# TOOLS FOR TEACHING HYDROLOGICAL AND ENVIRONMENTAL MODELING

Thorsten Wagener
Dept. of Civil and Environmental Engineering
The Pennsylvania State University
University Park, PA16802, USA,
e. thorsten@engr.psu.edu

Neil McIntyre
Dept. of Civil and Environmental Engineering,
Imperial College London
South Kensington, London
SW7 2AZ, UK
e. n.mcintyre@imperial.ac.uk

## Abstract

Being able to work with mathematical models of hydrological and environmental systems becomes increasingly important for students of civil and environmental engineering departments. Courses or course sections teaching the specific skills needed are slowly making their way into the curriculum. This paper discusses the requirements for tools to teach these topics. Two example tools for model identification and for model evaluation are presented in detail. The paper closes with a section on our experience of using these tools in undergraduate and graduate civil and environmental engineering education in the UK and in the USA.

## Introduction

Hydrology is the science of the occurrence and movement of water in all its forms in the natural and man-made environment. Environmental science is a broader study of environmental systems including and beyond the water aspect. Applied hydrology and environmental science include the fields of water resource engineering and environmental engineering. Common to all these disciplines is their increasing importance to society and the increasing demand for well-educated graduates and trained professionals.[3] Important current issues include the prediction and impact assessment of floods and drought, and the consequences of climate change on water resources and on ecosystems.

Hydrology and environmental science, and related engineering disciplines, are moving towards using an integrated systems approach to consider the complex water-environment-human interaction at space scales ranging from local to global, and at time scales ranging from minutes to decades or even centuries. The main approach to facilitate these complex studies is through the use of mathematical models. These models are simplifications of the real world that attempt to reproduce the relevant behavioral characteristics of the true natural system for the study at hand.

Modeling software is now routinely used in many areas of environmental/ hydrological practice, and building, applying and understanding these models has become an important aspect of research and operational work. Teaching the required skills to use models and computational tools is slowly becoming part of undergraduate and graduate engineering curricula.[15]

At this stage it is important to briefly clarify that the teaching of skills required to properly develop and use mathematical models of the natural world is not covered in standard statistics or time-series analysis classes. The complexity of the natural world, and its complex response behavior including thresholds, non-linearities and heterogeneities, as well as the requirement to predict effects of environmental change, often require models with an explicit representation of the physics of the underlying system, beyond standard statistical approaches. Building reliable and robust models of natural systems requires a solid scientific (qualitative) understanding of natural processes - something to which many engineering students have limited exposure. We will discuss these aspects in greater detail in later sections.

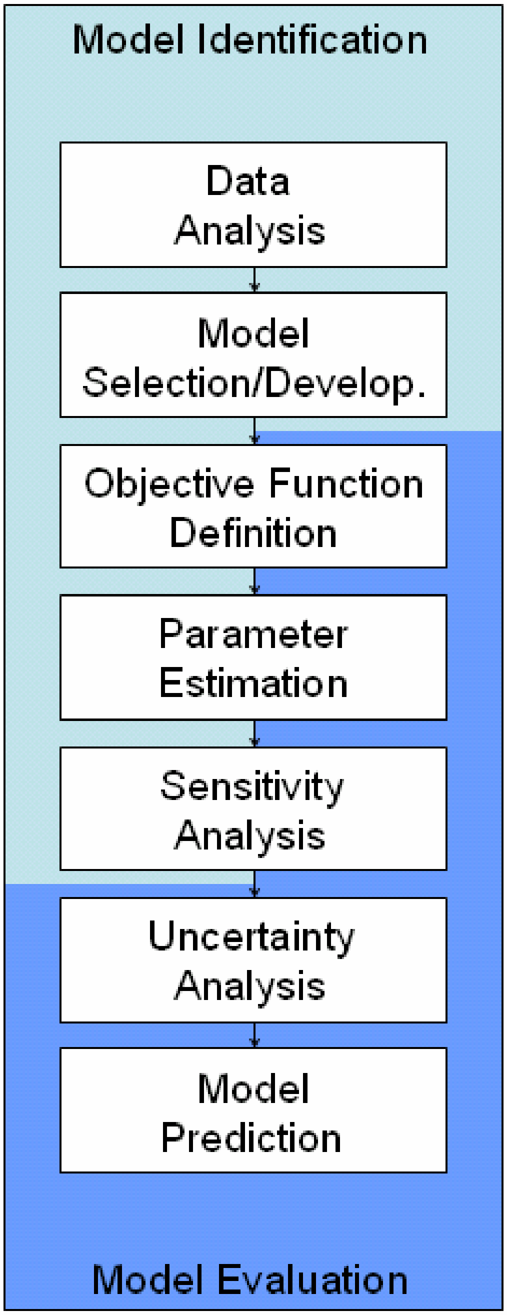
Consequences of the above discussion are that specific courses are required to teach modeling skills, or at least sections of courses have to be devoted to this task, and appropriate tools are required to facilitate this teaching. In this paper we explore these two issues utilizing examples of tools developed and used in both undergraduate and graduate courses in civil and environmental engineering departments in the UK (Imperial College London) and the USA (Pennsylvania State University). The following sections introduce the elements that should be covered in such a class, examples of tools that have been proven to be successful in implementing such a class, and discuss experience in teaching this topic in general.

## Modeling Curriculum/Protocol

The basic of good practice is a proper modeling protocol which is followed during model development, identification, application and evaluation. Reasons for following such a structured and transparent approach include the need for establishing credibility of the model results for (often non-expert) users and for the acceptance of the chosen model.[5] It also includes a thorough assessment of the reliability of the model results in the face of unavoidable uncertainties when modeling natural systems.[1] Limited time and limited resources might sometimes preclude the full implementation of a modeling protocol in engineering practice. However, this only increases the need for students to develop a thorough understanding of the important aspects of a modeling protocol.

In the context of this paper we define a modeling protocol as the *formalized process of applying mathematical models* of hydrological or environmental systems. Different modeling protocols have been suggested in the literature.[14],[5] While these protocols differ in the steps they include and in their order, they nonetheless have common basic elements. Here we focus on these common elements, accepting that there can be differences in the approach chosen. We separate the modeling protocol into two main components: *model identification and model evaluation* (Fig. 1). Model identification deals with the task of finding the appropriate model and parameters. Model evaluation deals with the task of estimating how well the model describes the real world system for the task at hand, and how uncertain the model and its

Figure 1. Flow chart depicting typical steps of a modeling protocol.



predictions are. There are multiple stages or steps within each of these components. Figure 1 shows that some of the steps can be part of either model identification or model evaluation. Assuming that the modeling problem is properly defined beforehand, the following steps should be included:

- *Data analysis* – collecting all available data and performing a quality analysis

- *Model selection/development* – selecting one or more models that could be used to represent the system at hand. This decision is typically based on the following criteria: modeling purpose, available data and specific characteristics of the natural system to be modeled. Sometimes a new model (new code) has to be written if no current model is applicable. There is no objective way of performing this step and it is often based largely on the experience of the modeler.

- *Objective function definition* – the definition of a numerical (or visual) measure of how to measure the performance of the model. This is usually a mathematical function summarizing the difference between observed and simulated system behavior.

- *Parameter estimation* – mathematical models of natural systems have parameters (within the mathematical functions of the model structure) that describe the specific characteristics of the location to be modeled. Examples of parameters include conductivities, storage capacities or decay rates. Many of these parameters cannot be measured in the field and have to be adjusted by matching observed and simulated system behaviors. This process is usually called calibration. In practice, considerable emphasis is usually placed on this stage for achieving a satisfactory model.[2]

- *Sensitivity analysis* – this is an analysis to estimate which parameters are most important to represent the system response correctly. Sometimes, unimportant parameters can be eliminated at this stage to simplify the parameter estimation stage (in an iterative process).

- *Uncertainty analysis* – this is a step of increasing importance since decision makers request more and more that the reliability of the model and its predictions be quantified.

- *Model prediction* – the final step is then to use the properly built and evaluated model for prediction of the variable of interest (e.g. streamflow). This step can include the estimation of uncertainty bounds on the prediction in addition to stating the best prediction only.

The recognition of the importance and intricacy of these stages has generated numerous lengthy textbooks and thousands of scientific papers/discussions. Hence, one can easily see the complexity of educating students on proper modeling procedures within a limited time. Each task involves significant amounts of understanding. Carefully designed lectures, supported by appropriate, easy-to-use tools are necessary to facilitate this.

## General Tool Requirements

General requirements for tools for hydrological and environmental modeling depend on how the tool is used in teaching. An overall requirement is, however, that the tool is representative of general model characteristics and modeling methods, and that it allows the lecturer to teach all stages of the modeling process. One important criterion for tool selection in this context is for how long the students are expected to work with the tool. If this time is short (say, over maximally 3-4 classes) then simplicity and a short learning curve are very important. If the students are expected to work with the tool over a longer period of time (say a semester) then flexibility of the tool becomes more of an issue. We have used the tool (described in the next section) in three different settings:

1. In an undergraduate class (e.g. for a case study or extended homework assignment). Undergraduate students generally have a more limited ability in programming, which means that an easy to understand graphical user interface (GUI) is very important.[16] If the class size is large, an additional aspect is the ease with which students can understand the tool and a good user manual, since spending time with many students on an individual basis is often (unfortunately) infeasible. The software underlying the specific tool should preferably be one which the students encounter in different classes so that they feel comfortable using it.
2. Short-term use of the tool in a graduate class. – The requirements here are similar to those under point 1. The focus for the students has to be on applying the tool, rather than learning how to use it, if only very few lectures are spent on teaching the students how to use hydrological/ environmental models. This requires a good GUI and some pre-formatted data sets that the students can use. In some cases, it may be possible to link the teaching to other parts of the graduate curriculum, for example by employing environmental data sets generated in another class. Furthermore, new modeling skills may be exploited and developed in subsequent research projects or in term projects as part of the class.
3. Long-term (throughout semester) use of the tool in a graduate class. – If the class focus is hydrological-environmental modeling for a longer period of time than the students will often be expected to write code themselves. This could take the form of individual programs, but a better approach – from our point of view – is the cumulative building of a toolbox. This way, new students benefit from the contributions and creativity of students that took the class before. It also allows them to critically review and subsequently improve previous contributions. The requirements this translates into include that the underlying programming language is easy to learn, even for students who had little or no exposure to this particular language. It should probably be an interpreted language (e.g. Matlab) that already includes many functions that can be used. The structure of the tool should be modular so that new contributions can be easily integrated and connected to existing elements. This requires a fixed input-output format for all modules.

Table 1. Requirements for a tool for the identification and evaluation of hydrological/environmental models.

| **Requirement** | **Explanation** |
|---|---|
| Graphical User Interface (GUI) | A GUI enables students who do not have the necessary programming skills to utilize the software nonetheless. |
| Modular structure | It should be easy to extend the software for more capable students who can write additional components. At the same time the software should be cumulative so that it continuously grows. Each module type should have a fixed input-output structure. |
| Based on multipurpose software | The software underlying the tool should be generic in the sense that it can also be used in other courses for programming tasks. |
| Based on easy to learn language | The software underlying the tool should be an easy to learn language with many build in functions to be utilized. The focus should be on gaining understanding, not on the actual programming. |

The requirements discussed above are summarized in Table 1. Note that commercially available modeling software is not generally flexible enough to meet these requirements, in terms of providing insight into the modeling stages (e.g. exploring model uncertainty), and providing a range of levels of use (e.g. access to underlying code as well as simple GUIs). We will discuss how these requirements can be implemented, using examples of tools specifically designed for teaching and research, in the following two sections.

## Example 1 - A Tool for Model Identification

Rainfall-runoff modeling is one of the main tasks in hydrology and models for this purpose often underlie environmental models where water is a key driver of environmental processes. The objective of these models is typically to relate the rainfall falling within a watershed (a topographically defined area) to the runoff leaving this watershed. This could for example be for flood forecasting, for low flow predictions as part of an environmental or ecological study, or to estimate available volumes for water resources planning. The rainfall-runoff modeling procedure is well described by the general procedure given above. A teaching tool for rainfall-runoff modeling is therefore a good example in the context of this paper.

The rainfall-runoff modeling toolbox (RRMT; Wagener et al.,[11],[14]) is a spatially lumped modeling framework developed to facilitate the identification of model structures for research and teaching applications. The RRMT is a generic modeling framework or shell that facilitates the development and implementation of different models in a modular fashion (Fig. 2). RRMT therefore represents a modeling concept, rather than a specific rainfall-runoff model. It has so far been used widely for research applications,[10],[14],[6]) and for teaching in different classes (see examples below).

The RRMT is implemented in the Matlab programming environment therefore providing access to a vast array of already implemented functions and strong visualization capabilities. This environment allows GUIs to be implemented and provides easy access to the rainfall-runoff models and analysis functions (Fig. 3 & 4). Also, Matlab is often taught to engineering students as part of programming and/or numerical analysis classes. Therefore, the Matlab basis allows classroom use of RRMT at a range of levels. For example, at the higher level of complexity, the user can write a Matlab script to automatically load data, changes model settings, run the model and store the results for many different applications (within the script files, all the components of the modeling process can be varied). Another advantage is that all the data related to one case study (incl. time-series data, definitions regarding which modules have been selected, etc.) are stored in a single file. This file is a Matlab structure array which allows for the storing of different types of numerical and textual information in a tree-type structure. The main benefit of this single file storage is that it simplifies the model set-up tremendously, so that the students do not loose any significant amounts of time during this task.

## Example 2 – A Tool for Model Evaluation

The Monte Carlo Analysis Toolbox (MCAT) is a collection of Matlab analysis and visualization functions integrated through a GUI, which extends the capabilities of the RRMT with respect to evaluating model behavior[9], including evaluation of model uncertainties, model performance and underlying model assumptions (Fig. 5). The stochastic analysis of predictions of models of natural systems is of increasing importance in research[8], and in education.[4] The MCAT does so by analysis and visualization of the outputs of Monte Carlo random sampling or a parameter population evolution algorithm. For example, simple ‘dotty’ plots of model response against parameter values allow the important

Figure 2. Schematic representation of suggested tool architecture.

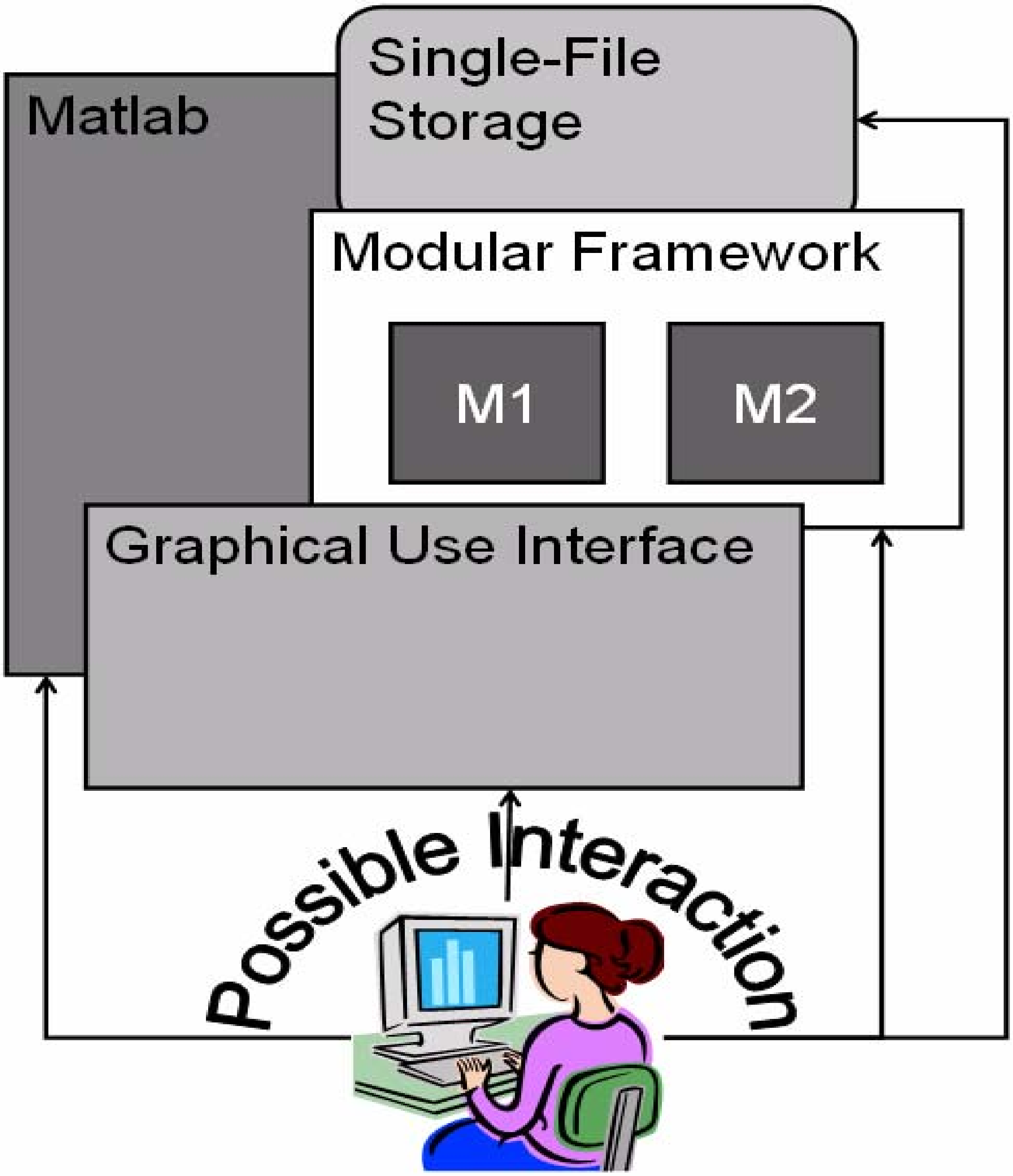
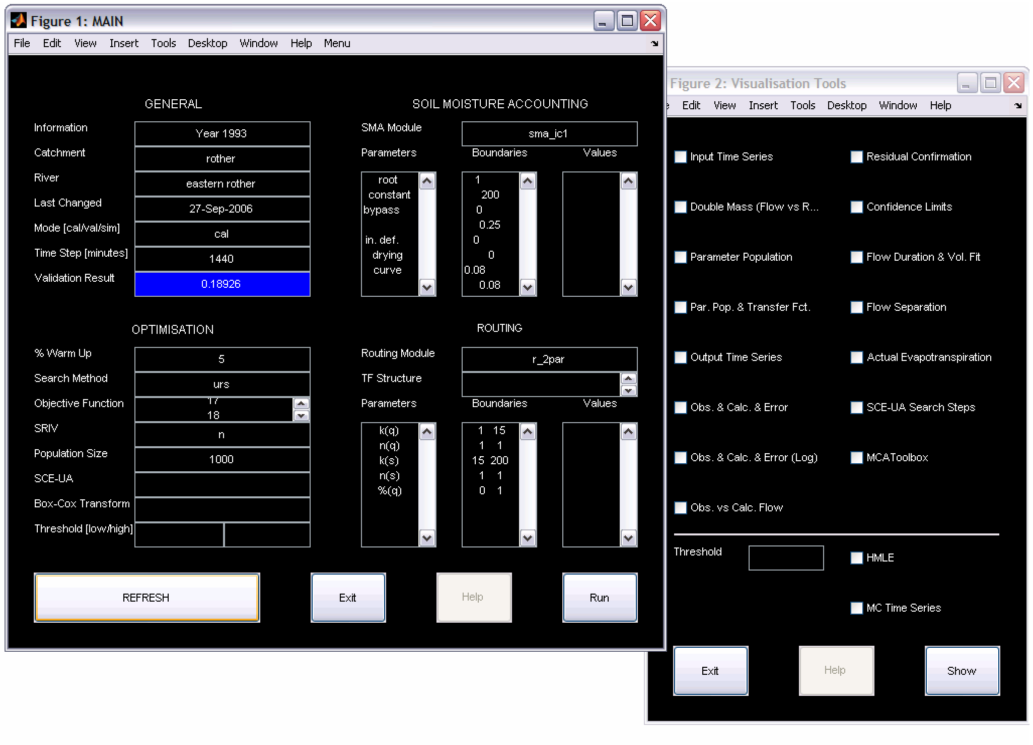


Figure 3. GUI of the Rainfall-Runoff Modeling Toolbox (RRMT) consisting of two parts, the main GUI and a visualization GUI.



Information: Year 1993
Catchment: rother
River: eastern rother
Last Changed: 27-Sep-2006
Mode [cal/val/sim]: cal
Time Step [minutes]: 1440
Validation Result: 0.18926

% Warm Up: 5
Search Method: urs
Objective Function: 17 18
SRIV: n
Population Size: 1000
SCE-UA:
Box-Cox Transform:
Threshold [low/high]:

SMA Module: sma_ic1

| Parameters | Boundaries | Values |
|---|---|---|
| root constant | 1<br>200 | |
| bypass | 0<br>0.25 | |
| in. def. drying curve | 0<br>0<br>0.08<br>0.08 | |

Routing Module: r_2par
TF Structure:

| Parameters | Boundaries | Values |
|---|---|---|
| k(q) | 1 15 | |
| n(q) | 1 1 | |
| k(s) | 15 200 | |
| n(s) | 1 1 | |
| %(q) | 0 1 | |

Figure 4. Example plot of the RRMT showing the rainfall, effective rainfall, observed and simulated streamflow, and soil moisture storage time-series.

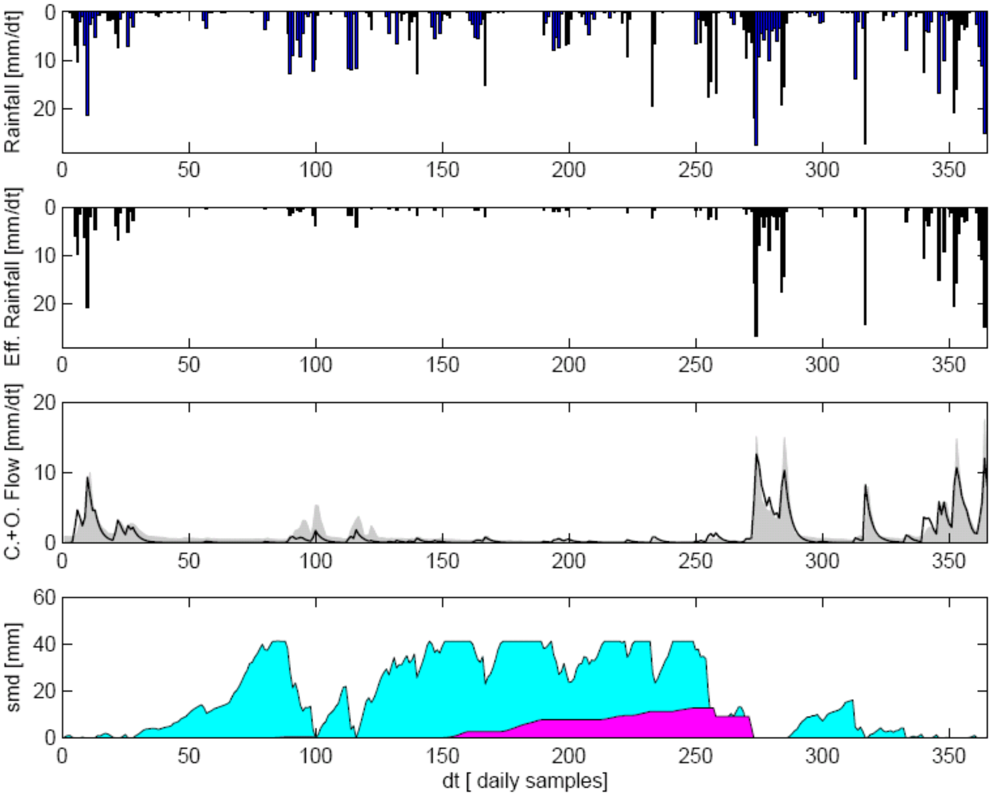


Figure 5. GUI of the Monte Carlo Analysis Toolbox (MCAT).

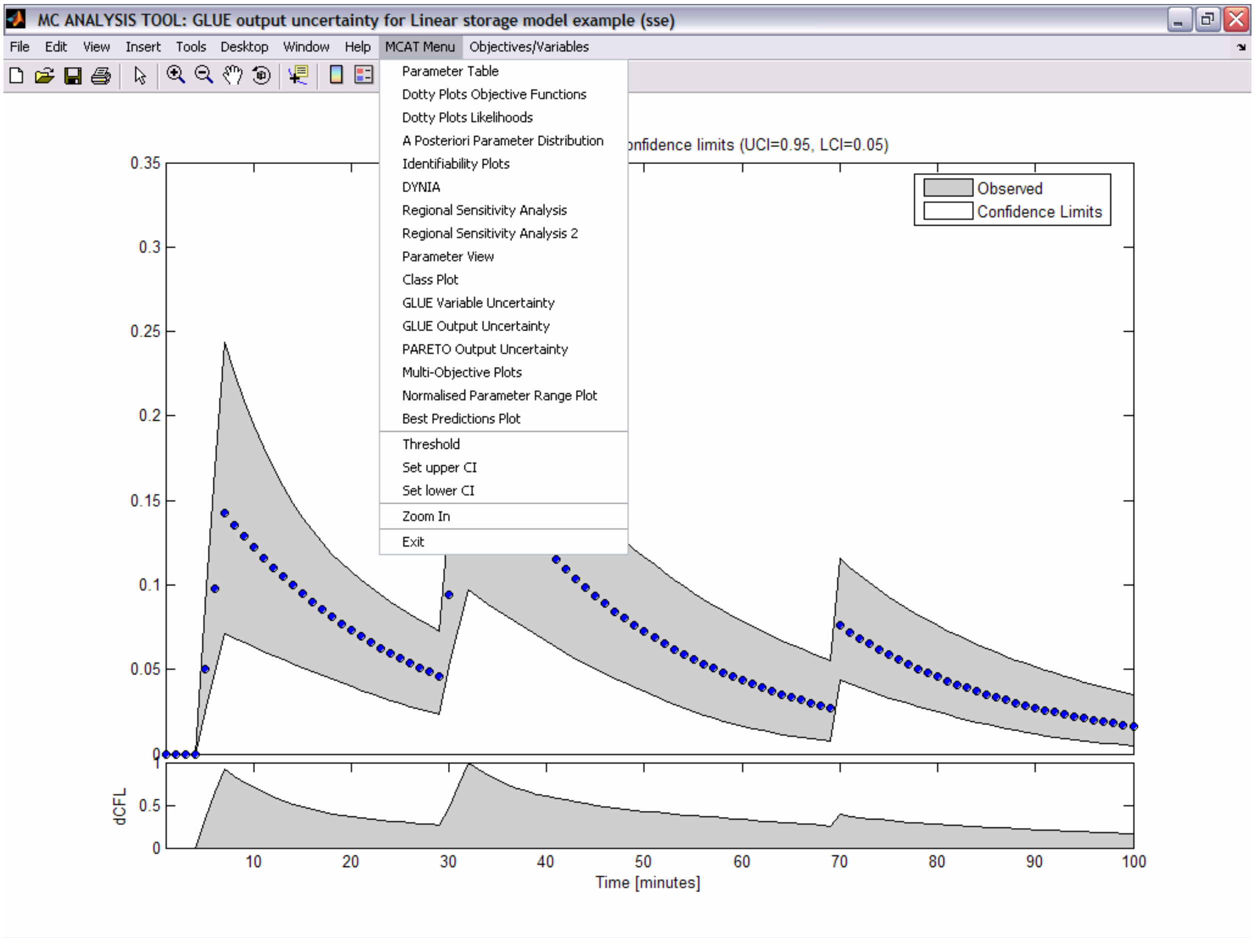
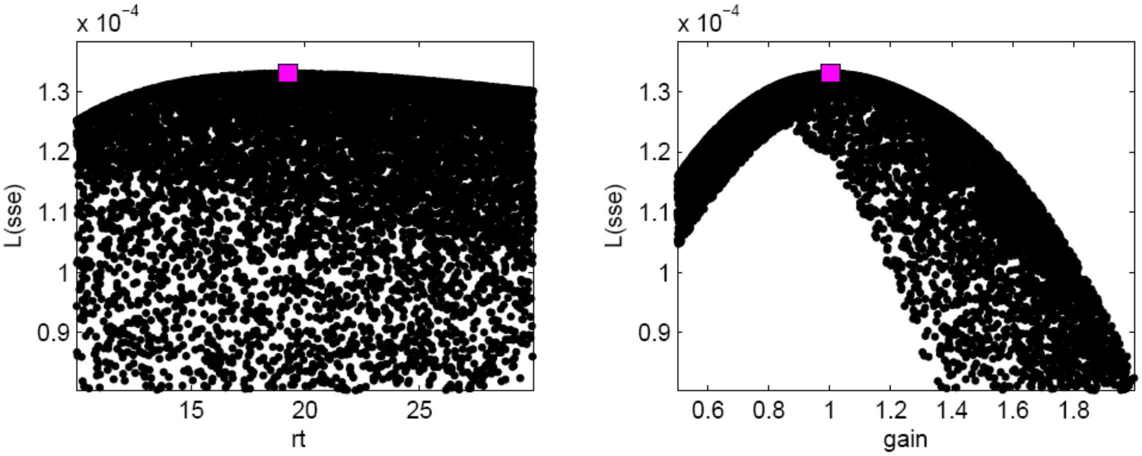


Figure 6. Example plot of the MCAT showing the popular dotty plots as introduced by Beven[1].



parameters to be isolated, and act as the basis for uncertainty analysis [1] (Fig. 6). As well as supporting practical modeling studies, the MCAT enables students to gain insight into model behavior and awareness of how models can be evaluated. Like RRMT, this may be done at a simple level (e.g. review of the dotty plots), or advanced graduate level (e.g. time-series analysis[12], or by adding new functions). While MCAT can be accessed directly from the RRMT interface, it is not specifically designed for rainfall-runoff modeling, but can be applied (off-line) to the evaluation of any dynamic mathematical model.

In summary, the RRMT and MCAT are examples of computational modeling tools, designed to meet the challenges of teaching environmental and hydrological modeling principles, procedures and skills to engineering students.

## Experience with RRMT and MCAT with BSc, MEng, MSc and PhD Students

In the Department of Civil and Environmental Engineering of Imperial College London, RRMT and MCAT are used in both undergraduate (MEng Civil Engineering) and postgraduate (MSc Hydrology) teaching. The MSc Hydrology is a specialist hydrology and water resources qualification, focusing on quantitative hydrology. Towards the end of the 10-week Rainfall-Runoff Modeling module, there are 3 lessons based around RRMT and MCAT, including a 3 hours workshop in the computer laboratory, supplemented by lectures and homework. These workshops are also usually attended by several 1st year PhD students.

The learning objectives of the RRMT and MCAT workshop is to (1) gain experience and understand the context of taught theory; (2) to explore and understand hydrological response and uncertainty via modeling; (3) to develop awareness of key modeling issues through experience with real data; (4) to develop transferable skills and knowledge (in data analysis, computing and environmental modeling). The students download the RRMT and MCAT from an internet site, as well as a written tutorial, and a library of supporting data. The tutorial teaches the use of the software by example, and requires the students to answer a series of questions on model evaluation, through modeling, data analysis and reference to theory. The students mainly use the GUIs, but are exposed to the full potential of the tools using additional Matlab functions.

The MEng Civil Engineering teaching is to final (4th) year students, taking an elective in Water Resource Engineering. These students have a more limited knowledge of hydrological science than the MSc students. The teaching is of a similar format, but only half the number of hours, and teaching objectives focus on understanding the civil engineering significance of hydrological science and modeling (e.g. climate change impacts on water resources and flooding) using the GUIs.

Ambitious students have appreciated the ability to get behind the interface, to explore the flexibility of the tools, to be both investigative and creative. Many MSc, MEng and PhD students chose to use the software in their research projects.

The MCAT and RRMT toolboxes have so far been used in two graduate hydrology courses in the Department of Civil and Environmental Engineering at the Pennsylvania State University. A future application within an undergraduate course in water resources engineering in the same department is planned.

The two graduate courses show how differently the software can be utilized. The first course is a Surface Hydrology course in which the topic of modeling covers only a short section of one week (3 hours). The short learning curve to understand and work with the toolboxes allows for this section to be taught in a computer lab. Students work in groups of 2 or 3 and learn the modeling process by implementing it step by step. The GUIs allow them to almost immediately focus on understanding concepts and methods, rather than having to spend time on understanding how to load data and use a specific software.

The second graduate course is called Systems Approach to Hydrologic Analysis and provides a semester long introduction to using mathematical models to understand and predict the behavior of natural systems. The students move beyond simply executing the model, but are actually trained on implementing the methods discussed in Matlab code themselves. The role of the toolboxes therefore changes to being a framework that can host new modules that the students add throughout the semester. It is possible to introduce new methods on real examples before the students go off and implement the methods themselves. It also allows them to compare their results. An important element of this course is to help the students to be creative and develop their own methods and visualization tools, rather than simply reproducing what others have done.

## Conclusions and Future Developments

The increasing importance of mathematical models of hydrological and environmental systems in engineering practice and in research requires a formalized training on how to build and use these models. Such training is slowly becoming part of curricula in civil and environmental engineering departments. This training should establish a thorough understanding of the modeling process, including an understanding of uncertainties, and model capabilities and limitations. To achieve this objective, it is necessary to have flexible software tools that focus the attention on understanding concepts and procedures, rather than on a particular software. This paper breaks the modeling process into two main components – model identification and model evaluation – describes requirements for tools for teaching these components, and introduces two Matlab toolboxes as examples. The RRMT and MCAT toolboxes have been used in graduate and undergraduate education in the UK (Imperial College London) and in the USA (Pennsylvania State University), and have proven to be easy to use and flexible teaching tools. Both toolboxes are available for download free of charge from the Hydroarchive website at http://www.sahra.arizona.edu/software/index_main.html.[7]

## Acknowledgements

We thank the following individuals for different contributions to the development of the methods, and tools and other teaching material discussed in this paper: Howard Wheater, Josh Kollat, Barbara Orellana, Hyosang Lee, Peter Young, Keith Beven, and Matthew Lees.

## References

1. Beven, K. J. 2005. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320 (1-2), 18-36.

2. Gupta, H.V., Beven, K.J. and Wagener, T. 2005. Model calibration and uncertainty estimation. In Anderson, M.G. et al. (eds.) Encyclopedia of hydrological sciences. John Wiley & Sons Ltd., Chichester, UK, 1-17.

3. Kirshen, P.H., Vogel, R.M. and Rogers, B.L. 2004. Challenges in Graduate Education in Integrated Water Resources Management. *Journal of Water Resources Planning and Management*, 130(3), 185-186.

4. Hossain, F. and Huddleston, D. A proposed computer-assisted graphics-based instruction scheme for stochastic theory in hydrological sciences. *Computers in Education Journal*, in Press.

5. Jakeman, A.J., Letcher, R.A. and Norton, J.P. 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modeling & Software*, 21, 602-614.

6. McIntyre, N., Lee, H., Wheater, H.S., Young, A. and Wagener, T. 2005. Ensemble prediction of runoff in ungauged watersheds. *Water Resources Research*, 41, W12434, doi: 10.1029/2005WR004289.

7. Wagener, T. 2003. Evaluation of catchment models. *Hydrological Processes*, 17, 3375-3378.

8. Wagener, T. and Gupta, H.V. 2005. Model identification for hydrological forecasting under uncertainty. *Stochastic Environmental Research and Risk Assessment*. doi: 10.1007/s00477-005-0006-5.

9. Wagener, T. and Kollat, J. 2006. Visual and numerical evaluation of hydrologic and environmental models using the Monte Carlo Analysis Toolbox (MCAT). *Environmental Modeling and Software*. In Press.

10. Wagener, T., Boyle, D.P., Lees, M.J., Wheater, H.S., Gupta, H.V. and Sorooshian, S. 2001. A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, 5(1), 13-26.

11. Wagener, T., Lees, M.J. and Wheater, H.S. 2002. A toolkit for the development and application of hydrological models. In Singh, V.P. and Frevert, D.K. (eds.) Mathematical models of large watershed hydrology. Water Resources Publications LLC, USA, 91-140.

12. Wagener, T., Gupta, H.V., Carpenter, K., James, B., Vázquez, R., Sorooshian, S. and Shuttleworth, J. 2003. A hydroarchive for the free exchange of hydrological software. *Hydrological Processes, Software Review*, 18, 389-391.

13. Wagener, T., Wheater, H.S. and Gupta, H.V. 2003. Identification and evaluation of watershed models. In Duan, Q., Sorooshian, S., Gupta, H.V., Rousseau, A. and Turcotte, R. (eds.) Calibration of watershed models. AGU Monograph, 29-47.

14. Wagener, T., Wheater, H.S. and Gupta, H.V. 2004. Rainfall-runoff modelling in gauged and ungauged catchments. Imperial College Press, London, UK, 332pp. ISBN 1-860944-66-3

15. Whiteman, W., and Nygren, K.P. 2000. Achieving the right balance: properly integrating mathematical software packages into engineering education. *Journal of Engineering Education*, 89(3).

16. Whitney, R.E. and Urquhart, N.S. 1990. Microcomputers in the mathematical sciences: Effect on courses, students, and instructors. *Academic Computing*, 4(6), 14.

## Biographical Information

Thorsten Wagener is an Assistant Professor of Hydrology in the Civil and Environmental Engineering Department of the Pennsylvania State University. He obtained civil engineering degrees from the University of Siegen (Dipl.-Ing. FH), Delft University of Technology (ir.) and Imperial College London (PhD). His main interests are hydrological modeling, integrated assessment and the integration of learning and research in engineering education.

Neil McIntyre is a Senior Lecturer of Hydrology in the Department of Civil and Environmental Engineering, Imperial College London. He has degrees from the University of Edinburgh (BEng) and Imperial College London (MSc and PhD). He has also spent 6 years working in the water industry in Scotland. Research interests include hydrological modeling, surface water quality, and flood and drought risk evaluation.