Changing the Rules for Regulating Lake Ontario Levels

Bill Werick
Werick Creative Solutions

Overview

In September 2007, the International Joint Commission (IJC), a joint U.S. - Canadian organization created by the Boundary Waters Treaty of 1909, is expected to promulgate new rules for the regulation of Lake Ontario water levels. Barring extreme difficulties from public review, the rules would most likely go into effect in 2008. This would be the first time (to my knowledge) that the rules for regulating releases on a major North American water system have been changed in the last 30 years, despite the fact that changes to the rules have been under study on almost every major basin. The IJC used shared vision planning to develop and vet these rules, and the Lake Ontario case study now stands as the most technically ambitious and successful shared vision planning application. This paper outlines how the shared vision planning effort unfolded, and highlights the innovations, strengths and weakness in this particular study.

Background

The International Joint Commission (IJC) issued an Order of Approval in 1952 to build the St. Lawrence River Hydropower Project, including a dam across the St. Lawrence River that allows the IJC to regulate Lake Ontario water surface elevations and flows and elevations in the St. Lawrence River. The IJC has since 1963 used a written set of regulation rules called "Plan 1958-D", but about half the weekly regulation decisions are considered "deviations" from the plan. These deviations have been necessary for many reasons, most importantly, because the written plan does not work well when water supplies are much less or much more than the 1860 to 1954 supplies that were used to design and test the plan.

In 1993, the multi-year Levels Reference Study recommended that the "Orders of Approval for the regulation of Lake Ontario be revised to better reflect the current needs of the users and interests of the system." That study did not address environmental impacts, a use of water not identified or explicitly protected in the treaty, nor did it precipitate a consensus on how the current needs could be addressed while protecting traditional uses. In April 1999, the International Joint Commission informed the governments that it was becoming increasingly urgent to review the regulation of Lake Ontario levels and outflows. A plan of study was endorsed in 1999 and the study began late in 2000.

The original plan of study did not define how plans would be formulated, evaluated and ranked, or how researchers would design their work to fit into an overall evaluation scheme. Late in the first year of the study, I made a presentation of how shared vision planning could be used on this study. The presentation included an Excel model based on a STELLA model developed for the five Great Lakes by Phil Chow and Hal Cardwell of the Corps' Institute for Water Resources. Thereafter the Board agreed that all subsequent planning work would be done using shared vision planning. A Plan Formulation and Evaluation Group (PFEG) was formed soon after, and PFEG began to restructure the study with the aim of linking research, public input and decision making. The PFEG reported to the Study Board and the Study Directors. The original membership was made up of those who had pushed the Board to make the formulation and evaluation of alternatives – planning – an identifiable and managed task rather than a natural happenstance of the technical studies. In some cases, PFEG had to realign research that had already begun and assist in the design of studies that were not yet underway. In other cases, work was well along and PFEG used what was done.

The study

Figure 1, below, shows how the shared vision model (in blue) fit with the study research (in yellow). Research was conducted by seven technical working groups (TWGs), managing water information (the Hydrology and Hydraulics TWG) and six impact areas (coastal, navigation, hydropower, municipal and industrial water, recreational boating and the environment). The blue boxes collectively were the SVM in design and use, not just the STELLA model. The shape of the research and models mimicked in many ways the relationships between the TWGs and the Study Board. For example, all TWG work products had to contribute directly or indirectly to the shared vision model. PFEG had no authority over the TWGs but PFEG advised the Board on how TWG research proposals would or would not support the Board's decision making process. The shared vision planning framework connected decision makers to experts and stakeholders:

- **Experts-decision makers.** Our planning process required all TWGs to conduct research that would support a quantitative connection between water levels and economic, environmental or social impacts. For the hydropower and recreational boating TWG's, this was a foregone conclusion and in fact, work along these lines was well along before shared vision planning was in place. But it took a substantial effort to shape environmental studies this way, and considerable tuning to re-shape the navigation and coastal studies.

- **Stakeholders-decision makers.** We asked the Study Board to hold six "practice" decision workshops to iteratively refine the criteria the Board would use to make its decision. Those workshops were conducted with stakeholders and often with

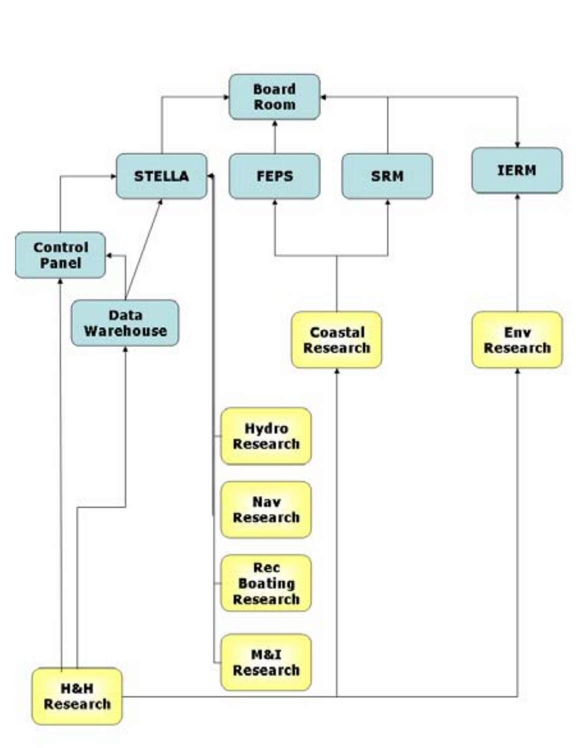

Figure 1. SVM Structure

commissioners present. These “fire drills” helped make sure that the Board understood what stakeholders wanted and helped stakeholders understand why the decisions were made the way they were.

- **Experts-stakeholders.** The study had already allowed stakeholders to participate in technical working groups even before starting the shared vision planning process. In the public TWGs -navigation, hydropower and M&I - stakeholders were represented by paid technical staff; in other impact areas, the stakeholder representatives were not as technically proficient. The shared vision planning process, especially the collaborative model building, had two primary impacts on the expert-stakeholder connection. First, it allowed experts to make sure they understood how stakeholders were impacted. Working with experts and stakeholders, we developed over one hundred hydraulic attributes such as seasonal water level ranges that were used to evaluate plans (especially in the early part of the study, before economic or environmental impact functions were complete). PFEG met with groups of stakeholders around the study area and worked with them to design their own section of the shared vision model that contained the information they told us they would use to rank plans, with tables and graphs they helped design. Second, it gave stakeholders a better understanding of how the impact measurements were linked to water levels, not just in their own areas of interest but also for issues that stakeholders with conflicting interests supported.

PFEG worked with the Environmental Technical Working Group chairs to review over two dozen environmental research scopes and to help establish mathematical relationships between water levels and a biological result. Dr. Joseph Depinto and Mr. Todd Redder of Limno-Tech, Inc. then developed a dynamic model relating water levels to the potential environmental impacts identified in the existing body of research subjects. Although they initially opposed the Integrated Ecological Response Model (IERM), environmental researchers eventually embraced it as their own and in workshop exercises, began to question their intuition when it differed from model results, rather than vice versa.

The model

There was considerable debate about what software to use to build the shared vision model (SVM). The final structure was a compromise that (in retrospective judgment) worked well, but was bent a little too much in the direction of researchers’ preferences. For example, the FEPS model was proprietary and impervious to casual review. PFEG found substantial errors in the FEPS model by close review of its documentation and results, but no one reviewed the code. In the end, many of its processes could have been programmed in STELLA or Excel where review would have been easier. Similarly, the IERM modeled the wetland algorithm in an essentially opaque code. After the study it became apparent that there were small differences between the researcher’s coding of the algorithm, the IERM coding, and subsequent attempts to model the algorithms in Excel. While the mathematical differences were not great and the resulting conclusions identical, in retrospect there are three good reasons for modeling the wetlands in easily accessible code. First, it would have allowed us to resolve small differences between modelers’ interpretations of the English language version of the algorithm. Second, it would have made it easier to use the model post study in adaptive management because it would have been much easier to modify the code. Finally, the argument that convinced people to use C++ during the study was the looping required to calculate non-wetland environmental performance indicators such as the northern pike model. In the end, the pike model meant it took more than an hour to run the IERM but the pike performance indicator did not help distinguish among plans. The final SVM was a system of models, not just one software or file, but all the results were captured in a sophisticated Excel spreadsheet that became the face of the SVM for most study participants. That spreadsheet came to be known as the Board Room. The PFEG led the development of the model, with STELLA and Excel coding being done primarily by Bill Werick and Mark Lorie of IWR, and David Fay and Yin Fan of Environment Canada. A few other agency experts added elements to the STELLA and Excel models. Lay stakeholders sometimes were engaged in modeling workshops, but by their choice, none did any coding. Stakeholders such as David Klein of the Nature Conservancy trusted the models because they were very familiar with the modeling effort, not because they performed it, and because they knew there was no censoring or significant time delay in reporting modeling news. When we found a big mistake, everyone knew about it the next day because the modeling process was very public and the model results were used directly in activities that stakeholders and decision makers took part in.

The planning process percolated through various models in this fashion:

- Researchers developed algorithms connecting impacts to water levels or flows using field data and their own analytic procedures. For instance, stage-damage relationships in the lower St. Lawrence River were developed using GIS that estimated the level of flooding on individual homes at a range of water levels. Information from these models was then used to develop damage functions in the shared vision modeling system.

- Board members, stakeholders, experts in various fields other than regulation plans and paid plan formulators would propose new regulation plans in conceptual terms and then the plan formulation team members would code the concepts. There were four formulation teams that experimented in four categories of plans: modifying the existing rules; optimization schemes; “natural” regulation, and coding of plan concepts offered by others. Each team would use whatever software they wanted to code the rules. The four teams would meet every few months to share successes and challenges; they were competitive but they were part of PFEG and ultimately wanted to see a great alternative produced more than they wanted the alterative from their team to be the best of a mediocre lot. Each team’s model output, a 4,848 quarter-month time series of releases, was then pasted into an Excel model, a part of the shared vision model called the “Control Panel.” That release set defined a unique alternative plan.

- Plan formulation was also used to explore the potential to solve problems, even with plans that would be impossible to implement. “Fence post” plans were also developed, with each fencepost defining a plan that was designed to serve one interest no matter the effect on other interests. These fence posts defined the decision space, and showed the limits of our ability to control water level related impacts. Most importantly, we showed that we could not reduce damages to Lake Ontario shoreline properties much more than we already had. In a similar fashion, we formulated "perfect forecast" versions of alternative plans so that we could quantify the potential benefits of better forecasts.

- Water levels and most impacts would then be calculated in a STELLA model dynamically linked to two spreadsheet input models, the Control Panel and Data Warehouse. After the STELLA model was run, tables from that model would then be copied and pasted into a third Excel model called the "Post Processor." The post processor included macros and tables that could be used to call external models that did the rest of the impact evaluations including Lake Ontario coastal impacts (FEPS), St. Lawrence River shore protection damages (SRM) and the environmental impacts (IERM). Those three models are described very briefly below.

- FEPS (Flood and Erosion Prediction System) is a proprietary C++ model developed by Baird Engineering during previous investigations into Great Lakes erosion and flooding research. FEPS uses water level erosion relationships developed using a very data intensive erosion model called COSMOS at several representative cross-sections around the lake and then applies the results over and over using reach specific parameters around the entire Lake Ontario coastline. Flooding damages are based on water levels and wave heights, capturing both inundation and wave impact damages, and shore protection structure damages are assessed using erosion and flooding models. Erosion at any moment in time is serially dependent on the water levels experienced in the years preceding that moment. Hence, a shore protection structure becomes more vulnerable to damage as erosion eliminates protective beachfront, and it may fail in the eighteenth year of simulation under one plan and in the twenty-fifth year under a different plan. Run time for the FEPS model was about three minutes.

- SRM (Shoreline Response Model) was a proprietary model developed by Pacific International Engineering to assess the effects of different releases on shore protection built along the banks of the St. Lawrence River. Our evaluations showed that all regulation plans being seriously considered had about the same amount of river shoreline damage. Once this was established, this model had little additional relevance in the evaluation process. Runtime was about a minute.

- IERM (Integrated Environmental Response Model) a Visual Basic model, was itself a collection of sub-models. When called from the post processor, the IERM would present a window announcing which sub-model was running. Run time was about 80 minutes on a modern laptop.

While there were four primary formulators, several more PFEG members had the model suite on their (personal?) computers and used it to evaluate models and to check the evaluations other people had done. This work was done methodically and on an ad hoc basis. As an example of the former, a non-formulator might question the results of a formulator and re-run the evaluation checking that all the agreed conditions (for example, was the FEPS model set to use the agreed application of the wave data, did the formulator use the recent revision to Plan 1958DD to define the baseline) were being honored. All of the modeling described above was used to evaluate plans using 101 years of quarter-monthly data. All these evaluations were designed around the 101 year, 4,848 quarter-month structure. When we first tested the alternatives with climate change and stochastic information, we had to manipulate the hydrologic input datasets to this structure. Twenty-nine year climate change datasets had been developed using the 29 years of historic data for which we had enough collateral information, such as precipitation and evaporation, to downscale and interpret global circulation model outputs. We simply repeated these 29-year datasets until we had 101 years. The study developed a 50,000-year stochastic hydrology, and at first we snipped particular 101 year "centuries" from this large data file to form four 101-year quarter-monthly datasets that represented extremes in the stochastic data, and put these snippets in the Data Warehouse spreadsheet so they could be used in this same way to evaluate plans with alternative hydrologic assumptions. Later, a full stochastic analysis using 495 101-year sequences was also done using FORTRAN code translated from STELLA equations and a variation of the FEPS code.

The four plan formulation teams compared results and benchmarked each others' plans, both over the internet and in face to face workshops. This developed a rich understanding of how the system worked, and allowed us to share breakthroughs wherever they were made. Stakeholders had complete access to these sessions, and while few took part in them, stakeholders who did take part helped spread news of plan formulation, and this helped people trust the process. Hundreds of alternatives were tested with the historic evaluations, which could take from two minutes (STELLA only) to ninety-minutes (STELLA, FEPS, SRM and IERM). The evaluations produced economic benefits as traditionally calculated for navigation (shipping cost changes); coastal (changes in expected damages); recreational boating (changes in the value of recreation-day values); hydropower (changes in the value of energy at marginal market based rates) and municipal and industrial water (changes in operating costs). The environmental impacts of each plan were calculated as the ratio of the score achieved by an alternative for a particular parameter to the score achieved under the current regulation plan (in Corps parlance, the "without project" condition). For example, the wetland model produced the acres of meadow marsh present each year after a specifically defined low water supply condition. Those acreages were averaged over the entire 101 year run for each alternative and then divided by the number of acres produced by Plan 1958DD, the baseline plan. In addition to the performance indicators, statistics on over one hundred stakeholder designed "hydrologic attributes" was calculated for each plan and displayed automatically in the Board Room in both central locations with each or all attributes, and in "Interest" corners designed based on focus group like meetings with several stakeholder groups, For example, the navigation industry had a place in the Board Room with graphic comparisons they helped design of the hydrologic attributes they said they would use in deciding which plans were their favorites.

The full stochastic analyses took over a day of computing time to run and these runs were done only for plans that were of particular interest. But the final economic benefit analyses were based on discounted values using the full stochastic evaluations. The discounting captured the reality that erosion happens no matter the regulation plan, so the only difference was how fast it happened (plans that slowed erosion down had positive economic benefits). Had we simply discounted damages using the 20th century "historic" hydrology, the differences between plans would have been muted and distorted, since the wettest and most damaging period came in the last three decades of the century. Instead, the stochastic version of the model recorded damages for each quarter-month of the 4,848 quarter-months in each of 495 101-year "centuries" and so was able to produce an average expected damage for each quarter-month into the future. These average damages were then discounted. A sensitivity analysis allowed various planning horizons and interest rates, but the final report was based on 4% discount rate and 30-year evaluation period. Figure 2, below, shows that Lake Ontario water levels could be nearly three feet higher and lower than recorded levels even under the current regulation plan, which seeks to compress lake level variation.

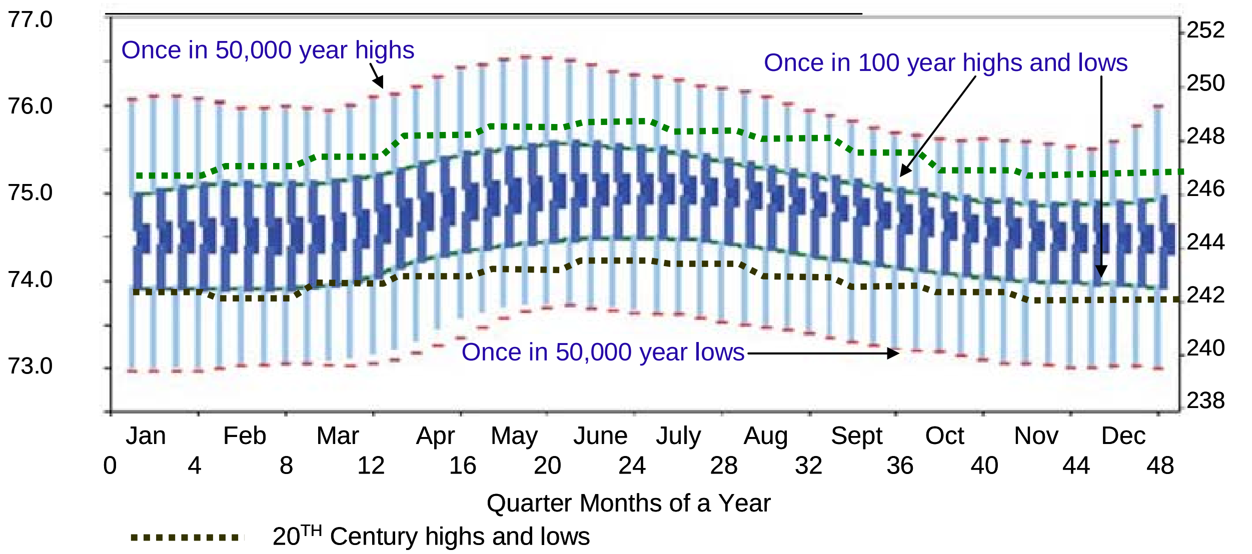

Figure 2. Range of possible Lake Ontario levels through the year under the current regulation rules

The Essential Conflicts

The IJC receives a fairly reliable stream of complaints from some stakeholders because they live or keep their boat in a place which cannot be made satisfactory through regulation. As is true in most places, people have built along the coast based on recent water levels, not on the inevitably higher and lower levels that will come after building. There are a few hundred homes along the Ontario and St. Lawrence coast that will receive at least nuisance flooding no matter how Lake Ontario is regulated. Similarly, there are a few hundred boat slips that will not offer enough draft when water levels are merely normal. This was probably exacerbated by the generally high levels in the last few decades, which coincided with the increase in boating ownership and use. On the other hand, drought management plans that held water on Lake Ontario as long as possible worked for people around the lake and along the river; large short term releases to create normal depths in the river often hurt people along the river because those releases drained Lake Ontario so much that severe release restrictions were needed when natural flows were even lower.

The main conflict that could be affected by regulation was between shoreline property damage and wetland plant diversity along Lake Ontario. Compressing lake level variations helped property owners but created a narrow band of transition between submerged and upland species. There is also a conflict between coastal damage above and below the dam. The damage risk on the river is by far the greatest when winter ice and snow in Quebec melts. If Lake Ontario is high at the same time, the release decision must balance the near certain river damage from higher releases against much larger potential damages along the lake if wind storms occur while lake levels are high.

The Results

The IJC asked the Study Board to provide options, not one recommendation for a new regulation plan. In their final report, the Board gave the IJC three regulation plans labeled A+, B+ and D+. Of all the plans that met the Study Board requirements, A+ maximized economic benefits, B+ maximized environmental benefits, and D+ minimized sectoral losses. All three plans created millions of dollars per year in net economic benefits, but Plan B+ created more positive and more negative benefits. The implementation costs for any of the plans will be relatively small, with all plans about the same, so no benefit-cost ratio was calculated, No plan was found that improved on the current plan in every sector; tradeoffs, sometimes fairly small, seemed unavoidable. We tried but could not reduce coastal damages from the B+ plan; it would cause an average of about $2.5 million per year in damages, an average created by no damage in most years but tens or hundreds of millions of dollars of damage every 20-30 years. We showed that we could eliminate these damages with perfect forecasting in the fall of local spring runoff into Lake Ontario (that is, not the flow from the Upper Lakes, which is fairly predictable). That creates hope that better forecasting, even if not perfect, would allow us to develop a risk management strategy for fall levels that would keep most of the environmental benefits and not cause more coastal damage than we would expect to experience under the current plan.

Lessons Learned

The Lake Ontario study, in my opinion, is the best water resources study ever done. A six page paper cannot do it justice. But it was not perfect. The study was the subject of some criticism from a National Academy peer review. While I believe most of the criticisms were the result of a lack of communication and the limited time the reviewers had to review this work, we did agree that we had failed to do some traditional documentation and had not communicated our risk and uncertainty analysis well. For example, reviewers asked why we had not used a hydraulic model to calculate water levels in the river rather than the regressed stage-discharge curves, and why we did not apply an uncertainty distribution to each element in the stage calculation so that a Monte Carlo analysis could calculate the probability distribution around the estimated stage which could then be applied probabilistically to stage damage curves. These were questions that had been addressed by the study's hydraulic engineers before the methods we used were selected. The relationship between the release decision and subsequent elevations in the St. Lawrence River was quite complex, dependent on channel roughness, tributary inflows below the dam, tides and wind. Yet the hydraulic engineers had forty years of data to base their regressions on, and could vary those assumptions in a Monte Carlo like manner in the stochastic simulation over 50,000 years, which varied not just the releases but the roughness factors and tributary flows. We were also able to compare our results to forty years of actual results – we were not building anything, so there was no need to speculate how these things would interplay together as there would be had we been channelizing or building levees. But the study team did not expect the review team to be as interested I the hydraulics as the peer reviewers were, and had not prepared a justification for their methods (although they had had to defend the methods before the Study Board, outside consultants and stakeholders before getting funding. The peer reviewers made their comments without discussing the issue with the study team and without any document that would allay their concerns. We've taken steps to address that since.

In retrospect, I believe the FEPS and IERM models could have been modeled in Excel or STELLA. Subtle misunderstandings and deeply buried errors in the FEPS modeling – an otherwise impressive modeling effort – caused a complete mid-study shift in plan evaluation, as we found that alternatives that seemed promising based on the flawed FEPS evaluations but were not when properly evaluated. And while there were iterative models in the IERM that could not have been done efficiently in STELLA or Excel, these sub-models did not play a significant role in decision making. The wetland plant diversity model, which was crucial in decision-making had been modeled successfully in STELLA. At study's end, we found discrepancies between the IERM and the original model developed by the biologist who had developed the algorithm linking water level sequences to plant diversity. The differences were small, caused by subtly different interpretations in code of concepts like "the highest three quarter-month elevations during the summer." After the study was completed, shared vision modelers developed Excel coding that we believe is faithful to the concept, but it has not had the broad review and endorsement it would have had if it had been done during the study. Throughout the history of shared vision modeling, I have never replicated a black box model in STELLA or Excel without finding errors or misdirection in the black box. Self-interested modeling technicians are quick to condemn the admittedly limited ability of STELLA and Excel to do iterative "do-loop" logic, but this has to be balanced against the fact that most black box models rarely get one good peer review, while the typical shared vision model is closely scrutinized by dozens of people and is used so often that mistakes are more prone to show up and get cleared up.

Finally, the Lake Ontario study gave us a chance to implement the “informed consent” decision process, a formalization of decision steps that grew out of experimentation with decision processes in shared vision planning studies over the last dozen years. I think the Lake Ontario informed consent process will serve as a model for future studies, although it might need to be scaled down for shorter studies with more modest budgets. The informed consent process is premised on the notion that decision makers, experts and stakeholders must practice the decision from about mid-study, using available information. Each practice decision on this study was observed and criticized both by the Study Board itself and by a consultant, Frank Lupi of Michigan State University.

Conclusion

We will not know for months whether the IJC will be able to implement the new regulation plan, but I am optimistic that this will happen. It may seem faint praise to say that a study this long and expensive and recent is the best ever done, but I think it is a useful provocation to make. For me, and I hope for others, it will serve as a standard against which we measure the effectiveness of the processes and tools of our planning and dispute resolution trade.