

A two-way calibration of the SWAT and OneLay/PolTra models using integrated modelling approach for the Lake Winnipeg Basin

W.G. Booty^a, I.W.S. Wong^a, R.C. McCrimmon^a, L.F. Leon^a and P. Fong^a, C.L. Richard^a

^a *Environment Canada, National Water Research Institute, 867 Lakeshore R., Burlington, Ontario, Canada,
Email: bill.booty@ec.gc.ca*

Abstract: Lake Winnipeg is Canada's sixth largest lake and the world's 10th largest freshwater lake with a surface area of 24,500 km². The Lake experiences eutrophication and increasing algal blooms due likely to nutrient loadings from agricultural runoff. Lake Winnipeg receives surface water discharges from the Lake Winnipeg Basin which spans across 106 km²; including four provinces and four states. Major rivers that flow into Lake Winnipeg include the Saskatchewan, Winnipeg, Red and Dauphin Rivers. These four rivers account for more than 80% of the Lake inflow. The Winnipeg River has the largest inflow at 45% of the total Lake inflow but the Red River dominates nutrient loadings with 63% of the Lake's phosphorus loading.

An integrated modelling approach has been used to analyze and provide solutions to the Lake Winnipeg eutrophication problem. The Soil and Water Assessment Tool (SWAT) is the watershed runoff model being used to simulate sediment and nutrient loadings from the La Salle River watershed (2,400 km²), which will then be up scaled to represent the entire Red River watershed, and then these loadings will be used as input to the lake model, OneLay/PolTra (O/P), to simulate the Lake's water quality. This integrated modelling approach has been found to provide a more flexible and efficient means to examine various remediation methods in the watershed. Only the Red River nutrient loads are being modelled due to its dominance of nutrient loads to the lake. The La Salle River empties into the Red River near Winnipeg, is 80% agricultural, and contributes approximately 1% of the nutrient load to the Red River watershed. The O/P model includes the 4 major inflowing rivers but only the Red River will vary during calibration and scenario analysis. The other 3 river loadings are based on observed data. Using artificial intelligence to determine the most optimal calibrations for loading provides for better calibration for SWAT and O/P through integrated modelling.

A genetic algorithm is used to auto calibrate the integrated models. As a first attempt, model parameters being calibrated include those related to total suspended sediment (TSS) loadings from the SWAT model and in-lake suspended sediment concentrations from the O/P model. The objective functions used in the genetic algorithm is the normalized root mean square error (NRMSE). The integrated model calibration incorporates back calibration or two-way calibration to improve the overall objective.. Initial results from the two-way calibration of the integrated SWAT and O/P models show the NRMSE for the watershed and lake TSS are 0.18 and 0.12, respectively. With a Bayesian network, the uncertainty propagation of the integrated models is estimated to be 29.2%. This uncertainty is reasonable given the lack of precise input data, the simplicity of the Upscaler model from the LaSalle Watershed to Lake Winnipeg; and the significant spatial resolution change from the SWAT model to the O/P model. With the successful implementation of this integrated modelling approach, we plan to populate this approach to other watersheds in the Lake Winnipeg Basin and beyond where data is available.

Keywords: *Integrated modelling, two-way calibration, Lake Winnipeg*

1. INTRODUCTION

Lake Winnipeg is the sixth largest freshwater lake in Canada and the 10th largest by surface area (24,500 km²) in the world. For many years water quality in Lake Winnipeg has been deteriorating. Lake Winnipeg receives surface water discharges from a large basin area (> 1M km²), which spans four provinces and four states (Figure 1). Nonpoint source pollution (NPS) from agricultural practices has been associated with an increased rate of eutrophication in the lake (Salvano *et al.*, 2009, Jones and Armstrong, 2001). The major Canadian sources for contributing loads of sediment (TSS), nitrogen (N) and phosphorus (P) to the lake originate on the agricultural regions of southern Manitoba, in particular from the Red-Assiniboine basin, as described in Bourne *et al.* (2002); who based their load estimates on long-term water quality and stream flow data. These excessive nutrients are also contributing to the growth of toxic blue-green algae in this lake. The blue-green algae also deplete oxygen content of the lake and affect fishing as well as recreational sports. Given the serious situation, the Canadian government decided to invest \$18 million in the Lake Winnipeg Basin Initiative (LWBI), to clean up the lake. It has been designed to allow collaboration with provincial agencies, binational agencies and regional and local partners to identify and address current data and knowledge gaps and develop a science-based nutrient and bloom management program within a viable socioeconomic framework based on ecologically relevant targets

This paper focuses on the integrated modelling research that is being carried out to predict reductions in sediment and nutrients to the lake and the impact of those reductions on lake water quality and the overall health of the lake ecosystem. In particular, we will focus on the integration of the SWAT (Arnold *et al.*, 1995) non-point source water quality model, and a lake model, OneLay/PolTra (O/P) (Simons and Lam, 1986), that are integrated in order to be able to carry out a two-way optimization of the loadings to Lake Winnipeg to meet government legislated water quality objectives.

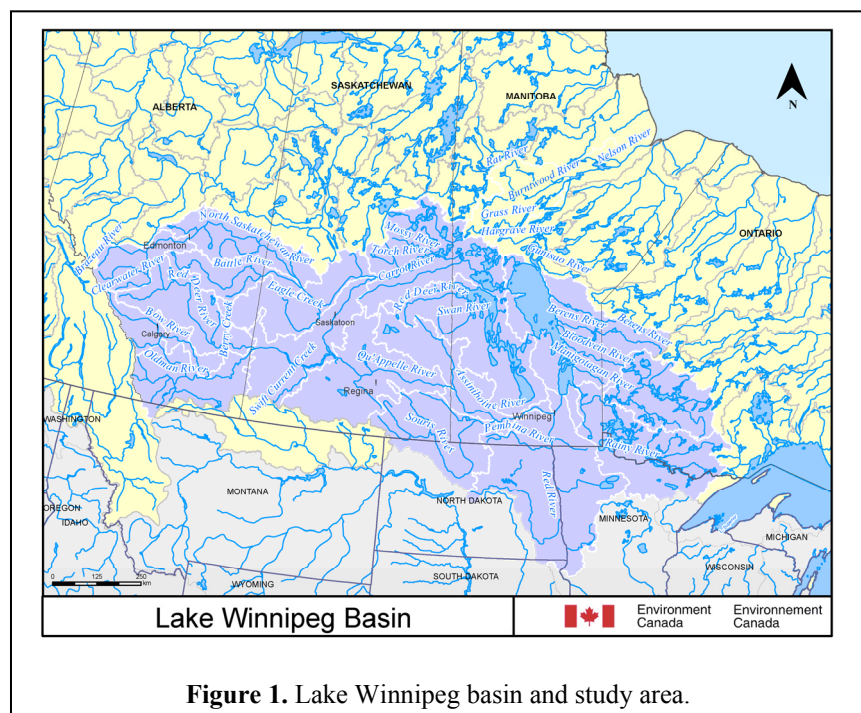


Figure 1. Lake Winnipeg basin and study area.

2. METHODS

In order to be able to carry out an integrated solution to the Lake Winnipeg eutrophication problem, the various models being employed also need to be integrated. Under the integrated modelling framework and as a long term goal, we plan to simulate how a given set of watershed management options, including agricultural land-use practices, may affect non-point source sediment and nutrient loadings and subsequently in-stream and in-lake water quality condition. We also plan to simulate, conversely, how we may determine the optimal or near-optimal watershed management options for a given set of in-lake water quality objectives i.e. two-way optimization.

2.1. SWAT Watershed Model

The SWAT model has been used to investigate the types of reductions of loadings of nutrients to Lake Winnipeg using three representative watersheds including the La Salle River (Leon *et al.*, 2010). They were selected to determine the impacts of land use change and Beneficial Management Practices (BMPs) implementation on nutrient loadings and the water quality of streams. The La Salle watershed (2,400 km²) is a sub-basin of the Red-Assiniboine River as shown in Figure 2. The Soil and Water Assessment Tool

(SWAT) is a well known watershed model (Di Luzio *et al.*, 2001; Neitsch *et al.*, 2002) which provides estimates of runoff, sediment yield, and nutrient loads at a sub-basin level. SWAT is being used as the watershed scale hydrologic model in this project, due in fact to its capacity to simulate loadings to contributing streams across a wide range of scales; it also provides tools for scenario testing in agricultural landscapes. The SWAT model output is linked with the integrated O/P lake model to examine the effects of loadings reductions for Lake Winnipeg.

SWAT requires input of Digital Elevation Model data or DEM, soil type distribution, and land use coverage. The La Salle River watershed is dominated mainly by clays (74%) and agricultural land (80%). The flatness of the La Salle River watershed required DEM manipulation to get a good delineation. The model setup includes additional point sources to simulate water diversions from the Assiniboine directly into the headwaters of the basin at three locations, and the discharges from a dozen sewage treatment plants.

2.2. Upscaler – Scale-up Loads from Sub-watersheds to Red River watershed

The Red River is approximately 287,000 km² near the Lake at Lockport, contributes 11% of flow to Lake, but is the dominant source of sediment and nutrients (Lake Winnipeg Stewardship Board, 2006). The effects of watershed BMPs on Lake Winnipeg's water quality requires the scaling-up of the loadings from the sub-watershed to represent the whole Red River watershed loadings in this land-lake model integration because the Red River watershed is too large to model in its entirety at this stage of the project. An OpenMI component model, Upscaler, was created for scaling up flow and sediment from one model representing one area to another as shown in Figure 3. La Salle river watershed is approximately 2,400 km² which is 0.8% of the Red River watershed. Based on linear regression, the form of the model is shown in equation 1 where slope and offset are determined through linear regression of La Salle watershed outlet values compared to observed Red River concentrations. The Upscaler can also be used in the reverse to downscale Red River values to the La Salle watershed outlet.

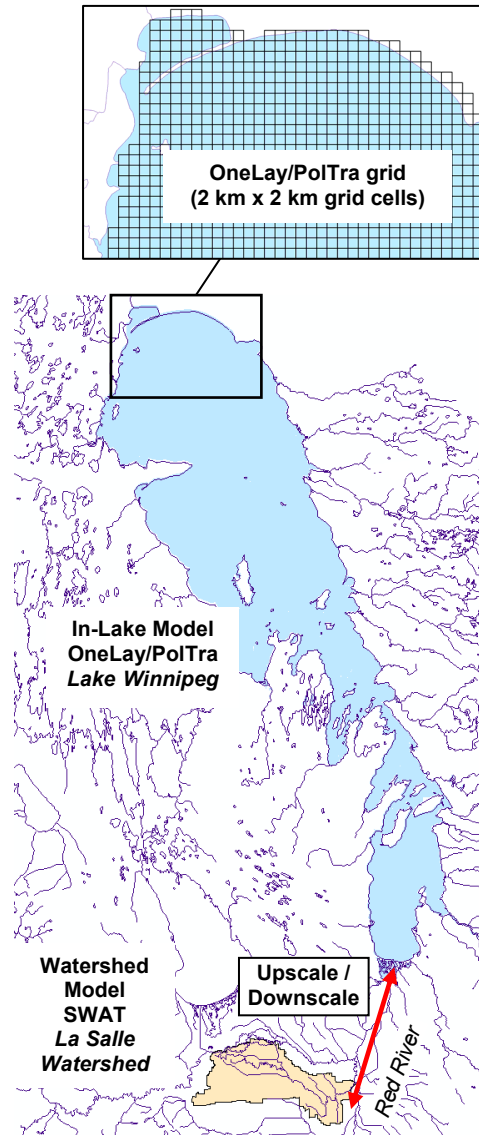


Figure 2. Integrated modelling connection.

$$(\text{Upscale to Red River TSS}) = \text{slope} \times (\text{La Salle watershed outlet TSS}) + \text{offset} \quad (1)$$

TSS data is more sporadic spatially and temporally and a linear regression produced a correlation coefficient of 0.72. There is possible backwater influence from the Red River affecting observations in the La Salle. The

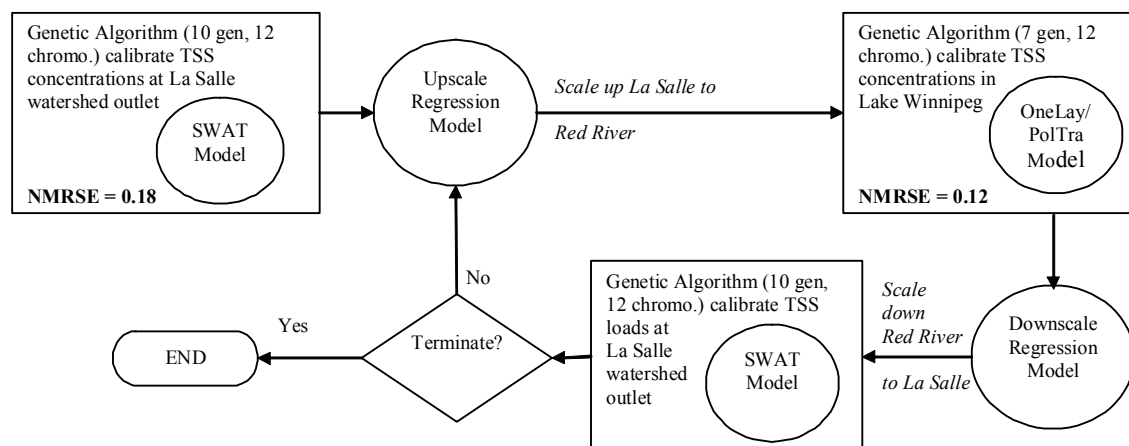


Figure 3. OpenMI Integrated Modelling of SWAT, Upscaler, and OneLay/PolTra.

main consideration regarding scales is the practicality of being able to manipulate model inputs to generate and simulate different scenarios. Large scale applications are more useful for large forcing scenarios such as in climate change.

2.3. O/P Lake Model

The first level of model integration involved combining the OneLay and PolTra models (Simons and Lam, 1986) to form a 2-D horizontal, vertically mixed lake model. The models are based on a rectangular grid representation of a lake. OneLay is the hydrodynamic program that uses lake depths, river inflow/outflow, and wind vector to simulate horizontal currents and water level surges. PolTra is the pollutant transport model that uses the bathymetry and water transport computed by OneLay to simulate water and sediment concentrations. OneLay and PolTra for this study included the addition of resuspension from the sediments based on orbital velocity.

We have applied OpenMI (Open Modelling Interface) standards for the OneLay and PolTra lake models to improve interactions among the models and for model calibration. The OpenMI is a software standard that allows compliant models to exchange data at run time (www.OpenMI.org). Using OpenMI techniques, wrappers were written in C# for the 2 programs. This enables the programs to exchange data at each time step, as opposed to running them in sequence, which improves model integration and allows integration with other models such as SWAT to form a whole catchment modelling system (Figure 2). Originally the hydrodynamic computations were completed before the water quality model started and each model was calibrated separately.

3. AUTOMATED CALIBRATION

In this paper, the integrated model calibration incorporates back calibration or two-way calibration to improve the overall objective and to try to reduce the impact of uncertainty and different levels of model complexities. In general, automated calibrations for SWAT and O/P are performed using a genetic algorithm (GA) (Goldberg, 1989). The steps for the two-way calibration are as follows:

- Step 1) Run GA for SWAT calibrating on in-stream TSS at the watershed outlet; objective function based on comparing SWAT simulated vs. observed (Observed TSS between 1996 and 2002); 10 generations with 12 chromosomes each.
- Step 2) Get Upscaler coefficients by fitting linear regression through SWAT watershed outlet TSS (best run from Step 1) versus Red River at Lockport observed daily TSS.
- Step 3) Use step 2 results to scale-up the best SWAT run TSS concentration (flow weighted average) for the O/P simulation period. This will give the Red River TSS to use in O/P.
- Step 4) Run GA to attempt to calibrate O/P on Lake TSS. Calibrate on 5 model parameters and the Red River TSS inflow concentration, whose range is based on the observed and Upscaler value +/- 50%; 7 generations with 12 chromosomes.
- Step 5) Downscale Red River TSS from best O/P run in step 4.
- Step 6) Run GA for SWAT but calibrate only on the monthly load based on the downscaled TSS from step 5;
- Step 7) Continue to cycle through and repeat Steps 3 to 6 until convergence;

The SWAT model for the La Salle watershed was setup to run from 1996 to 2007. The O/P model for Lake Winnipeg was setup to run from June 23 to October 31 of 2002 due to the availability of observations. Automatic calibration of SWAT, Upscaler and O/P for TSS was made using the OpenMI versions. In previous work, normally the Nash Sutcliffe efficiency (NSE) is used as the fitness function for SWAT and the root mean square error (RMSE) for O/P. However, some future automated calibration options may require a combined fitness from both models. To accommodate this, the normalized RMSE (NRMSE) is used for all models since it can be easily combined. We have adopted the following fitness function:

$$\text{Combined NRMSE} = 0.5 \text{ SWAT NRMSE} + 0.5 \text{ O/P NRMSE} \quad (2)$$

The SWAT model inputs used here were previously calibrated for flow (Leon *et al.*, 2010). Only a few selected parameters related to TSS are part of the automated calibration during these initial options testing and include peak rate factor (PRF), sediment transport coefficient (SPCON), sediment transport exponent (SPEXP), channel erodibility (CH_EROD), and channel cover (CH_COV). Other parameters plan to be added in future work. Parameter ranges for the GA shown in Table 2 are based on previous work, specifically the SWAT reference manual and other literature reviews.

The O/P simulations used the spring observations for the initial conditions so that the NRMSE is calculated for the summer and fall observations. The calibration parameters for O/P included drag coefficient (Drag) and bottom friction (BFric) in OneLay and settling rate (Wset), resuspension coefficient (Kr), critical shear velocity (Ucr), and sometimes Red River TSS concentration (RivTss), in PolTra. Future calibrations will consider other parameters coincidentally such as water levels and nutrients. Parameter ranges used in the GA are shown in Table 3.

The integrated model calibrations were done on a server containing 12 processors. To maximize use and reduce computation time, the GA for each model was setup to use 12 chromosomes; one on each processor, basically doing 12 runs at the same time. This is advantageous especially for the O/P model which typically takes 40 minutes for one simulation.

Table 2. SWAT GA Parameters

Parameter	Lower range	Upper range
PRF	0	2
SPCON	0.0001	0.01
SPEXP	0.5	2
CH_EROD	0	1
CH_COV	0	1

Table 3. O/P GA Parameters

Parameter	Lower range	Upper range
Drag	0.001	0.02
BFric	0.01	0.1
Ucr	0	2
Kr	1e-14	1e-7
Wset	1e-7	1e-2

4. CALIBRATION RESULTS AND UNCERTAINTY PROPOGATION

The two-way calibration had NRMSE for SWAT and O/P of 0.18 and 0.12, respectively. The combined NRMSE is 0.15. Figure 4 shows the SWAT monthly average simulated TSS with the observed data for the La Salle Watershed. Figure 5 shows the O/P simulated TSS with the observed data for Lake Winnipeg. The results are not unexpected given the complexity of the integrated models.

For each model, there is a degree of uncertainty associated. It is of interest to assess the uncertainty of the integrated model. A Bayesian network (Reckhow, 2003) is proposed to analyze the uncertainty propagation. We used the correlation coefficients (r) of the simulated and observed results of the three models as inputs to the Bayesian network. The uncertainty of the model is represented by $(1 - r)$. The correlation coefficients of the SWAT, Upscaler and O/P models are 0.61, 0.72 and 0.79, respectively. As a result, we can set up the Bayesian network using the Netica software (www.netica.com).

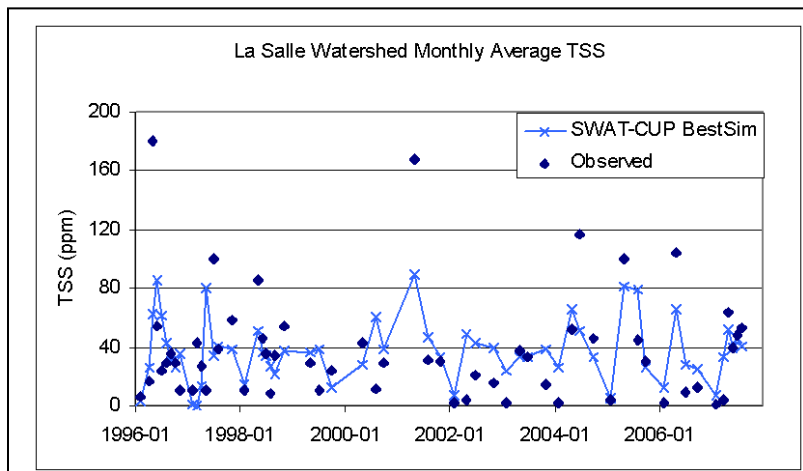


Figure 4. SWAT Simulated and Observed TSS.

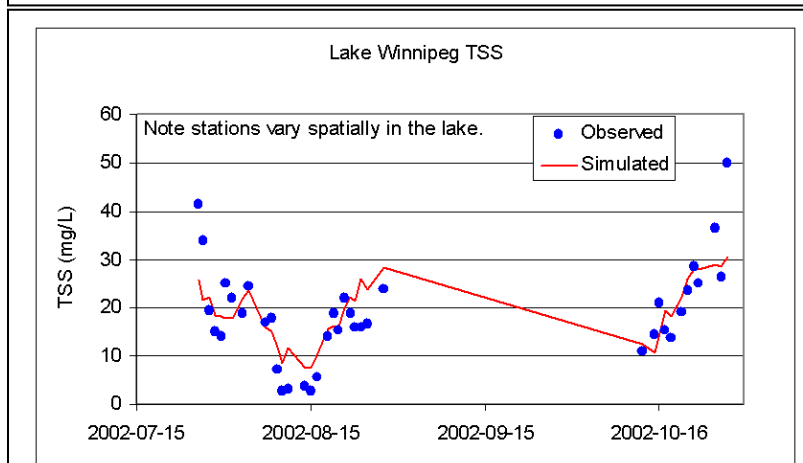
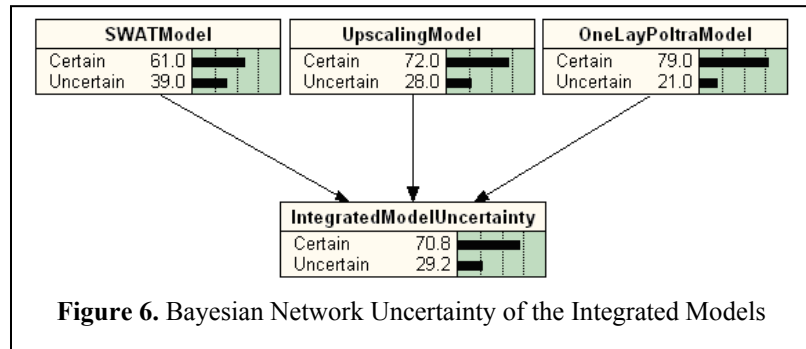


Figure 5. O/P Simulated and Observed TSS.

The uncertainty propagation of the integrated model is achieved by activating the network. In this case, the

overall uncertainty of the system is 29.2% (Figure 6). The estimated uncertainty is reasonable given the lack of input data to the SWAT and O/P models. With improved monitoring data, it is expected the uncertainty propagation of the integrated models should be reduced.



5. DISCUSSION

A two-way calibration of land – lake models was made fully automated using OpenMI combined with a genetic algorithm. The back and forth calibration has the advantage of improving the overall results of the entire system, trying to minimize the uncertainties of each model. Also, manual calibration of such a system would be impractical. A disadvantage of this system is that, on an individual basis, the results of a single model would likely be better if they were calibrated by themselves.

There is an uncertainty associated with the linking of the SWAT, Upscaler, and O/P models. SWAT has the normal uncertainties associated with input data and calibration data gaps but additionally has uncertainty related to the important spring snow melt which SWAT was not originally designed for, which affects flow and subsequently TSS simulations. Another cause of uncertainty is possible backwater affects from the Red River because the La Salle River has a flat slope at the downstream end. The backwater would affect the observations used for calibrating SWAT. The Upscaler is probably the weakest link in the integrated models since it uses a simple regression and is coarse in that it uses modelling of 1% of the Red River basin to scale up for the whole Red River. The La Salle watershed has its characteristics which undoubtedly differ from the whole Red River. O/P also has uncertainties such as using monthly inputs for loading and therefore is more suitable for looking at longer term trends as opposed to SWAT which can be suitable for shorter time periods such as daily. When we sum up all of these uncertainties, the uncertainty propagation is estimated to be 29.2% based on the Bayesian network approach. Due to these uncertainties of the integrated models, a relative comparison, not absolute, for scenario comparison is more realistic. For example, a land use change scenario would be run to find relative improvements in the system.

The best NRMSE obtained for the two-way calibration of Lake TSS for 2002 was 0.12. This value seems reasonable given the levels of uncertainty but there may be room for improvement, possibly by just increasing the number of generations and chromosomes of the GA, because previous calibration of just O/P achieved an NRMSE of less than 0.1 (Leon *et al.*, 2010).

The TSS simulation in SWAT is dependant on the flow. If more rigorous recalibration of flow for SWAT is done then this would affect TSS, so there is room to improve the results. If there was more data in the lake and for longer period simulations and/or used more detailed, shorter time period Red River loading then the NRMSE should go down. Future work of integrated modelling to identify hot spots in land to reduce TSS in the lake could involve connecting to another model like AGNPS (Agriculture Non-Point Source) (Young *et al.*, 1986). which can be used to do source tracing. Also, splitting the O/P model to simulate the north and south basins separately could improve results as the resolution of the lake model, with 2 km grid cells, is relatively coarse for the narrows between the basins: the narrowest point is modelled with only two grid cells wide.

6. CONCLUSIONS AND RECOMMENDATIONS

We have successfully linked the OneLay and PolTra models together using OpenMI software to enable parallel operation of the models. This has allowed a more complete and accurate calibration of the models than could be obtained when they were run in series as the OneLay outputs were used as PolTra input every time step. Also, the Genetic Algorithm (GA) calibration is automated except for setting the calibration parameter ranges, thereby saving a lot of manual effort.

The SWAT model, previously calibrated for flow, has been calibrated for sediments in order to begin the evaluation of the various BMP's that are being developed by Agriculture and Agri-Food Canada in collaboration with the local farmers to reduce loadings to Lake Winnipeg. The next step is to add nutrients such as phosphorous and nitrogen components to the calibration of each model. Also, implementing SWAT in OpenMI would improve the flexibility and efficiency of the integrated system.

Booty *et al.*, A two-way calibration of the SWAT and OneLay/PolTra models using integrated modelling approach for the Lake Winnipeg Basin

The Upscaler is a coarse method of representing the whole Red River. Future work will increase the complexity by including more sub watersheds, such as the Little Saskatchewan and Boyne rivers to better represent the diversity of land use and land forms. Also, a land use factor to address the whole Red River land use may be incorporated in the Upscaler.

It is also recommended that the O/P model be run for longer time periods to see the longer term impacts of land use changes and to calibrate further for other years where data is available. Possibly the model can be modified to simulate ice cover periods to enable simulation matching time periods provided by SWAT. Also, adjusting the time scales to the lowest common denominator, such as changing the O/P model inputs (monthly) to closer match the output of SWAT (daily) would be an improvement. The next step is to attempt the two-way calibration for nutrients such as phosphorus which is currently being upgraded in O/P.

The OpenMI allows us to further expand to integrated environmental modelling to both watershed and lake models, and beyond. The results of the watershed models will affect the calibration of the lake models, and vice versa. Two-way calibration within this type of model integration is required for overall optimal model calibration.

ACKNOWLEDGMENTS

The authors acknowledge the efforts of G. Benoy, J. Vanrobaeys, Q. Yang, S. Melles, M. Sloboda, D.C.-L. Lam, D.A. Swayne, J. Lawrence, M. Conley, S. Hall, N. Hnatiuk, R. Yerubandi, Manitoba Water Stewardship, City of Winnipeg, and O. Resler who contributed to different aspects of this study.

REFERENCES

- Arnold, J.G., R. Srinivasan, R.S. Muttiah, and R. Williams (1998). Large area modeling and assessment. Part 1 Model development. *Journal of the American Water Resources Association*, 34, 73-89.
- Bourne, A., N. Armstrong, and G. Jones 2002. A preliminary estimate of total nitrogen and total phosphorus loading to streams in Manitoba, WQM Section. Manitoba Conservation Rep#. 2002-04.
- Di Luzio, M., R. Srinivasan, and J.G. Arnold, (2001). Integration of watershed tools and SWAT model into BASINS. *Journal of the American Water Resources Association*, 38(4), 1127-1141.
- Goldberg, D.E. (1989). Genetic algorithms in search, optimization and machine learning. Kluwer Academic Publishers, Boston, MA.
- Jones, G., and N. Armstrong (2001). Long-term trends in total nitrogen and total phosphorus concentrations in Manitoba streams. Water Quality Management Section, Water Branch, Manitoba Conservation, Winnipeg, MB. Manitoba Conservation Report No. 2001-07, 154 pp.
- Leon, L.F, Booty, W., Wong, I., McCrimmon, C., Melles, S., Benoy, G., Vanrobaeys, J. (2010). Developments in the Integration of Watershed and Lake Modeling in the Lake Winnipeg Basin. Proceedings of the International Environmental Modelling and Software Society (iEMSs), 2010 International Congress on Environmental Modelling and Software, Ottawa, Canada.
- Lake Winnipeg Stewardship Board (2006). Reducing Nutrient Loading to Lake Winnipeg and its Watershed. http://www.gov.mb.ca/waterstewardship/water_quality/lake_winnipeg/lwsb2007-12_final_rpt.pdf
- Neitsch, S.L., J.G. Arnold, J.R. Kiniry, R. Srinivasan, and J.R. Williams (2002). Soil and Water Assessment Tool user's manual version 2000. GSWRL Report 02-02, BRC Report 02-06. Texas Water Resources Institute TR-192: College Station, TX. 438 pp.
- Open Modelling Interface (OpenMI): www.OpenMI.org.
- Reckhow, K.H. (2003). Bayesian Approach in Ecological Analysis and Modeling, Models in Ecosystem Science, Chapter 9, C.D. Canham, J.J. Cole and W.K. Lauenroth (ed.), Princeton University Press, New Jersey, 168-183
- Salvano, E., D.N. Flatten, A.N. Rousseau, and R. Quilbe (2009). Are current phosphorus risk indicators useful to predict the quantity of surface waters in Southern Manitoba, *Canada Journal of Environmental Quality*, 38(5), 2096-2105.
- Simons, T.J. and D.C.L. Lam (1986). Documentation of a 2-Dimensional X-Y Model Package For Computing Lake Circulations and Pollutant Transports in Physics-based modeling of lakes, reservoirs, and impoundments. Report on Environmental Effects of the Energy Division, ASCE, U.S. EPA through the Oak Ridge National Laboratory; edited by William G. Gray, 258-308.
- Young R., Onstad, C., Bosch, D., Anderson, W. (1986). Agricultural Nonpoint Source Pollution Model: A Watershed Analysis Tool, Model Documentation, Agricultural Research Service, U.S. Department of Agriculture, Morris, MN.