2004.
- Saltelli A., Ratto M., Tarantola S., Campolongo F.. 2005. Sensitivity analysis for chemical models. *Chemical Reviews* 105(7), 2811-2828.
- Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S., 2008. *Global Sensitivity Analysis: The Primer*. Wiley and Sons, Sussex, England.
- Simlab, 2011. Simlab sensitivity analysis software Retrieved March, 2011 from <http://simlab.jrc.ec.europa.eu/>
- Singhal, B.B.S., Gupta, R.P., 1999. *Applied hydrogeology of fractured rocks*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 400 pp.

Stiegliz M., Rind, D., Famiglietti J., Rosenzweig, C., 1997. An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. *Journal of Climate*, 10, 118-137.

Walters, C. J., 1986. Adaptive management of renewable resources. MacMillan, New York, New York, USA.

Walters, C., Holling, C.S., 1990. Large-scale management experiments and learning by doing. *Ecology*, 71, 2060-2068.

Wilk, J., Kniveton, D., Andersson, L., Layberry, R., Todd, M. C., Hughes, D., et al. 2006. Estimating rainfall and water balance over the Okavango River Basin for hydrological applications. *Journal of Hydrology*, 331(1-2), 18-29.

Wolski, P., 2009. Assessment of hydrological effects of climate change in the Okavango Basin. Okavango River Basin Transboundary Diagnostic Analysis Technical Report. EPSMO. Capetown, South Africa.

Xu, Y., Beekman, H.E., 2003. Groundwater recharge estimation in southern Africa, UNESCO IHP Series No. 64, published by UNESCO: Paris.

Tables

Table 1. Input factor PDF's for the two regional simulations. [Low $\pm 15\%$, Medium $\pm 30\%$, High $\pm 45\%$ of calibrated values.]

Parameter	Description	Uncertainty
RDF ^(*)	Rainfall distribution function	Medium
PI1 (mm mo ⁻¹)	Interception for veg type 1	Medium
PI2 (mm mo ⁻¹)	Interception for veg type 2	Medium
AFOR ^(*)	Percent area covered by veg type 2	Low
FF ^(*)	Evaporation scalar for veg type 2	Medium
ZMIN (mm mo ⁻¹)	Minimum infiltration rate	High
ZMAX (mm mo ⁻¹)	Maximum infiltration rate	High
ST (mm)	Maximum soil moisture storage	Medium
POW ^(*)	Power function for unsaturated runoff	High
FT (mm mo ⁻¹)	Maximum unsaturated zone runoff	High
GW (mm mo ⁻¹)	Maximum groundwater runoff	High
R ^(*)	Actual versus potential evaporation	Medium
GPOW ^(*)	Power function for groundwater runoff	High
DDENS (km km ⁻²)	Drainage density	Medium
T (m ² d ⁻¹)	Transmissivity	Medium
S	Storativity	Medium
GWS ^(*)	Groundwater slope	Medium
RWL (m)	Rest water level	Medium
RSF (%)	Riparian strip factor	High

^(*) Denotes unitless input

Table 2. Important inputs for the 3 and 5 regional approaches. A ✓ indicates importance. Gray shading indicates agreement in importance/unimportance between the three and five region approaches.

Process	Input	Three Regions			Five Regions				
		West	East	South	West	East	South	middle	Omatoko
Rainfall	RDF	✓	✓	✓	✓	✓	✓		✓
Interception	PI1								
	PI2				✓				
Infiltration	ZMIN	✓					✓		
	ZMAX	✓	✓	✓	✓	✓	✓		✓
Evapotranspiration (ET)	RSF							✓	
	AFOR				✓				
	FF	✓	✓		✓	✓			
Actual/Potential ET	R		✓						
Unsaturated storage	POW	✓			✓				
	FT	✓			✓	✓			
Unsat. GW storage	ST	✓			✓	✓			
Groundwater (GW) storage	GW	✓	✓		✓	✓			
	GPOW	✓	✓		✓	✓			
	T								
	S								
	RWL								
GW discharge	DDENS								
Interbasin flow	GWS								

Table 3. Comparison of the three and five region uncertainty analyses for two objective functions (1) the coeff for monthly flow at Mohembo and (2) the Mean Annual Flow in MCM at Mohembo.

		Mean	Median	SD	SEM σ	Skew γ	Kur κ^T
Coefficient of Efficiency	3 Region	0.226	0.411	$9.77e^{-3}$	$4.98e^{-3}$	-3.16	13.6
	5 Region	0.252	0.405	$7.88e^{-3}$	$4.02e^{-3}$	-3.20	15.2
Mean Annual Flow	3 Region	8399	8327	1941	16.5	0.472	0.290
	5 Region	8365	8242	1791	13.6	0.398	0.275

SD = standard deviation; SEM = standard error of the mean; Skew = skewness, kurt = kurtosis

Table 4. Ranking of importance of input variables

Rank ^(*)	Input	Uncertainty of PDF ^(**)	
Important	1	RDF	Medium
	2	ZMAX	High
	3	GW	High
	4	FF	Medium
	5	GPOW	High
Medium Importance	6	ST	High
	7	R	Medium
	8	PI2	Medium
	9	FT	High
	10	ZMIN	High
	11	POW	High
	12	AFOR	Low
Less Important	13	PI1	Medium
	14	DDENS	Medium
	15	S	Medium
	16	T	Medium
	17	RSF	High
	18	GWS	Medium
19	RWL	Medium	

^(*) Importance designated by the Morris Method through a total sum of first and higher order interactions. Important greater than 10, 10 < Medium importance < 1, Low importance < 1.

^(**) Low $\pm 15\%$, Medium $\pm 30\%$, High $\pm 45\%$ of calibrated values.

Table 5. Stationary versus climate change confidence intervals (CI). Units in flow (MCM).

	Stationary	Climate Change	% Difference
Lower 95% CI	7,373	9,923	29%
Average	8,108	10,822	28%
Upper 95% CI	8,897	11,747	29%
CI width	1,523	1,824	18%

Figure Captions

Figure 1. Location Map

Figure 2. Model conceptualization.

Figure 3. (a) three and (b) five regional simulations of the Okavango Basin.

Figure 4. Morris Method three region GSA results for the ceff of monthly flow at Mohembo.

Figure 5. Morris five region sensitivity analysis for the ceff of monthly flow at Mohembo.

Figure 6. FAST uncertainty analysis for annual mean monthly flow at the outlet of each of the regions. Flow is shown as (a) three region regular, (b) three region normalized by catchment area, (c) five region regular, and (d) five region normalized by catchment area.

Figure 7. Uncertainty analysis of ceff's for the three and five region approaches.

Figure 8. FAST three region sensitivity analysis for the ceff of the monthly outflow at Mohembo. First order sensitivity in grey, higher orders in black, and the total sensitivity as the entire bar.

Figure 9. FAST five region sensitivity analysis for the ceff of the monthly outflow at Mohembo. First order sensitivity in grey, higher orders in black, and the total sensitivity as the entire bar.

Figure 10. Cobweb plot of best fit model trajectories. Input values are normalized (0-1) by the ranges of their assigned PDF's. Each line represents one of the top 30 best fit lines from the three region simulation. These runs are chosen only from those that have a ceff greater than 0 in all regions.

Figure 11. Map of the model structure and the identification of the most important processes. Black fill indicates the most important inputs, dark grey is medium importance, light grey is low importance, and white indicates no importance (boundary conditions or outputs). Structure adapted from Kapangaziwiri (2008) and Hughes et al. (2006).

Figure 12. Total sensitivity of ceff for outflow at Mohembo for (a) 3 and (b) 5 regional scenarios.



Figure 1.

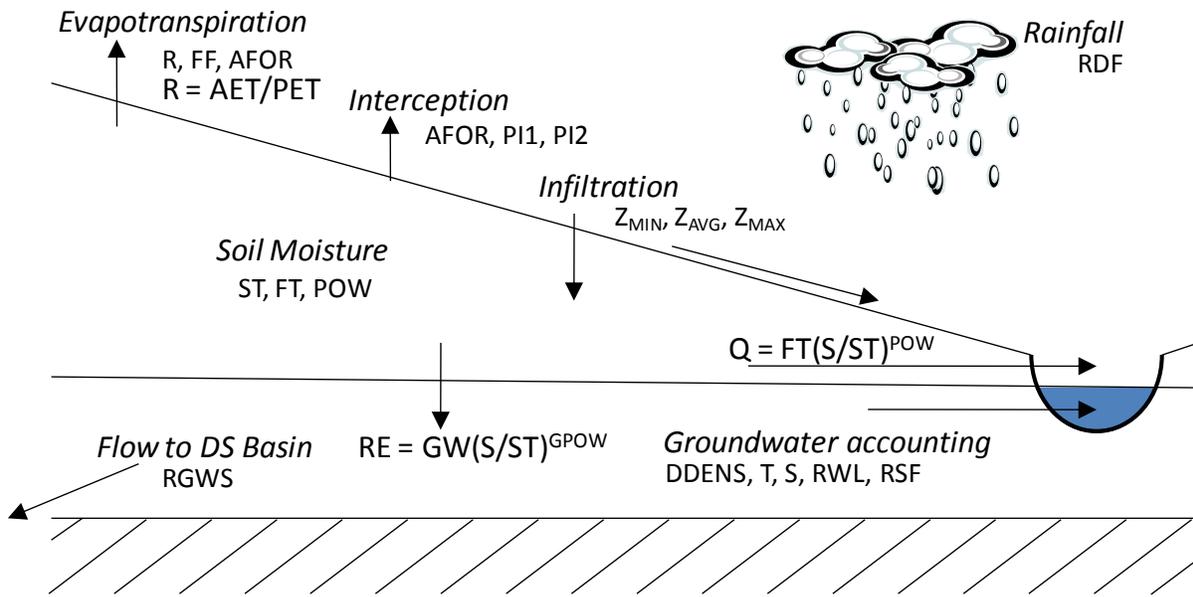


Figure 2.

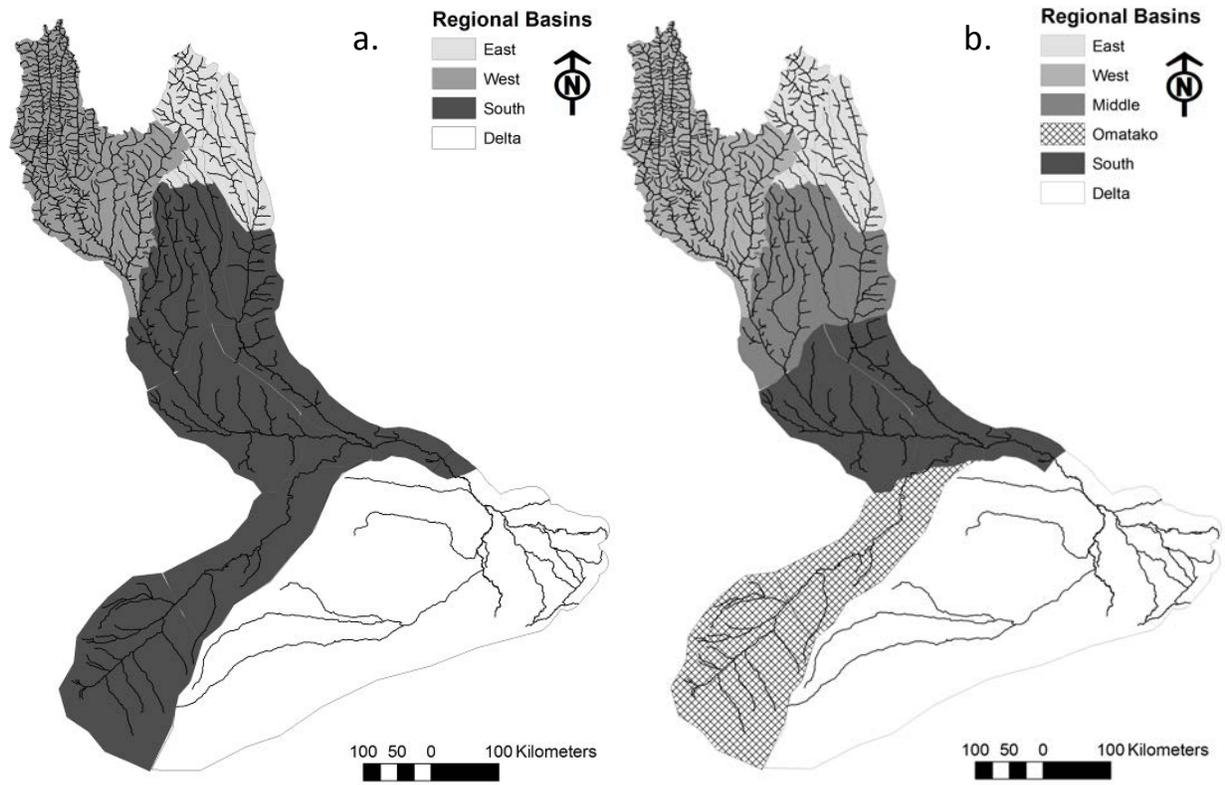


Figure 3.

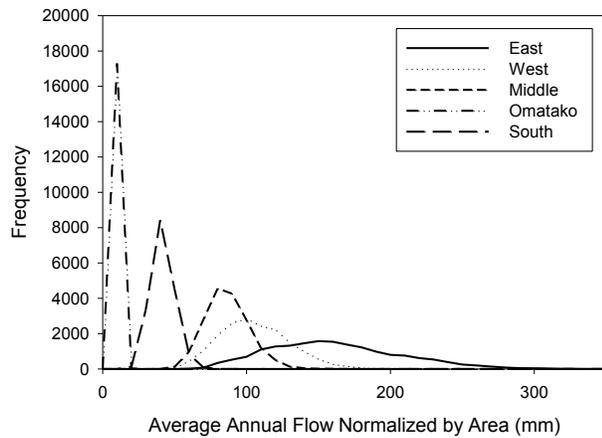
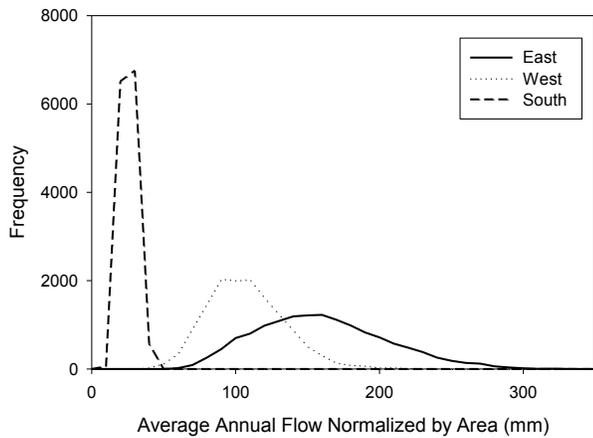
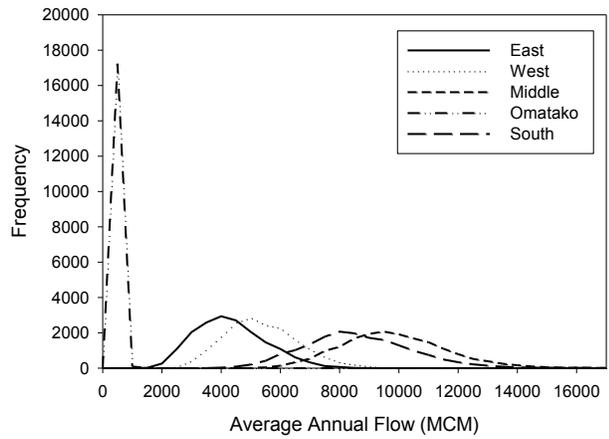
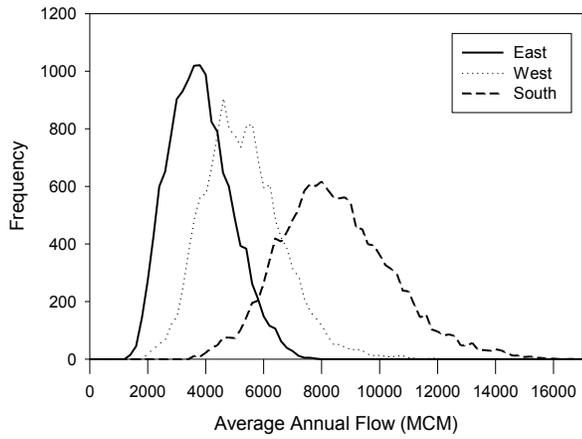


Figure 6.

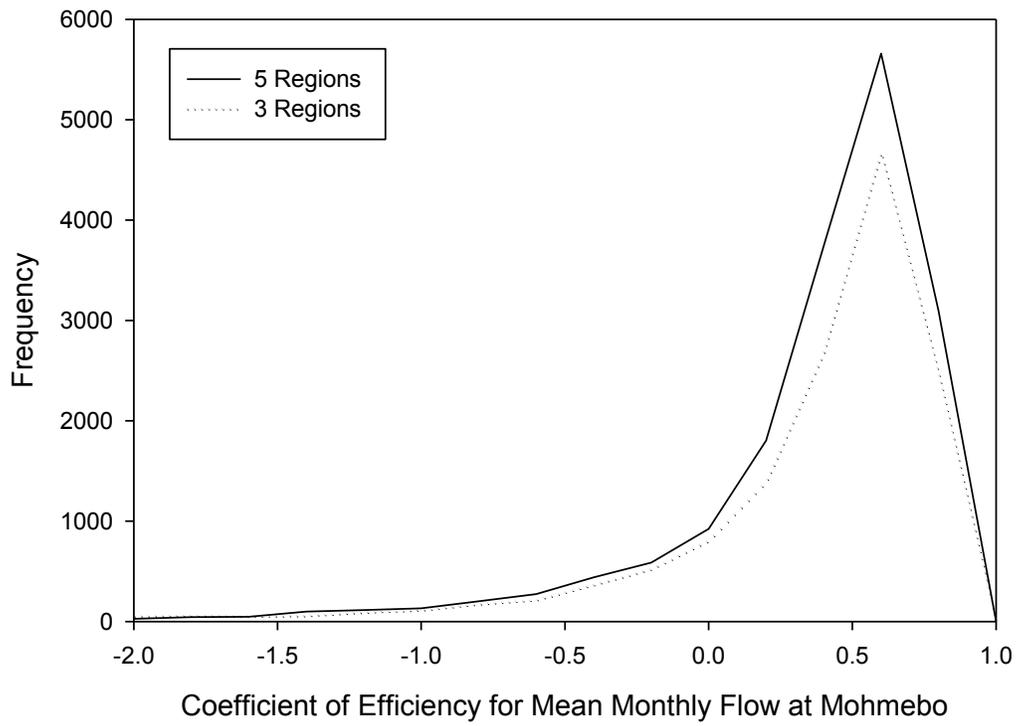


Figure 7

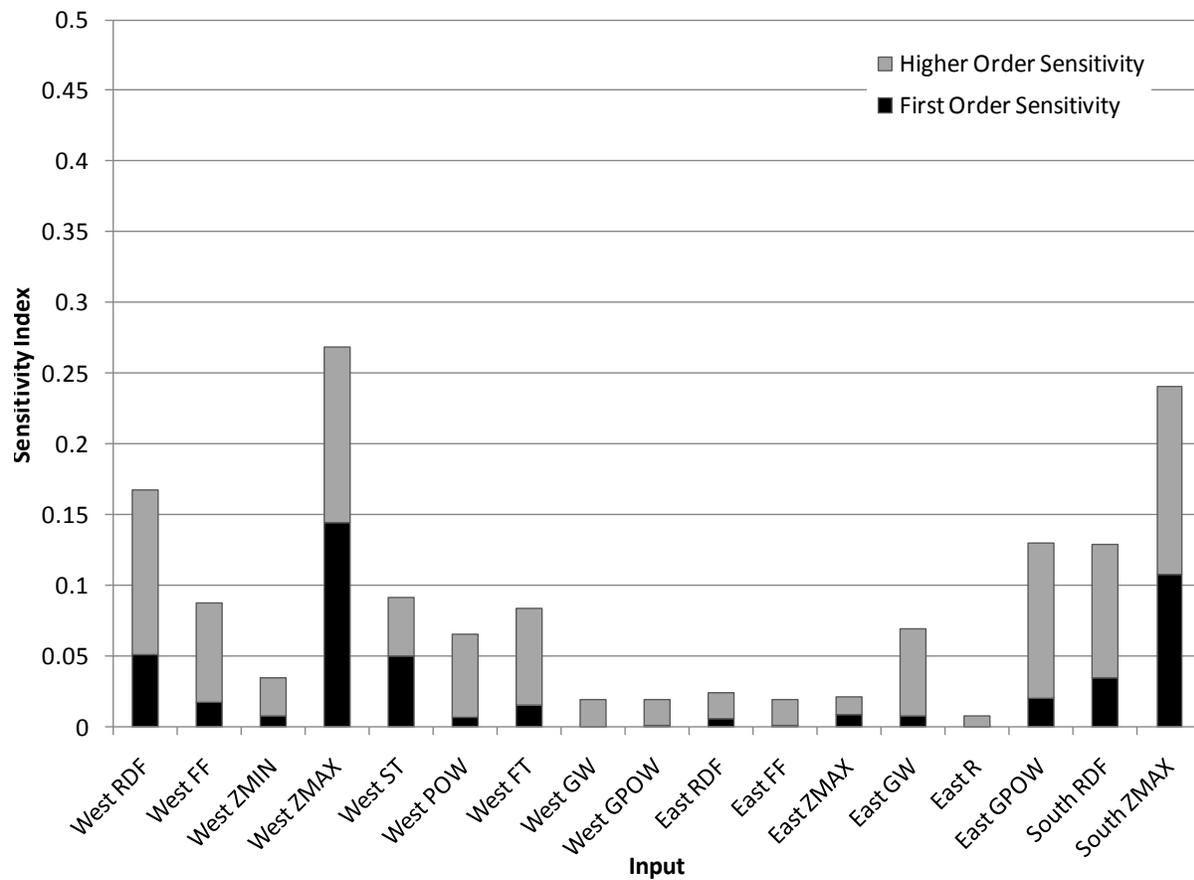


Figure 8.

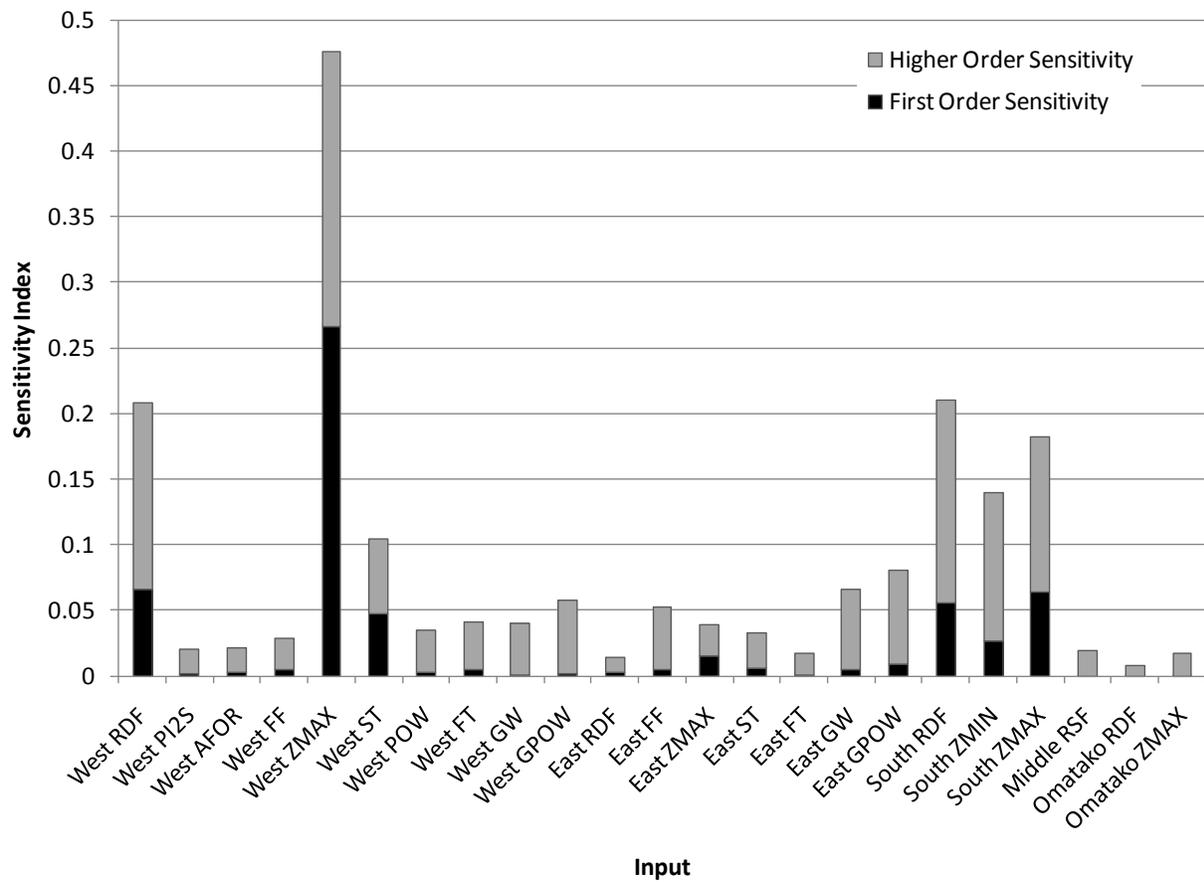


Figure 9.

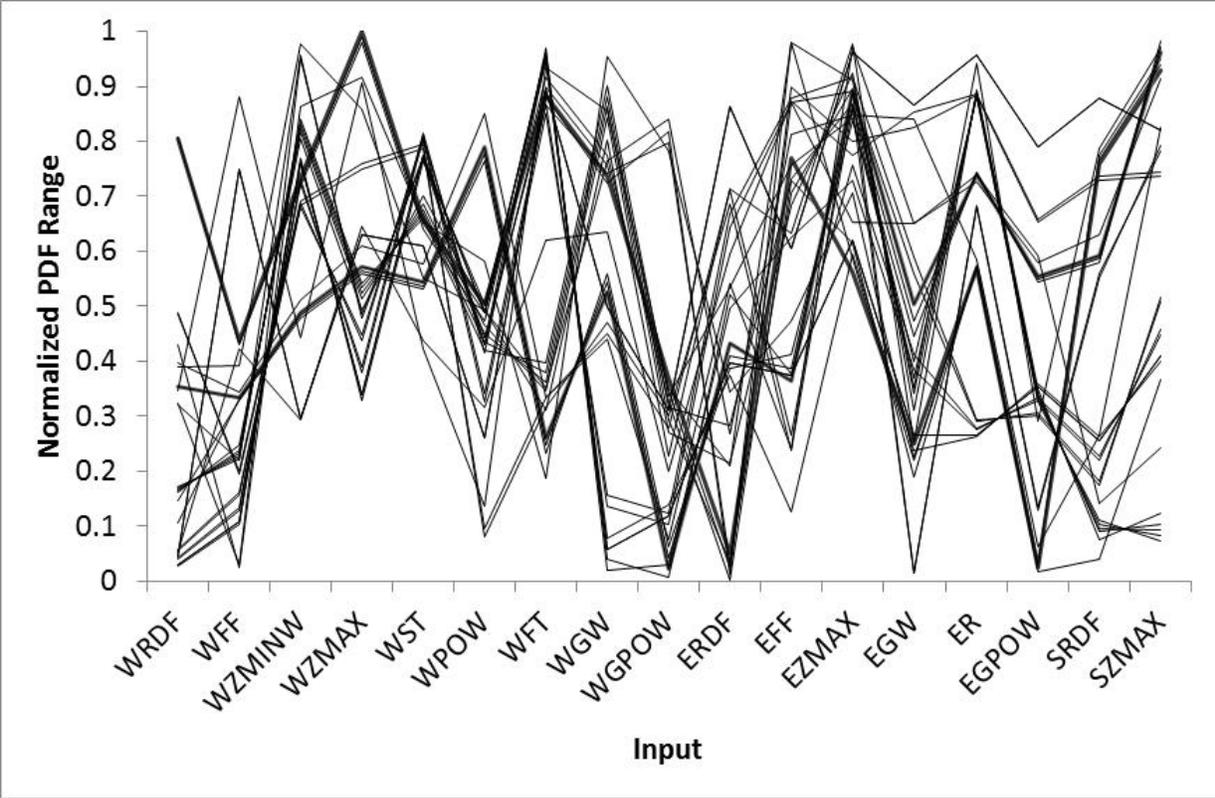


Figure 10.

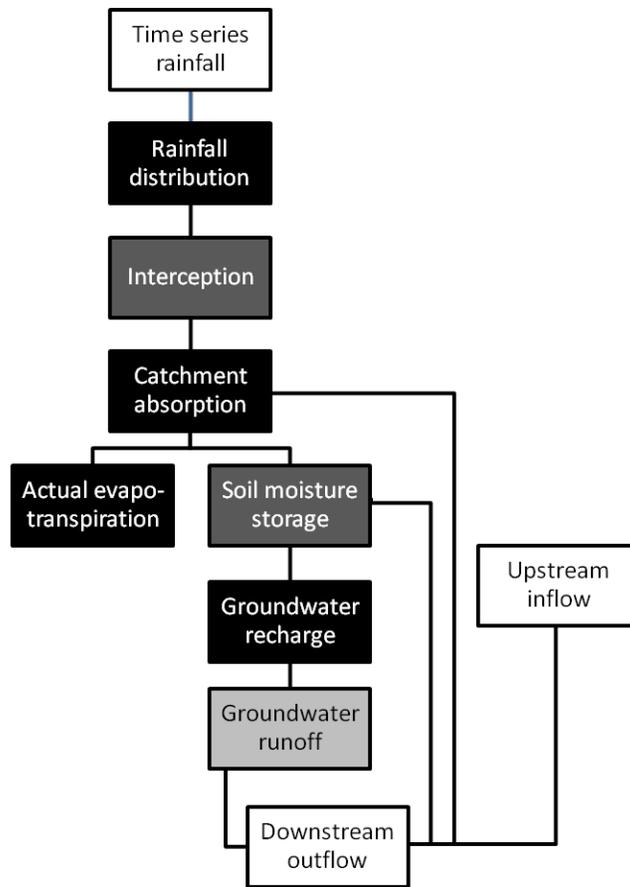


Figure 11.

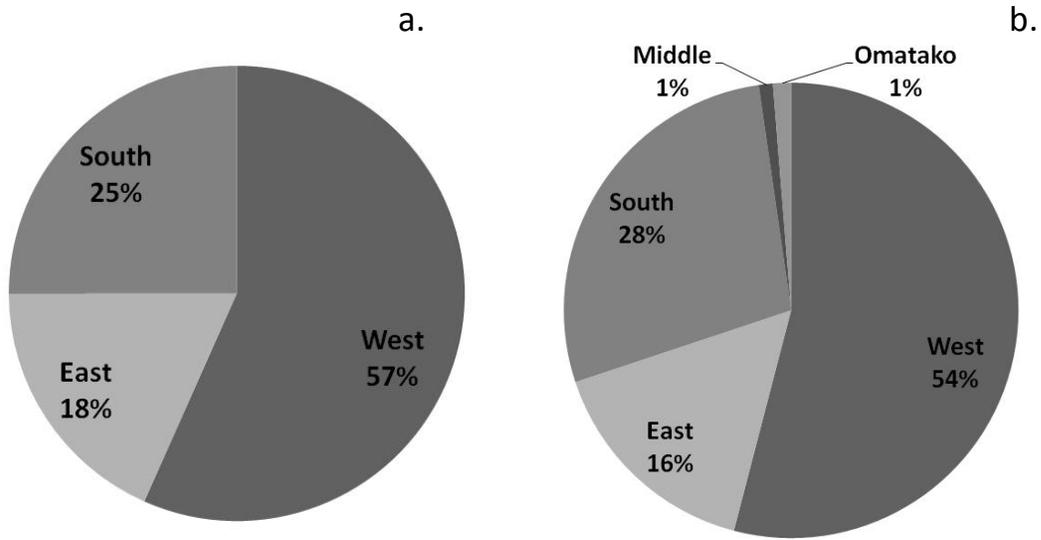


Figure 12.

Chapter 3. Objectively defining uncertainty in the face of data scarcity using a reservoir model of the Okavango Delta, Botswana

Abstract

Sensitivity and uncertainty analysis are tools that give managers quantifiable confidence in model results and give modelers and researchers guidance for optimizing models. The Okavango Delta in Botswana is a large and unique inland Delta and one of the most pristine large wetlands in Africa. The Okavango Research Institute (ORI) hydrologic model is a linked reservoir model that is overlaid by a GIS grid model that was developed to simulate the affect that the annual flood pulse has on the Delta. This model has been calibrated and validated in the Delta however; a formal sensitivity and uncertainty analysis has not yet been conducted on the model. The objective of this research is to conduct a global sensitivity and uncertainty analysis (GSA/UA) on the ORI hydrologic model of the Okavango Delta in order to identify the most important model inputs, to quantify the uncertainty in the model output, and objectively define prior probability distributions through Monte Carlo Filtering. This research applies two GSA/UA methods to the ORI model: the modified Method of Morris and the extended FAST technique. The Morris method is a qualitative GSA screening technique and the extended FAST method is a quantitative variance based GSA/UA technique. One of the largest criticisms of the GSA process is the rather arbitrary methods for establishing probability distributions (PDFs). Through this work, Monte Carlo Filtering is shown to be an objective method for redefining prior PDFs and reducing model uncertainty. Results indicate that the most important inputs are the volume thresholds, which determine the volume at which water will begin to spill out of

each reservoir, followed by the floodplain porosity, which is important for infiltration, and the island extinction coefficient, which is used to simulated transpiration. There is one reservoir, Nqoga1, which is highly interconnected and may be considered a keystone reservoir. Monte Carlo filtering was used as an objective method for reassigning the PDFs and reducing model uncertainty. These results identify Nqoga1 as a critical area for monitoring and managing within the Delta, refine and reduce model uncertainty, and present a general method, useful beyond this work, for objectively developing input PDFs especially where data is scarce.

1. Introduction

Because models are simplifications of reality they are inherently uncertain. Ignoring model uncertainty undermines the value of a model for its use in the decision making process (Beven, 2006a). Understanding this uncertainty is important for several reasons. Understanding uncertainty allows decision makers to acknowledge the reliability of models when weighing the risk of various decisions (Saltelli et al., 2008). In addition, identifying sensitive factors allows for the optimized allocation of resources for future data collection leading to the refinement of the input values and a reduction of model uncertainty. Sensitivity analysis can also provide an understanding of how input factors interact within the model structure. In a complex model, these interactions may not be obvious and may have important impacts on the model output due to their non-additive nature. (Saltelli et al., 2008)

Adaptive Management (AM) acknowledges uncertainty through a cyclical process that adapts and improves based on continuous monitoring and assessment of goals and is advocated for complex problems (Walters and Holling, 1990). Modeling is an integral part of the AM process

as models are able to simulate large scale experiments and scenarios. As a cyclical process, through AM models are continuously improved based on new information.

Global Sensitivity and Uncertainty Analysis (GSA/UA) and Monte Carlo Filtering (MCF) are two tools that can be used to systematically understand and reduce the uncertainties within a model and thus, assess and improve the model's reliability (Scott, 1996; Saltelli et al., 2008).

GUA quantifies model uncertainty and GSA apportions the total model uncertainty to each of the inputs. One of the biggest criticisms of GSA/UA is the rather arbitrary setting of prior PDFs.

MCF can be used to objectively redefine prior PDFs to make the results of the uncertainty analysis more realistic when there is no physical or empirical data to base the PDFs on

otherwise. This is particularly an issue in data scarce areas where there is little data available

for defining the PDFs. Using MCF to redefine prior PDFs based on realistic or acceptable model results gives an objective way to set prior PDFs. This is particularly useful for management

purposes because only realistic model results are investigated. Redefining these PDFs allows a GUA that more accurately represents the calibrated values and gives insights into the inner working of the model.

This paper presents the work of a GSA/UA and MCF on a reservoir model in the Okavango Delta, Botswana. The Okavango Basin is a large transboundary watershed in southern Africa.

Within the Basin, the Okavango River feeds the Okavango Delta, which is a large inland delta.

The Delta is an alluvial fan that spreads out on the edge of the Kalahari Desert. The water from the Basin is evapotranspirated in the Delta and never reaches the sea. This inland Delta and its

hydrology are particularly important and unique because it supports a diversity of wildlife and people who otherwise exist in an extremely dry environment (Kgathi et al., 2006).

From 1975 to 2002 Angola, the largest landholder in the basin, has been in a civil war and as a result economic development in the area has been limited (Kgathi et al., 2006). Current peace in the area has removed previous barrier for development and commercial expansion and the impacts on natural resources that may follow may be imminent. Additionally, climate change projections show various scenarios for the future of the area which could alter the current flooding patterns of the Delta (Milzow et al., 2010, Wolski 2009). As policy makers in the area face management decisions, the availability of a hydrologic model that can simulate flows within the Delta with some known degree of certainty will give managers the appropriate confidence in predictive modeling scenarios.

The relatively empirical approach of the reservoir model has persisted due to the Delta's remote location, its large size, and the fact that the topography has little relief. The Delta is located in northwest Botswana, no roads traverse the area, and trips to the internal portion are possible only by air or boat. Furthermore, the system is quite large. Flooding extents in the Delta range from 6,000 – 12,000 km² (Wolski et al., 2006) and the entire geologic alluvial fan is approximately 40,000 km² (Gumbrecht et al., 2005). The area also is extremely smooth and flat. Maximum local relief is generally less than 2 or 3 m (McCarthy et al., 2003; Gumbrecht et al., 2001) and the standard deviation of transects are approximately 5⁻⁶/1 km/km (Gumbrecht et al., 2005). Therefore, as in most wetlands, flow direction and inundation area is determined by very slight changes in elevation. Traditional techniques that measure changes in elevation over

this scale do not capture the shallow relief. Furthermore, temporal variations in flow are caused by poorly understood and hard to model processes such as hippopotami movements, sedimentation, and vegetation. Overall, data scarcity is a significant issue in this area.

A reservoir modeling approach has often been used to simulate flows in the Delta (Dincer et al., 1987; SMEC, 1990; Scudder et al., 1993; Gieske, 1997; WTC, 1997; Wolski et al., 2006). A review of these models can be found in Kiker et al. (2008) and Wolski et al. (2006). In the ORI model, flows in the Delta are routed through linked reservoirs. The ORI links the reservoir model to a GIS grid model which floods individual 1 km² pixels. The inundated status of each pixel is based on probabilities gathered from historical satellite imagery of the flooding extents. This model was developed specifically for the Okavango Delta and the outputs include reservoir flooding extents as well as outflows from the distributaries. The ORI model has been calibrated and tested (Wolski et al., 2006); however, neither a formal uncertainty nor sensitivity analyses have been conducted on the model. Such an uncertainty and sensitivity analysis will give reliability in the model providing managers with confidence in model results. This process will also identify key areas and processes for managing and monitoring.

Outputs from the ORI model have been used by policy makers in the area (DEA, 2008). The Okavango Delta Management Plan (2008) references the use of this model for predicting the impacts of abstractions, seismic activity, climate change, and channel clearing. This model is currently being used to predict climate change scenarios (Wolski, 2009).

The objectives of this paper are to assess the ORI model's to model uncertainty, provide confidence in model results, and identify areas where future work may enhance model

performance. This is done by conducting a GSA/UA and MCF on the model. Prior probability density functions (PDFs) will be developed to represent the range of plausible values for any given input. Inputs factors here will be defined as non-time series variables in the model. The GSA/UA will use a framework in which the Modified Morris GSA (Morris, 1991; Campolongo et al., 2007) screening method will be employed followed by the quantitative extended FAST GSA/UA technique (Cukier et al., 1978; Koda et al., 1979; Saltelli et al., 1999). Monte Carlo Filtering (Saltelli et al., 2008) will be used to objectively redefine PDFs and refine model uncertainty. Through this methodology, the sensitivity and uncertainty of the model will be quantified, model uncertainty will become more accurate and precise, and a deeper understanding of the model structure and reliability will be gained. For the purposes of this research and in accordance with Wolski et al.'s original calibration (2006), the average inundation area of the Delta as well as the individual reservoirs will be the objective function. This takes into account not only the entire Delta, but also the individual reservoirs and their behaviors. The inundation area is an appropriate objective function because in the absence of good topographic data it is a proxy for volume, and is an important social and ecological factor with regards to floods and droughts.

2. Methods

2.1 The Study Site: Okavango Delta, Botswana

The Okavango Basin is a large transboundary [maximum flooding extents 12,000 km² (Wolski et al., 2006)] watershed located in southern Africa that is shared between three countries: Angola, Namibia, and Botswana (Figure 1). River flow originates in the Angolan portion of the Basin where rainfall ranges from 1600 mm yr⁻¹ in the north to 600 mm yr⁻¹ in the south. Ninety-five

percent of the runoff generating area is contained within Angola (Hughes et al, 2006). The two downstream basin nations: Namibia and Botswana are much more arid with rainfall averaging approximately 450 mm yr⁻¹. Angola is also more mountainous, with elevations ranging between 1,200 and 1,800m, while the Botswana portion of the Basin is extremely flat, with elevations ranging between 900 and 1,000m.

Water flows in a delayed annual flood pulse through the Okavango River from Angola, through Namibia, and then into Botswana, which contains the Okavango Delta. Rainfall in Angola occurs between December and March. It takes two months for this flood pulse to reach the top of the Delta and another four to five months to attenuate through the Delta, finally reaching the distal ends from August through October (Wolski et al., 2006). This water never finds the sea and instead spreads out over the flat alluvial fan that is the Okavango Delta and is then evaporated on the border of the Kalahari desert.

2.2 The ORI model

Wolski et al.'s ORI model (2006) is the most recently developed linked reservoir model created specifically for the Okavango Delta (Figure 2 and 3). The ORI model represents flood duration, flood frequency, flooding extents, and outflow from the Boro River at Maun (Figure 2). The model operates on a monthly time step. Flow is input at the top of the Delta according to gauge data. The volume of water in each reservoir is then simulated by the continuity equation where V is the volume of water, I is inflow, ET is evapotranspiration, P is precipitation, Q is outflow, and Q_{inf} is infiltration to groundwater (Wolski et al., 2006).

$$\frac{dV}{dt} = I - ET + P - Q - Q_{inf} \quad (1)$$

The inundated area is related to the volume of water through a power relationship. In this equation A is the inundated area, V is the volume of water, and n and b are empirical coefficients which are based on topographic data.

$$A = nV^b \quad (2)$$

Upstream reservoirs may have several outlets and can feed more than one downstream reservoir. Model inputs for each reservoir include area, topography, evaporation, and flow variables (Table 1 and Figure 3). Some of these inputs are constant throughout the model (extinction depths, porosities, and n in the area/volume relationship), and some vary between the reservoirs (b in the area/volume relationship, delay, rainfall ratio parameter, and groundwater reservoir areas), and some vary between the reservoir connections (flow resistance and volume threshold). Outputs from the reservoir model include monthly inundation area and outflow from the Boro distributary.

Groundwater flow and infiltration processes are represented by a series of sub-reservoirs including the surface reservoir, two floodplain groundwater reservoirs, and two island groundwater reservoirs. Water in the surface reservoir infiltrates into the floodplain groundwater reservoirs. Water in the floodplain groundwater reservoir then flows to the island groundwater reservoirs. The second island and second floodplain groundwater reservoirs are activated based on the volume of the flood.

Actual evaporation is based on potential evapotranspiration. Transpiration is not simulated with specific regard to types of plants. Evapotranspiration in non-inundated areas uses the

concept of an extinction coefficient which simulates a linear decrease in the rate of evapotranspiration with depth from the surface. There is little meteorological data available for the interior of the Delta. However, a weather station in Maun, on the border of the Delta, has time series data from 1970 to the present. The model uses the Penman-Montieth equation (Allen et al., 1998) to calculate the reference crop transpiration. These calculations were adjusted based on weather station measurements made inside the Delta and also with measurements from an eddy covariance system located at Maun and Mxaraga (Wolski et al., 2006).

Rainfall is input over the inundated areas of the Delta based on an inverse distance-weighted relationship between two weather stations in the area. The rainfall ratio parameter is used to vary the rainfall in the reservoirs based on this relationship. In the model, rain within the Delta is not simulated as falling on the non-inundated areas and cannot directly raise the water table

The ORI model links this reservoir model to a GIS model to simulate the spatial distribution of the flood. The GIS model is based on satellite imagery of flooding extents. A Gaussian probability density function (PDF) was assigned to each pixel describing the likelihood of the pixel being inundated given the size of a flood. Then, given a volume of water in the reservoir model each pixel is assigned an inundation/non-inundation status based on the more likely probability.

The model was calibrated manually through trial and error adjustments of the inputs (Wolski et al., 2006). The objective functions for calibration were the flooding extents and outflow from the Boro River. The estimation of flooding extents is derived from satellite imagery available

from McCarthy et al. (2003). The Delta was gridded and in each image cells were assigned a flooded/non-flooded status. Outflow from the Boro River is gauged in the river.

Results from the ORI model show good correlation with observed data. The model produces a monthly inundation area that compares to observed data with a root mean squared error (RMSE) of 528 km² and a correlation coefficient of 0.90 for the entire Delta (Wolski et al., 2006). Additionally, the outflow at Maun shows a RMSE of 11.8 M m³/month and a correlation coefficient of 0.91 when compared to observed flows (Wolski et al., 2006).

2.3 Global Sensitivity and Uncertainty Analysis

There are a number of different approaches available for conducting a sensitivity analysis on models ranging from simpler one-at-a time (OAT) variational methods to advanced global techniques (Saltelli et al., 2008). With OAT techniques the variation of the model output is investigated through changing a single model input at a time. This traditional sensitivity analysis method is limited since it usually explores a small parametric range, can only consider a few inputs, and does not account for interactions between inputs (Saltelli et al., 2008).

When model responses are non-additive and non-linear, as with many complex models, simple OAT techniques are not appropriate and global techniques that evaluate the input factors of the model concurrently over the whole parametric space (described by probability density functions) are more appropriate. Different types of global sensitivity methods can be selected based on the objective of the analysis (Cacuci, 2003; Saltelli et al., 2008). This study uses a model evaluation framework around two such modern global techniques, a screening method (Morris, 1991) and a quantitative variance-based method (Saltelli et al., 1999). This two-step

process has been used in the examination between input output relationships in several environmental investigations (Muñoz-Carpena et al., 2007; Muñoz-Carpena et al., 2010; Fox et al., 2010; Jawitz et al., 2008; Chu-Agor et al., 2011). The screening method allows an initial reduction in the number of inputs to use in the more computationally intensive quantitative FAST sensitivity and uncertainty analyses. Global sensitivity methods can yield four main products for the Okavango application. (1) GSA can be used to explore model behavior and give an assurance of the absence of errors (Saltelli et al., 2008). (2) GSA provides a ranking of the importance of the inputs. This is useful in prioritizing future data collection and calibration techniques (Saltelli et al., 2008). (3) GSA also identifies the type of influence of the important inputs (first order or higher interactions) (Saltelli et al., 2008). (4) An uncertainty assessment of the model can also be used as the basis for the risk assessment of proposed management scenarios.

2.3.1 Modified Method of Morris

The GSA/UA of the ORI model begins with a qualitative screening assessment using the Method of Morris (Morris, 1991). The Morris method is less computationally intensive than other quantitative methods and is therefore used to screen for the most sensitive or important inputs. The least important inputs are then set to constants and the most important inputs are further examined using the FAST quantitative GSA/UA. Each model input (X_i ; $i = 1, 2, \dots, k$) is assigned a PDF which is divided into p levels. An input matrix is formed that describes the region of experimentation with k factors and p levels. Values are selected from this matrix for model simulations according to an efficient sampling trajectory. When performing the Morris

method, the model is run for $n(k+1)$ iterations, where k is the number of inputs and n may vary between 4 and 10.

The probability distribution function (PDF) of elementary effects (F_i) is obtained through randomly sampling values in the parameter space (X_1, X_2, \dots, X_i) and then computing according to Eq. 3.

$$F_i \approx d_i(X) = \frac{y(X_i, \dots, X_{i-1}, X_{i+\Delta}, X_{i+1}, \dots, X_k) - y(x)}{\Delta} \quad (3)$$

Morris proposes two sensitivity indices: μ is the mean of F_i representing the general magnitude of effect of the uncertainty of an input on the model output and σ is the standard deviation of F_i , which estimates higher-order relationships such as nonlinear and interaction effects. This standard deviation represents the spread of the importance that each input is responsible for based on values chosen for the other inputs.

The Modified Method of Morris (Campolongo et al., 2007) has a number of improvements over the original Method of Morris. It allows an analysis of models with multiple outputs, allows factors to be grouped, and has a more effective sampling strategy at no additional computation cost. Under the Morris method elementary effects are local OAT measures but this method may actually be considered global because μ^* is an average of the elementary effects.

Therefore, the sensitivity index μ^* is approximately as good as indices based on variance methods. On the Morris scatter plots the x axis ranks the direct or first order of input factor importance and the y axis ranks the higher order interactions of the input factors. These higher

order interactions represent the range of the importance of an input that depends on the values for the other inputs.

2.3.2 Quantitative Variance-Based FAST Global Sensitivity and Uncertainty Analysis

Once the most sensitive factors are identified using the Modified Method of Morris, a more quantitative sensitivity analysis is conducted on the most sensitive factors using the Fourier Amplitude Sensitivity Test (FAST) (Cukier et al., 1978; Koda et al., 1979). FAST uses Fourier analysis to decompose the variance of model outputs into first order variances for each input. The Extended FAST technique (Saltelli et al., 1999) also allows for the quantification of higher levels of variance. These higher levels of variance describe the interactions between inputs. $V(Y)$ describes the summation of the first and higher order interactions as the total variance of a single input according to Eq. 4.

$$V(Y) = \sum_i V_i + \sum_{i < j} V_{ij} + \sum_{i < j < l} V_{ijl} + \dots + V_{123\dots k} \quad (4)$$

V is then the summation of the first order variance for all of the inputs (V_i) and also the residual which is the variance attributed to all interactions ($V_{i,j}$, etc). FAST also defines S_i as a measure of global sensitivity. S_i is the ratio of the variance that is attributed to a single input divided by the total model variance. In a model in which there are no interactions the sum of all S_i 's across all inputs is equal to one. In models where there are interactions this sum will be greater than one. One limitation of this variance-based technique is that it cannot be used on factors that are interdependent (Crosetto and Tarantola, 2001). When performing the extended FAST GSA/UA, a model is run for $C \approx fM$ iterations, where M is a number between 100's and 1000's and f is the number of inputs (Saltelli et al., 1999).

Both the Modified Method of Morris and Extended FAST rely upon sampling values from PDFs for each uncertain input factor and running the model based on these sampled data sets. SimLab v3.2.6 (SimLab, 2011) software was used to sample input PDFs and construct sets of model inputs. The model is then run multiple times based on the input sets created by SimLab. The model results are combined into SimLab with the inputs to obtain to Method of Morris statistics and FAST indexes.

2.4 Monte Carlo Filtering

Monte Carlo Filtering (MCF) is an additional method that can be used to decrease uncertainty (Saltelli et al., 2008; Rose et al., 1991). The Okavango is an area where data is particularly scarce and as a result the prior PDFs are wide and rather arbitrary. MCF can be used to objectively redefine and truncate prior PDFs to reduce and refine uncertainty. In MCF, a threshold is established such that model outputs are classified as either behavioral (B) or non-behavioral (\bar{B}). Based on whether or not the output exceeds the behavioral threshold, two subsets of the input set X_i are defined as X_i/B and X_i/\bar{B} .

To determine if X_i/B is statistically different from X_i/\bar{B} the two-sided Smirnov test is used on each input factor independently. The null hypothesis (H_0) is that the distribution of the inputs producing B is equal to the distribution of inputs producing \bar{B} . Following, hypotheses with $f_m(X_i | B)$ and $f_n(X_i | \bar{B})$ as the PDFs for input factor X_i are:

$$H_0 : f_m(X_i | B) = f_n(X_i | \bar{B}) \quad (5)$$

$$H_1 : f_m(X_i | B) \neq f_n(X_i | \bar{B}) \quad (6)$$

The Smirnov test statistic is:

$$d_{m,n}(X_i) = \sup_y \| F_m(X_i | B) - F_n(X_i | \bar{B}) \| \quad (7)$$

F is the cumulative distribution function for input factor, X_i , \sup is the maximum vertical distance between the behavioral and non-behavioral cumulative distributions, and $\|$ is the absolute value. A large value for $d_{m,n}$ implies that $f_m(X_i | B)$ and $f_n(X_i | \bar{B})$ are statistically different from each other and also that X_i is an important factor in defining the behavior of the model. (Saltelli et al., 2008)

2.5 Selection of Prior Input Probability Distributions

A careful and methodical selection of the prior PDFs for each input is especially important because it can be the most subjective step in this process and has significant impacts on the results. Probability distributions are determined based on literature values and expert opinion. When the data for the inputs shows no apparent distribution such as normal or triangular, a uniform PDF allows for equal probability of selection across the defined range (Muñoz-Carpena et al., 2007). Because of the lack of data that would suggest otherwise and in the spirit of conservatism, uniform distributions are used as the default prior distribution type. In an area where more data is available autocorrelation tests could be done to show dependence between data. However, here no such data exists and inputs are assumed independent.

The area volume relationships for the reservoirs, n and b_i , were based on 2 foot contour topographic data. The inputs n and b_i are empirically fit according to Eq. 2. In the calibrated model, n is set to 25 and b_i ranges between 0.62 and 0.76 depending on the reservoir, i .

Because of the absence of literature values regarding these inputs and the empirical nature of their derivation, the PDFs for n and b_i are each designated as uniform, and vary by some percentage around the calibrated value. The uniform PDF for n is set to Uniform ($\pm 20\%$) of the calibrated model. Setting PDFs for b_i is more problematic since, as a power exponent, varying b_i by $\pm 20\%$ of the calibrated values makes the model fail occasionally. As a result, the uncertainty associated with these inputs is set to a more limited distribution: Uniform ($\pm 10\%$) of the calibrated value. In the absence of literature values or other expert opinion the default PDFs for the remainder of the inputs are assumed to range between $\pm 20\%$ of the calibrated value.

The volume threshold input factors, v_{az} , where a is the upstream reservoir and z is the downstream reservoir, determine at what level water flows from one reservoir to another. These inputs are empirically derived and were used to calibrate the model (Wolski et al., 2006). There are 20 volume threshold inputs. Due to the absence of literature values describing ranges of probably values for these inputs, the PDFs are set to uniform, varying $\pm 20\%$ of their calibrated values.

Each reservoir is divided into an area of floodplain and island so that groundwater flow between the two areas can be represented. There are 16 floodplain, fa_i , and 16 island area inputs, ia_i . The PDFs for the area for floodplains and islands within each reservoir is set to uniform, varying $\pm 20\%$ of the calibrated values.

The island and floodplain extinction depths, $idet$ and $fdet$, are important for evapotranspiration, empirically represent rooting depth, and one value for each is used to represent the entire

Delta. The calibrated value was 20m for the floodplains and 5m for the islands. There are no literature values for extinction depths in the Okavango Delta or surrounding areas. However, literature values of modeled and measured values in for extinction depths in other areas vary between 1.5 and 60m (Shah et al. 2007, Niswonger et al., 2006, and Owe and Brubaker, 2000). This large range limits the capability of generating a probability distribution based on good physical data. Consequently, uniform distributions of $\pm 20\%$ of the calibrated values are assigned for the island and floodplain extinction depths.

The reservoir time constant input factor, k_{az} , is equivalent to surface flow resistance. As with the volume thresholds, there are 20 reservoir time constant inputs in the model. These values vary between 0.02 and 5. This input factor was used to originally calibrate the model (Wolski et al., 2006). In the absence of literature data regarding values for the reservoir time constant in the Okavango Delta, the prior PDFs are designated as uniform, varying $\pm 20\%$ of the calibrated value for each reservoir.

Soil porosity is represented as two input factors in the model: island ($ipor$) and floodplain ($spor$) soil porosity. In the model, these factors are constant throughout the Delta, are not reservoir dependant, and are each set to 0.3. Literature values show soil porosity varying between 0.29 and 0.59 in the Okavango Delta (Wolski and Savenije, 2006) and between 0.43 and 0.49 in Kalahari Sands (Wang et al., 2007). Based on these literature values, the PDF for both the island and floodplain soil porosities are set to be uniform varying between 0.29 and 0.59.

The delay input factor, $delay_i$, is an empirically derived discrete switch that slows the water flow and is turned on in four of the reservoirs in the calibrated model. This input was used in

calibrating the model. For the sensitivity analysis, the delay inputs which are turned on the calibrated model are set to turn on or off with a probability of 50%.

Rainfall is input into the model using meteorological data collected from Maun and Shakawe. An inverse distance-weighted coefficient was used to allow for proportional differences in rainfall amounts in each of the 16 reservoirs, m . In the calibrated model, values for m range between 0 and 0.9. To investigate the full range behaviors that this input may have on the model these inputs are assigned uniform PDFs that range between 0 and 0.9.

3. Results and Discussion

3.1 Modified Morris Global Sensitivity Analysis

With 114 inputs, the model was run 1150 times using the Morris method (Table 1). Two general conclusions were drawn from the Morris method sensitivity analysis (Figure 4). First, because the inputs are scattered on both the σ and μ^* axes, both direct (first order) and indirect (higher order) effects are important. Second, a ranking of input significance on the inundated area output is determined.

The results showed that there are nine inputs which stand out as the most sensitive, each falling above 300 on the direct and indirect axis of the Morris graph (Figure 4). These most sensitive inputs are: the island extinction coefficient ($idet$), the floodplain soil porosity ($fpor$), the exponential coefficient in the area volume relationship for Xudum (b_5), and the volume thresholds for the linkages between the Nqoga1 and: Boro (v_{26}), Manuachira1 (v_{27}), Xudum (v_{25}), Thaoge (v_{24}), Nqoga2 (v_{23}), and Selinda (v_{29}). From the strong sensitivity of these six volume thresholds, the importance of this type of model input is apparent. There are a total of

20 volume thresholds in the model. In the Morris analysis, six of the 20 volume thresholds stood out as highly sensitive inputs and they all flow out of the Nqoga1 reservoir (Figure 4). There are a total of seven volume thresholds that flow out of Nqoga1. The one volume threshold that originates from Nqoga1 but is not highly sensitive is linked to the Boro reservoir where there are duplicate links to the Nqoga reservoir. This double link in the Boro, which splits the flow between Nqoga1 and the Boro, is unique in the Nqoga1 reservoir and may explain why this volume threshold did not display the same level of sensitivity. Nqoga1 is a keystone reservoir that has important influence in determining how the flood is routed between the reservoirs and is responsible for much of the sensitivity in the model. The identification of Nqoga1 as a keystone reservoir and the volume thresholds as important inputs has management implications. This area should be focused on for monitoring and managing because it is likely that changes in this area could affect how the water is apportioned to the downstream reservoirs.

3.2 FAST Variance Based Global Sensitivity

The subset of the most sensitive inputs identified in the Morris method was selected for further investigation in using the Extended FAST method (Table 2). For this GSA/UA, the model was run 8,289 times, while varying the nine most sensitive inputs that were identified in the Morris method (*idet*, *fpor*, *b₅*, *v₂₆*, *v₂₇*, *v₂₅*, *v₂₄*, *v₂₃*, and *v₂₉*), throughout the parameter space that was defined by the probability distributions. All other inputs were set to constants. As with the Morris analysis, the output of interest is the flooding extents for each reservoir as well as for the entire Delta.

The FAST GSA corroborated the Morris results showing that while the higher order interactions within the model are important, the first order interactions dominate the model's behavior (Figure 5 and 6). The higher order interactions account for approximately 20% of the totally model sensitivity. In the FAST GSA the first order interactions are all relatively similar while the higher order interactions vary somewhat between the inputs. The fact that each of the important volume thresholds flowing out of the Nqoga1 reservoir has similar levels of importance shows that, for management purposes, this entire reservoir should be monitored for changes and protected from disturbance.

3.3 Monte Carlo Filtering

Using the results from the FAST GSA, MCF was conducted to objectively reduce model uncertainty. For MCF, the FAST results were filtered to find simulations that matched within $\pm 50\%$ of the calibrated output (average inundation area) for each reservoir. Behavioral solutions were sought where each reservoir's average inundation was within this range. No single GSA result produced a simulation in which all of reservoirs were within $\pm 50\%$ of the calibrated values. The inputs were examined to find the source of this issue. The important volume thresholds, v_{23} , v_{24} , v_{25} , v_{27} , and v_{29} are all for spillover from Nqoga1 into various reservoirs. All prior PDFs for these inputs were set to $\pm 20\%$ based on the empirical nature of the model structure and the lack of data providing physical evidence. The original calibrated values for these thresholds range from 580 to 700. For each GSA simulation, each of the volume thresholds was randomly set within the same range of $\pm 20\%$ and water was diverted between the reservoirs. It was determined that no behavioral solutions were found because the PDFs were too large and when the volume thresholds in v_{24} , v_{25} , v_{26} , and v_{29} were set to

opposite ends of their distributions, the water was spilling into one of the reservoirs and drying out the others. For this reason, the PDFs for the important volume thresholds were truncated to $\pm 10\%$ and the GUA was rerun.

Originally, $\pm 50\%$ was chosen as the behavioral threshold goal but this resulted in a minimum number of behavioral solutions even after the truncation of the important volume threshold PDFs. Increasing the behavioral threshold to $\pm 60\%$ of the calibrated value increased the number of behavioral solutions thus, increasing the likelihood of finding statistical differences between behavioral and non-behavioral distributions. The truncation of the important volume threshold PDFs resulted in 103 out of 8,289 simulations meeting the behavioral criteria. The two-sided Smirnov test showed that the behavioral distributions for v_{23} , and v_{27} were significantly different from the non-behavioral distributions. These are two highly linked reservoirs (Figure 3). The new filtered PDFs for v_{23} , and v_{27} displayed somewhat triangular distributions (Figure 7). New normal PDFs were assigned with $v_{23}T \sim (585, 645, 715)$ and $v_{27}T \sim (522, 616, 638)$. The FAST GSA/UA was rerun with the new prior distributions.

3.4 FAST Variance Based Global Uncertainty Analysis

The FAST MCF GUA results were normalized by the maximum inundation area to be able to compare the results across reservoirs. The results revealed a range of responses to the uncertainty analysis amongst the internal reservoirs (Figure 8). The upstream reservoirs and Maunachira1 demonstrated the least amount of uncertainty. This uncertainty increases in the middle and downstream eastern reservoirs and reaches its maximum in the western reservoirs. This trend mirrors the physical dynamics in the Delta where the upstream areas are more

confined, permanently flooded, and display little variation, while the downstream areas are seasonally flooded and exhibit more variation in their annual flooding extents. Furthermore, in the FAST results some of the reservoirs display a Gaussian distribution (Mbarga and Khwai) and some have long tails (Selinda).

After MCF, confidence intervals, standard deviations, standard error of the mean, skewness, and kurtosis for the flooding extents for each reservoir and the entire Delta were constructed to examine the variance (Table 2). This data shows the varying levels of uncertainty in the reservoirs. As was shown in Figure 8, the upstream reservoirs vary minimally and the downstream reservoirs displaying much more uncertainty. As can be seen in the standard deviation of the total Delta, the uncertainty in the total Delta is not the summation of the uncertainty in the individual reservoirs. This may be due to the fact that the inputs for this uncertainty analysis have little to do with the water balance and are more related to where the water is moving around within the different reservoirs. In the various simulations the total volume of surface water is changed less than which reservoirs the water is routed to.

The MCF GUA was compared to the original GUA (Figure 9 and Table 3). MCF reduced uncertainty in all of the reservoirs except two (Panhandle and Nqoga1) (Table 3). These two reservoirs are both upstream from where the effects of the important volume thresholds should be felt. Reduction of uncertainty downstream from Nqoga1 varied between 13 and 189%. Thus, MCF was used to objectively truncated the prior PDFs and reduce the model uncertainty improving the precision of the results. Additionally, the MCF produced modeled results that more closely represented the calibrated ranges (Figure 9). After the MCF, the

results cluster around the calibrated values more so than before this additional analysis.

Therefore, MCF enhanced both the accuracy and the precision of the uncertainty analysis.

4. Conclusions

This GSA/UA of the ORI reservoir model provides important insights into the sensitivity and uncertainty of the ORI model and into reservoir models in general. These insights are useful for modelers when in interpreting and defining model uncertainty and for policy makers through providing insights into model reliability and identifying important areas of the system.

Managers may use the results of a sensitivity analysis to strategically monitor and identify and optimize model reliability. The Adaptive Management process is a cyclical process that is constantly modeling, monitoring, learning and improving. This work has identified important areas for study that will optimize the learning process. The greatest hindrance to modeling the hydrology of the Delta, as shown by this work, is the lack of topographic data. This sensitivity analysis highlights the importance of volume thresholds in the ORI reservoir model and particularly the keystone reservoir Nqoga1. Future monitoring efforts, which are aimed at improving model performance, should be focused on understanding the topography of this area and how water is routed from this area to the distributaries. Managers should be aware that any changes in the Nqoga1 area may have dramatic downstream consequences.

Additionally, the MCF reduced uncertainty which improves model reliability and adds confidence to model results which is helpful for managers when making decisions based on these results.

This exercise was useful for modelers because it reduced model uncertainty and also identified the importance of model inputs. During this sensitivity analysis a number of inputs were shown to be relatively insensitive and have little impact on the model outputs. Insensitive inputs include k_{az} , ia_i , and m_i . In future modeling efforts these inputs may be ignored or set to constants in order to simplify modeling and calibration without compromising accuracy. Conversely, the inputs: v_{az} , $idet$, and $fpor$ were found to be particularly sensitive. Additionally, Nqoga1 was identified as a keystone reservoir. Future calibration efforts should start by adjusting these inputs within their likely bounds. In particular v_{az} was found to be quite sensitive and future efforts may focus on obtaining physically determined values for v_{az} . Doing so may not only improve the model but will also result in a more physically representative model.

Ivanović and Freer (2009) point out that it is especially difficult to objectively test the uncertainty of inputs for empirical models and that most methods proposed thus far have been subjective (Beven, 2006b; Beven et al., 2007; Goldstein and Rougier, 2009). Pappenberger and Beven (2006) also make the point that uncertainty analyses may be too subjective to be of practical use. In this work MCF was used in conjunction with GSA/UA to explore how PDFs can be objectively set and model uncertainty reduced. This is especially an issue in data poor areas. MCF was used to objectively redefine prior PDFs based on behavioral outputs and decrease uncertainty. This is a function of the structure of the reservoir model and the uniform distribution of the inputs. In the GUA, changing the volume thresholds shifted water from one reservoir to another in a physically unrealistic manner. When the thresholds were given

uniform distributions and assigned values on opposite ends of their distributions, one reservoir would fill up while the others remained empty. Thus, the narrow uncertainty bounds clustered around the origin in the original GUA for these inputs. Changing these input PDFs based on the behavioral distributions narrowed their ranges and forced the majority of the inputs toward the calibrated values. This reduced uncertainty and it moved the outputs closer toward the calibrated values (Figure 9). Therefore, this MCF method successfully redefined the behavioral PDFs so that the model outputs more closely resembled the calibrated values in precision and accuracy.

In conclusion, this work has technical and management implications by presenting an objective way of defining prior PDFs, reducing model uncertainty, and identifying important areas for study. As the ORI model is used and refined in the future, these points can optimize how the model is improved. The MCF method for defining prior PDFs is a generic tool that can be used to improve any GSA/UA application, especially in data poor areas.

Acknowledgements

The authors would like to thank the University of Botswana and the Okavango Research Institute for their assistance in this research as well as the National Science Foundation, the University of Florida's IGERT AM-W3 program, and Mark Brown for their support. Also, the authors acknowledge the University of Florida's High Performance Computing Lab, for providing computational resources and support URL:<http://hpc.ufl.edu>.

References

- Allen R.G., Pereira L.S., Raes D., Smith M., 1998. Crop evapotranspiration: Guidelines for computing crop requirements. Irrigation and Drainage. Paper No. 56, FAO, Rome, Italy.
- Beven, K., 2006a. On undermining the science? Hydrological Processes, 20, 3141-3146.
- Beven KJ. 2006b. A manifesto for the equifinality thesis. Journal of Hydrology, 320: 18–36.
- Beven KJ. 2007. Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. Hydrology and Earth System Sciences 11: 460– 467
- Cacuci, D.G., Mihaela, I.B. , Ionel, M.N., 2003. Sensitivity and uncertainty analysis. Chapman & Hall/CRC Press, Boca Raton.
- Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environmental Modeling & Software, 22, 1509-1518.
- Chu-Agor, M.L., Munoz-Carpena R., Kiker G., Emanuelsson A., Linkov I., 2011. Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses, Environmental Modelling & Software, 26(5), 593-604.
- Crosetto, M., Tarantola, S., 2001. Uncertainty and sensitivity analysis: tools for GIS-based model implementation. International Journal of Geographical Information Science, 15(5), 415-437.
- Cukier, R.I., Levine, H.B., Shuler, K.E., 1978. Nonlinear sensitivity analysis of multiparameter model systems. Journal of Computational Physics, 26, 1-42.

- Department of Environmental Affairs (DEA), 2006. Okavango Delta management plan. Okavango Delta management plan project. Gaborone, Botswana.
- Dincer, T., Child, S., Khupe, B., 1987. A simple mathematical model of a complex hydrologic system--Okavango Swamp, Botswana, *Journal of Hydrology*, 93, 41-65.
- Fox, G.A., Muñoz-Carpena R., Sabbagh G.J., 2010. Influence of flow concentration on parameter importance and prediction uncertainty of pesticide trapping by vegetative filter strips, *Journal of Hydrology*, 384, 164–173.
- Gieske, A., 1997. Modelling outflow from the Jao/Boro River system in the Okavango Delta, Botswana. , *Journal of Hydrology*, 193, 214-239.
- Goldstein M, Rougier J. 2009. Reified Bayesian modelling and inference for physical systems. *Journal of Statistical Planning and Inference* 139: 1221– 1239, DOI:10.1016/j.jspi.2008.07.019.
- Gumbrecht, T., McCarthy, T.S., Bauer, P., 2005. The micro-topography of the wetlands of the Okavango Delta, Botswana. *Earth Surface Processes and Landforms*, 30, 27-39.
- Gumbrecht, T., McCarthy, T.S., Merry, C.L., 2001. The topography of the Okavango Delta, Botswana, and its tectonic and sedimentological implications. *South African Journal of Geology*. 104, 243-264.
- Jawitz, J.W., Muñoz-Carpena, R., Muller, S., Grace, K.A., James A.I., 2008. Development, Testing, and Sensitivity and Uncertainty Analyses of a Transport and Reaction Simulation Engine

- (TaRSE) for Spatially Distributed Modeling of Phosphorus in South Florida Peat Marsh Wetlands: U.S. Geological Survey Scientific Investigations Report 2008-5029. 109 p.
- Kgathi, D.L., Kniveton, D., Ringrose, S., Turton, A.R., Vanderpost, C.H.M., Lundqvist, J., et al., 2006. The Okavango: A river supporting its people, environment and economic development. *Journal of Hydrology*, 331, 3-17.
- Kiker, G.A., Munoz-Carpena, R., Wolski, P., Cathey, A., Gaughan, A., Kim, J., 2008. Incorporating uncertainty into adaptive, transboundary water challenges: a conceptual design for the Okavango River basin. *International Journal of Risk Assessment and Management*, 10, 312-338.
- Koda, M., Mcrae, G.J., Seinfeld, J.H., 1997. Automatic sensitivity analysis of kinetic mechanisms. *International Journal of Chemical Kinetics*, 11, 427-444.
- McCarthy, J.M., Gumbricht, T., McCarthy, T., Frost, P., Wessels, K., Seidel, F., 2003. Flooding patterns of the Okavango wetland in Botswana between 1972 and 2000. *Ambio*. 32, 453-457.
- Milzow, C., Kgotlhang, L., Burg, V., Kinzelbach, W., 2008. Setup and calibration of a distributed hydrological model of the Okavango Delta using remote sensing data. *Proceedings of the ModelCARE 2007 Conference, Denmark, IAHS Publ. 320, IAHS, Wallingford, UK, pp 169–175*
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics*. 33, 161-174.

- Muñoz-Carpena, R., Zajac Z., Kuo Y.M., 2007. Global sensitivity and uncertainty analysis of the water quality model VFSSMOD-W. *Transactions of the American Society of Agricultural and Biological Engineers* 50, 1719-1732.
- Muñoz-Carpena, R., Fox, G.A., Sabbagh, G.J., 2010. Parameter importance and uncertainty in predicting runoff pesticide reduction with filter strips, *Journal of Environmental Quality*, 39(1), 1–12.
- Niswonger, R.G., Prudic, D.E., Regan, R.S., 2006. Documentation of the unsaturated-zone flow (UZFI) package for modeling unsaturated flow between the land surface and the water table with MODFLOW-2005, *USGS Techniques and Methods 6-A19*. Reston, Virginia: USGS.
- Owe, M., Brubaker, K., 2000. *Remote Sensing and Hydrology 2000: A Selection of Papers Presented at the Conference on Remote Sensing and Hydrology 2000*. Santa Fe New Mexico.
- Pappenberger F, Beven KJ. 2006. Ignorance is bliss: 7 reasons not to use uncertainty analysis. *Water Resources Research* 42(5).
- Rose, K. A., Smith, E. P., Gardner, R. H., Brenkert, A. L., Bartell, S. M., 1991. Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. *Journal of Forecasting*, 10, 117-133.
- Saltelli, A., 2004. *Sensitivity analysis in practice: A guide to assessing scientific models*. Wiley, Hoboken, NJ, 2004.

Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S.,
2008. *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, West Sussex England.

Saltelli, A., Tarantola, S., Chan. K.P., 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*. 41, 39-56.

Scott, E.M., 1996. Uncertainty and sensitivity studies of models of environmental systems. In:
Charnes, J.M., Morrice, D.J., Brunner, D.T., Swain J.J., (Eds.). *Winter Simulation
Conference: Proceedings of the 28th conference on Winter simulation, Coronado,
California*. 255-259.

Scudder, T., Manley, R., Coley, R., Davis, R., Green, J., Howard, G., et al., 1993. The IUCN review
of the Southern Okavango Integrated Water Development Project. IUCN, Switzerland.

Shah, N., M. Nachabe, M. Ross. 2007. Extinction depth and evapotranspiration from ground
water under selected land covers. *Groundwater*, 45, 329-338.

Simlab, 2011. Simlab sensitivity analysis software. Retrieved March, 2011 from
<http://simlab.jrc.ec.europa.eu/>

Snowy Mountains Engineering Corporation (SMEC). 1990. Southern Okavango integrated water
development. Unpublished Report, Department of Water Affairs, Gaborone, Botswana.

Walter, C.J., C.S. Holling. 1990. Large-scale management experiments and learning by doing.
Ecology 71, pp. 2060–2068.

- Wang, L., D'Odorico, P., Ringrose, S. Coetzee, S. Macko, S.A., 2007. Biogeochemistry of Kalahari sands. *Journal of Arid Environments*, 71, 259-279.
- Water Transfer Council (WTC). 1997. Feasibility study on the Okavango River to Grootfontein link of the Eastern National Water Carrier. Unpublished Report, Department of Water Affairs, Ministry of Agriculture, Water and Rural Development. Windhoek, Namibia.
- Wolski, P., M. Murray-Hudson. 2006. Recent changes in flooding in the Xudum distributary of the Okavango Delta and Lake Ngami. Botswana, *South African Journal of Science*, 102(3-4), 173–176.
- Wolski, P., H.H.G. Savenije. 2006. Dynamics of floodplain-island groundwater flow in the Okavango Delta, Botswana, *Journal of Hydrology*, 320, 283-301.
- Wolski, P., H.H.G. Savenije, M. Murray-Hudson, T. Gumbricht. 2006. Modelling of the flooding in the Okavango Delta, Botswana, using a hybrid reservoir-GIS model, *Journal of Hydrology*, 331, 58-72.
- Wolski, P., 2009. Assessment of hydrological effects of climate change in the Okavango Basin. Okavango River Basin Transboundary Diagnostic Analysis Technical Report. EPSMO. Capetown, South Africa.

Tables

	Input	Units	Prior PDF	CV	Description
Spatially Uniform	fdet	m	U(4,6)	5	Floodplain extinction depth
	idet	m	U(16,24)	20	Island extinction depth
	fpor	%	U(0.29,0.5)	0.3	Floodplain soil porosity
	ipor	%	U(0.29,0.5)	0.3	Island soil porosity
	n	Unitless	U(20,30)	25	Area/volume coefficient (Eq. 2)
1	b ₁	Unitless	U(±20%*)	0.62	Area/volume coefficient (Eq. 2)
2	b ₂	Unitless	U(±20%*)	0.71	
3	b ₃	Unitless	U(±20%*)	0.76	
4	b ₄	Unitless	U(±20%*)	0.73	
5	b ₅	Unitless	U(±20%*)	0.76	
6	b ₆	Unitless	U(±20%*)	0.76	
7	b ₇	Unitless	U(±20%*)	0.72	
8	b ₈	Unitless	U(±20%*)	0.72	
9	b ₉	Unitless	U(±20%*)	0.73	
10	b ₁₀	Unitless	U(±20%*)	0.73	
11	b ₁₁	Unitless	U(±20%*)	0.73	
12	b ₁₂	Unitless	U(±20%*)	0.73	
13	b ₁₃	Unitless	U(±20%*)	0.71	
14	b ₁₄	Unitless	U(±20%*)	0.72	
15	b ₁₅	Unitless	U(±20%*)	0.72	
16	b ₁₆	Unitless	U(±20%*)	0.73	
1	v _{1,1a}	Mm ³	U(±20%*)	500	Volume threshold
2	v _{1,1b}	Mm ³	U(±20%*)	2000	
3	v _{2,1}	Mm ³	U(±20%*)	580	
4	v _{2,2}	Mm ³	U(±20%*)	700	
5	v _{2,3}	Mm ³	U(±20%*)	700	
6	v _{2,4}	Mm ³	U(±20%*)	400	
7	v _{2,5}	Mm ³	U(±20%*)	700	
8	v _{2,6}	Mm ³	U(±20%*)	700	
9	v _{2,7}	Mm ³	U(±20%*)	650	
10	v ₃	Mm ³	U(±20%*)	250	
11	v ₅	Mm ³	U(±20%*)	350	
12	v _{6,1}	Mm ³	U(±20%*)	450	
13	v _{6,2}	Mm ³	U(±20%*)	500	
14	v _{6,3}	Mm ³	U(±20%*)	210	
15	v _{7,1}	Mm ³	U(±20%*)	300	
16	v _{7,2}	Mm ³	U(±20%*)	300	
17	v _{7,3}	Mm ³	U(±20%*)	300	

	Input	Units	Prior PDF	CV	Description
18	v_8	Mm^3	$U(\pm 20\%^*)$	300	Volume threshold
19	v_{15}	Mm^3	$U(\pm 20\%^*)$	700	
20	v_{16}	Mm^3	$U(\pm 20\%^*)$	700	
1	$k_{1,1a}$	Months	$U(\pm 20\%^*)$	0.3	Resistance to flow (residence time)
2	$k_{1,1b}$	Months	$U(\pm 20\%^*)$	5	
3	$k_{2,1}$	Months	$U(\pm 20\%^*)$	1.5	
4	$k_{2,2}$	Months	$U(\pm 20\%^*)$	1	
5	$k_{2,3}$	Months	$U(\pm 20\%^*)$	4	
6	$k_{2,4}$	Months	$U(\pm 20\%^*)$	0.12	
7	$k_{2,5}$	Months	$U(\pm 20\%^*)$	3.8	
8	$k_{2,6}$	Months	$U(\pm 20\%^*)$	1.5	
9	$k_{2,7}$	Months	$U(\pm 20\%^*)$	1	
10	k_3	Months	$U(\pm 20\%^*)$	3	
11	k_5	Months	$U(\pm 20\%^*)$	0.1	
12	$k_{6,1}$	Months	$U(\pm 20\%^*)$	0.09	
13	$k_{6,2}$	Months	$U(\pm 20\%^*)$	0.02	
14	$k_{6,3}$	Months	$U(\pm 20\%^*)$	0.02	
15	$k_{7,1}$	Months	$U(\pm 20\%^*)$	1.7	
16	$k_{7,2}$	Months	$U(\pm 20\%^*)$	0.8	
17	$k_{7,3}$	Months	$U(\pm 20\%^*)$	1	
18	k_8	Months	$U(\pm 20\%^*)$	5	
19	k_{15}	Months	$U(\pm 20\%^*)$	0.1	
20	k_{16}	Months	$U(\pm 20\%^*)$	0.3	
1	$delay_1$	Binary	$D(0,1)$	1	Delay input for units
2	$delay_2$	Binary	$D(0,1)$	1	
3	$delay_3$	Binary	$D(0,1)$	1	
4	$delay_4$	Binary	$D(0,1)$	1	
1	m_1	%	$U(0.1,0.9)$	0.1	Rainfall ratio input
2	m_2	%	$U(0.1,0.9)$	0.5	
3	m_3	%	$U(0.1,0.9)$	0.8	
4	m_4	%	$U(0.1,0.9)$	0.4	
5	m_5	%	$U(0.1,0.9)$	0.8	
6	m_6	%	$U(0.1,0.9)$	0.8	
7	m_7	%	$U(0.1,0.9)$	0.5	
8	m_8	%	$U(0.1,0.9)$	0.4	
9	m_9	%	$U(0.1,0.9)$	0.4	
10	m_{10}	%	$U(0.1,0.9)$	0	
11	m_{11}	%	$U(0.1,0.9)$	0	
12	m_{12}	%	$U(0.1,0.9)$	0	
13	m_{13}	%	$U(0.1,0.9)$	0	
14	m_{14}	%	$U(0.1,0.9)$	0.9	
15	m_{15}	%	$U(0.1,0.9)$	0.9	

	Input	Units	Prior PDF	CV	Description
16	m ₁₆	%	U(0.1,0.9)	0	Rainfall ratio input
1	fa ₁	km ²	U(±20%*)	184	Floodplain groundwater reservoir area
2	fa ₂	km ²	U(±20%*)	100	
3	fa ₃	km ²	U(±20%*)	152	
4	fa ₄	km ²	U(±20%*)	138	
5	fa ₅	km ²	U(±20%*)	321	
6	fa ₆	km ²	U(±20%*)	348	
7	fa ₇	km ²	U(±20%*)	100	
8	fa ₈	km ²	U(±20%*)	140	
9	fa ₉	km ²	U(±20%*)	92	
10	fa ₁₀	km ²	U(±20%*)	1	
11	fa ₁₁	km ²	U(±20%*)	100	
12	fa ₁₂	km ²	U(±20%*)	100	
13	fa ₁₃	km ²	U(±20%*)	100	
14	fa ₁₄	km ²	U(±20%*)	206	
15	fa ₁₅	km ²	U(±20%*)	84	
16	fa ₁₆	km ²	U(±20%*)	300	
1	ia ₁	km ²	U(±20%*)	41.4	Island groundwater reservoir area
2	ia ₂	km ²	U(±20%*)	7.8	
3	ia ₃	km ²	U(±20%*)	18	
4	ia ₄	km ²	U(±20%*)	133.2	
5	ia ₅	km ²	U(±20%*)	176	
6	ia ₆	km ²	U(±20%*)	141	
7	ia ₇	km ²	U(±20%*)	25.2	
8	ia ₈	km ²	U(±20%*)	18	
9	ia ₉	km ²	U(±20%*)	47.4	
10	ia ₁₀	km ²	U(±20%*)	0.6	
11	ia ₁₁	km ²	U(±20%*)	6	
12	ia ₁₂	km ²	U(±20%*)	6	
13	ia ₁₃	km ²	U(±20%*)	6	
14	ia ₁₄	km ²	U(±20%*)	160.8	
15	ia ₁₅	km ²	U(±20%*)	48.6	
16	ia ₁₆	km ²	U(±20%*)	50	

Table 1. ORI model input factors. The factors that are reservoir dependent have individual PDFs for each reservoir.

U = uniform continuous distribution, D = uniform discrete distribution, CV = calibrated value

*PDF Varies between ±20% of the calibrated model

Table 2. Uncertainty analysis statistics of aerial flooding extents after MCF.

Reservoir	Mean	Median	Lower 95% CI*	Upper 95% CI*	SD*	SEM*	Skew*	Kur*
Panhandle	788.9	789.1	785.3	792.2	2.1	0.0	-0.2	-0.9
Nqoga1	492.7	492.7	489.7	495.8	1.9	0.0	0.1	-1.0
Nqoga2	488.2	499.3	178.9	693.2	157.1	1.7	0.5	-0.7
Thaoge	340.1	239.1	83.3	821.7	246.5	2.7	0.9	-0.1
Xudum	775.5	780.5	86.8	1883.0	550.5	6.1	0.6	-0.5
Boro	912.5	791.1	373.4	1621.5	414.1	4.6	0.4	-1.2
Maunacharia1	483.6	483.8	477.0	488.8	3.5	0.0	-0.3	-0.4
Maunacharia2	319.7	306.8	188.0	474.6	81.6	0.9	0.3	-0.5
Selinda	243.1	78.0	10.9	872.2	274.5	3.0	1.2	0.3
Mborga	455.4	426.4	340.4	411.9	143.1	1.6	0.6	-0.1
Khwai1	220.7	206.6	117.8	360.3	69.8	0.8	0.5	-0.4
Total Delta	5504.8	5485.0	4920.0	6171.0	357.5	3.9	0.3	-0.2

* CI = confidence interval, SD = standard deviation; SEM = standard error of the mean; Skew = skewness, kurt = kurtosis

Table 3. Width of 95% Confidence Interval for Average Flooding Extents (km²)

	Original Prior PDF	MCF PDF	% Difference
Panhandle	5.5	6.9	21.2%
Nqoga1	5.4	6.1	12.5%
Nqoga2	815.1	514.3	-45.2%
Thaoge	1863.3	738.3	-86.5%
Xudum	2210.0	1796.1	-20.7%
Boro	1428.5	1248.1	-13.5%
Maunacharia1	411.0	11.8	-188.8%
Maunacharia2	526.5	286.6	-59.0%
Selinda	1621.1	861.3	-61.2%
Mborga	884.3	71.6	-170.0%
Khwai1	384.4	242.5	-45.3%
Delta	2167.8	1251.0	-53.6%

Figure Captions

Figure 1. Okavango Basin Location Map.

Figure 2. The ORI model. Adapted from Wolski et al., 2006.

Figure 3. Diagram of the ORC reservoir model nodes and links. Adapted from Wolski et al., 2006.

Figure 4. Morris method GSA results for the inundation of the entire Delta. Volume/area coefficient for Xudum (b_5), island extinction coefficient ($idet$), floodplain porosity ($fpor$), volume threshold between Nqoga1 and: Nqoga2 (v_{23}), Thaoge (v_{24}), Xudum (v_{25}), Boro (v_{26}), Maunachira1 (v_{27}), and Selinda (v_{29}). Labels for the less important inputs are omitted for clarity.

Figure 5. FAST GSA first and grouped higher order (R) proportional results for the inundation of the entire Delta. Volume/area coefficient for Xudum (b_5), island extinction coefficient ($idet$), floodplain porosity ($fpor$), volume threshold between Nqoga1 and: Nqoga2 (v_{23}), Thaoge (v_{24}), Xudum (v_{25}), Boro (v_{26}), Maunachira1 (v_{27}), and Selinda (v_{29}).

Figure 6. FAST first order and higher order, and total order GSA results for the inundation of the entire Delta. Volume/area coefficient for Xudum (b_5), island extinction coefficient ($idet$), floodplain porosity ($fpor$), volume threshold between Nqoga1 and: Nqoga2 (v_{23}), Thaoge (v_{24}), Xudum (v_{25}), Boro (v_{26}), Maunachira1 (v_{27}), and Selinda (v_{29}). The height of the entire bar is the total sensitivity.

Figure 7. Frequency histograms for values for a) v_{27} and b) v_{23} resulting in behavioral outputs. New triangular distributions based on these results are also shown.

Figure 8. FAST GUA grouped by upstream, middle, eastern, and western reservoirs. Values are normalized by the maximum flooding extents in each reservoir for comparison. Refer to Figures 2 and 3 for the exact locations.

Figure 9. Comparison of original GUA and MCF GUA: Nqoga1, Maunachira1, and Mborga. In each graph the calibrated average inundation for each reservoir is also indicated.

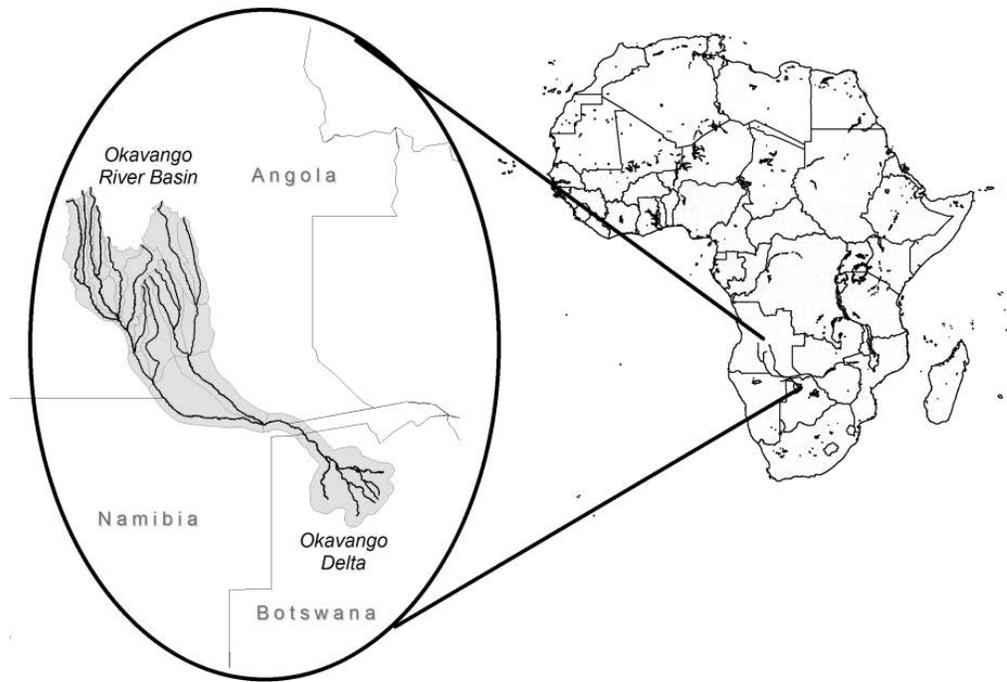


Figure 1.

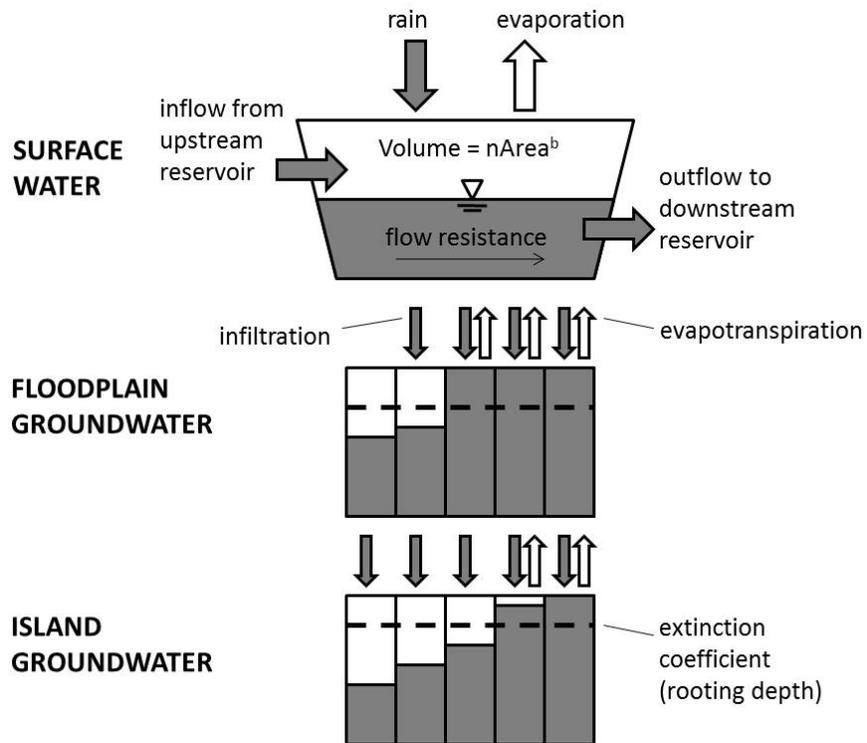


Figure 2.

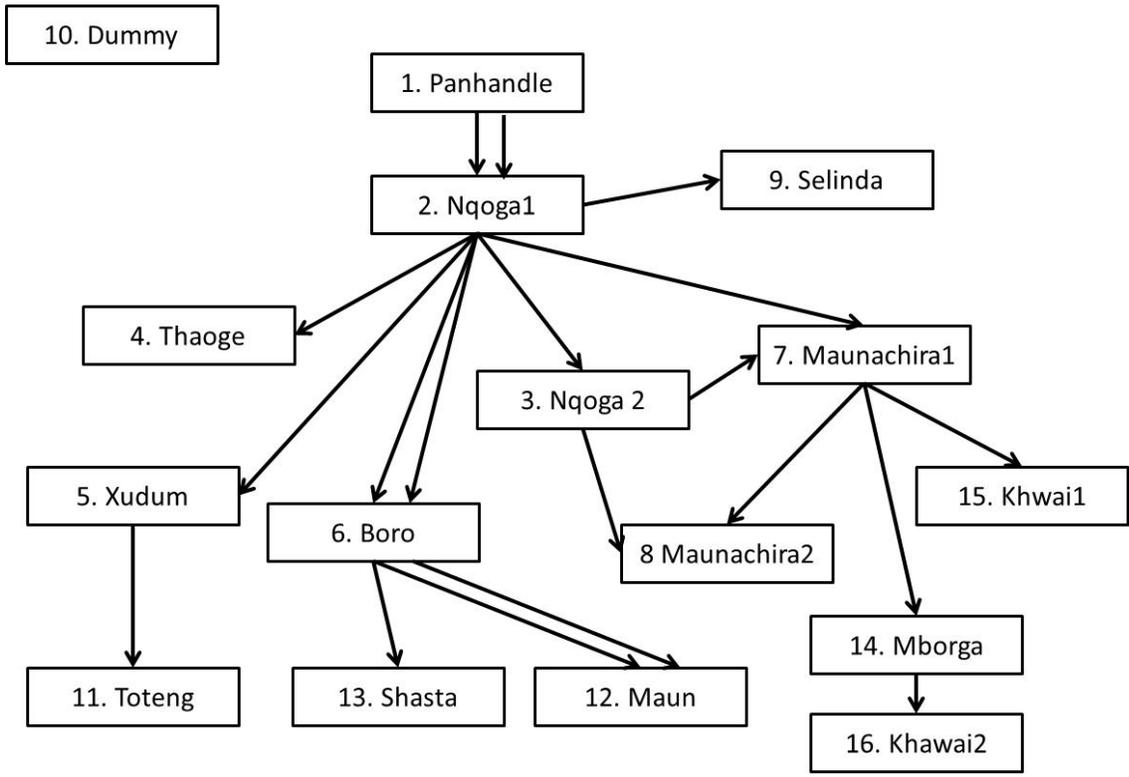
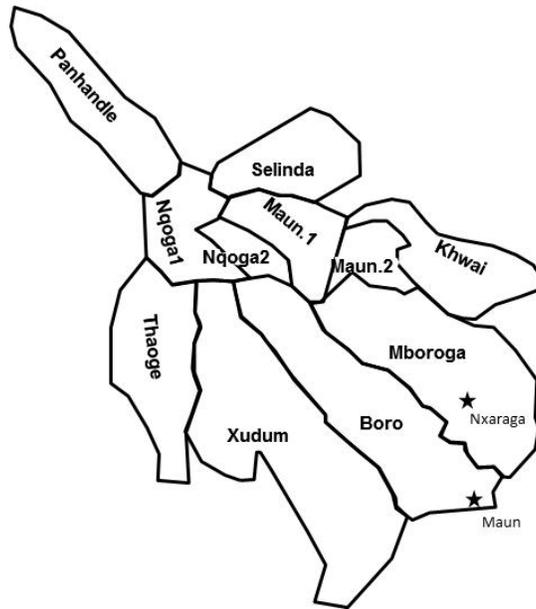


Figure 3.

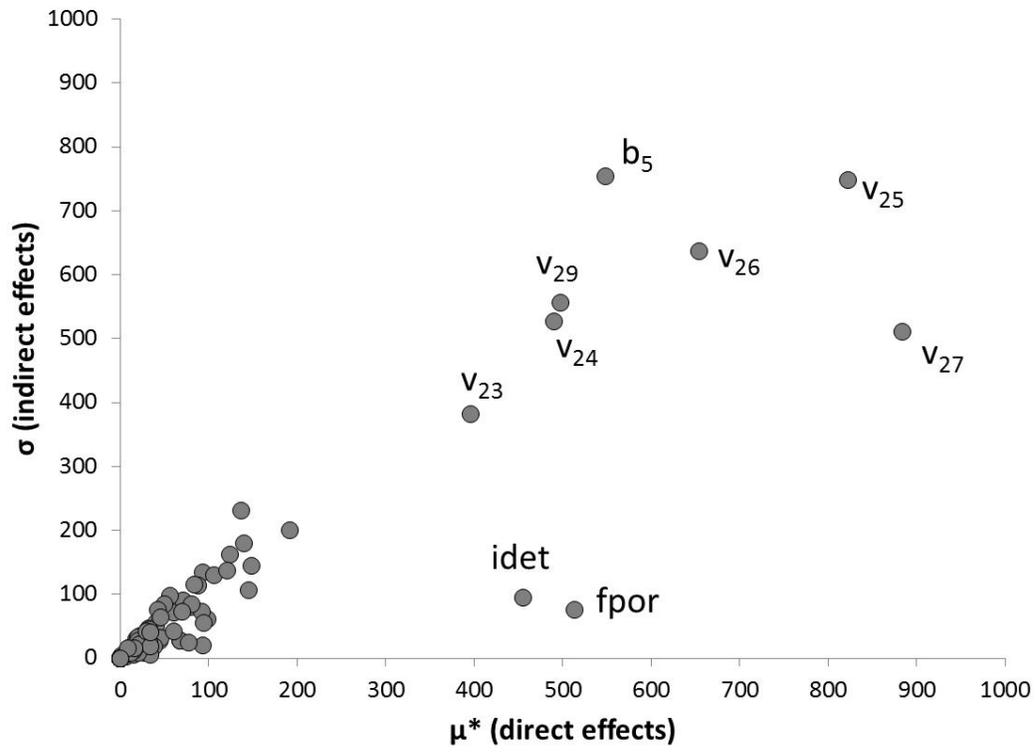


Figure 4.

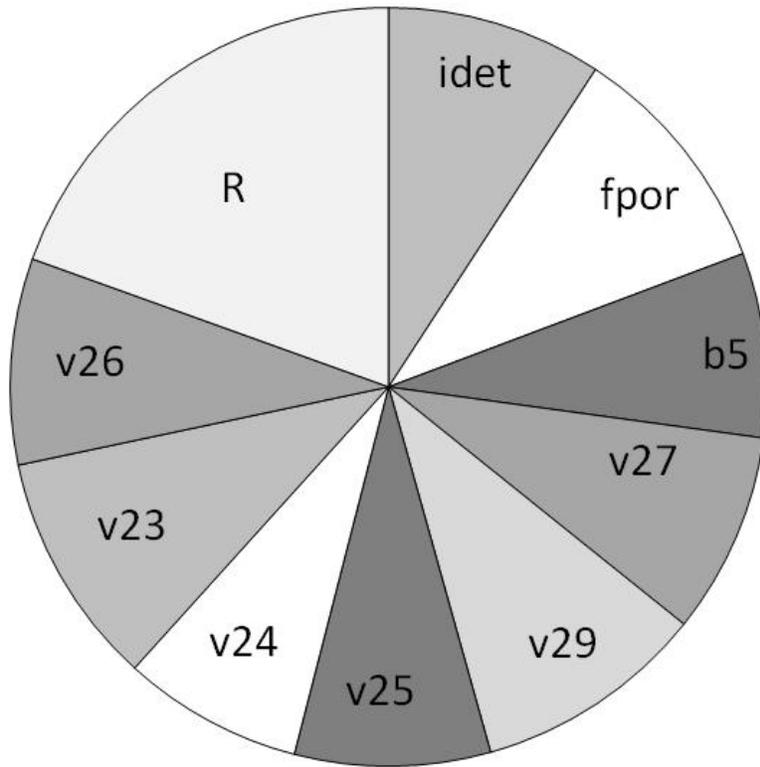


Figure 5.

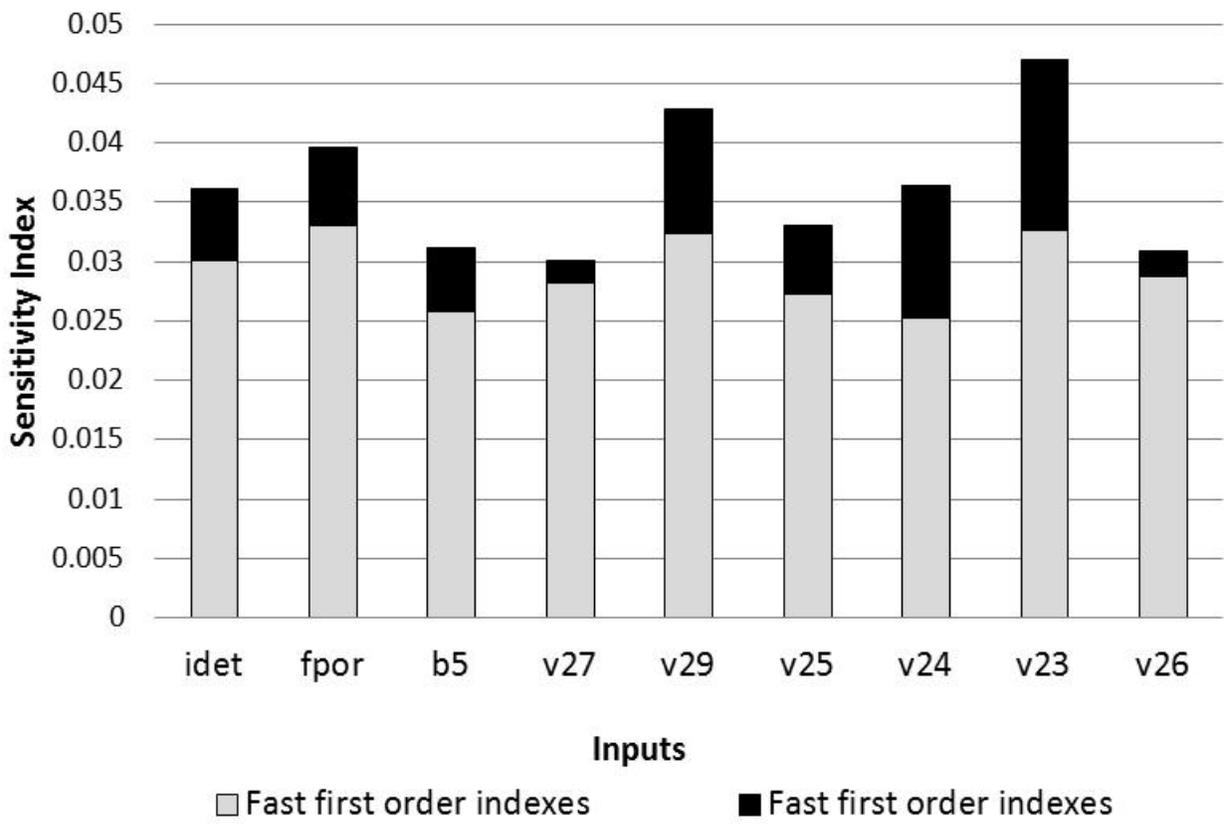
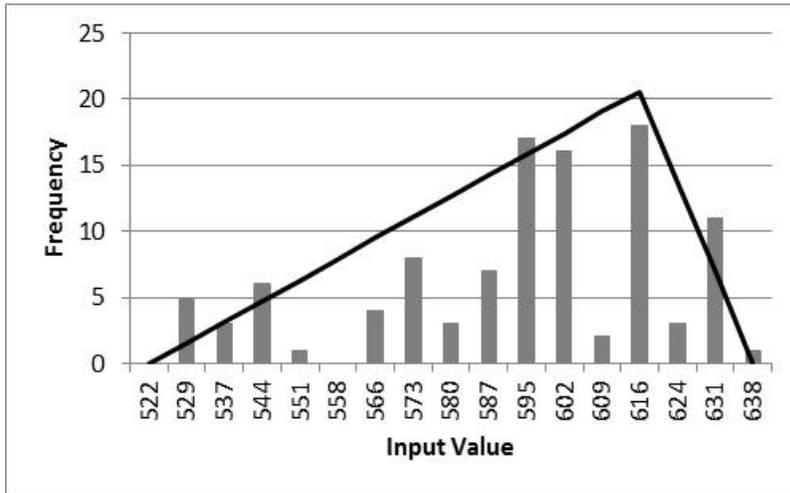


Figure 6.

a) v_{27}



b) v_{23}

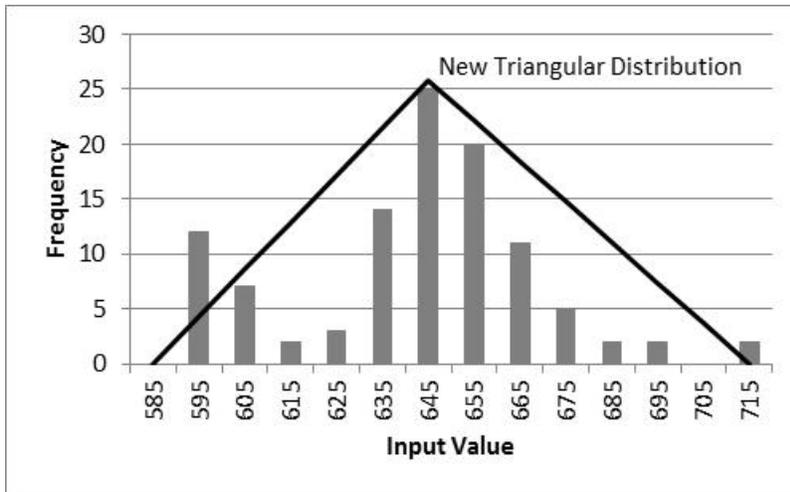


Figure 7.

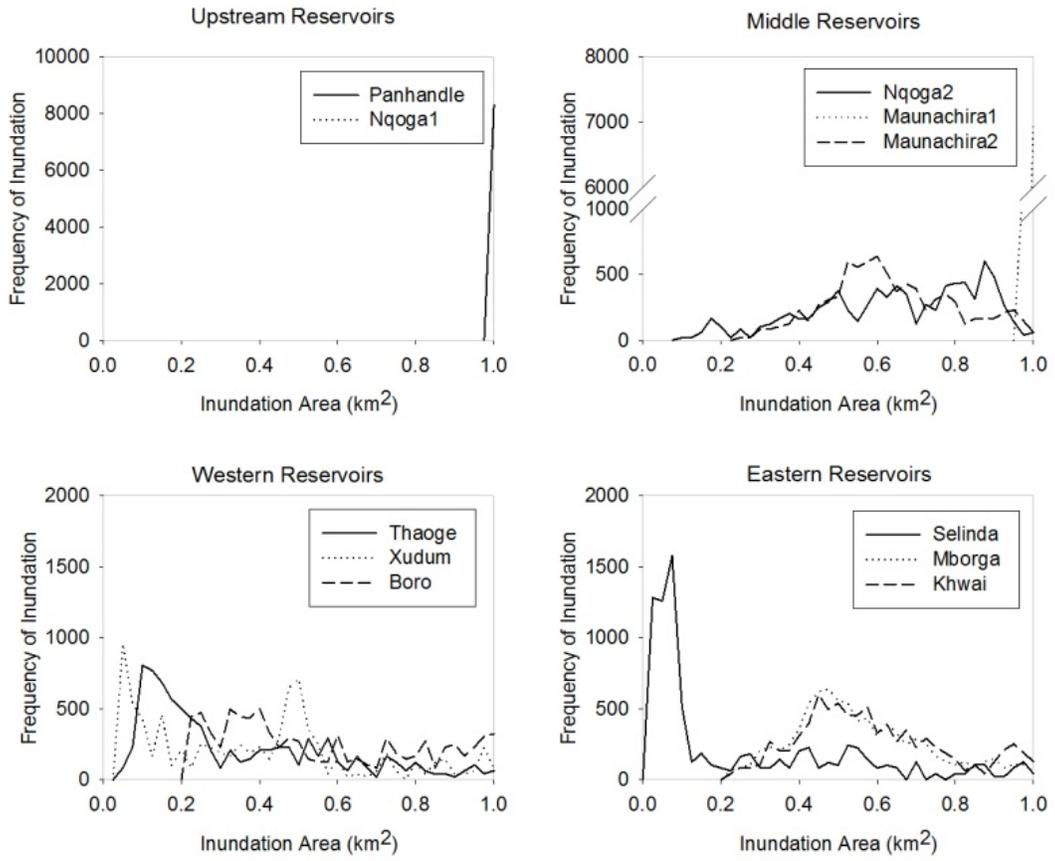


Figure 8.

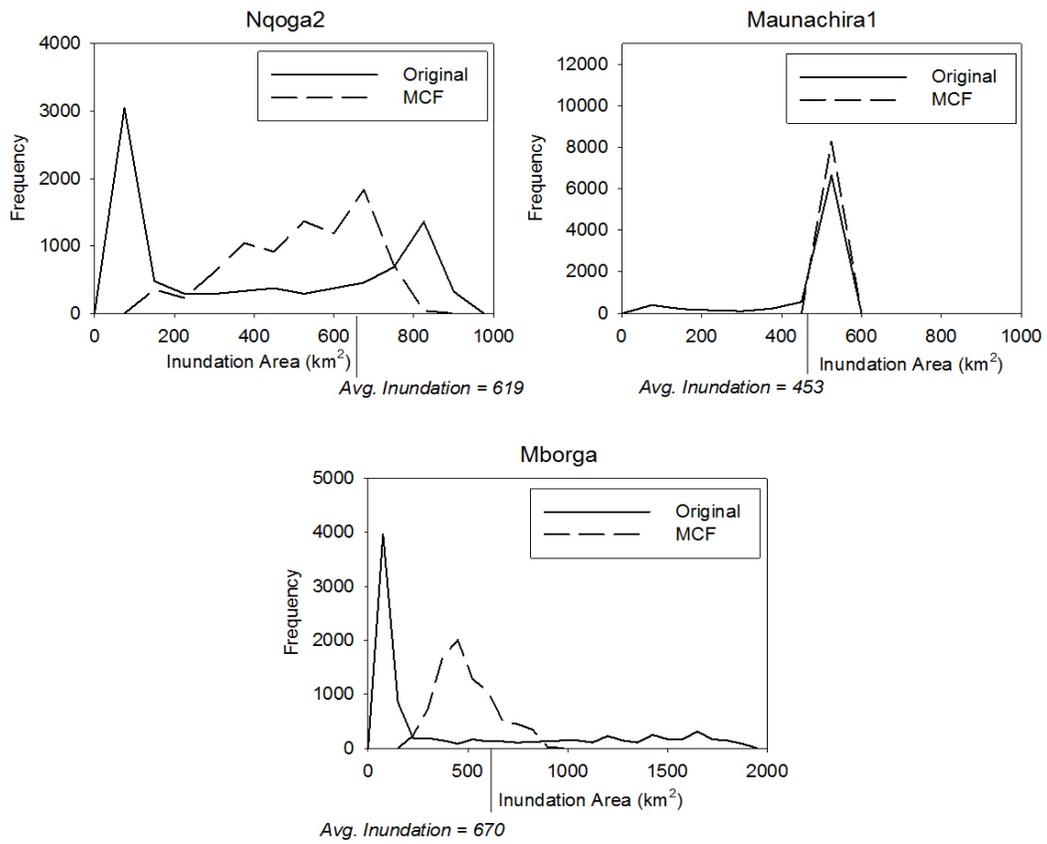


Figure 9.

Chapter 4. A fish population model for the Okavango Delta, Botswana: How the flood pulse drive population dynamics

Abstract

The hydrology of the annually flood pulsed Okavango Delta in Botswana is theorized to be a main driver for fish population dynamics. A common theory explaining this phenomenon is the Flood Pulse Concept which describes a lateral connection between the channel and floodplain and a biological response to floodplain inundation corresponding to the release of terrestrial nutrients. This work tests that theory in the Okavango Delta by developing and applying a quantitative model of fish population dynamics with the flood pulse as the major driver. Because of their commercial importance, tilapia and in particular *Oreochromis andersonii*, *Tilapia rendalli*, and *Oreochromis macrochir* are used as indicator species. The model runs on an annual time step where age cohorts are tracked over time. Density dependant recruitment, mortality, and vulnerability to catch are simulated. Model results are compared to commercial gill net data.- In this data poor environment there is a high level of uncertainty and resulting equifinality and so an optimal input set is not sought. However, clues to the flood as a driver of fish population dynamics are explored. This work uses an unbiased inverse optimization routine that searches throughout the parametric space for the best model solutions. For model evaluation, prior probability density functions (PDFs) were established for each input. Global sensitivity analysis was then used to identify the inputs that were the most important in determining the model outcome. It can be argued that the definition of prior PDFs is arbitrary. Therefore, Monte Carlo filtering was used to redefine and truncate some of the important prior PDFs. Model results give a best fit of a coefficient of efficiency of 0.64 and a coefficient of

determination (R^2) of 0.64. The most important inputs relate to mortality, catch vulnerability, and growth. There is a high level of equifinality in the inputs that set the baseline population for the fish. However, the only input that directly relates the inter-annual dynamics of fish recruitment to the flood pulse converges to an optimal value. At this optimal value modeled recruitment is almost doubled between low and high flood years. This provides evidence that the flood is indeed a major driver of fish population dynamics in the Delta.

1. Introduction

The Okavango Delta in Botswana is a large inland delta located in an arid climate that experiences an annual flood pulse from its upstream watershed (Figure 1). Flooding extents in the Delta range from 6,000 – 12,000 km² (Wolski et al., 2006) and the entire alluvial fan is approximately 40,000 km² (Gumbricht et al., 2005). The Okavango Delta is home to over 70 fish species (Kolding, 1996) and there are three main fishing industries in the Delta: subsistence, small scale commercial, and recreational (Merron 1991, Mosepele and Kolding 2003).

Mosepele (2001), states that 65% of the people living in northern Ngamiland (which encompasses the entire Delta) depend on fish as a source of livelihood. Monitoring and preserving the health of this resource is important both economically for the people who fish here and ecologically for the fish and other wildlife that depend on the annual flow of water during the dry season.

Environmental flows are criteria that describe the degree to which a natural hydrologic regime must be maintained in order to preserve valuable features of an ecosystem (Tharme and King, 1998; King et al., 1999, Tharme, 2003). Environmental Flows have been developed for the

Okavango Delta and a special report was published specifically on fish (Mosepele, 2009). A major recommendation in the report is for the development of a quantitative relationship between the flood pulse and fish population dynamics in the Delta so that managers can understand feedbacks between the hydrology of the Delta and fish population dynamics. Fish population models are often used to track populations and predict responses to management decisions in other locations (Walters and Martel, 2004; Rogers et al., 2010; Walters et al., 2008). A fish population model has yet to be developed for the Okavango Delta.

The Flood Pulse Concept (Junk et al., 1989) (FPC) describes an ecological response to flood pulsed hydrology where nutrient availability is linked to the inundation of the floodplain. In the FPC the aquatic/terrestrial transition zone (ATTZ) is the area where the aquatic environment meets the terrestrial environment. On an incoming and outgoing flood the ATTZ is a 'moving littoral.' On an incoming flood, as the water inundates the floodplain, the ATTZ has high inputs of nutrients from terrestrial sources such as vegetation and detritus. This leads to high primary productivity in the ATTZ. The FPC, the inputs of nutrients, and high primary productivity in the ATTZ iterates with each flood and it is hypothesized that fauna can adapt to take advantage of the increased food availability (Junk et al., 1989).

Research investigating the influence of the flood pulse on fish populations throughout the world has been conducted with a variety of results (Mérona and Gascuel, 1993; King et al., 2003; Gaff et al., 2004; DeAngelis et al.1997). Much of this research shows that these relationships are complex and difficult to quantify. King et al. (2003) investigated floodplain usage by fish in the Murray Darling Basin, Australia where there is annual inundation via snow

melt as well as a historical notion that the flood pulse is a major driver for fish populations. Through sampling, these authors noted that floodplain utilization by fish was not as pronounced as expected. They proposed a more complex system and suggested a model based on optimum conditions for floodplain utilization including: temperature, flood pulse predictability, the rate of change in the hydrograph, and inundation duration and area. However, the flood pulse in the Murray Darling Basin may be less predictable than in the Okavango Delta implying that the fish in the Murray Darling may be more opportunistic and less consistent in their behavior. Additionally, the hydrology of the Murray Darling has been significantly altered and it is possible that the ecological interactions are somewhat unbalanced because of these anthropogenic changes.

The Everglades also experiences an annual flood pulse during the rainy season which has been theorized to be a driver for fish populations (Gaff et al., 2004 and DeAngelis et al., 1997).

DeAngelis et al. (1997) constructed a mechanistic model, Across Trophic Level System Simulation Landscape Fish model (ALFISH), that spatially predicts fish populations based on the flood pulse. Here, a fish model is built on top of a spatially explicit hydrologic model that simulates the annual flood pulse. Periphyton, macrophytes, detritus, mesoinvertebrates, and macroinvertebrates are all modeled as well as two fish functional groups: big fish and small fish. As the flood rises, modeled fish move into the floodplain in response to increased food availability. Then, as the flood recedes, modeled fish move to find refugia and mortality increases as a result of crowding and predation. Four types of mortality are simulated: background mortality, density dependent mortality, predation by the large fish, and failure to

find refugia. The fish life cycle is simulated according to the von Bertalanffy age/length/weight relationships (von Bertalanffy, 1957). Gaff et al. (2004) critique ALFISH concluding that inundation area is not the only driver for fish populations and that there may be other parameters that are also important. They state that the best model fit that ALFISH is able to achieve is an R^2 of 0.88 for water depth and 0.35 for fish density. However, an R^2 of 0.35 between fish density and water depth reflects empirical data showing that the hydrology only accounts for 20-40% of the variability in the sampled fish population density.

Mérona and Gascuel (1993) investigated statistical relationships between commercial fish catch and the annual flood in the Amazonian floodplain. They found three interesting relationships. (1) There was a positive correlation between catch and the flood peak three years prior, which they speculated as associated with recruitment. (2) There was a correlation between catch and the water level during its rise two years prior that was possibly associated with competition. (3) There was a correlation between catch and severe low water stage two years prior that likely due to increased mortality. They were able to produce a statistical model with three variables that explained more than 83% of the variability in the annual fish abundance. Similar to the Okavango, this system experiences a regular and predictable flood pulse.

In the Okavango Delta no quantitative studies have been conducted to specifically show how fish respond to the flood pulse. However, there have been studies in the Okavango that show how the annual flood pulse produces an ecologic response in other ecological aspects (Hoberg et al., 2002; Merron, 1991). Hoberg et al. (2002) provide a food web conceptual model for ecological responses to the annual flood pulse in the Delta. They measured a 'first flush' effect

at the onset of the flood which results in a release of nutrients. During the rising flood there was a burst in nutrients, primary production, and phytoplankton. Concentrations of nitrogen rose from 1.5 to 3.5 mg l⁻¹ and phosphorus rose from 125 to 450 µg l⁻¹. Primary production reached its peak at 300 µg C l⁻¹ d⁻¹ and maximum chlorophyll *a* values were 24 µg chl*a* l⁻¹ (Hoberg et al., 2002). The authors went on to state that resting zooplankton eggs hatched when they were submerged by the floodwater and fed on the abundant phytoplankton and other food sources provided by the burst in primary production. Peak concentrations of zooplankton went from 0.1 to 10 mg DW l⁻¹ during the rise of the flood and reached up to 90 mg DW l⁻¹ at the extreme near-shore edges. A qualitative analysis of the fishes' response to the flood was also conducted (Hoberg et al., 2002). Of the tilapiine species: *O. andersonii*, *T. rendalli*, and *T. sparrmanii* were noted as following the rising flood into the study area. Juveniles were also seen with an increasing frequency just after the peak of the flood. Gut analysis of the fishes showed that smaller fish fed on more zooplankton indicating the importance of the 'first flush' effect for the juveniles. At the end of the flood season very few fry were observed with the conclusion that they migrated out of the area before the connection with the main river system was lost. Merron (1991) related spawning period to the flood pulse in the Okavango Delta. He proposed that the higher the magnitude of the annual flood, the longer the water is retained on the floodplain, leading to a longer spawning period and greater overall production of fish. Additionally, Mosepele et al. (2009) proposed that survivability for smaller fishes is increased in dense floodplain vegetation types because the vegetation provides protection from predators.

Thus, data shows that other aspects of the ecology of the Delta are driven by the flood pulse, there are qualitative indications that the fish depend on the flood pulse, other models have been built to describe fish population dynamics as affected by flood pulse, and there is a stated management need for a model that links the flood pulse to fish population dynamics. In response to these issues and needs, this work will produce a fish population of the Okavango Delta that is driven by the annual flood pulse. The model will be inverse calibrated to objectively investigate best fit model input sets. Traditionally, the goal of inverse optimization techniques is to find the optimal set of inputs for a given model to match some measured data (Mertens et al., 2006). Beven and Binley (1992) and Beven (1993) comment on the limitations of the concept of an optimal parameter set. Beven and Freer (2001) coined the term 'equifinality' to refer to the fact that there may be many input sets that all lead to an equally acceptable mode simulation. These inputs may come from different places in their defined parametric space and there may be no way to distinguish between the equally acceptable input sets. This work does not seek an optimal parameter set. Instead, it looks for evidence of the flood as a driver for fish population dynamics despite equifinality. It does this through investigating key inputs that are directly related to the inter-annual variability of fish population dynamics versus the baseline population size.

This work will use global sensitivity and uncertainty analysis (GSA/UA) in two ways. It will use GSA to identify unimportant inputs that can be set to constants without affecting the model, thus simplifying the model. One of the largest critiques of GSA/UA is the rather arbitrary methods for setting the prior probability distributions (PDFs). This work will use Monte Carlo

(MC) Filtering (Saltelli et al., 2008) to objectively redefine the prior PDFs within their predefined defined physical bounds based on realistic model results. Additionally, this method reduces model uncertainty (Saltelli et al., 2008).

The objective of this work is to produce a mechanistic fish population model for the Okavango Delta that is driven by its annual flood pulse. Because of their commercial importance, breams (tilapia) and in particular *O. andersonii*, *T. rendalli*, and *O. macrochir* are used as indicator species. Model results are compared to commercial gill net data from 1996 to 2005. Probability distribution functions (PDFs) are developed for each model input. For this work, inputs will be defined as any non-time series variable. Global Sensitivity Analysis (GSA) is used to identify inputs that are important for determining model results. Unimportant inputs are set to constants and Monte Carlo (MC) filtering is used to refine important input PDFs. This mechanistic model is used to test the conceptual model that fish population dynamics are driven by the flood pulse.

2. Methods

2.1 Fish data

Commercial catch data are available from January 1996 to December 2005. Fish were caught in gill nets and the daily catch per unit effort (CPUE) was recorded. The CPUE is defined as the number of fish caught per gill net per day. Because a relatively standard gear is used throughout this catch, CPUE can be used as a rough indicator of fish density. These data however, were not collected for experimental purposes, which can cause discrepancies. Since the fish data that is available only includes fish counts and does not contain age, weight, or

length values the data that is necessary for simulating monthly spawning dynamics was not sufficient and an annual model was deemed most appropriate for this situation. Additionally, the fishermen may set their nets where they believe the fish densities will be the highest biasing the data set. Furthermore, gill nets select for particular fish size fish because of the mesh size. Thus, fish of some sizes are more vulnerable to catch than others. Fish were also caught in the northern portion of the Delta in the Panhandle where flooding is more confined and the effect of the annual flood is less pronounced (Figure 1).

For the purposes of this model, the maximum annual CPUE is the objective function that will be used to compare measured data to modeled data. In order to compute this from the modeled data, the fish must be divided by an inundated area. This is done using in the Okavango Research Institute (ORI) hydrologic model of the Delta (Wolski et al., 2006). The ORI model is a linked reservoir model that simulates flooding extents in the Delta. One of the reservoirs in the model explicitly represents the Panhandle, where the fish data was collected. This reservoir is not the reservoir that was of focus for model calibration, nor is it a reservoir of high variability as it is an area that is more channelized and permanently inundated than the lower Delta proper. Because of these reasons, in each year the total modeled vulnerable population is divided by the average of the minimum flooding extents simulated in the Panhandle. This produces an overall better model fit than dividing by the annual minimum flooding extents.

The fish that were caught were not recorded to the level of species. Instead, breams, which are an informal category that make up a significant proportion of the commercial fishing industry, were lumped into one category. Breams include all of the major Cichlidae species. According

to Mosepele et al. (2003) Cichlidae is the principal family of fish in both subsistence and commercial gill net fisheries. The three species with the highest indexes of relative importance for the commercial gill net fishery are all tilapia (part of the Cichlidae family) and include in order of importance: (1) *O. andersonii*, (2) *T. rendalli*, and (3) *O. macrochir*. For the purposes of this work, these three fish species will be used as a representative species to base parameters such as maximum age and growth functions, with particular importance placed on *O. andersonii*.

2.2 Model structure

This work describes the development of an annual time step fish population model driven by the flood pulse in the Okavango Delta (Figure 2). Fish are divided into annual age classes which are tracked over time. The model structure uses the Beverton and Holt (1957) stock recruitment relationship (Eq. 1) to estimate recruitment, where $R_{t,n}$ is the number of recruits per year per age class, $P_{t,n}$ is the potential recruits per year per age class (a function of the stock size), and α and β are parameterized unitless von Bertalanffy parameters (1957). α/β describes the maximum recruitment at a high stock and α describes the maximum recruitment/stock at a low stock.

$$R_{t+1,n} = \frac{\alpha_t P_{t,n}}{1 + \beta_t P_{t,n}} \quad (1)$$

$P_{t,n}$ is given by Eq. 2 where $N_{t,n}$ is the number of fish per year and age class, e is fecundity (number of eggs hatched), and b is the number of broods per season.

$$P_{t,n} = N_{t,n} e b \quad (2)$$

Three constants are required to parameterize the Beverton and Holt equation: one point along the Beverton and Holt curve where recruitment is constant ($N_{t,0}$), survivability (S) (yr^{-1}), and the Goodyear Compensation Ratio (CR) (Goodyear, 1977). The steady state condition of the Beverton and Holt equation is given as Eq. 3:

$$N_{t,0} = N_{t,0}S + R_{t,0} \quad (3)$$

Thus, R_0 (constant recruitment, used for calibration) can easily be found by Eq. 4:

$$R_{t,0} = (1 - S)N_{t,0} \quad (4)$$

CR represents the maximum compensatory increase in juvenile survival as the stock size is decreased. α can be derived from CR and the steady state juvenile survival rate ($R_{t,0}/N_{t,0}$) (Walters and Martel, 2004) (Eq. 5) such that:

$$\alpha_t = (CR) \frac{R_{t,0}}{N_{t,0}} \quad (5)$$

And β can be derived from (Walters and Martel, 2004) (Eq. 6):

$$\beta_t = \frac{CR-1}{N_{t,0}} \quad (6)$$

The fish population's response to the flood is simulated through recruitment. In order to relate recruitment to the flood a modification was made to the Beverton and Holt density dependant recruitment relationship such that the number of fish is proportional to the annual maximum inflow (Eq. 7). In this equation, F_a is the maximum inflow in a given year ($\text{M m}^3 \text{ yr}^{-1}$), \bar{F} is the average of the annual maximum inflows ($\text{M m}^3 \text{ yr}^{-1}$), and c_t is a scaling coefficient ($\text{Fish yr}^{-1} \text{ age}^{-1}$)

¹flooding extents⁻¹). This does not change the actual population within the fish in the matrix, but does change the population for the calibration of α and β and calculations for recruits in each time step (Eq. 8, 9, 10, and 11). In these equations the (*) represents factors effected by the flood coefficient but which are not actually changed in the matrix.

$$N_{t,0*} = N_{t,0} + (F_a - \bar{F})c_t \quad (7)$$

$$R_{t,0} = (1 - S)N_{t,0*} \quad (9)$$

$$N_{t,0} = N_{t,0}S + R_{t,0} \quad (8)$$

$$\alpha_t = (CR) \frac{R_{t,0}}{N_{t,0*}} \quad (10)$$

$$\beta_t = \frac{CR-1}{N_{t,0*}} \quad (11)$$

The result is a change in the carrying capacity of recruits and rate of recovery between high and low flood years (Figure 3).

In each time step, after recruitment is calculated, mortality is calculated. Each age experiences mortality according to the Beverton and Holt (1959) relationship (Eq. 12). In this equation N is the number of fish, Z represents total mortality (yr^{-1}), k is a growth coefficient (unitless), t is the time step (yr), and n is the age class (yr). The exponential function results in higher mortality in younger age classes. Z is divided into two parts: natural mortality (M) and fishing mortality (F) where $Z = M + F$.

$$N_{t,n} = N_{t-1,n-1} \exp(-Zk\Delta t) \quad (12)$$

Various studies point to an allometric relationship between body weight and mortality (Peterson and Wroblewski, 1984; McGurk, 1986; Lorenzen, 1996). de Graaf et al. (2005) showed that this relationship can also be related to body length using the von Bertalanffy (1957) length/weight relationship where M_u is the mortality at unit length ($\text{yr}^{-1} \text{cm}^{-1}$), L is body length in cm, and a and β are the previously mentioned von Bertalanffy parameters (unitless) (Eq. 13).

$$M = M_u a^{-0.3} L^{-0.3\beta} \quad (13)$$

Length at a given age is calculated according to the von Bertalanffy equation, where L_n is the length at age n (yr) and k is the same growth coefficient in Eq. 10 (unitless), and L_∞ is the asymptotic length (Eq. 14).

$$L_n = L_\infty (1 - \exp^{-k\Delta n}) \quad (14)$$

Not all fish are equally vulnerable to catch because of the size selective gear. Once the modeled fish population is calculated, the vulnerability to catch of each age class is computed. A dome shaped double logistic model (Allen et al., 2009) is used where V_n is the vulnerability (unitless) at age n (yr^{-1}), TL is the average length (cm) at age n , L_{low} is the lower length (cm) at 50% vulnerability, SD_{low} is the standard deviation of the distribution for L_{low} , L_{high} is the upper length (cm) at 50% vulnerability, and SD_{high} is the standard deviation of the distribution for L_{high} (Eq. 15).

$$V_n = \frac{1}{1 + \exp^{-\frac{(TL - L_{LOW})}{SD_{LOW}}}} - \frac{1}{1 + \exp^{-\frac{(TL - L_{HIGH})}{SD_{HIGH}}}} \quad (15)$$

The measured maximum annual CPUE generally occurs at the annual low flood when the fish are most concentrated in a smaller area. In order to get a similar measure of density, the vulnerable fish were divided by an area of inundation in the Panhandle where the fish were caught. Surprisingly, dividing the fish by the ORI modeled average minimum area of inundation in the Panhandle reservoir provided a better fit than the ORI modeled annual minimum area of inundation in the Panhandle. There may be several explanations for this apparent discrepancy. The ORI model was calibrated to the larger Delta and in particular the Boro reservoir and not specifically the Panhandle. So, there may be more model error in the flooding extents in the Panhandle as it was not the focus of the model calibration. The Panhandle is also less dynamic than the larger Delta and more channelized. Therefore, dividing the fish by a constant areal value is somewhat consistent with the physical characteristics of the site. This portion of the model could easily be made dynamic to investigate the impact of inter-annual inundation, if fish data for the larger Delta was made available. For each year, the modeled annual maximum CPUE is compared to the measured annual maximum CPUE using the coefficient of efficiency (Nash and Sutcliffe, 1970) which is the objective function of the model.

2.3 Global Sensitivity and Uncertainty Analysis

Global sensitivity and uncertainty analysis (GSA/UA) is used to apportion the variation of model outputs onto the model inputs based on prior distributions. The extended FAST GSA/UA method (Cukier et al., 1978; Koda et al., 1979) uses Fourier analysis to decompose the variance of a set of model outputs into first order variances for each input. For this method, the model is executed $C \approx Nk$ times, where k is the number of inputs and N is a number that ranges

between 100's and 1,000's (Saltelli et al., 1999). The extended FAST technique (Saltelli et al., 1999) allows for the additional computation of higher levels of variance which describe the interactions between the inputs (Eq. 14). Here, $V(Y)$ describes the total variance of a single input including first and higher levels of variance.

$$\boxed{V(Y)} = \sum_i V_i + \sum_{i<j} V_{ij} + \sum_{i<j<l} V_{ijl} + \dots + V_{123\dots k} \quad (14)$$

FAST also defines S_i as an index of global sensitivity. S_i is the ratio of the variance that is ascribed to a single input divided by the total model variance. In a model where there are no interactions, the sum of the S_i 's across all of the inputs is equal to one. In models where there are interactions this sum is greater than one. Note that inputs used in this method must be independent and are assumed so for this work.

2.4 Monte Carlo filtering

Using the results of the FAST GSA/UA, MC filtering (Saltelli et al., 2008) was used to filter out the unrealistic outputs or non-behavioral results and redefine the prior distributions based on the realistic outputs. MC Filtering divides the outputs into 'behavioral' (B) and 'non-behavioral' (\bar{B}) based on a threshold that is defined by the user. The B / \bar{B} status is mapped back to the inputs and two subsets of each model input, X_i , were defined as $X_i|B$ and $X_i|\bar{B}$ based on their behavioral /non-behavioral status. The behavioral subset contains n elements and the non-behavioral subset contains \bar{n} elements such that $n + \bar{n} = N$ where N is the number of model simulations. The PDFs $f(X_i|B)$ and $f(X_i|\bar{B})$ are then assigned. The two-sided Smirnov test is performed to check the significance of the difference between the two

distributions $f(X_i|B)$ and $f(X_i|\bar{B})$. In the Smirnov test the test statistic $d_{n,\bar{n}}$ is defined by Eq. 15:

$$d_{n,\bar{n}}(X_i) = \sup \|F_n(X_i|B) - F_{\bar{n}}(X_i|\bar{B})\| \quad (15)$$

The null hypothesis for this test is $f(X_i|B) = f(X_i|\bar{B})$. The null hypothesis is rejected at a significance level, α . A small α for a particular input indicates a high importance of that input for driving the behavior of the model (Saltelli et al, 2008).

For this work, the two-sided Smirnov test was also used to narrow the range of the prior distributions for each of the important inputs as defined in the GSA. If the null hypothesis was rejected, the prior distribution was reassigned based on $f(X_i|B)$.

2.5 Model Optimization

PDFs were developed for each input based on literature values and date when available. The model was run iteratively sampling inputs from the PDFs using the extended Fourier amplitude sensitivity test (FAST) sampling routine (Cukier et al., 1978; Koda et al., 1979; Saltelli et al., 1999) with SimLab software (SimLab, 1999). This sampling routine is an unbiased method that samples throughout the parametric space and is able to highlight the variety of input sets that result in good fit model simulations. Sensitive inputs were identified and uncertainty was measured. MC filtering (Saltelli et al, 2008) was used to filter out unacceptable model simulations as defined by the coefficient of efficiency and prior distributions were refined based on the inputs that generated the acceptable simulations. Insensitive inputs were set to constants. The GSA was rerun with the new prior distributions and varying only the important inputs.

2.6 Probability density functions

Inherent to these methods is the importance of the selection of the prior PDFs. Regardless of the internal mathematics within the model, assigning a wider PDF to an input makes it more sensitive. Accordingly, much effort was devoted to assigning appropriate PDFs for each input based on best knowledge for each factor. In an area where more data is available autocorrelation tests could be done to show dependence between data. However, here no such data exists and inputs are assumed to be independent. These inputs include mortality parameters, the Goodyear compensation ratio, vulnerability parameters, age at maturity, fecundity, broods per season, inflow, and the inflow scaling coefficient (Table 2). When the data for the inputs shows no apparent distribution such as normal or triangular, the PDF can be set to uniform (Muñoz-Carpena et al., 2007). The uniform distribution allows for equal probability of selection across the defined range.

Within the model, mortality is handled in two ways (1) overall mortality for the parameterization of the Beverton and Holt Stock recruitment equation (Eq. 3) and (2) the allometric relationship between fish length and mortality for each age class (Eq. 15).

In the Beverton and Holt stock recruitment relationship (Eq. 1) mortality is given as a constant, $Z = M + F$ where Z is total mortality (yr^{-1}), F is fishing mortality (yr^{-1}), and M is natural mortality (yr^{-1}). Estimates of the total number of fishermen in the Okavango Delta in the 1990s was approximately 5,000 with 300 of those being gill net fishermen, and only 40 full-time commercial fishers (Mosepele, 2001; Kgathi et al., 2005). Because of the small scale of commercial and subsistence fishing in the Okavango Delta and the low efficiency of the gear

fishing pressure in the Delta is fairly light (Mosepele and Kolding, 2003; Kgathi et al. 2005).

Thus, for the purposes of this research F will be considered negligible.

Neglecting F , the prior distribution for M is defined based on literature values (Table 1). Values of M for tilapiine species were compiled from a number of studies including the three indicator species in the Okavango Delta. These values range between 0.67 and 1.39 for the indicator fishes in the Okavango Delta. Based on these data, the PDF for Z was set to Uniform (0.67, 1.39).

The growth coefficient k varies from 0.25 to 1.0 from a number of studies including the three indicator species in the Okavango Delta (Table 1). Based on these data, the PDF for k was set to Uniform (0.25, 1.0).

The allometric relationship between mortality and fish body length (de Graaf et al., 2005) calculates a decreasing rate of mortality with increasing body length. The four inputs required for calculating mortality according to the allometric relationship are the mortality at unit weight ($\text{y}^{-1}\text{g}^{-1}$), M_u , the von Bertalanffy parameters α and β parameters, and asymptotic length (cm), L_∞ . The original value for M_u was given as a constant 3 (Lorenzen, 1996). More recently, it has been determined that M_u should be estimated for each species (de Graaf et al., 2005; Lorenzen, 2001). de Graaff et al. (2005) used values of M_u ranging between 1 and 4.5 for species with survival rates between 50 and 80%. Based on the values of Z chosen for the PDF survivability for these fishes ranges between 25 and 50%. The PDF for M_u was set so that ranges similar to the values that were defined for Z through the literature could be achieved. Based on this analysis the PDF for M_u was set to Uniform (3, 8)

Mosepele and Nengu (2003) show values for the von Bertalanffy α and β parameters for the three indicator species specific to the Okavango Delta. For *O. andersonii* α is given as 0.004 and β is 3.424, for *O. macrochir* α is 0.014 and β is 3.106, and for *T. rendalli* α is 0.026 and β is 2.911. Based on Mosepele and Nengu's (2003) ranges for the von Bertalanffy parameters, PDFs for α and β were established: α Uniform (0.004, 0.026) and β Uniform (2.911, 3.424).

According to Mosepele and Nengu (2003) in the Okavango Delta, L_∞ for *O. andersonii* is 53cm, for *T. rendalli* it is 47cm, and for *O. macrochir* it is 40cm. From these values, the PDF for L_∞ was set to Uniform (40, 53).

The Goodyear compensation ratio (CR) (Goodyear, 1977) describes the rate at which fecundity changes following a depletion in stock (usually from exploitation). Smaller values of CR result in a higher rate of recruitment in a depleted stock. According to Walters et al. (2008), when recruitment compensation is not especially strong the CR is less than 20. Walters et al. (2007) state that long lived benthic species likely have CR's in the range of 10 to 50. In Myers et al.'s (1999) meta-analysis of a variety of fish (mostly pelagic species) values for CR ranged from 1.4 to 123.5 with an average of 18.6. A meta-analysis of stock-recruitment data by Goodwin et al. (2006) showed CR is in the range of 5 to 100. Most specifically, Goodwin et al.'s (2006) analysis showed that values for CR in perciformes varied between 3 and 50. Based on these analyses, the PDF for CR was set to Uniform (3, 50).

A dome shaped double logistic vulnerability catch curve (Allen et al., 2009) was used to model the vulnerability of fish at different lengths and corresponding age classes (Figure 4). This mimics the size selective nature of the gill nets. The catch curve was parameterized from

inspection of a smaller dataset of commercial catch from the Okavango Delta where fish length was available (Mosepele, 2009). A histogram of caught fish length was produced and a corresponding catch curve was developed with PDFs for each of the parameters. The PDFs for the lengths and standard deviation at the upper end of the curve were set to include a wider and higher distribution to account for the mortality that is occurring and also to test for a logistic vulnerability shape versus a dome shape. Based on inspection, the PDFs were set as follows: L_{low} is Uniform (23, 25), SD_{low} is Uniform (1, 3), L_{high} is Uniform (28, 60), and SD_{high} is Uniform (10, 30).

The model calculates recruitment per time step based on the number of mature fish in that year. Several studies have investigated the age at which cichlids become mature in southern Africa. Dudley (1974) measured the total length and sexual maturity of *O. andersonii* in the Kafue floodplains. He found that that during the years of his study, no fish under 26 cm were in spawning condition, three out of 64 fish from 26-29 cm were in spawning condition, and more than 30% of larger males and 40% of larger females were in spawning condition. Dudley also aged the fish with annual ring formations. He concluded that *O. andersonii* usually spawn after the age of four and very rarely under the age of three. Similarly, Van der Waal (1976) found that in the Zambezi River the minimum size for sexual maturity in *O. andersonii* was 25 – 27 cm. Hay et al. (2000) also measured the minimum size of for sexual maturity in *O. andersonii* in the Okavango River Namibia which they found to be 13 cm for males and 26 cm for females. Based on these literature values, with emphasis on the ring formation as better measure of age than length, and a PDF for age a sexual maturity was set to Normal (4.25, 0.5).

Fecundity refers to the number of eggs hatched per brood. According to Mortimer (1960) *O. andersonii*, between 17 and 25 cm in length, lay 349 to 567 eggs in ponds. Additionally, Chandrasoma and De Silva (1981) found intraovarian egg counts in *T. rendalli* ranged between 760 and 6,160 in a lake in Sri Lanka. And Marshall (1979) found that *O. macrochir* can produce 1,000 to 5,000 eggs within their ovaries and may mouthbrood up to 800 eggs in Lake Mchilwaine, Zimbabwe. According to these data and, the PDF for fecundity was set to Uniform (350, 800).

Several sources state that these indicator species may lay more than one brood per season. Skelton (1993) stated that *T. rendalli* and *O. andersonii* both raised several broods each summer. Naesje et al. (2004) described that *T. rendalli* may lay several broods each season in the Kwando River, Namibia. Mortimer (1960) examined *O. andersonii* for physiological indications of having multiple broods per season. This study did not find physiological indications of multiple broods. However, the same study also observed two instances in ponds where one breeding pair spawned twice in one season. From this literature, the PDF for broods per season in the model was set to Uniform (1, 2).

Any measured data is inherently flawed, including the gauged inflow. The ORI hydrologic model results give a correlation coefficient of 0.90 (Wolski et al., 2006). Based on this, a Uniform PDF was assigned to the annual maximum inflow as $\pm 10\%$ of the measured values.

The flood coefficient (c) is a scaling factor with the units $\text{Fish yr}^{-1} \text{age}^{-1} \text{flooding extent}^{-1}$ that describes how recruitment changes with the magnitude of the flood (Eq. 7). There is no literature to support values for this coefficient and this research is the first investigation into

the quantitative effects of the flood on fish populations in the Okavango Delta. A trial and error investigation into the appropriate ranges for this coefficient was conducted prior to the GSA to get a sense of the values that would drive the model into a behavioral fit. Values for this coefficient ranging between 5 and 25 create acceptable model outputs. Therefore, the PDF for the flood coefficient was set to Uniform (5, 25).

3. Results

With 16 inputs, the model was run 39,952 times for the GSA/UA using the FAST methodology before MC Filtering. In this set of simulations, when comparing annual measured and modeled CPUE, the model achieved a maximum coefficient of efficiency of 0.64 and an R^2 of 0.77 (Figure 5). These results are relatively promising compared to Gaff et al.'s (2004) study which achieved an R^2 of 0.35 in the Everglades where fish density was modeled in response to water depth.

The PDFs in the dome shaped double logistic curve for catch vulnerability were set wide enough so that a dome shape as well as a logistic shape could be approximated. The MC filtering showed L_{high} tending toward the upper ranges of the established PDF (42–60cm) which makes the shape of the vulnerability catch curve approach a logistic and tend away from a dome shaped curve. This should be verified using field data. Additionally, the commercial catch histogram shows that the majority of fish caught are between 27 and 32 cm long. According to Mosepele and Nengu's (2003) growth coefficient (k) for *O. andersonii*, a fish of this length would be approximately 1 year old. However, according to Dudley's (1974) growth coefficient (k) for *O. andersonii*, a fish of this length would be approximately 4 years old. This is a rather wide range in ages. Also according to Dudley, *O. andersonii* usually reach maturity after the age

of 4. This has important management implications because fish should be protected up to the age of maturity to promote recruitment. Because of the implications for recruitment and the sensitivity of k in the model, it is important to get a better indication of the value of k for the indicator species in the Delta. This would aid in setting size limitations and promoting recruitment of fishes in the Delta.

3.1 Global Sensitivity Analysis

According to the FAST GSA results (Figure 6), The most important factors in this model in order of importance, are: k (growth factor), M_u (mortality at unit length), and L_{high} (upper length at 50% vulnerability). These are also highly interactive inputs (Figure 7). Inputs that contributed less than 3% of the total model sensitivity include: maturity, fecundity, broods per season, L_{low} , SD_{low} , SD_{high} , flood coefficient, and inflow coefficient (Figure 7). These unimportant inputs were set to constants in the next round of MC Filtering. Additionally, in future modeling efforts less emphasis may be placed on these inputs because of their relative lack of importance.

The flood coefficient, which determines the relationship that flow has on recruitment, is not one of the most sensitive inputs. However, this is the only input that directly relates recruitment and population size to the flood. All of the other inputs are related to the overall size of the population. To get density, which is the objective function, the total population is divided by the area of inundation for each year. Therefore, density is going to be dynamic in each model run even if the population is static. So, even if the flood coefficient was set to zero, density would still be dynamic. However, the flood coefficient is the only input that relates to the inter-annual variability of the population size to the flood size through recruitment. A

scatter plot of the flood coefficient and the coefficient of efficiency shows that the coefficient of efficiency converges at 0.64 at a flood coefficient of 15 (Figure 8). This is where the model obtains an optimal fit. At a value of 15, the number of recruits varies between 7,300 and 14,200, approximately doubling the number between low and high flood years (Figure 9). Thus, there are many ways to get a baseline population level by manipulating all of the other inputs and a high level of equifinality in this portion of the model. But there is only one way to change the dynamics which is through the flood coefficient and this value converges to an optimal value and less equifinality in this portion of the model.

3.2 Monte Carlo filtering

All of the outputs were mapped to their corresponding inputs so that the inputs that created the best fit outputs could be better understood. A threshold of a coefficient of efficiency of 0.50 was set with any model outcome greater than or equal to 0.50 defined as behavioral and any output less than 0.50 defined as non-behavioral. 0.50 was strategically chosen as a threshold as a tool for model optimization. Using this threshold, of the 39,952 runs 329 were shown to be behavioral. This optimized the solutions that fell into the behavioral category while at the same time ensured that there were enough values in the behavioral range to employ the two sided Smirnov test. The two sided Smirnov test showed that of the 8 important inputs, 7 had distributions where the behavioral inputs were significantly different from the non-behavioral inputs: L_{high} , M_u , k , α , β , L_{∞} , and Z . The 7 significantly different distributions did not display uniform shapes but rather triangular shapes. Triangular posterior input PDFs were developed for each of the inputs that were shown to be significantly different between the behavioral and non-behavioral inputs through the two-sided Smirnov test (Figure 10).

3.3 Global Uncertainty Analysis

The model was rerun through the FAST GSA/UA with the unimportant inputs set to constants and the posterior triangular PDFs were assigned to the inputs: L_{high} , M_u , k , α , β , L_∞ , and Z . The model results showed similar optimization from before the MC Filtering with a maximum coefficient of efficiency of 0.64 and a maximum R^2 of 0.64. Through MC filtering model uncertainty was reduced from an average coefficient of efficiency of -19.4 to -5.4. The minimum coefficient of efficiency was reduced from -306.5 to -202.4. The 95% confidence interval was reduced from 0.10 to -48.4 in the original simulation to 0.59 to -23.01 in the MC Filtered simulation (Figure 11 b and Figure 12).

4. Conclusions

The objective of this work was to test the conceptual model that fish population dynamics in the Okavango Delta is driven by the annual flood pulse through a mathematical simulation. It was found that there is one input in the model (the flood coefficient) that directly drives the inter-annual variability in fish population. The inter-annual variability in the real population is relatively small. However, simulating this inter-annual variability is the foundation of this work (Figure 8). The flood coefficient's sensitivity is low and may be regarded as a relatively 'unimportant' input. However, this does not mean that the effect of the flood and the flood coefficient is not important for producing the best fit model simulations. The most important inputs in the model are all related to the average or baseline vulnerable population size, whereas the flood coefficient is the only input that creates variability in the population between years. All of the other inputs (besides k) exhibit major issues of equifinality. They do not converge and are able to vary throughout their ranges while still achieving best fit results.

Thus, the average population size can be modeled in various ways. However, the value for the flood coefficient converges toward a single solution. Thus, the inter-annual variability can only be modeled in one way. In order to achieve a model fit that approximates the average population (coefficient of efficiency of 0.5), you only need to get the baseline population correct. However, in order to get the best model fit, and reach a coefficient of efficiency above 0.5, both the baseline population and the inter-annual variability must be modeled correctly. The fact that the model is able to simulate a coefficient of efficiency of 0.64 provides evidence that the flood is an important driver in the system despite the fact that there is equifinality in portions of the model.

To reduce uncertainty, MC filtering was applied to the initial inverse optimization routine. The important prior distributions were re-set to triangular distributions where the behavioral inputs were found to be significantly different from the non-behavioral inputs based on the two sided Smirnov test. L_{high} , M_u , k , α , β , L_{∞} , and Z were all found to have significantly different behavioral inputs and were assigned triangular distributions. This method reduced model uncertainty by reducing the 95% confidence interval from a width of 48.5 to 23.6. This is useful and allows the modeler in a data poor environment to focus in on likely values within the ranges of physically acceptable PDFs. The modeler can then work in tandem with the biologist to ensure that the new PDFs makes sense in a real biological setting providing converging lines of evidence for a more accurate depiction of the system and its interactions.

Several limitations can be identified in this work. The fish data was from commercial catch and not experimental data and so fisherman preferences, knowledge, and other human variables

may play into the data. The data contained only counts of tilapia fish and was not accompanied with age, length, or weight and so seasonal interpretations of responses to flooding were impractical. The data came from the Panhandle which is more permanently flooded and is less dynamic than the larger Delta. Additionally, the ORI hydrologic model which simulates the inundation area was not specifically calibrated to the Panhandle reservoir and was more focused on the larger Delta. Finally, the variability of the annual fish population is not exceptionally dynamic. The maximum annual CPUE only fluctuates between 26 and 39. In heavily fished areas, fish populations are often much more dynamic, lending to more variability to model. In spite of these limitations, this model was still able to simulate fish population dynamics to within a coefficient of efficiency of 0.64. In addition to this the flood coefficient, which is the only input that affect the inter-annual variability in the population size, converges toward a single value. At a value of 15, this flood coefficient is responsible for approximately doubling recruitment between low and high flood years. Thus, this mechanistic model corroborates the conceptual model (Merron, 1991; Mosepele et al., 2009) and qualitative observations (Hoberg et al. (2002) made about fish response to the flood pulse in the Okavango Delta.

Acknowledgments

This work was funded in part by the National Science Foundation's AM-W3 IGERT at the University of Florida. The authors would like to thanks to the researchers and staff at the ORI for their assistance with data, on-site logistics, and generous hospitality. Special thanks to Piotr Wolski and Mike Murray Hudson from the University of Botswana for their guidance and Mark Brown for his support.

References

- Allen, M.S., Brown, P., Douglas, J., Fulton, W., Catalano, M., 2009. An assessment of recreational fishery harvest policies for Murray cod in southeast Australia. *Fisheries Research*, 95, 260–267.
- Beven K., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, 16, 41– 51.
- Beven, K.J., Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, 6, p.279–298.
- Beven, K.J. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems. *Journal of Hydrology*, 249, 11–29.
- Beverton, R.J.H., Holt, S.J., 1957. On the Dynamics of Exploited Fish Populations. *Fishery Investigations, Series II Volume XIX*, Ministry of Agriculture, Fisheries and Food.
- Beverton, R.J.H., Holt, S.J., 1959. A review of the lifespans and mortality rates of fish in nature and the relation to growth and other physiological characteristics. *Ciba Foundation Symposium - The Lifespan of Animals (Colloquia on Ageing)*, 5: 142–177.
- Booth, A.J., Merron, G.S., 1996. The age and growth of the greenhead tilapia *Oreochromis macrochir* (Pisces: *Cichlidae*) from the Okavango Delta, Botswana. *Hydrobiology*, (321), 29-34.

Booth, A.J., Merron, G.S., Buxton, C.D., 1995. The growth of *Oreochromis andersonii* (Pisces: *Cichlidae*) from the Okavango Delta, Botswana, and a comparison of the scale and otolith methods of ageing. *Environmental Biology of Fishes*, 43(2), 171-178.

Chandrasoma, J., De Silva, S.S., 1981. Reproductive biology of *Puntius sarana*, an indigenous species, and *Tilapia rendali (melanopura)*, an exotic in an ancient man-made lake in Sri Lanka. *Fisheries Management*, 12, 17-28.

Crerar, S., 1997. Hydrology of the Okavango River system upstream of the Okavango Delta. Appendix A – Initial Environment Evaluation. In: Feasibility study on the Okavango River to Grootfontein Link of the Eastern National Water Carrier, vol. 4, part 3: Appendices, 31 pp.

Cukier, R.I., Levine, H.B. and Schuler, K.E., 1978. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26, 1–42.

de Graaf, G.J., Dekker, P.J., Huisman, B., Verreth J.A.J., 2005. Simulation of Nile tilapia culture in ponds, through individual-based modeling, using a population dynamic approach. *Aquaculture Research*, 36, 355–471.

DeAngelis, D.L., Loftus, W.F., Trexler, J.C., Ulanowicz, R.E., 1997. Modeling fish dynamics and effects of stress in a hydrologically pulsed ecosystem *Journal of Aquatic Ecosystem Stress and Recovery*, 6, 1–13.

Dudley, R. G. 1974. Growth of Tilapia of the Kafue floodplain, Zambezi: predicted effects of the Kafue Gorge Dam. *Transactions of the American Fisheries Society* 103:281–291.

- Gaff, H., Chick, J., Trexler, J., DeAngelis, D., Gross, L., Salinas, R., 2004. Evaluation of and insights from ALFISH: a spatial-explicit, landscape-level simulation of fish populations in the Everglades. *Hydrobiologia*, 520, 73–87.
- Garrod D.J., 1963. An estimation of the mortality rates in a population of *Tilapia esculenta* Graham (Pisces, *Cichlidae*) in Lake Victoria, East Africa *ournal of the Fisheries Research Board of Canada*, 20, 195–227.
- Getabu, A., 1992. Growth parameters and total mortality in *Oreochromis niloticus* (L) from Nyanza Gulf, Lake Victoria. *Hydrobiologia*, 232, 91–97.
- Goodwin N.B., Grant A., Perry A.L., Dulvy N.K., Reynolds J.D., 2006. Life history correlates of density-dependent recruitment in marine fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, 63, 494–509.
- Goodyear, C.P., 1977. Assessing the impact of power plant mortality on the compensatory reserve of fish populations. In: *Proceedings of the Conference on Assessing the Effects of Power Plant induced Mortality on Fish Populations* (ed. W. Van Winkle). Permagon Press, New York, pp. 186–195.
- Gumbricht, T., McCarthy, T.S., Bauer, P., 2005. The micro-topography of the wetlands of the Okavango Delta, Botswana. *Earth Surface Processes and Landforms*, 30(1), 27–39.
- Hay, C.J., Næsje, T.F., Breistein, J. Hårsaker, K., Kolding, J., Sandlund, O.T. van Zyl, B., 2000. Fish populations, gill net selectivity, and artisanal fisheries in the Okavango River, Namibia. Recommendations for a sustainable fishery. NINA.NIKU Project Report No. 010. 105p.

- Hoberg, P., Lindholm, M., Ramberg, L., Hessen., D.O., 2002. Aquatic food web dynamics on a floodplain in the Okavango delta, Botswana. *Hydrobiologia*, 470, 23–30.
- Junk, W.J., Bayley, P.B., Sparks, R.E., 1989. The flood pulse concept in river-floodplain systems. *Canadian Special Publications of Fisheries and Aquatic Sciences*, 106, 110–27.
- Kapangaziwiri, E., 2008. Revised parameter estimation methods for the Pitman monthly rainfall-runoff model. Master's thesis, Rhodes University, South Africa.
- Kapangaziwiri, E., Hughes, D., 2008. Towards revised physically-based parameter estimation methods for the Pitman monthly rainfall-runoff model. *Water South Africa*, 32(2), 183-191.
- Kgathi, D.L., Mmopelwa, G., Mosepele, K., 2005. Natural resource assessment in the Okavango Delta: case studies of some key resources. *Natural Resources Forum*, 29(1), 70–81.
- King, J.M., Tharme, R.E., Brown, C.A., 1999. Definition and Implementation of Instream Flows. Thematic Report for the World Commission on Dams. Southern Waters Ecological Research and Consulting, Cape Town, South Africa.
- King, A.J., Humphries, P., Lake, P.S., 2003. Fish recruitment on floodplains: the roles of patterns of flooding and life history characteristics. *Canadian Journal of Fisheries and Aquatic Sciences*, 60, 773–786.
- Koda, M., Mcrae, G.J., Seinfeld, J.H., 1979. Automatic sensitivity analysis of kinetic mechanisms, *International Journal of Chemical Kinetics*, 11, 427-444.

Kolding, J., 1996. Feasibility study and appraisal of fish stock management plan in Okavango.
University of Bergen, Bergen, Norway.

Lorenzen, K. (2001) Using population models to assess culture-based fisheries: a brief review with an application to the analysis of stocking experiments. In: Reservoir and Culture-Based Fisheries: Biology and Management. ACIAR Proceedings No. 98, Canberra, Australia (ed. by S.S.De Silva), pp. 257–265.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in fish: a comparison of natural ecosystems and aquaculture *Journal of Fish Biology*, 49, 627-647.

Marshall, B.E., 1979. Observations on the breeding biology of *Sarotherodon machochir* (Boulenger) in Lake Mcllwaine, Rhodesia. *Journal of Fish Biology*, 14, 419-424.

McGurk, M.D., 1986. Natural mortality of marine pelagic fish eggs and larvae: role of spatial patchiness. *Marine Ecology Progress Series*, 34, 227-242.

Mérona B., Gascuel D., 1993. The effects of flood regime and fishing effort in the overall abundance of an exploited fish community in the Amazon floodplain. *Aquatic Living Resources*, 6, 97–108.

Merron, G.S., 1991. The ecology and management of the fishes of the Okavango Delta, Botswana, with particular reference to the role of the seasonal flood. PhD dissertation. Rhodes University, Grahamstown, South Africa.

Mertens, J., Stenger, R., Barkle, G.F., 2006. Multiobjective inverse modeling for soil parameter estimation and model verification. *Vadose Zone Journal*, 5,917–933.

Mortimer, M.A.E., 1960. Observations on the biology of *Tilapia andersonii* (Castelnau), (*Pisees, Cichlidae*), in Northern Rhodesia. Joint Fisheries Research Organization Annual Report (Lusaka), 9, 41-67.

Mosepele K., 2009. Okavango River Basin Technical Diagnostic Analysis: Environmental Flow Module Specialist Report, Country: Botswana, Discipline: Fish. OKACOM 2010 Project Management Unit, Luanda, Angola. Sponsored by GEF.

Mosepele, K., 2001. Preliminary description of the Okavango Delta fishery. Unpublished technical report. Fisheries Section, Ministry of Agriculture, Gaborone, Botswana.

Mosepele, K., Nengu, S., 2003. Growth, Mortality, Maturity and Length-Weight Parameters of Selected Fishes of the Okavango Delta, Botswana. Published in ACP-EU Fisheries Research Report No. 14 Fish Biodiversity: Local Studies as Basis for Global Inferences.

Mosepele, K., Kolding, J., 2003. Fish stock assessment in the Okavango Delta: preliminary results from a length based analysis. In: Bernard, T., Mosepele, K., Ramberg, L., (Eds.), Proceedings of the Environmental Monitoring of Tropical and Subtropical Wetlands Conference, 4–8 December, 2002, Harry Oppenheimer Okavango Research Centre, University of Botswana, Maun, Botswana, pp. 363–390.

Mosepele, K., Mmopelwa, T.G., Mosepele, B., 2003. Characterization and monitoring of the Okavango Delta artisanal fishery. In: Bernard, T., Mosepele, K., Ramberg, L. (Eds.),

Environmental Monitoring of Tropical and Subtropical Wetlands. University of Botswana, Maun, Botswana and University of Florida, Gainesville, FLA.

Mosepele, K., Moyle, P.B., Glen, M.S., Purkey, D.R, Mosepele, B., 2009. Fish, floods and ecosystem engineers: Aquatic conservation in the Okavango Delta, Botswana. *Bioscience*, 59 (1), 53 – 64.

Muñoz-Carpena, RZ. Zajac, Yi-Ming Kuo, 2007. Evaluation of water quality models through global sensitivity and uncertainty analysis techniques: Applications to vegetative strip filter model VFSMOD-W. *Transactions of the ASABE* 50(5):1719-1732.

Myers, R. A., K. G. Bowen, N. J. Barrowman., 1999. Maximum reproductive rate of fish at low population sizes. *Canadian Journal of Fisheries and Aquatic Sciences*, 56, 2404–2419.

Naesje, T.F., Hay, C.J., Nickanor, N., Koekemoer, J.H., Strand, R., Throstand, E.B., 2004. Fish populations, gill net catches, and gill net selectivity in the Kwando River, Namibia. NINA Project Report 27: 1-64.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3), 282–290.

Peterson, I., Wroblewski, J.S., 1984. Mortality rate of fishes in the pelagic ecosystem. *Canadian Journal of Fisheries and Aquatic Sciences*, 41, 1117-1120.

- Rogers, M.W., Allen, M.S., Brown, P., Hunt, T.L., Fulton, W., Ingram B.A., 2010. A simulation model to explore the relative value of stock enhancement versus harvest regulations for fishery sustainability. *Ecological Modelling*, 221, 919–926.
- Saltelli A., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., Tarantola S., 2008. *Global Sensitivity Analysis: The Primer*. Wiley and Sons, Sussex, England.
- Saltelli, A., Tarantola, S. Chan, K.P., 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, Vol. 41, No. 1, pp.39–56.
- SimLab. 2009. Simlab sensitivity analysis software. Retrieved November 2009 from simlab.jrc.ec.europa.eu/
- Skelton, P., 1993. *A Complete Guide to the Freshwater Fishes of Southern Africa*. , Southern Book Publishers, Halfway House, South Africa.
- Tharme, R.E., King, J.M.. 1998. Development of the Building Block Methodology for Instream Flow Assessments, and Supporting Research on the Effects of Different Magnitude Flows on Riverine Ecosystems. Water Research Commission Report No. 576/1/98.
- Tharme, R.E., 2003. A global perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. *River Research and Applications*, 19,397–442.
- Tweedle, D., Turner, J.L., 1977. Age, growth and natural mortality rates of some cichlid fishes of Lake Malawi. *Journal of Fisheries Biology*, 10, 385-98.

- Van der Waal, B.C.W. 1976. 'n Visekologiese studie van die Liamnezimeer in die Oos- Caprivi met verwysing na visontginning deur die Bantoebevolking. - Ph.D. dissertation, Rand Afrikaans University, South Africa. 192 pp. (In Afrikaans).
- von Bertalanffy, L., 1957. Quantitative laws in metabolism and growth. *The Quarterly Review of Biology*, 32(3), 217–231.
- Walters C. J., Hilborn R., Christensen V., 2008. Surplus production dynamics in declining and recovering fish populations. *Canadian Journal of Fisheries and Aquatic Sciences*, 65, 2536-2551.
- Walters C. J., Hilborn R., Parrish R., 2007. An equilibrium model for predicting the efficacy of marine protected areas in coastal environments. *Canadian Journal of Fisheries and Aquatic Sciences*, 64, 1009-1018.
- Walters, C. J., 1986. Adaptive management of renewable resources. MacMillan, New York, New York, USA.
- Walters, C., Holling, C.S., 1990. Large-scale management experiments and learning by doing. *Ecology*, 71, 2060-2068.
- Walters, C.J., Martell, S.J.D., 2004. Fisheries ecology and management. Princeton University Press, Princeton, N.J.

Wilk, J., Kniveton, D., Andersson, L., Layberry, R., Todd, M. C., Hughes, D., et al. 2006. Estimating rainfall and water balance over the Okavango River Basin for hydrological applications. *Journal of Hydrology*, 331(1-2), 18-29.

Wolski, P., 2009. Assessment of hydrological effects of climate change in the Okavango Basin. Okavango River Basin Transboundary Diagnostic Analysis Technical Report. EPSMO. Capetown, South Africa.

Wolski, P., Savenije, H.H.G., Murray-Hudson, M., Gumbricht, T., 2006. Modelling of the flooding in the Okavango Delta, Botswana, using a hybrid reservoir-GIS model. *Journal of Hydrology*, 331, 58-72.

Tables

Table 1. Natural mortality (M) and growth coefficients (k) for selected tilapiine species.

Species	M(yr ⁻¹)	K(yr ⁻¹)	Location	Reference
<i>Oreochromis andersonii</i>	1.39	1.0	Okavango Delta, Botswana	Mosepele and Nengu (2003)
<i>Oreochromis andersonii</i>	0.67	0.25	Okavango Delta, Botswana	Booth et al. (1995)
<i>Oreochromis macrochir</i>	1.5	1.0	Okavango Delta, Botswana	Mosepele and Nengu (2003)
<i>Oreochromis macrochir</i>	0.95	0.42	Okavango Delta, Botswana	Booth et al. (1996)
<i>Tilapia rendalli</i>	1.22	0.78	Okavango Delta, Botswana	Mosepele and Nengu (2003)
<i>Serranochromis angusticeps</i>	1.46	1.0	Okavango Delta, Botswana	Mosepele and Nengu (2003)
<i>Serranochromis robustus</i>	1.21	0.83	Okavango Delta, Botswana	Mosepele and Nengu (2003)
<i>Oreochromis niloticus</i>		0.254	Lake Victoria, Kenya	Getabu (1992)
<i>Haplochromis anaphyrmus</i>	1.45	0.671	Lake Malawi, Mozambique	Tweedle and Turner (1977)
<i>Haplochromis mloto</i>	0.92	0.55	Lake Malawi, Mozambique	Tweedle and Turner (1977)
<i>Lethrinops longipinnus</i>	1.69	0.571	Lake Malawi, Mozambique	Tweedle and Turner (1977)
<i>Lethrinops parvidens</i>	1.20	0.487	Lake Malawi, Mozambique	Tweedle and Turner (1977)
<i>Tilapia esculenta</i>	1.75	0.28	Lake Victoria, Kenya	Garrod (1963)

Table 2. List of inputs and initial distributions.

	Input	Units	Distribution
1	Z	yr ⁻¹	U(0.67, 1.39)
2	K	unitless	U(0.25, 1.0)
3	M _u	yr ⁻¹ cm ⁻¹	U(3, 8)
4	α	unitless	U(0.004, 0.026)
5	β	unitless	U(2.911, 3.424)
6	L _∞	cm ⁻¹	U(40, 53)
7	CR	unitless	U(3, 50)
8	L _{low}	cm ⁻¹	U(1, 3)
9	SD _{low}	cm ⁻¹	U(23, 25)
10	L _{high}	cm ⁻¹	U(28,60)
11	SD _{high}	cm ⁻¹	U(10, 30)
12	Maturity	yr	N(4.25, 0.5)
13	Fecundity	eggs	U(350, 800)
14	Broods	yr ⁻¹	U(1, 2)
15	Inflow	M m ³ yr ⁻¹	U(-10%, +10%)
16	Flood coefficient	unitless	U(5, 25)

Figure Captions

Figure 1. Site location. The Okavango Delta, with fish sampling sites marked with stars.

Figure 2. The conceptual model.

Figure 3. A demonstration of the response in recruitment as a result of the change in annual maximum flood. The rate of recovery and the carrying capacity increases during smaller floods. $Flood_1 > Flood_2 > Flood_3$.

Figure 4. An example of the dome shaped double logistic curve (Allen et al., 2009) for catch vulnerability. $L_{low} = 2$, $SD_{low} = 2$, $L_{high} = 45$, $SD_{high} = 10$. Histogram data from Mosepele, 2009. (a) shows the dome shaped curve and (b) shows the curve approaching a logistic shape. Both shapes are possible within the bounds of the defined PDFs.

Figure 5. The best fit model simulation for the initial inverse optimization. Coefficient of efficiency of 0.64.

Figure 6. First order sensitivity for the coefficient of efficiency of the modeled fish density compared to the measured CPUE from the first inverse optimization.

Figure 7. First and higher order sensitivities for coefficient of efficiency of modeled fish density compared to measured CPUE from first inverse optimization.

Figure 8. Ranges of values for the flood coefficient (c) creating behavioral outputs. Note that c converges toward a single value in the highest values for the coefficient of efficiency.

Figure 9. The effect of the flood and the flood coefficient on recruitment. (c = 15)

Figure 10. Important inputs whose behavioral distributions are significantly different from the non-behavioral distributions. These graphs show histograms of the behavioral distributions and their newly defined triangular prior distributions.

Figure 11. All GUA results. (a) unfiltered (b) MC filtered. Because of the scale on the x axis (-400 – 1 with 50 unit increments labeled), coefficient of efficiencies greater than 1 are undetectable. Therefore, Figure 12 shows a blowup of results with coefficient of efficiencies greater than 0.

Figure 12. Uniform and MC Filtered GUA outputs with a coefficient of efficiency greater than 0.

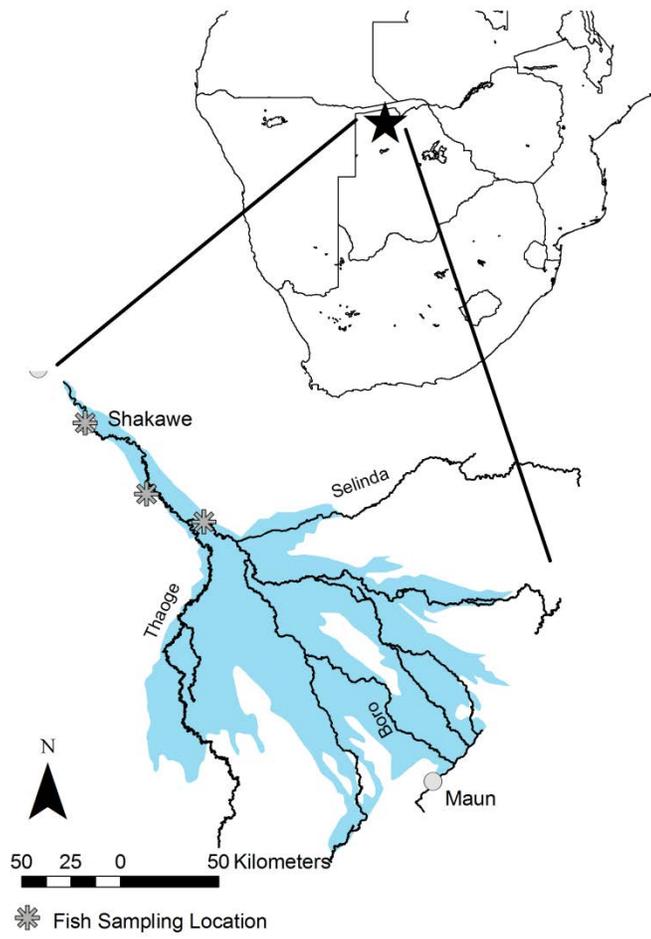


Figure 1.

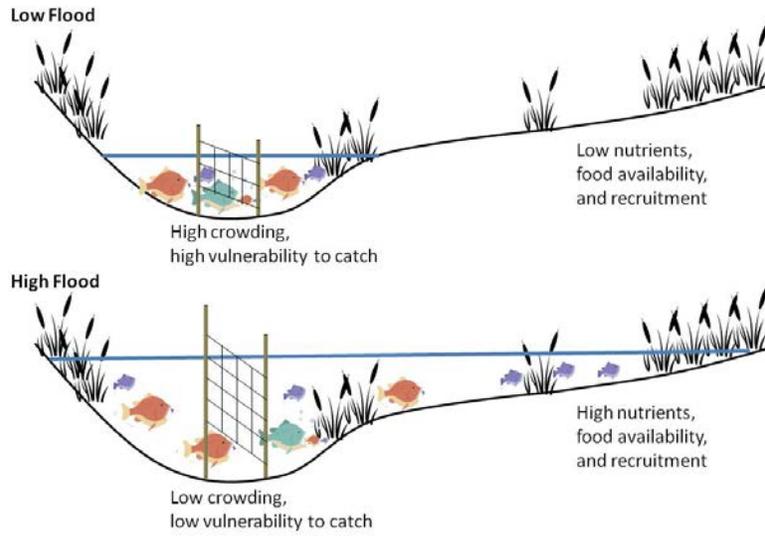


Figure 2.

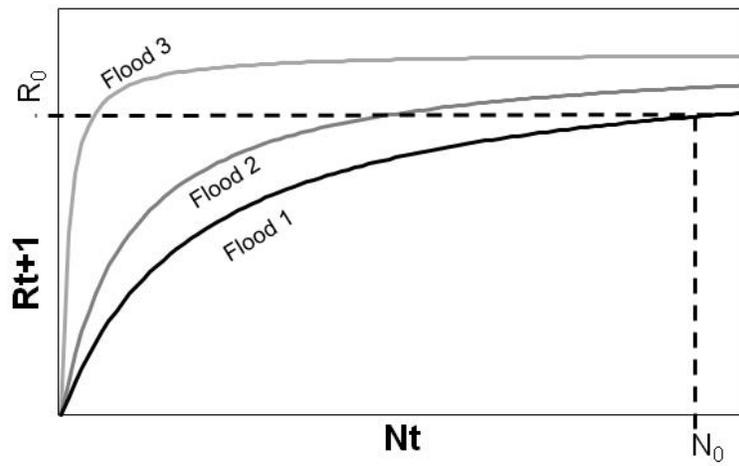


Figure 3.

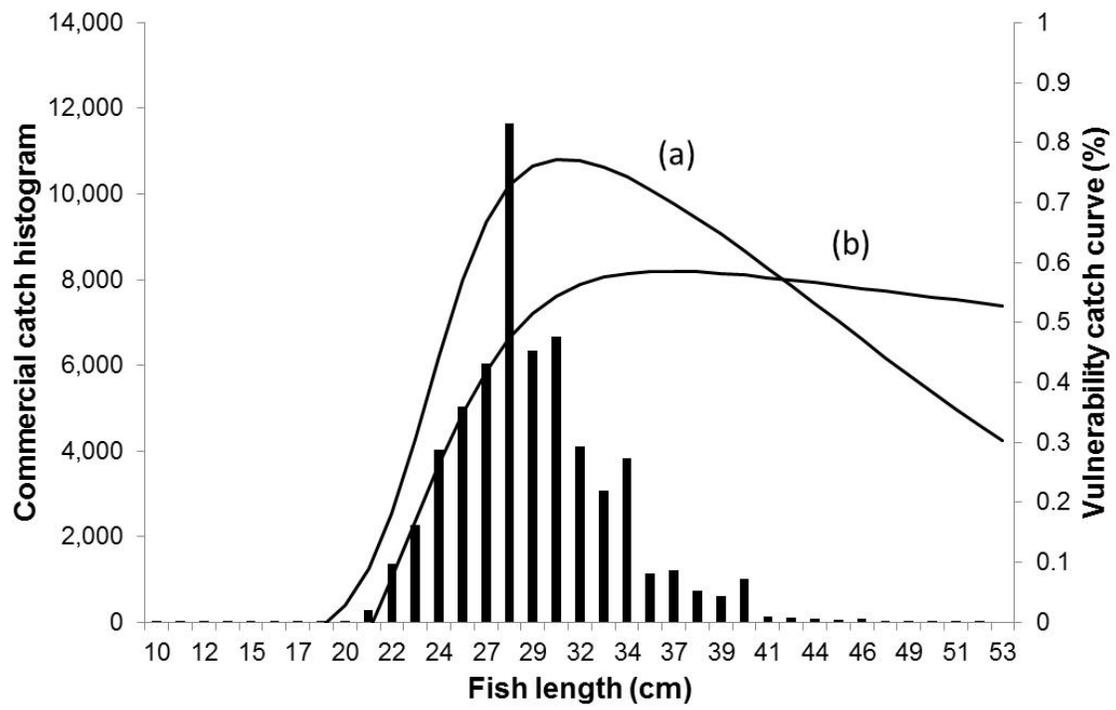


Figure 4.

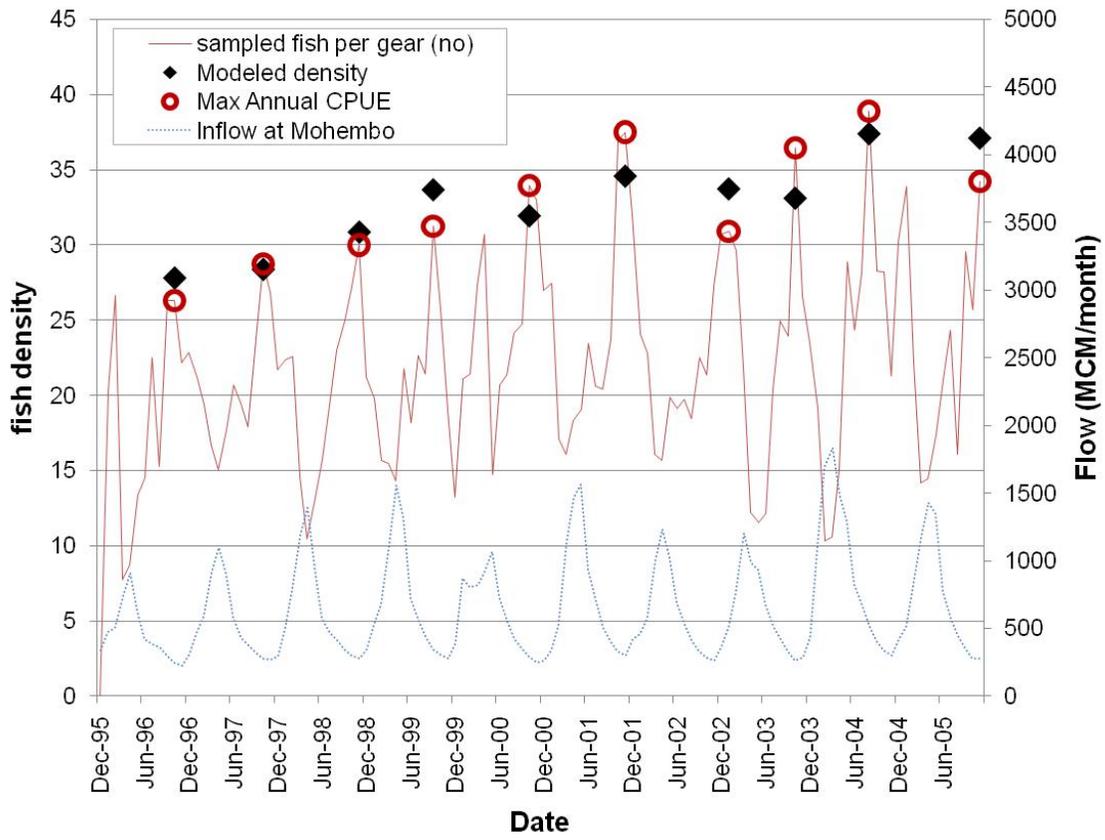


Figure 5.

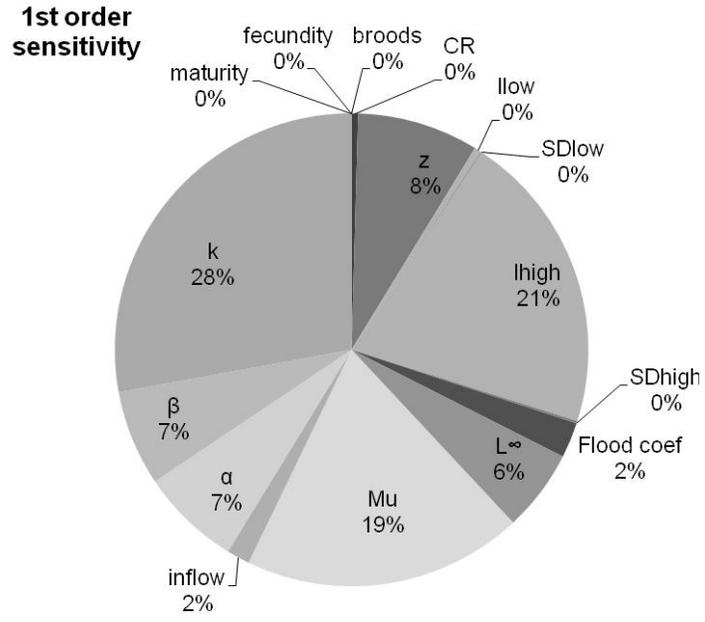


Figure 6.

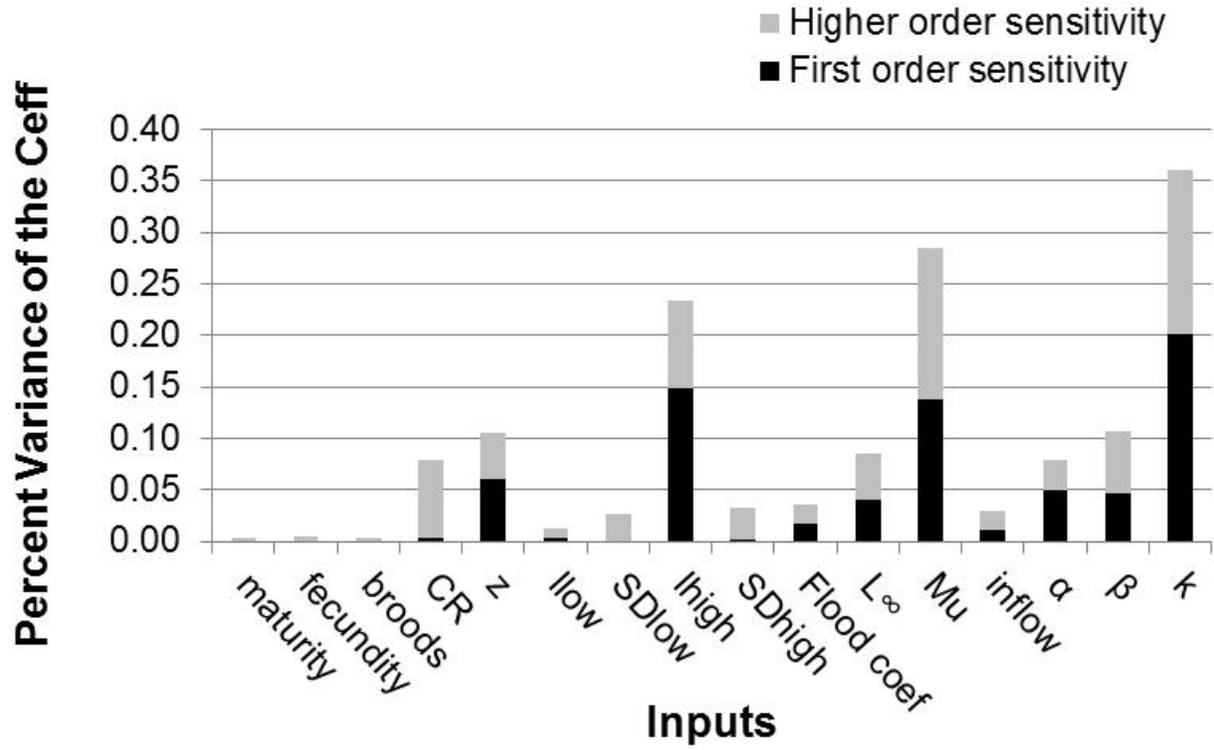


Figure 7.

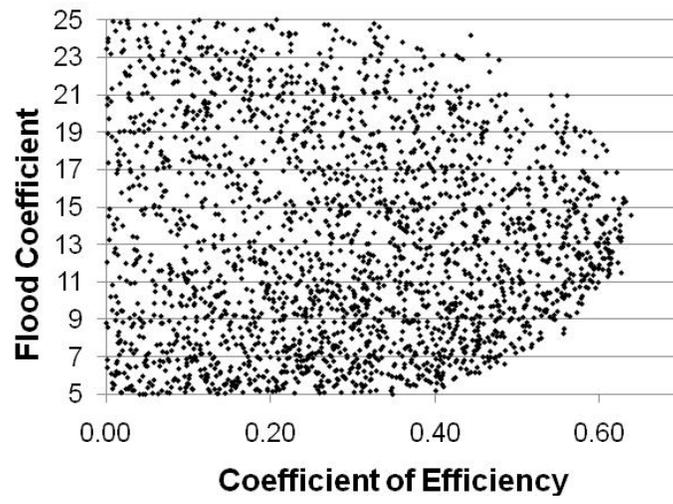


Figure 8.

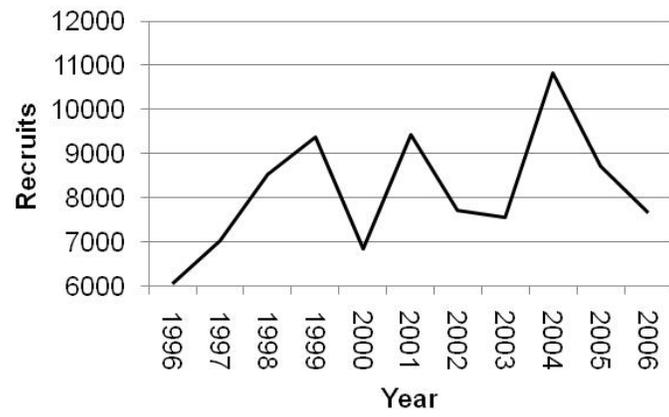


Figure 9.

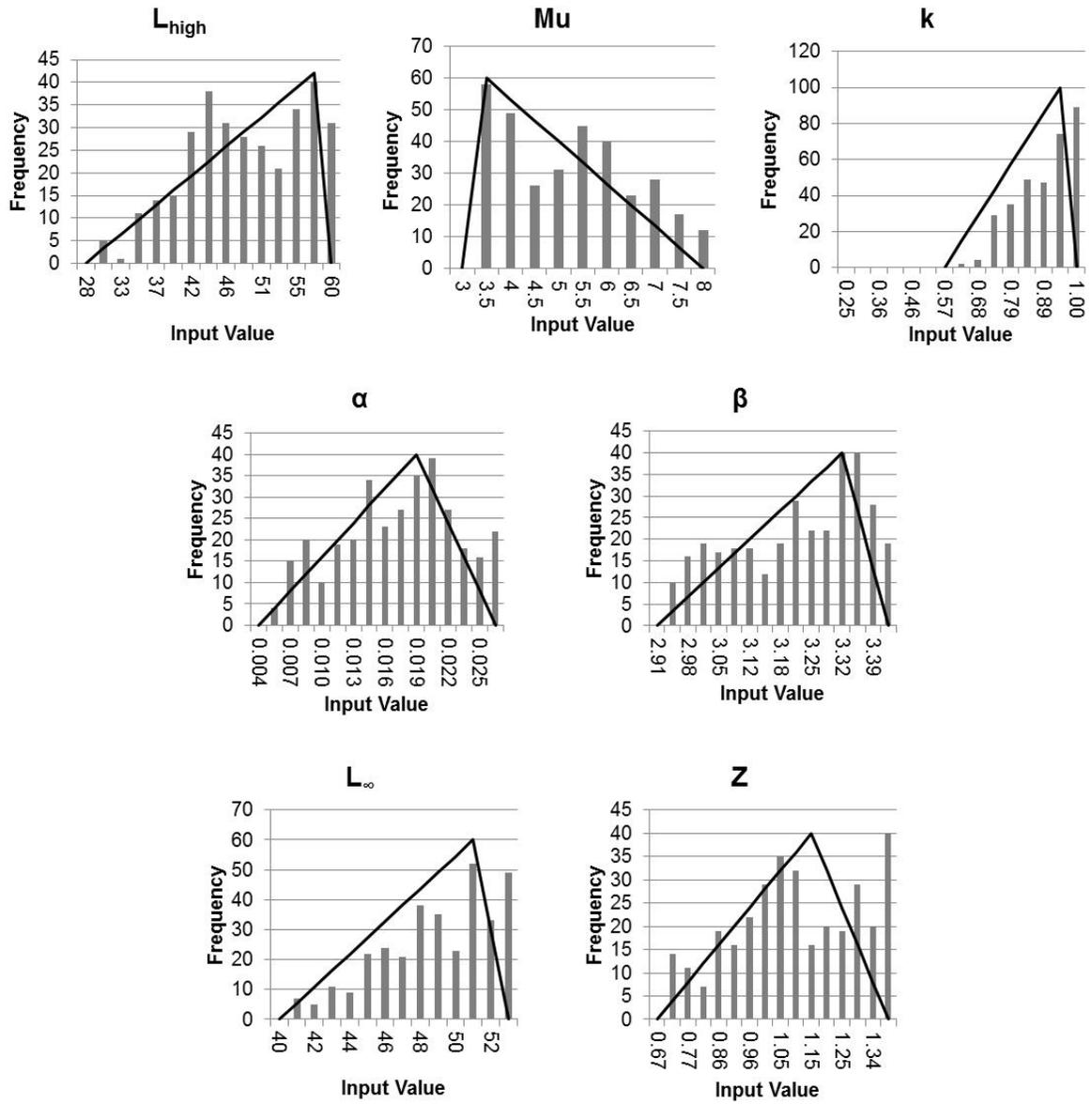
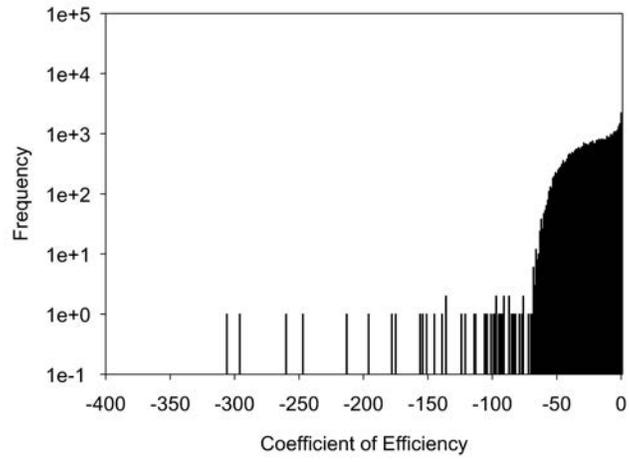
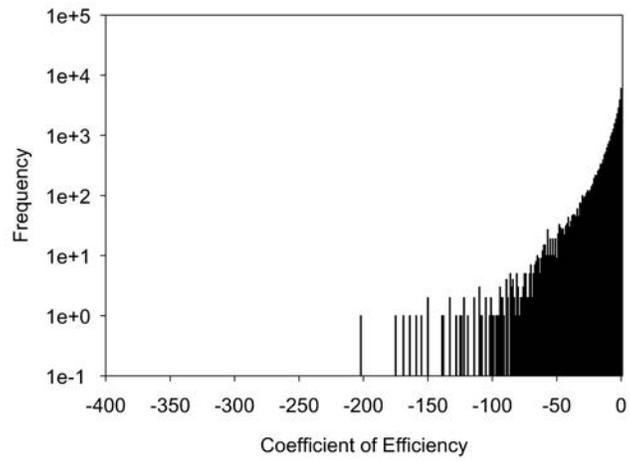


Figure 10.



a.



b.

Figure 11.

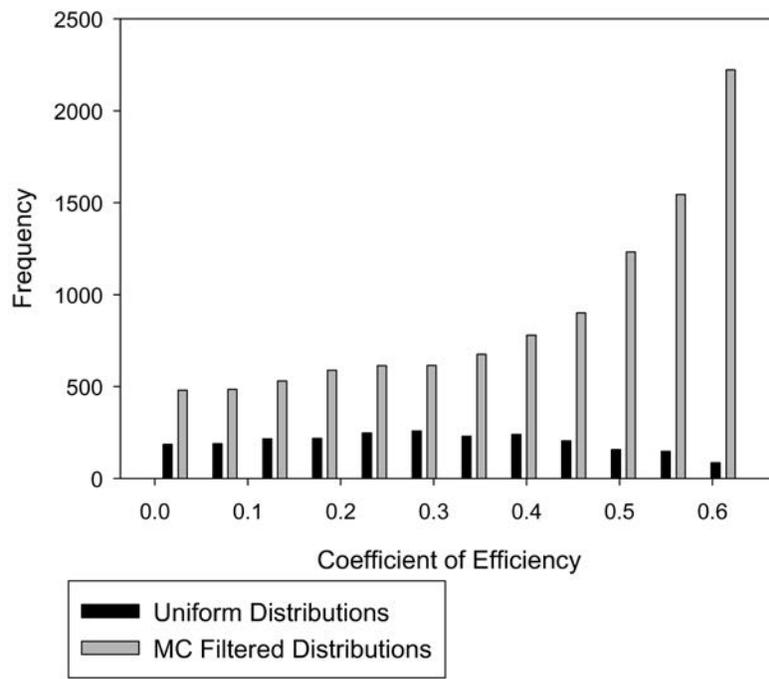


Figure 12.

Chapter 5. Conclusions

This research involving the GSA/UA and MC filtering of the ORI, Pitman, and Okavango Delta fish models provides important insights into the use of these tools not only for model building and simplifying but also as tools for reducing uncertainty, measuring the quality of models for management purposes, and corroborating ecological theories in the presence of model equifinality. GSA/UA is a tool that is traditionally used to aid modelers by identifying important inputs and setting unimportant inputs to constants. This allows the modeler to avoid over-parameterization and simplify models and also enables the modeler to focus on the most important inputs for further study. GSA/UA also often provides important insights into model inner workings that may aid in model development. All of these aspects of GSA/UA were accomplished in this work. However, a broader application of the use of these tools allowed additional insights to be garnered from these methods.

The GSA/UA and MC filtering of the ORI reservoir model provided valuable insights into the sensitivity, uncertainty, and inner workings of the ORI model and into reservoir models in general. The importance of volume thresholds was recognized; the keystone reservoir was identified; PDF's were refined; and a monitoring strategy for decreasing model uncertainty and increasing predictive capabilities was presented. MC filtering also decreased model uncertainty and brought the outputs closer to the calibrated values. These insights are useful for both modelers and policy makers in understanding the system and making decisions regarding model scenarios of climate change and development.

The Morris screening coupled with the FAST quantitative two tiered regionalization GSA/UA of the Pitman model highlighted the most important and least important inputs, quantified model uncertainty, compared regional groupings, and explored the uncertainty of an alternative model scenario. The conclusions are useful especially within the context of Adaptive Management when interpreting model reliability, determining the usefulness of model predictions, identifying gaps in knowledge, and focusing monitoring efforts. The most sensitive processes in the model have to do with rainfall distribution, catchment absorption, evapotranspiration, and groundwater recharge, and groundwater runoff is the least important process. The most important regions are the more humid headwater regions where the majority of the flow originates. The importance of each input is also compared to the uncertainty that was defined for the PDF's (low, medium, high). Generally, the inputs that were assigned high uncertainties were more important than those with medium or low uncertainties. This highlights the fact that the width of the input's PDF has an influence on its importance because inputs with a wider PDF are able to vary over a larger parameter space.

Because of the data scarcity in the Okavango Basin, and resulting wide PDFs that were set for the inputs in the Pitman model, there are multiple input sets that can simulate the same goodness of model fit. The issue of equifinality was evaluated to understand the effects that it may have on using the model in a predictive capacity for climate change scenarios. The most practical best fit model simulations were run under a climate change scenario and the 95% confidence interval of the average annual flow increased by approximately 18% between the original simulations and the climate change simulations.

A fish population model driven by the flood pulse was developed for the Okavango Delta. The fish model integrated GSA/UA and MC filtering in the development stages. The most important factors in this model were μ (mortality at unit length), L_{high} (upper length at 50% vulnerability), and k (growth factor). These were also highly interactive inputs. The model's high level of interaction was derived from both the interplay between the inputs in each time step and an internal inter-annual feedback loop. The least important inputs in the model were maturity, fecundity, the number of broods per season, the flood coefficient, SD_{low} , and L_{low} . These inputs were set to constants for model simplification. The flood coefficient is an unimportant input. However, this does not mean that the effect of the flood and the flood coefficient is not important for producing behavioral outputs. The most important inputs in the model are related to the overall population size, whereas the flood coefficient is the only input that relates solely to inter-annual population dynamics. The fact that the flood coefficient converges toward a value that produces the best fit in the model provides evidence that the flood is an important driver in the system. Moreover, all of the other inputs (besides k) show major issues of equifinality. They do not converge and are able to vary throughout their PDF ranges while still achieving best fit solutions.

After the GSA/UA was conducted and the unimportant inputs were set to constants in the fish model, MC filtering was performed. The prior distributions were re-set where the behavioral inputs were found to be significantly different from the non-behavioral inputs based on the two sided Smirnov test. L_{high} , μ , k , α , β , L_{∞} , and Z were all found to have significantly different behavioral inputs and were assigned triangular distributions based on the distribution of the

behavioral inputs. This method reduced uncertainty and is useful in a data poor environment allowing the modeler to focus on likely values within the ranges of physically acceptable PDFs. The modeler can then work with ecologists to ensure that the model makes ecological sense and provide converging lines of evidence for theories of system interactions.

Data scarcity presents many challenges for modelers and policy makers when developing, calibrating, interpreting model results, and applying models in predicative and dynamic capacities. However, there are ways of managing for these challenges. Adaptive Management is specifically designed for determining a course of action within the context of uncertainty. The tools presented here are couched within the context of Adaptive Management and quantitatively analyze model uncertainty in the light of practical management issues. The issue of data scarcity is acknowledged upfront and at the same time, uncertainty is reduced, model quality is assigned, and evidence of ecological theory is supported in spite of equifinality.

Limitations

Models are by nature simplifications of reality. Thus, as they are useful in understanding the world in a simpler fashion, they also introduce uncertainty. The human brain in fact views reality through a conceptual filter and every model, measurement, and connection that we perceive is inherently uncertain. GSA/UA is intended to shed light on this uncertainty. There are two major limitations associated with the GSA/UA in the work contained in this body of research. This work only investigates the uncertainty and sensitivity of model inputs. It does not address the uncertainty in model structure or time series data. In addition, there is inherent uncertainty in the PDF's that are developed for these inputs. As was shown in the

Pitman model, the width of the PDF has an impact on the importance of the input. In a data poor environment it is particularly difficult and often impossible to define PDF's for inputs based on evidence and data. In the absence of data, PDF's can be based on \pm some percentage and also given a ranking such as high, medium, and low uncertainty. This uncertainty in the PDF's introduces uncertainty in the GSA/UA. Accordingly, an uncertainty analysis of the uncertainty analysis could be conducted; however that is beyond the scope of this work.

MC filtering was shown to be a useful tool in reassigning prior PDF's and reducing uncertainty. However, this method is based on setting a rather arbitrary threshold for behavioral outputs. Arguments can be made for setting this threshold. For example in the ORI case, the threshold was set by balancing inputs and outputs. The input PDF's were truncated within reason to get a number of behavioral outputs. In the fish model the threshold was set to calibrate the model to an optimal range. Nonetheless, the threshold is a cutoff, a line drawn in the sand. A continuum based on the degree of behavior would be a better filter.

Appendix A. ORI Reservoir Model High Performance Computing GSA/UA Script and Instructions

The following script is used to run the ORI model (Wolski et al., 2006) in batch on the UF High Performance Computing (HPC) system. Essentially, for the GSA the ORI program is run in batch. The batch is broken up into smaller jobs to run more quickly. Each smaller job has a .job file associated with it. The program, Wine is used to run the Windows based Visual Basic program on the HPC Unix machine.

File name: Delta.job

Begin script:

```
#!/bin/sh
#PBS -N testjob
#PBS -o testjob.out
#PBS -e testjob.err
#PBS -M acathey@ufl.edu
#PBS -l nodes=1:ppn=1
#PBS -l walltime=30:10:00

date
hostname

echo "Changing directory to work area"
cd /scratch/ufhpc/acathey/ORISA/MCF/delta400
pwd

##
## Create dummy display
##
echo "Picking a unique display"
ncpus=`grep vendor_id /proc/cpuinfo | wc -l`
display=1
while [ $display -le $ncpus ] ; do
    if [ ! -e /tmp/.X${display}-lock ] ; then
        break
    fi
    display=`expr $display + 1`
done
echo "Unique display found: $display"
```

```
echo "Firing up dummy X server"
Xvfb :${display} > /dev/null 2>&1 &
jobs
export DISPLAY=:${display}

##
## Actual program commands
##
echo "Invoking wine"
wine delta.exe
echo "Done"
date

##
## Clean up after ourselves
##

echo "Killing dummy X server"
kill -TERM %1
# This should already be cleaned up, but just in case
rm -f /tmp/.X${display}-lock
rm -f /tmp/.X11-unix/X${display}
```

Appendix B. Fish model data and code

Table 1. Modeled fish populations and measured data. Measured data from Mosepele, 2009.

Date	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Total Modeled Population	19,467	19,866	21,600	23,565	22,352	24,198	23,172	23,172	26,164	25,965
Modeled Fish Density	28	28	31	34	32	35	34	33	37	37
Sampled CPUE	26	29	30	31	34	38	31	36	39	34
Modeled yr 1	6,052	7,021	8,541	9,375	6,856	9,432	7,567	7,567	10,837	8,717
Modeled yr 2	4,463	3,937	4,567	5,556	6,099	4,460	5,016	5,016	4,922	7,049
Modeled yr 3	2,525	2,977	2,626	3,046	3,706	4,068	4,093	4,093	3,346	3,283
Modeled yr 4	1,742	1,704	2,009	1,772	2,056	2,501	2,007	2,007	2,762	2,258
Modeled yr 5	1,975	1,181	1,155	1,362	1,201	1,394	1,861	1,861	1,361	1,872
Modeled yr 6	869	1,341	802	784	925	816	1,151	1,151	1,264	924
Modeled yr 7	740	591	911	545	533	628	643	643	782	859
Modeled yr 8	531	503	401	619	370	362	377	377	437	532
Modeled yr 9	368	361	342	273	421	252	290	290	256	297
Modeled yr 10	201	250	245	232	185	286	167	167	197	174

Table 2. Monthly sampled catch per unit fish data used for validation (Mosepele, 2009)

Date	CPUE	Date	CPUE	Date	CPUE	Date	CPUE
Jan-96	20.3	Aug-98	23.0	Mar-01	16.1	Oct-03	36.5
Feb-96	26.7	Sep-98	25.0	Apr-01	18.3	Nov-03	26.6
Mar-96	7.8	Oct-98	27.3	May-01	19.1	Dec-03	23.5
Apr-96	8.7	Nov-98	30.0	Jun-01	23.5	Jan-04	19.1
May-96	13.4	Dec-98	21.3	Jul-01	20.6	Feb-04	10.3
Jun-96	14.6	Jan-99	19.8	Aug-01	20.4	Mar-04	10.6
Jul-96	22.5	Feb-99	15.7	Sep-01	23.8	Apr-04	15.1
Aug-96	15.3	Mar-99	15.4	Oct-01	36.9	May-04	28.9
Sep-96	26.3	Apr-99	14.3	Nov-01	37.5	Jun-04	24.3
Oct-96	26.3	May-99	21.8	Dec-01	32.0	Jul-04	28.1
Nov-96	22.2	Jun-99	18.2	Jan-02	24.1	Aug-04	38.9
Dec-96	22.8	Jul-99	22.6	Feb-02	22.8	Sep-04	28.3
Jan-97	21.2	Aug-99	21.4	Mar-02	16.1	Oct-04	28.2
Feb-97	19.4	Sep-99	31.3	Apr-02	15.7	Nov-04	21.3
Mar-97	16.7	Oct-99	25.7	May-02	19.9	Dec-04	30.3
Apr-97	15.1	Nov-99	19.0	Jun-02	19.1	Jan-05	33.9
May-97	17.7	Dec-99	13.2	Jul-02	19.8	Feb-05	22.2
Jun-97	20.7	Jan-00	21.1	Aug-02	18.4	Mar-05	14.2
Jul-97	19.4	Feb-00	21.4	Sep-02	22.5	Apr-05	14.5
Aug-97	17.9	Mar-00	27.3	Oct-02	21.4	May-05	17.4
Sep-97	23.8	Apr-00	30.7	Nov-02	27.3	Jun-05	21.0
Oct-97	28.7	May-00	14.7	Dec-02	30.7	Jul-05	24.4
Nov-97	26.9	Jun-00	20.7	Jan-03	30.9	Aug-05	16.1
Dec-97	21.7	Jul-00	21.3	Feb-03	29.6	Sep-05	29.6

Jan-98	22.4	Aug-00	24.1	Mar-03	21.9	Oct-05	25.7
Feb-98	22.6	Sep-00	24.8	Apr-03	12.2	Nov-05	34.3
Mar-98	14.6	Oct-00	34.0	May-03	11.5	Dec-05	20.3
Apr-98	10.4	Nov-00	33.0	Jun-03	12.1		
May-98	13.2	Dec-00	27.0	Jul-03	20.4		
Jun-98	15.8	Jan-01	27.4	Aug-03	25.0		
Jul-98	19.4	Feb-01	17.1	Sep-03	24.0		

Fish Model Code (Visual Basic)

Main Module

```
Public do_hydro_model As Boolean
Public do_veg_model As Boolean
Public workdir, mapdir As String
Public noftsteps As Integer
Public evapdata As Integer
Public file_suffix As String
```

```
Sub Main()
```

```
  docalcflag = True
```

```
  workdir = "H:\HOORCfishModel\annual\annual SA for Cal MC constantkz3" 'CurDir
```

```
  mapdir = workdir & "maps\"
```

```
  evapdata = 2
```

```
  file_suffix = "_2008"
```

```
  'file_suffix = "_baseline"
```

```
  'file_suffix = "_csiro"
```

```
  'file_suffix = "_had"
```

```
  'file_suffix = "_ccc"
```

```
  'file_suffix = "_gldf"
```

```
  'file_suffix=""
```

```
  'Call hydro_model.hydro_model_main
```

```
  Call fish_model.fish_model_main
```

```
End Sub
```

Fish Module

```
Public file_flood, file_inflow, file_out, file_fish, d_file, GSA_out As String
Public flood() As Integer
Public inflow() As Variant
Public total_flood(), maxflood(), minflood() As Double
Public recdate(), fl(), infl() As Variant
Public unitcode() As Integer
```

```
Public population_matrix() As Double
Public caught_pop() As Double
Public y, ts, month, start, nofyears, lifespan As Integer
Public recK(400000), lhigh(400000), deGraafM(400000), deGraafa(400000), deGraafb(400000),
k(400000), z(400000), Linf(400000) As Double
Public runs, G As Variant
Public recruits(400000) As Variant
Public Spawners() As Double
Public maxinflow(), mininflow(), minet(), maxet() As Double
Public et() As Double
Public stop_error As Integer
Public EndSub As Label
```

```
Sub fish_model_main()
```

```
DoEvents
```

```
file_flood = workdir & "\out_area_2008.csv"
file_inflow = workdir & "\input_2008.csv"
```

```
Call getnoftsteps
nofyears = Int(noftsteps / 12) ' this is the number of years to analyse
```

```
start = 5
MaxLength = 53
lifespan = 10
ReDim recdate(noftsteps)
ReDim total_flood(noftsteps), maxflood(noftsteps), minflood(noftsteps) As Double 'flood in
km^2
ReDim inflow(noftsteps) As Variant 'inflow in MCM at mohembo
ReDim fish_biomass(lifespan, noftsteps) As Double 'biomass is in tons
ReDim fish_length(lifespan) As Double 'fish length in cm
ReDim population_matrix(lifespan, noftsteps) As Double
ReDim total_pop(noftsteps) As Double
ReDim et(noftsteps) As Double
ReDim caught_pop(noftsteps), write_pop(noftsteps) As Double
ReDim Spawners(noftsteps) As Double
```

```
runs = 39944
```

```
Call read_inflow
Call read_hydro_results
Call GSA
```

```

'Call noGSA

End Sub

Sub GSA()
file_out = workdir & "\\Data\\fish_GSA" & ".csv"
Open file_out For Output As #2

file_GSAinputs = workdir & "\\Data\\inputs2.sam"
Open file_GSAinputs For Input As #3
Input #3, buzz
Input #3, buzz
Input #3, buzz
Input #3, buzz

For G = 1 To runs
    Input #3, reK(G), lhigh(G), deGraafM(G), deGraafa(G), deGraafb(G), z(G), k(G), Linf(G)
    'maturity(G) = Round(maturity(G))
Next G

Close #3
For y = start To nofyears
    If y > 41 And y < 52 Then
        Write #2, "ts " & y + 1954,
    End If
Next y
Write #2,

lflow = 23.5
sigmalow = 1.8
sigmahigh = 20

For G = 1 To runs
    stop_error = 0
    For y = start To nofyears
        Call make_fish

        For N = 1 To lifespan
            L = Linf(G) * (1 - Exp(-N) * k(G))
            vulnerable = 1 / (1 + Exp(-(L - lflow) / sigmalow)) - 1 / (1 + Exp(-(L - lhigh(G)) /
sigmahigh))
            If vulnerable < 0 Then vulnerable = 0
            caught_pop(ts) = caught_pop(ts) + population_matrix(N, ts) * vulnerable
        
```

```

Next N

If y > 41 And y < 52 Then
  If stop_error = 1 Then
    Write #2, 99999,
  Else
    Write #2, caught_pop(y) / 700, '(minflood(y) * sflood(G)), '700,
  End If
End If
End If

Next y
Write #2,
ReDim caught_pop(noftsteps) As Double
Next G
Close #2
Close #99
End Sub
Sub noGSA()
file_out = workdir & "\fish_biomass_2008" & ".csv"
Open file_out For Output As #2

'ceff = 0.637
maturity(1) = 4.189057665
fecundity(1) = 735.7842698
broods(1) = 1.359830489
reck(1) = 14.9926638
z(1) = 1.298389041
lflow(1) = 23.8242803
sigmalow(1) = 1.48082475
lhigh(1) = 36.67651807
sigmahigh(1) = 26.46092243
coef(1) = 12.89298458
Linf(1) = 47.09325197
deGraafM(1) = 4.844473799
inflow10(1) = 1.086765657
deGraafa(1) = 0.022612467
deGraafb(1) = 3.35601894

maturity(G) = Round(maturity(G))

'set headers in excel file
Write #2, "",
For ts = 1 To lifespan
  Write #2, "age " & ts,

```

```

Next ts
Write #2,

G = 1
For y = start To nofyears
    Call make_fish
Next y

For y = start To nofyears
    For m = 1 To 12
        ts = (y - 1) * 12 + m
        Write #2, "ts " & ts,
        For N = 1 To lifespan
            L = Linf(G) * (1 - Exp(-N * k(G)))
            vulnerable = 1 / (1 + Exp(-(L - llow(G)) / sigmalow(G))) - 1 / (1 + Exp(-(L - lhigh(G)) /
sigmahigh(G)))
            Write #2, population_matrix(N, y) * vulnerable,
        Next N
        Write #2,
    Next m
Next y
Close #2

Call write_to_excel

End Sub
Sub make_fish()

    ts = y '(y - 1) * 12 + month
    ReDim Spawners(noftsteps) As Double

    'Average max annual flooding extents in the panhandle = 876.2 km2, average = 801.6, min
= 704.4

    'Average max annual inflow for all years = 1529, for fish data years = 1328
    'Average min annual inflow for all years = 350.7, for fish data years = 269.8
    'Average avg annual inflow for all years = 774

    'Average max ET for all years = 190.5, avg = 126.1, min = 63.4
    Nt1 = 25000
    broods = 1.5
    maturity = 4.5
    fecundity = 600

```

coef = 13

No = Nt1 + (maxinflow(ts) - 1529) * coef

Ro = (1 - Exp(-z(G) * k(G))) * No

reca = reK(G) * Ro / No

recb = (reK(G) - 1) / No

If ts = 5 Then 'calculate the age structure for the first year

For N = 1 To lifespan

If N = 1 Then

population_matrix(N, ts) = Ro

Else

L = Linf(G) * (1 - Exp(-N * k(G)))

zL = deGraafM(G) * deGraafa(G) ^ -0.3 * L ^ (-0.3 * (deGraafb(G))) 'ge Graaf et al
(2005) Simulation of Nile Tilapia in Ponds, used Mu of 1

population_matrix(N, ts) = population_matrix(N - 1, ts) * Exp(-zL * k(G))

End If

Next N

Else 'calculate the age structure for the rest of the years

For N = 1 To lifespan

If N = 1 Then

For n2 = maturity To lifespan

Spawners(ts) = Spawners(ts) + population_matrix(n2, ts - 1)

Next n2

On Error GoTo EndSub

Spawners(ts) = (Spawners(ts) * coef) * fecundity * broods

population_matrix(N, ts) = reca * Spawners(ts) / (1 + recb * Spawners(ts))

Else

L = Linf(G) * (1 - Exp(-N * k(G)))

zL = deGraafM(G) * deGraafa(G) ^ -0.3 * L ^ (-0.3 * (deGraafb(G))) 'ge Graaf et al
(2005) Simulation of Nile Tilapia in Ponds, used Mu of 1

population_matrix(N, ts) = population_matrix(N - 1, ts - 1) * Exp(-zL * k(G))

End If

Next N

End If

Exit Sub

EndSub:

stop_error = 1

End Sub

Sub getnoftsteps()

Open file_flood For Input As #1

' check number of time steps

noftsteps = -1

Do While Not EOF(1)

Line Input #1, temp

noftsteps = noftsteps + 1

Loop

Close #1

End Sub

Sub read_inflow()

Dim infl(6) As Variant

ReDim maxinflow(noftsteps)

ReDim mininflow(noftsteps)

ReDim maxet(noftsteps)

ReDim minet(noftsteps)

Open file_inflow For Input As #11

For t = 1 To noftsteps

Input #11, infl(1), infl(2), infl(3), infl(4), infl(5), infl(6)

inflow(t) = infl(2)

et(t) = infl(5)

Next t

Close #11

For y = 1 To nofyears

maxinflow(y) = inflow((y - 1) * 12 + 1)

mininflow(y) = inflow((y - 1) * 12 + 1)

maxet(y) = inflow((y - 1) * 12 + 1)

minet(y) = inflow((y - 1) * 12 + 1)

For m = 1 To 12

ts = (y - 1) * 12 + m

If inflow(ts) > maxinflow(y) Then

maxinflow(y) = inflow(ts)

End If

```

    If inflow(ts) < mininflow(y) Then
        mininflow(y) = inflow(ts)
    End If

    If et(ts) > maxet(y) Then
        maxet(y) = et(ts)
    End If

    If et(ts) < minet(y) Then
        minet(y) = et(ts)
    End If
Next m

Next y

End Sub

Sub read_hydro_results()

ReDim fl(11, noftsteps) As Variant
ReDim recdate(noftsteps) As Variant

Open file_flood For Input As #10
Input #10, a, b, c, d, e, f, q, h, i, J, k, L

For t = 1 To noftsteps
    Input #10, buzz, recdate(t), fl(1, t), fl(2, t), fl(3, t), fl(4, t), fl(5, t), fl(6, t), fl(7, t), fl(8, t), fl(9, t),
    fl(10, t), fl(11, t)
    total_flood(t) = fl(1, t) + fl(4) + fl(5) + fl(6) + fl(11) + fl(2) + fl(3) + fl(9) + fl(7) + fl(8) + fl(10)
Next t
Close #10

For y = 1 To nofyears - 1
    maxflood(y) = total_flood((y - 1) * 12 + 1)
    minflood(y) = total_flood((y - 1) * 12 + 12)

    For m = 1 To 12
        ts = (y - 1) * 12 + m

        If total_flood(ts) > maxflood(y) Then
            maxflood(y) = total_flood(ts)
        End If
    
```

```
If m = 1 And total_flood(ts + 12) < minflood(y) Then
    minflood(y) = total_flood(ts + 12)
End If
```

```
If m = 2 And total_flood(ts + 13) < minflood(y) Then
    minflood(y) = total_flood(ts + 13)
End If
```

```
If m = 12 And total_flood(ts) < minflood(y) Then
    minflood(y) = total_flood(ts)
End If
```

```
If m = 11 And total_flood(ts) < minflood(y) Then
    minflood(y) = total_flood(ts)
End If
```

```
Next m
minflood(y) = (total_flood((y - 1) * 12 + 11) + total_flood((y - 1) * 12 + 12) + total_flood((y) *
12 + 1) + total_flood((y) * 12 + 2)) / 4
```

```
Next y
```

```
End Sub
```

```
Sub write_to_excel()
```

```
'this writes the bucket model to excel
```

```
Set objExcel = CreateObject("Excel.Application")
```

```
Set objWorkbook1 = objExcel.Workbooks.Add(workdir & "\fish_biomass_2008.csv")
```

```
Set objWorkbook2 = objExcel.Workbooks.Open(workdir & "\Data\results.xlsx")
```

```
objExcel.Visible = True
```

```
objWorkbook1.Worksheets("fish_biomass_2008").Range("A1:k645").Copy
```

```
objWorkbook2.Worksheets("fish_biomass_2008").Range("A1:k645").PasteSpecial Paste =
xlValues
```

```
objWorkbook1.Application.CutCopyMode = False
```

```
objWorkbook1.Close
```

```
Set objExcel = Nothing
```

```
End Sub
```