International Environmental Modelling and Software Society (iEMSs)
2012 International Congress on Environmental Modelling and Software
Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany
R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)
http://www.iemss.org/society/index.php/iemss-2012-proceedings

Model Management for watershed practitioners: Options for a rural watershed in Ontario, Canada

Thorsten Arnold
Drinking Water Source Protection, Saugeen, Grey Sauble, NBP Region
thorsten.r.arnold@gmail.com

Abstract: Watershed management (WM) organizations use quantitative numerical modeling as a tool to evaluate environmental measurement data, to predict the impact of environmental change on the existing system, and to assess the impact of changes to the existing water cycle, for example after engineering projects. This paper makes an argument that model lifecycles, as the period from when the need for a new numerical model was identified, until the last time that a numerical model is accessed, have changed dramatically when the paradigm of WM has started integrating multiple perspectives and adaptive practices, while model lifecycle management has not received sufficient attention. As a result, WM organizations are struggling to apply numerical modeling efficiently and cost-effectively under their obligation to translate adaptive management into praxis.

Model management can be defined as the process of organizing model maintenance. It includes the setting of standards on how consultants deliver models to practitioners, ensures that relevant knowledge is available to watershed practitioners, and minimizes requirements for technical knowledge, for example by automating data processing with an improved human-model interface. We demonstrate the use of a visualization and software tool for model management, as applied within the Drinking Water Source Protection program of the Province of Ontario (Canada), which relied heavily on numerical modeling for delineating vulnerable areas around municipal water supplies.

Keywords*:* Model management; Adaptive watershed management; Data management; Knowledge management

1 INTRODUCTION

Watershed management (WM) aims at protecting and managing natural resources and their functions today and into the future [Conservation Ontario, 2010]. WM is adaptive, if practices "integrate project design, management and monitoring to provide a framework for testing assumptions, adaptation and learning" (IFAD Guide on Project M&E, http://www.ifad.org/evaluation/guide/annexa/a.htm). Unlike other approaches, adaptive WM accepts that any attempt to completely understand the managed system is bound to fail, partly because of its complexity, partly because any decision is based on multiple and conflicting value judgments, and partly because the system permanently evolves over time. In response, adaptive watershed management relies on continuous, long-term, multi-level learning [Pahl-Wostl 2010] that simultaneously provides safeguards against hasty, short-viewed decisions in a sometimes less-than-perfect political framework.

A paradigmatic shift that integrates adaptive management into practice fundamentally changes the responsibilities of WM organizations. Iterative learning and cyclic project design requires a rethinking of several procedures within WM organizations to overcome institutional barriers [Medema et al., 2008]. This paper focuses on the organisational practices to make numerical models tools for such continuous learning. No scientific literature could be found about how to institutionalize model maintenance over sustained periods of time [Arnold 2012].

1.1 Modeling for adaptive watershed management

Historically, environmental properties are assumed constant and WM responsibilities could be formulated as confined project outputs, for example the identification of a spatial area impacted by 100-years flood events or concentration limits of chemical pollutants. If precise objectives are defined for one-time products, these products could be designed comprehensively. Operationally, transaction cost economics affirms that comprehensively-defined, one-time products can be efficiently provided through a market mechanism [Williamson 2002]. Model building should thus be commissioned to specialized third parties, based on complete contracts.

Under the adaptive management paradigm, modeling products are understood as intermediate summaries of the best science available that most be updated and improved within during a long-term learning process. A model update requires re-visiting and questioning the knowledge embedded in the model, and maintaining only the good knowledge, while revising outdated information. At the time of updating, current staff must access the embedded knowledge in a cost-effective manner, even after staff turnover or other changes within the organization. Hence, the management of knowledge becomes a new task that challenges established practices.

Quantitative numerical models condense local knowledge as well as specialized knowledge from academic disciplines, such as flood hydrology, land surface hydrology, hydrogeology, hydrodynamics of open water processes, forestry, meteorology, and the description of soil-plant interactions. In addition, models embed technical knowledge about the modeling software and its formats, information technology (IT) knowledge about data management principles, and knowledge about the practices established within an organization, such as documentation and storage. In adaptive management, models reflect this incomplete knowledge that increments over time. Access to the knowledge embedded in numerical models becomes a prerequisite for a cost-effective translation of adaptive WM principles into operational praxis.

WM organizations are struggling to apply numerical modeling effectively and cost-efficiently under their obligation to translate adaptive management into praxis. Unfortunately, the practical side of managing models and embedded knowledge cost-effectively has received little attention by academic research as well as in the public management realm. As a consequence, the lifecycle of a model too often ends with the project funding, creating an inherent barrier to the adaption of adaptive WM practices. In this paper, we aim to formulate the challenge of model lifecycle management for adaptive WM, and present a technical solution to model lifecycle management with integrated model visualization software.

Model Management (MM) is the procedures of an organization to manage models, the knowledge and data required along this life cycle. Efficient MM must provide access to relevant software, data and knowledge over the model lifecycle, while minimizing the costs for administration, staff, monitoring, model building, and IT infrastructure provision. Next to technical solutions, MM includes the development, establishment and enforcement of organizational policies, practices, guidelines and procedures (adapted from the International Data Management Association, http://www.dama.org/i4a/pages/index.cfm?pageid=3339 accessed on Nov 29, 2011). What are technical design options to reduce the entry barrier to modeling in order to improve the (re-)usability of models within the adaptive watershed management process?

1.2 Model lifecycle and knowledge requirements

From an organizational perspective, the lifecycle of a model starts with the moment when the need for a new numerical model was identified, until the last time that a numerical model is accessed. Academia and public WM organizations have mostly established practices for creating single-use modeling products. In this paper, it is argued that model lifecycles have changed dramatically when the paradigm of WM has started integrating multiple perspectives and adaptive practices, while model lifecycle management has not received sufficient attention. To better understand the exigencies of model management, we describe the model lifecycle and the knowledge required at each step.

At the beginning of a model's lifecycle, the need for a model is established and its purpose is defined. Then, partners for model building must be identified and the anticipated product must be described comprehensively. If third-parties are contracted as partners, this design typically entails a contract and cost estimates. As lead agency, the WM organization supervises the model building and provides feedback, eventually modifying the model's design. When model building is finished, model building products must be finalized at the required quality, and a contract with third parties is closed. The deliverables of the model building (data, software, knowledge, and reports) must then be stored, published, and maintained for future use.

Knowledge requirements during each lifecycle step vary significantly. When the need for a model is established, the watershed manager requires goal knowledge about what this product shall achieve. Goal knowledge may be established by governmental staff as experts, by a legal authority, or together with a wider group of stakeholders. The identification of appropriate partners requires knowledge of the community and how it operates, knowledge to translate modeling objectives into a detailed product description (a request for proposals), and the review of project proposals.

For clear division of labor with project partners, deliverables must be specified in comprehensive contracts. Within organizations, schedules must identify milestones, resources, and responsible staff. Both options require knowledge about how to establish procedures that are legal, technically feasible, scientifically sufficient, and cost-effective.

The supervision of the model building process requires administrative capacity, local knowledge about the study area, access to the best and most recent relevant data, access to model building capabilities, and to technical and scientific knowledge for quality control. Before the model building is finalized, the quality of all model-based products must be endorsed by relevant stakeholders and all reports, data and necessary software products must be handed over. Model management must apply scientific and local knowledge to ensure that this transfer is of good quality, complete, and compliant with local good practices.

After model building, deliverables must be conditioned for storage, maintenance and updating, which requires technical and procedural knowledge. When updating a model, improved or new raw data must be translated into the model input format (pre-processing), the model must be re-run, and model results must be updated from these new outputs (post-processing). Furthermore, model quality must be reviewed, model assumptions may be refined, and new scenarios may be requested within the decision process. Eventually, the model manager requires knowledge for deciding when a model is out-dated and should be replaced with a new model.

1.3 Exigency on software tools for adaptive model lifecycle management

Transparency and reproducibility of model results are new components in the model lifecycle, partly to enhance trust by stakeholders and partly to enable regular updating. Transparency and reproducibility are difficult design criteria for IT architecture, as it must support incremental learning and tolerate staff turnover in a cost-effective manner. In this process, tools required for data processing and viewing can impair transparency and reproducibility, if they (a) are proprietary with third parties, (b) use non-standard software which entails prohibitive learning efforts for WM staff, or (c) relies on commercial products where license costs are prohibitive. Instead, IT architecture should circumvent these three technical barriers to transparency and reproducibility: Proprietary disclosure and prohibitive learning and licensing cost.

One IT design option translates all data into a standardized data format and uses a single front-end for data viewing and analysis from multiple models (Figure 1). Rather than relying on multiple tools for each model and contract, this design gathers data conversion tools within a flexible toolbox that is backed with standards on data formats and documentation, and with coding conventions. Ideally, this one-stop software package would handle a wide range of model structures (regular and irregular grids, finite elements) and data of multiple dimensionalities, in combination with GIS functionalities and options for data processing and analysis. The software package and data standard should be flexible enough for integrating data from any numerical model, as well as new analysis routines that may be contract-specific. Conveniently, the software would offer templates tailored to the WM organization for reporting data and creating maps. It may offer a meta data framework that documents sources of measurement and model data. Version handling could be applied to document data processing routines.

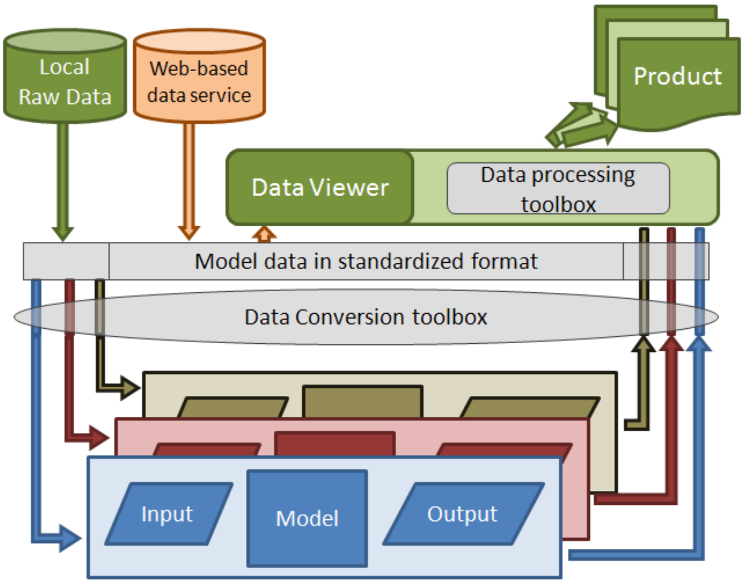

Figure 1. Technical architecture for model management

Model management software must handle enormous amounts of data simultaneously, such that multidimensional time series of 3D outputs can be viewed conveniently, for example from meteorological and hydrodynamic model output at hourly time steps. A single run of a spatial-temporal model can easily create several Gigabytes of output, and users must access several hundred Gigabytes of model output for multi-scenario analysis. Convenience and stability requirements exceed those offered by conventional GIS software and most high-level languages for script-based data processing used within these packages (SQL, Visual Basic).

Last but not least, as core knowledge management tool, this software package must be financially viable over the long-term, even when watershed managers deal with fluctuating governmental budgets for environmental services.

The exigencies for adaptive model management may seem excessive. However, GIS and database technology has evolved in less than two decades into a de-facto operational standard, and redefined how organizations handle spatially referenced data. While spatial time series of model outputs and spatial monitoring may produce an enormous data quantity, most architecture requirements remain comparable.

2 ONTARIO'S ADAPTIVE APPROACH TO WATER PROTECTION

2.1 Drinking water source protection under the Clean Water Act (2006)

Within a multi-layer approach to ensure safe drinking water, the Province of Ontario has initiated the Drinking Water Source Protection (DWSP) program to protect drinking water sources such as groundwater aquifers and surface water. DWSP was initialized in 2006 with the passing of the *Clean Water Act* ("the Act") and its regulations, which also defined source protection regions as jurisdiction boundaries. Source protection authorities were formed in 2007, and committees of local stakeholders were appointed. In each region, this committee directs staff to prepare (1) terms of reference, (2) a scientific assessment report which identifies vulnerable areas and assesses risk for water quality and quantity, and (3) a plan with policies that address these risks. These three components form the source protection plan, which come into effect if the Minister of the Environment of Ontario approves the plan and sets a date for review.

Technical Rules supplement the Act and prescribe numerical modeling, as generally accepted code of practice, for the delineation of areas where contamination (or the alteration of physical properties) could have negative impact on municipal drinking water sources. The most important “vulnerable areas” are those areas from which municipal wells or surface water intakes draw water. These areas are defined by the time-of-travel to the well or intake. Other vulnerable areas are those where aquifers have little protection against contamination from the surface, and significant groundwater recharge areas. Water budgets, time-of-travel, and groundwater recharge are quantified and mapped using numerical models.

Based on a risk rating, activities located within vulnerable areas must then be regulated through local policies, if the quality and quantity of municipal drinking water is in jeopardy. Policies are developed by a committee of local stakeholders, within the legal scope of the Act and technical scope provided by the Ministry.

In the Province of Ontario, municipal official plans, such as the source protection plan, are typically revised every 5 to 10 years, or as deemed by the Minister according to the Act (Section 36.(1)). These regular reviews will define the framework for implementation and monitoring over the next years (Figure 2). While details on such review may depend on the governing political party at that time, the Technical Rules to the Assessment Report [2009] outline the framework under which the review will proceed. Situations when models should be revised include if (i) data or information becomes available such that existing models no longer reflect good scientific practices, (ii) the physical system changes such that the existing vulnerable area delineation is outdated (especially water use quantity or the water cycle) (iii) the local committee requests additional analysis, for example assessing the risk posed by specific activities during extreme events, or (iv) scientific methods are challenged legally. Such revision should build upon the existing knowledge base and successively improve assumptions where necessary and feasible.

Across the province of Ontario, watersheds are characterized by high diversity with respect to population density and urbanization, remoteness, water scarcity, water quality. Conservation Authorities (“CAs”) are local WM organizations who provide environmental protection services within these watersheds on behalf of local municipalities. Accordingly, CAs vary considerably in their resources, tasks, and in their level of technical staff. In partnerships that are determined locally in each Source Protection Region, Source Protection Authorities are part of or collaborate closely with these CAs.

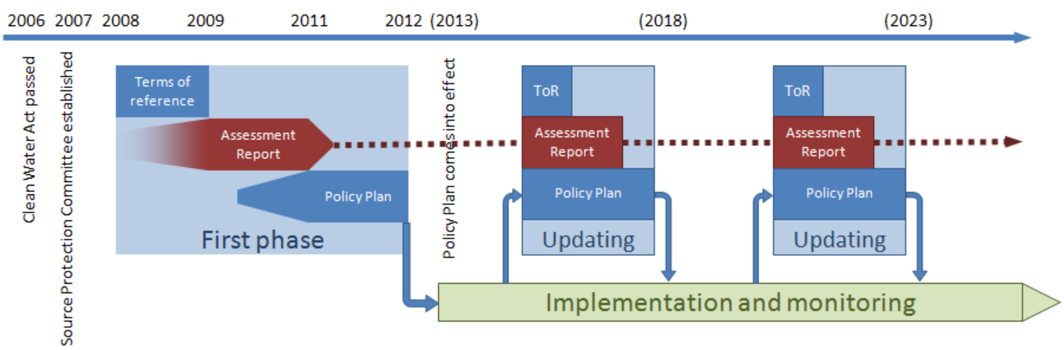

Figure 2 Drinking water source protection as adaptive process

The rural source protection region “Saugeen, Grey Sauble, Northern Bruce Peninsula” is inhabited by approximately 150,000 permanent residents, about half of which are connected to municipal water lines. Source Protection staff hosted regular meetings of the committee and were responsible for technical writing of reports, for stewardship programs, and for communication. Many technical works were contracted to third parties, for example engineering consultants. For this region alone, 55 models were created, including nested Great Lake current models with curvilinear grids and hourly time steps, a response-unit based regional surface runoff model with daily time steps, a regional finite-element groundwater model and local finite differences groundwater models. Key staff are expected to continuously work on implementation, beyond the finalization of the Source Protection Plan. However, many contracts with technical staff will run out and knowledge will be lost.

2.2 Model management objectives

Within DWSP, model management objectives are to provide access to software and data, such that governmental investment into modeling will be saved for future use, such that the benefits to the public will be maximized within organizational circumstances. In this way, continuity can be guaranteed, cost for model updating is minimized, and public institutions are prepared to defend their decision transparently, e.g. during potential legal litigations.

Objectives for model management are, inter alia:

- Perform data analysis of hydrodynamic model output, as part of vulnerable area delineation and risk assessment, such as particle tracking (forward and backward) and water quality concentration analysis
- Perform quality assurance of models and comparing outputs with measurement data
- Visualization of inputs and outputs
- Access to full model input data and software
- Provide a framework that minimizes the needs for technical knowledge on data formats, data processing routines, and data structure with respect to models and measurements
- Document and store models, such that future staff can easily access models and the knowledge embedded within these models

2.3 The SDA software

As a tool for model management, the Saugeen, Grey Sauble, Northern Bruce Peninsula Source Protection Region has adapted the Spatial Data Analyzer© (SDA) software [Lu and Arnold, this conference], which provides many of the capabilities outlined in Section 1.3 and meets model management objectives defined in Section 2.2. SDA was developed by Baird & Associates for the analysis of high-dimensional model data. It has since been improved into a model management tool used for multiple applications within the company, and has been marketed outside of this company. It provides a generic platform for the visualization of GIS data, multi-dimensional spatial time series (“temporal data”), and tracking data. Shown in one or multiple views, SDA offers convenient and simultaneous visualization of measurement stations and dynamic model outputs in 1D, 2D and 3D. Using a time slider, multiple model outputs can be visualized at impressive speed, even with hourly 3D model outputs with high spatial resolution and hourly measurement station data that covers decades (compare Figure 3). SDA also offers template-based mapping and the creation of animations.

Embedded in SDA is a data analysis toolbox that – next to data extraction and basic statistic analysis – offers hydraulic frequency analysis, particle tracking, balancing, and the translation of data between varying grids. Using a separate toolbox, several data formats can be imported conveniently: GIS polygons and images with multiple projections, output from several hydrological and hydrodynamic and wave models, NetCDF files, and data from web-based hydrological services. In addition, a Text Import Wizard allows custom definition for importing ASCII files. SDA can access column data provided with GIS polygon files, e.g. for labeling, and data editors. Most functions are well documented in the Help wizard. Overall, the software uses drag-and-drop functions that are intuitive and easy to learn.

Two hydrological models were integrated into this framework for quality control (GAWSER, Section 2.4) and scenario analysis (Delft3D-FLOW/WAQ/PART, not described in this paper), and data was stored in project format such that technical staff has rapid access to review model outputs for future updating decisions.

2.4 The GAWSER model: Quality control and reusability

The regional surface-runoff model GAWSER [Schroeter, 1989, Schroeter et al., 1996] is based on hydrological response units (hru). Model input files, the model software, model documentation, and GIS files for spatial referencing of input and output files were provided by the project consultant AquaResources. The ASCII control file specifies parameters for 21 zones of uniform meteorology, 47 sub watersheds, 84 routing channels and 1200 hrus, and measurement data time series, as boundary conditions and for calibration.
As part of quality control, model outputs were compared with daily and hourly measurements by the Saugeen Valley Conservation Authority (Figure 3). Time series show flow measurement and modeled data at the mouth of the Saugeen River during the spring freshet of 1988 (bottom left in Figure 3). While flow quantities are well-represented, the timing of the modeled freshet lags behind about one week. With respect to the main model objective (providing long-term balances of sub watersheds), model quality is sufficient. However, for model reuse for other purposes, such as flood forecast, observed deficiencies need to be addressed.

Significant staff time was required to gain understanding of the GAWSER control file syntax, and especially for extracting the routing structure and parameters from this file. The ASCII control file was modified manually, in order to obtain model outputs for individual sub watersheds or hrus. The model was re-run and outputs were then translated from ASCII tabular format into SDA data format. Scripts were created for translating model outputs (ASCII tables) into SDA format. Reuse of these scripts requires considerable technical knowledge, such that the translation is a barrier to reproducibility.

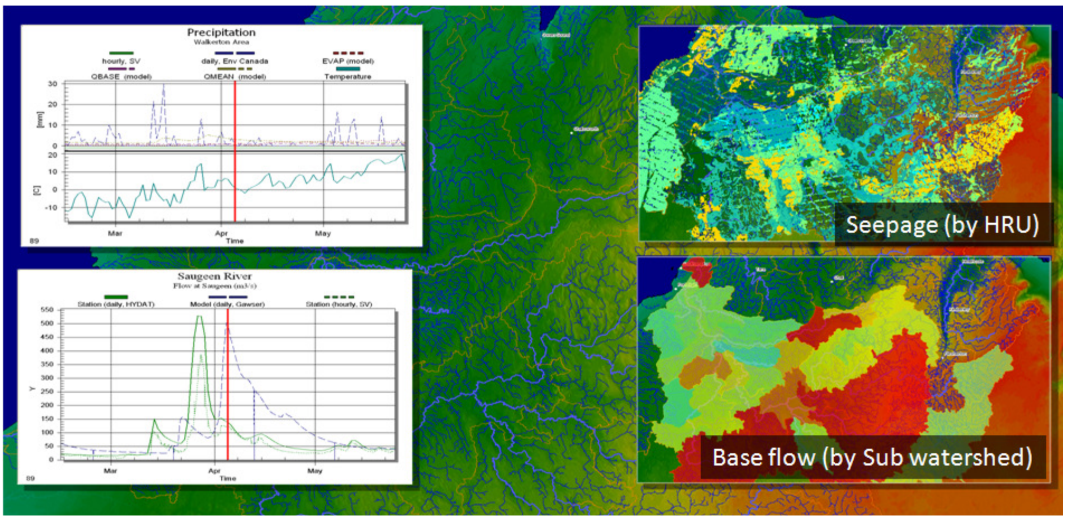

Figure 3 Visualization and quality control of GAWSER model output.

2.5 Storage of model output

SDA projects were embedded within a file structure that documents the full model lifecycle. This cycle includes the initial project proposal, contracts, background studies, input data used for model building, modeling software, model input files, model output files, data processing routines that translate model outputs into SDA data format, and a SDA project which enables viewing of input and output data. If multiple scenarios were created, then each of these was documented and imported into this project. Even though SDA provides no separate framework for file storage, it was found to be a convenient tool for providing rapid access to many data components.

3 DISCUSSION AND CONCLUSION

Compared to the specifications in Section 1.3, SDA meets many requirements but also leaves room for further improvements. While the strengths listed in Section 2.3 could be confirmed, several additions could improve its usability. For example, features that could be strengthened are exporting functions, meta data handling, and several small changes that would reduce the entry barrier for new users. Features that would increase the flexibility of the product include the integration of custom scripts into the import toolbox and the data analysis toolbox; and enhanced visualization options for groundwater modeling (cross-section visualization and carving).

The SDA framework provides a solid technical tool for model management. However, such technical tool can only be one component within the model management strategy of any organization. Institutional buy-in is required at each stage of the model lifecycle: transparent contracting that ensures full delivery of all model components, software selection that is consistent with reuse and avoids licensing traps, budgeting that takes into account longer model lifecycles, tools that minimize the need for technical knowledge, and a broader knowledge management strategy that ensures continuous access to relevant knowledge, provides opportunity for training, and room for learning.

Effective model management remains a practical barrier to adaptive watershed management. This paper provides some underlying thoughts on how to institutionalize MM, and it demonstrates an approach taken for the technical management of model outputs.

4 REFERENCES

Arnold, T., in press. Procedural knowledge for integrated modelling: Towards the Modelling Playground. Accepted for the thematic issue on the future of integrated modelling science and technology. *Environmental Modeling & Software*, 2012.

Conservation Ontario 2010. Integrated Watershed Management - Navigating Ontario's Future. Conservation Ontario, Canada, 2010

Lu, Qimiao and Thorsten Arnold. 2012. The Spatial Data Analyzer© as a model visualization and analysis tool. *International Environmental Modelling and Software Society (iEMSs). 2012 IEMSs Managing Resources of a Limited Planet,* Leipzig, Germany. R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.)

Medema, W., B. S. McIntosh, and P. J. Jeffrey. 2008. From premise to practice: a critical assessment of integrated water resources management and adaptive management approaches in the water sector. *Ecology and Society 13(2),* 29.

Pahl-Wostl, C., Holtz, G., Kastens, B. and Knieper, C. 2010. Analyzing complex water governance regimes: the Management and Transition Framework. Environmental Science & Policy 13 (7), 571-581. doi: 10.1016/j.envsci.2010.08.006

Schroeter, H. 1989, GAWSER Training Guide and Reference Manual - School of Engineering, Univerity of Guelph, 1989

Schroeter, HO, Van Vliet, DJ, Boehmer, K., Beach, D, 1996. Continuous in-stream temperature modeling: Integration with a physically-based subwatershed hydrology model. Ann Arbor Press, INC., Chelsea, MI 48118 (USA). pp. 27-48. 1996.

Williamson, OE. (2002). The Lens of Contract: Private Ordering. The American Economic Review Vol. 92, No. 2, *Papers and Proceedings of the One Hundred Fourteenth Annual Meeting of the American Economic Association (May, 2002),* pp. 438-443