International Environmental Modelling and Software Society (iEMSs)
2010 International Congress on Environmental Modelling and Software
Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada
David A. Swayne, Wanhong Yang, A. A. Voinov, A. Rizzoli, T. Filatova (Eds.)
http://www.iemss.org/iemss2010/index.php?n=Main.Proceedings

AutoCalibration of Integrated Lake Water Quality Models for the Lake Winnipeg Basin Initiative

Markiyan Sloboda[1], David Swayne[1], William Booty[2], Craig McCrimmon[2], Isaac Wong[2]
[1]Computing Research Laboratory for the Environment, University of Guelph, Guelph, N1G2W1, CANADA, *dswayne@uoguelph.ca*
[2]Aquatic Ecosystem Management Research Division, Canada Centre for Inland Waters, 867 Lakeshore Road, Burlington L7R4A6, *bill.booty@ec.gc.ca*

Abstract: Machine learning (ML) and genetic algorithm (GA) are well known techniques which can be used to calibrate environmental models. This paper investigates the calibration of the 2-D horizontal, vertically mixed lake models OneLay and PolTra using ML and GA routines. A GA was used jointly with the Open Modelling Interface (OpenMI) wrapper approach on a single powerful server. Explicit and implicit gridded approaches with the Probably Approximately Correct (PAC) learning were used as ML techniques. Monte Carlo simulations were used to generate the model input parameters for explicit and implicit gridding. Parallel computing using Shared Hierarchical Academic Computing Network (SHARCNET) was used for explicit and implicit gridded approach calibration.

Keywords: Machine learning, genetic algorithm, Monte Carlo simulation, OpenMI, parallel computing, MPI, SHARCNET.

1. INTRODUCTION

The growing understanding of environmental mechanics combined with dramatic increases in computing power has allowed for more realistic representation and simulation of lake dynamics and solving prediction problems using environmental modelling. These advantages have, however, increased the need for improved model calibration in order for models to make better predictions. In this paper we will be looking into calibrating OneLay and PolTra (Simons and Lam, 1986), as an integrated 2-D horizontal, vertically mixed lake model.

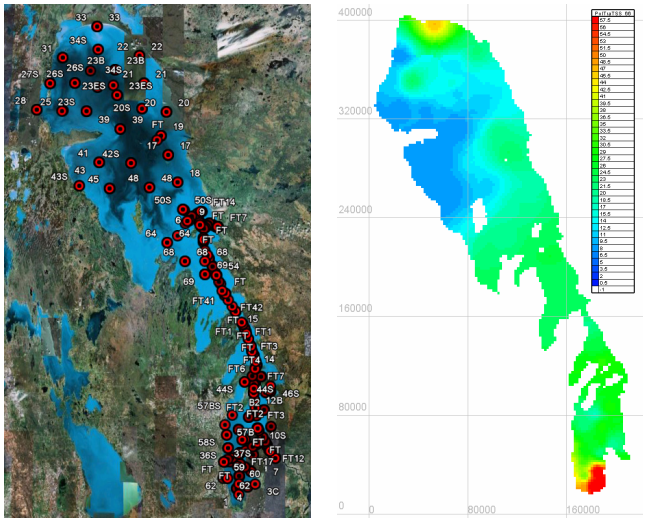
The Lake Winnipeg Basin Initiative has been set up to reduce the loadings of nutrients to Lake Winnipeg to improve lake water quality. This requires the integration of watershed and lake models to optimize the reductions to meet water quality guidelines and protect the overall lake ecosystem. With the Lake Winnipeg Basin covering almost 1 million square kilometers, there is the obvious problem of sparse data, scaling issues and extensive spatial modelling. OneLay and PolTra are water quality models that are being integrated into the overall modelling system. In order for the system to operate efficiently, the models must be calibrated as efficiently as possible (Booty et al., 2009). We describe attempts to calibrate the integrated OneLay and PolTra models against the available observed data from July 26 to October 28 of 2002 as preliminary experiments before carrying out the fully integrated watershed and lake modelling process.

In this paper three different calibration techniques were used and compared to calibrate OneLay and PolTra. The first technique uses a GA and OpenMI wrapper. The second and third techniques use ML with the "gridded" approach. One is an explicit gridded approach (Sloboda et al., 2009) with an application of PAC learning (Russell et al., 2003) with 95% and 99% accuracy and the other one is an implicit gridded approach.

For implementation tests of our techniques and for algorithm implementation and analysis, *PEAS* terminology (*Performance measure, Environment, Actuators, Sensors*) analysis will be used. *Performance measure* is characterized by finding an acceptable calibration parameter set, according to one or more appropriate statistical measures, with a significant speed gain over conventional methods. The *Environment* consists of the parameter space as described by the model developers, the physical process models and the outputs. The *Actuators* are the model's input parameters themselves and the *Sensors* are the measures of the output parameters (calibration parameters) by which the solution's acceptability is determined (Russell et al., 2003).

2. PROBLEM STATEMENT, DEFINITIONS AND METHODOLOGY

The main objective was to calibrate models based on the available observed data. A number of parameters have measurements for spring, summer and fall for 2002 in Lake Winnipeg. Observed values used for calibration were recorded at different time stamps and spatially distinguished locations. Figure 1 (a) shows a map with the stations locations for 2002, where observed data was measured and (b) shows the contoured TSS concentrations for fall 2002.



(a) (b)

Figure 1. Lake Winnipeg 2002 stations locations and contoured TSS concentrations fall 2002

Calibration was done using the GA with OpenMI (www.openmi.org) wrappers for OneLay and PolTra as well as the explicit gridded approach for summer to fall of 2002 for TSS and using 5 calibration parameters: OneLay – DRAG and BFRIC; PolTra – WSET, UCR and KR. OneLay and PolTra were set to produce output every 1.5 minutes.

2.2. OneLay and PolTra

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. Originally the hydrodynamic computations were completed before the water quality model started. The models are based on a rectangular grid representation of a lake. These programs were written in FORTRAN and designed to run in the DOS environment.

Using OpenMI techniques, wrappers were written in C# for OneLay and PolTra. This enabled programs to exchange data at each time step, as opposed to running them in sequence, which improved model integration and allowed integration with other models and data sources to potentially form a whole catchment modelling system. For example, a river or watershed model that empties into Lake Winnipeg can provide their output to OneLay and PolTra as input on the fly as opposed to running the models independently, reformatting the river model output as OneLay input then running OneLay etc.

Parameters that were adjusted for calibrations and their ranges are:

OneLay:

- DRAG (drag coefficient for wind stress); [0.005, 0.005] dimensionless
- BFRIC (bottom friction), [0.01,0.1] cm/s

PolTra:

- WSET (settling coefficient); [1E-08,1E-03] cm/s
- UCR (critical shear velocity for re-suspension); [0,2] cm/s
- KR (re-suspension coefficient);[1E-15, 1E-08] dimensionless

2.3. Genetic algorithm

The GA used for this study is an open source algorithm written by David L. Carroll (www.cuaerospace.com/carroll/ga.html). This program is a FORTRAN version of a genetic algorithm driver. His code initializes a random sample of individuals with different parameters to be optimized using the genetic algorithm approach, i.e. evolution via survival of the fittest. The selection scheme used is tournament selection with a shuffling technique for choosing random pairs for mating.

2.4. Machine learning

One possible alternative to a GA studied in more detail in this paper, is an application of machine learning using a "gridded" approach. If the input parameter space is gridded, it is possible to perform several simulations which are independent of any fitness function or observation. Initial consideration of this approach would rule it out on the grounds of the impossibly large numbers of simulations. For example, for 5 parameters, each uniformly distributed on [0,1], the numbers of simulations that would blanket the whole space, to a single decimal digit, is 10^5. To overcome this problem, we used a machine learning approach with the PAC learning, and accepted that we might miss a solution.

Using the PAC learning, it can be shown that with only 300 simulations we are 95% certain that we have a solution on the discretized space that is at most 5% in error. The total number of examples E required to approximately represent a hypothesis H with probability 1- δ to within a small positive ε has been shown to satisfy

$$E \geq (\ln(H) - \ln(\delta)) / \varepsilon \qquad (1)$$

The proof of the theorem concerning this lower bound can be found in (Sloboda et al., 2009).

2.5. Explicit gridded approach

In general, the explicit algorithm starts from a gridded set of actuators to generate, in the space of calibration sensors, point-sets of values. It differs from sensitivity analysis and from GA approaches in that it decouples calibration from validation and production runs. Using the gridding approach combined with the PAC learning it should be possible to autocalibrate the hydrological model using minimum run time. Monte Carlo simulations are used in this approach to generate the space of actuators, which are used by the model and produced the space of sensors.

The grid of actuators was generated with both 95% and 99% accuracy using PAC learning results. To achieve 95% accuracy it was necessary to run 300 simulations and 99% - 1650 simulations. Five actuators (DRAG, BFRIC, UCR, KR and WSET) were used for calibration. Table 1 displays range and step for each actuator used in the calibration. Figure 2 schematically represents the grid with all the possible values of actuators that can be used for calibration based on the ranges and steps defined in Table 1.

Table 1: Actuators values

Actuator	Range		Step
	Start	End	
DRAG	0.0005	0.005	0.00001
BFRIC	0.01	0.1	0.0001
UCR	0.0	2.0	0.001
KR	1E-15	1E-08	0.01
WSET	1E-08	1E-03	0.01

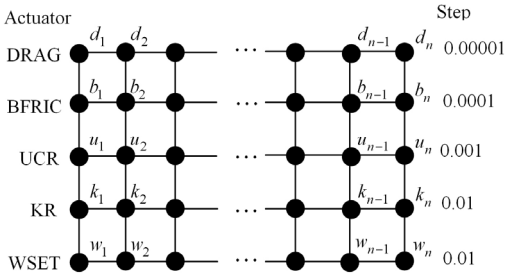


Figure 2: Grid of actuators

In Figure 2 values at each point for DRAG, for example, are:

DRAG: $d_1 = 0.0005$, $d_2 = 0.00051$, …, $d_{n-1} = 0.00499$, $d_n = 0.005$

Values for KR and WSET use a log scale and, for example, WSET is calculates as:

WSET: $a_1 = \log_{10}(1E-08)$; $a_1 = a_1 + 1$; $w_1 = 1E-08$,

$w_2 = 10^{(a_1+step)} = 10^{(a_1+0.01)} = 1.02E\text{-}07$, … , $k_n = 10^{(a_1+n*step)} = 1E-03$

After the grid of actuators is generated the algorithm picks a value for each actuator using uniform distribution. Actuators are not dependent on each other and their values are picked from the grid of values randomly and independently.

2.6. Implicit gridded approach

The other attempted approach to calibrate OneLay and PolTra is an implicit gridded approach. The first step is to generate a large set of actuators and there corresponding simulated values at each time stamp. The next step is to store all the actuators and simulated values into the database. The final step is to find actuator values that best match sensors space, in other words find the input calibration parameters which produce simulated values that best fit the observed values.

2.7. Shared Hierarchical Academic Computing Network (SHARCNET)

SHARCNET is a consortium of Canadian academic institutions who share a network of high performance computers. SHARCNET provides a wide array of high performance computer

systems on a dedicated, private high speed wide area network. In this particular study we used "zebra" cluster (Opteron @ 2.2 GHz, RAM 8.0 GB) and "narwhal" (Opteron @ 2.2 GHz, RAM 8.0 GB) (www.sharcnet.ca). Both clusters support Message Passing Interface (MPI) (www.mcs.anl.gov/research/projects/mpi).

3. RESULTS

The best calibrated result is achieved when following is satisfied

$$\sum_{Time\ stamp} RMSE(Time\ Stamp) \rightarrow \min \qquad (2)$$

In this particular time dependent dataset there are 196 time stamps with observed values against which simulated data was compared at each time stamp. These time stamps were collected from July 26 to October 28 of 2002.

For every 1.5 minute time step, OneLay calculated results which were then input to PolTra, then PolTra calculated TSS, which was then compared to the observed value at a specific time stamp. To run simulations for explicit and implicit gridded approaches on the SHARCNET, the program had to be modified to support MPI and be fully parallelized, in order to speed up the calculation process.

3.1. GA Results

A GA was used as one of our calibration techniques. OpenMI wrappers were written to run the GA and capture the output for each simulation into the database. The GA runs were run sequentially on a server. A previous, similar version of the models was run for over 60 generations and it was found that to successfully calibrate the model required about 35 generations.

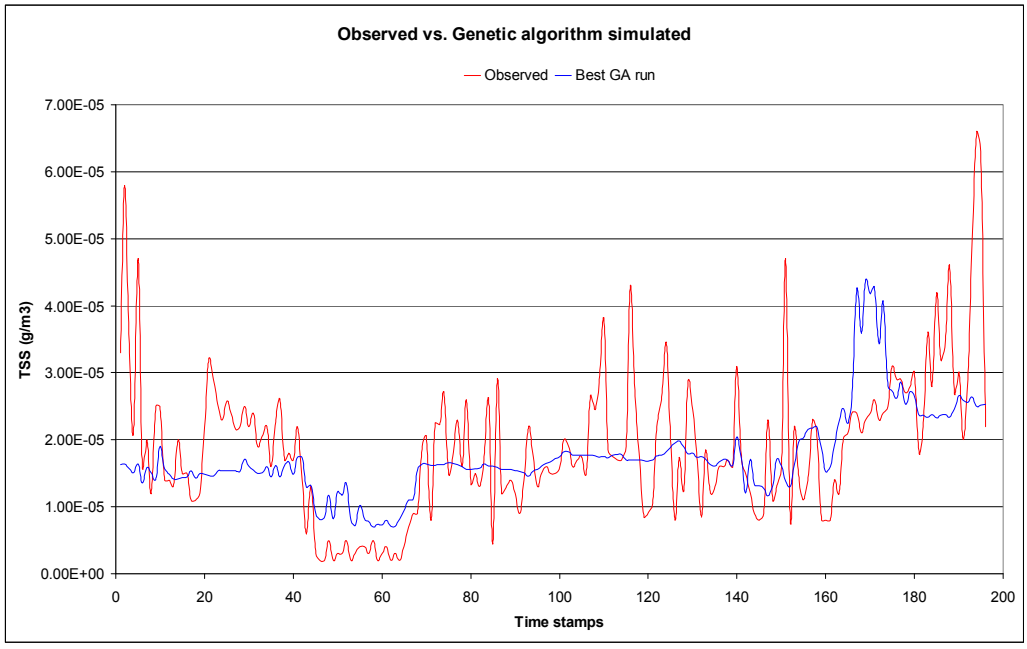


Figure 3. Best calibration using GA

Each generation requires 10 runs of the OneLay and PolTra models. To run each generation on the server (4 Intel Xeon E7330 (quad core) CPUs (2.4 GHz) 64 GB RAM) takes approximately

2 1/3 days. Therefore, to run all 35 generations, requires about 81 days. This speed could potentially be improved by running the 10 runs of each generation in parallel but this was not done for this study. Table 2 shows actuator values and RMSE produced. Using these previous version best GA run coefficients in the new model produces the results shown in Figure 3.

Table 2: Best calibration result using GA*

DRAG	BFRIC	KR	UCR	WSET	RMSE (g/m^3)	r^2
0.021591558	0.043751699	1.89E-10	0.580905995	3.19E-06	9.69E-06	0.237

*Using previous model version best GA coefficients in new model version.

3.2. Explicit Gridded

Table 3 shows actuator values for the best result from 360 simulations, which, according to PAC learning, correspond to the 5% uncertainty. The calibration was done on the SHARCNET taking about 9.5 hours using 60 processors.

Table 3: Best calibration result for 360 simulations

DRAG	BFRIC	KR	UCR	WSET	RMSE (g/m^3)	r^2
0.00476	0.0510	2.04E-10	0.956	2.04E-06	9.74E-06	0.314

Table 4 has actuator values for the best result from 1843 simulations, which, according to PAC learning, corresponds to the 1% uncertainty. Calibration was done on the SHARCNET using MPI and it took about 25 hours using 97 processors.

Table 4: Best calibration result for 1843 simulations

DRAG	BFRIC	KR	UCR	WSET	RMSE (g/m^3)	r^2
0.005	0.0244	1.62E-10	0.578	7.41E-07	9.45E-06	0.285

Figure 4 shows calibration results using 5% and 1% uncertainty.

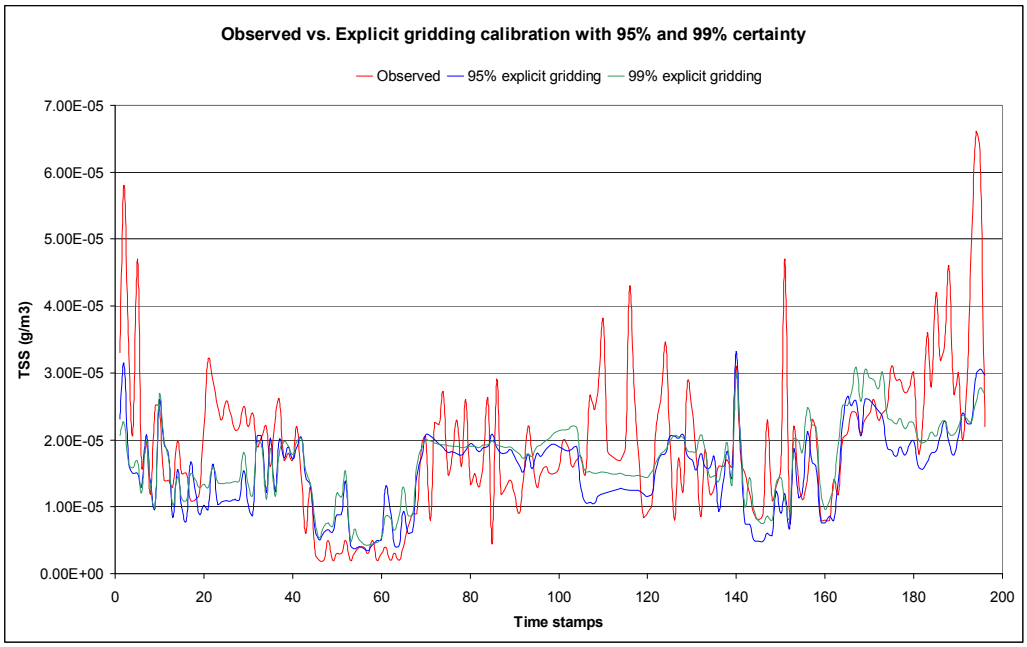


Figure 4. Explicit gridded approach

3.3. Implicit gridded approach

The total number of simulations that were run for this time period is 10201. Therefore there are 10201 distinct actuators which produced a total of 10201 * 196 = 1999396 simulated values, since there are 196 distinct observations at different time stamp and locations. Three tables (observed values, actuators and simulated values) are stored in the database and they uniformly cover the whole space of actuator values.

The first approach that was taken consists of finding the minimum mean square error (MSE) value at each time stamp and recording the actuators that produced that result. Then for each set of actuators which produced the best result at each time stamp, r^2 is calculated for the entire run.

Table 5 displays time stamps, actuators, MSE at specific time stamp, RMSE and r^2 for the whole calibration run using actuators for a specific time stamp using the approach described above. Results are sorted in the descending order of r^2 values. First column, *N*, shows at which time stamp the actuators set produced the closest value to the observed value.

Table 5: r^2 calculated for the entire run using best actuators at each time stamp

N	DRAG	BFRIC	KR	UCR	WSET	MSE at time stamp *N* (g/m^3)	RMSE (g/m^3)	r^2
160	0.00498	0.0117	1.74E-10	1.309	2.69E-06	1.49E-18	1.07E-05	0.404423
38	0.00466	0.0426	2.57E-11	0.144	5.01E-06	8.41E-18	9.84E-06	0.326589
…	…	…	…	…	…	…	…	…
150	0.00057	0.0159	5.62E-09	1.43	0.000912	2.89E-18	1.62E-05	4.42E-07

The best r^2 =0.404423 and is shown in Figure 5. However, if we use a different actuator set for each time stamp with minimum MSE, r^2 = 99.9% and simulated values overlap the observed values with only minor differences, which render to virtually identical graphs. This, however, is an extension to OneLay and PolTra which is yet to be explored.

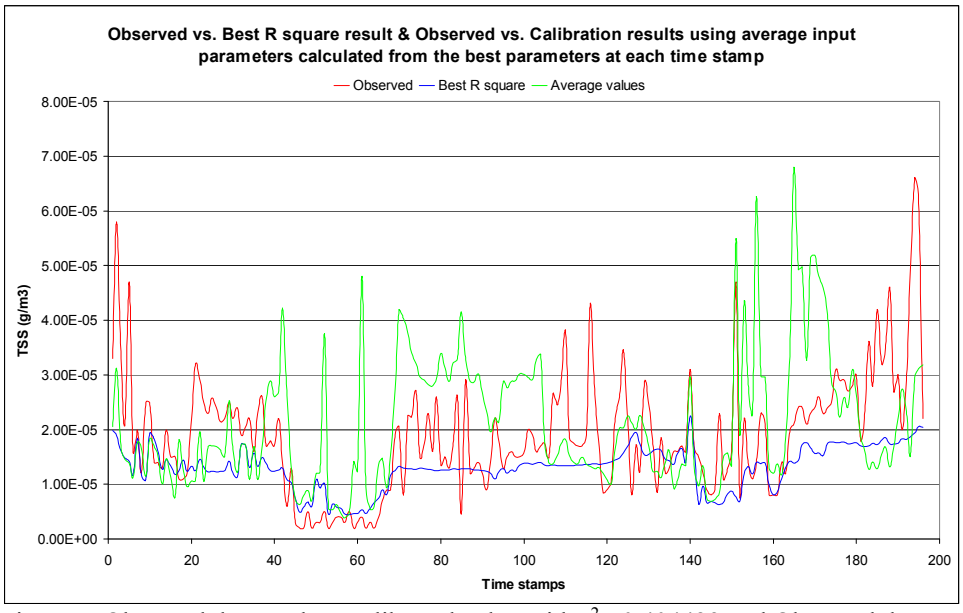


Figure 5. Observed data vs. best calibrated value with r^2 =0.404423 and Observed data vs. average actuators simulated values

The second approach is based on running the models using the average values over all time stamps, for each actuator which produced the best MSE at each time stamp. After average values were calculated, one simulation of OneLay and PolTra was run with these values as input calibration parameters. The average actuators, RMSE and r^2 values are displayed in Table 6. The results are also plotted in Figure 5. This approach produced a negative result (to low a value of r^2)

Table 6: Average actuators calculated from the best actuators at each time stamp and results

DRAG	BFRIC	KR	UCR	WSET	RMSE (g/m^3)	r^2
0.00273617	0.0557107	1.05E-09	1.02745	8.73E-05	1.41E-05	0.060316

4. CONCLUSIONS

It was shown that it is possible to calibrate the integrated 2-D lake model using a GA and an explicit gridded approach. The drawback of a GA is that it requires a lot of time for calibrating the model, and the explicit gridded approach requires a high performance environment or at least a network of a few fast computers. If one compares Figure 3 and Figure 4, it is easy to see that results from the explicit gridded approach fit the data better then results from a GA run. The implicit gridded approach was not appropriate for this particular model, since observation measurement stations are spread all over the lake. The implicit calibration approach would be inapplicable without a large number of observed values within the calibration period at the same station location.

5. REFERENCES

Booty, W.G., Wong, I.W., McCrimmon, R.C., Lam, D.C.L., and Fong, P. Improving Lake Winnipeg Integrated Environmental Modelling with OpenMI, International Symposium On Environmental Software Systems Proceedings, Venice, Italy, 2009.

Legates, D., and McCabe, J. Evaluating the use of "Goodness of Fit" measures in hydrological and hydroclimatic model validation, *Water Resources Research, 35*. 233-241, 1999.

Russell, S., and Norvig, P. *Artificial Intelligence: a Modern Approach*, 2nd Ed., by Stuart Russell and Peter Norvig, Prentice-Hall, USA, 2003

Simons, T.J. and Lam, D.C.L. Documentation of a Two-Dimensional X-Y Model Package For Computing Lake Circulations and Pollutant Transports in Physics-based modeling of lakes, reservoirs, and impoundments: a report / prepared under the auspices of the Committee on Environmental Effects of the Energy Division of the American Society of Civil Engineers under a grant from the U.S. Environmental Protection Agency through the Oak Ridge National Laboratory; edited by William G. Gray, 1986.

Sloboda, M., Sharma, V., Hood, S., Swayne, D., Booty, W., Lam, D., and Wong, I., Innovative autocalibration techniques using high performance computing, paper presented at 18th World IMACS/MODSIM Congress, Cairns, Australia, July 13-17, 2009.

www.cuaerospace.com/carroll/ga.html

www.mcs.anl.gov/research/projects/mpi

www.openmi.org

www.sharcnet.ca