Rationale and Strategy for a Community Modeling Platform in the Hydrologic Sciences

Report of the CHyMP Scoping Workshop held March 26-27, 2008
Washington, DC

James Famiglietti (University of California, Irvine)
Lawrence Murdoch (Clemson University)
Venkat Lakshmi (University of South Carolina)
Richard Hooper (Consortium of Universities for the Advancement of Hydrologic Sciences, Inc.)

CUAHSI Technical Report #8

A Report from the Hydrologic Community to the National Science Foundation

April 12, 2010

Table of Contents

This material is based upon work supported by the National Science Foundation under Grant No. 0814193.

The need for community modeling in hydrologic science

The hydrologic sciences community is faced with an unprecedented challenge. The need to understand the highly complex workings of the hydrological cycle and to predict and project its changes has never been greater. Worldwide changes in precipitation patterns, the frequency of flooding and drought, increases in humidity and cloudiness, the decay of snow and ice cover, and rising sea level are now the reality of climate change [IPCC, 2007]. Moreover, climate model simulations agree that these changes will continue, if not accelerate, in the foreseeable future. The problem of an accelerating or increasingly intense hydrological cycle is further exacerbated by population growth. Recent assessments estimate that one-third of the global population already lives in water-stressed regions [Giles, 2006]. Projections of water availability suggest that over 60% will face increased water stress by the year 2025 due to both climate change and population growth [Vörösmarty et al., 2002]. In many locations around the world, the utilization of available water resources will be pushed to the limits. Clearly, a comprehensive framework for observing, understanding, predicting and adapting to changes in the water cycle and water availability is critical for our national security in the areas of health, politics, socioeconomics and food.

An essential component of this framework is a conceptually and technologically advanced hydrologic simulation capability that, for reasons discussed below, simply does not exist at present. Such a system could be used not only to understand and predict water cycle change and its causes, but also to address a number of compelling questions of national and international significance. For example, the following issues could be substantively answered with advanced simulation capability:

- *What are the impacts of changing climate, population growth and land use change on the availability of freshwater resources? Will there be enough available fresh water for the U. S. and global populations in the decades to come?*

- *How can water management best adapt to changes in global and regional hydrology, for example, the decreasing snowpack in the Western U. S.? What are the local- to global-scale feedbacks of new management strategies?*

- *What are the full 'Earth system' implications (e.g. regional climate, ecological and food production changes) for large-scale energy production alternatives that are linked to the water cycle, e.g. biofuels? What are the water requirements of significantly increased feedstock production?*

- *Is increased water storage on land a credible component of a strategy to manage current rates of global sea level rise?*

Unfortunately, hydrological model development has lagged far behind the increasingly urgent need to address these and other similarly complex questions. No integrated modeling framework currently exists that is capable of both simulating all major components of terrestrial water storage and flow (i.e., snow and snowmelt, alpine glaciers, permafrost, lake and reservoir storage, floodplain dynamics, streamflow, soil moisture, groundwater), water management (i.e. irrigation, reservoir operations, water redistribution and transfer), water use and withdrawals, alpine glacier dynamics, urban hydrology, etc., and of assimilating a wide array of *in situ* and remotely sensed data. Instead, simulation tools remain fragmented by and even within hydrologic and related sub-disciplines. This traditional separation acts as a serious impediment to significantly advancing predictive understanding. It is imperative for society that we understand and characterize key linkages, e.g. between climate, land use change, and the water cycle, in order to propose viable and comprehensive environmental solutions. Until this is achieved, our ability to predict and project water cycle change will remain insufficient relative to our evolving needs.

This report documents the rationale for the development, distribution and support of a community-based models and modeling tools for hydrologic science. Essential aspects of the activity are the development of a Community Hydrologic Modeling Platform (CHyMP) [Famiglietti et al., 2008], regional- and continental-scale integrated hydrologic models, strong links to CUAHSI Data Services [Maidment, 2008], and access to high performance computing for model simulation. We envision that the CHyMP will be a platform of component models that can be linked together and implemented across scales, from local scales of measurement and process studies, to the large watershed and regional scales at which. Community engagement, at all levels of the modeling activity, is essential for the success of the effort that we describe here.

Enabling new science

As described in the previous section, a community modeling effort would enable new science, while eliminating the need to 'recreate the wheel' at the initiation of new research projects. The availability of models and modeling tools to the community would allow researchers to spend more time focusing on science, rather than on the repeated development of software that already exists but is not readily accessible. Today's challenging funding environment and the desire to maintain a competitive advantage in grant opportunities do little to ameliorate this counterproductive situation.

Beyond time savings, the availability of a community modeling framework greatly expands the range of questions that can be addressed in its absence. Simply stated, the natural and managed hydrologic cycles are too complex, and require too much data and expertise, to be modeled by individual PIs or small collaborative efforts. The development of comprehensive, integrated hydrologic models, at very high resolution, that exploit the availability of advanced information systems, that can assimilate *in situ* and remotely sensed data, requires a community of model developers and users committed to the development and support of the effort. The resulting community models would enable the next generation of researchers to tackle the pressing issues relevant to today's rapidly changing environment. For example, hydrologic scientists will soon be called upon to address questions at the water-energy nexus. Consequently, our models must have the capacity to accurately predict water availability in space and time in order to characterize the energy required to transport water to end users, and to heat it to the desired temperature. Arguably, this capacity is limited at present. Similarly, as the cryosphere continues to degrade, changes in permafrost, snow and alpine glacier extent all portend rapid changes in high latitude hydrology. Models will be required to accurately simulate these changes, as well as link to models of changing biogeochemistry, such as carbon dioxide and methane emissions in emerging Arctic lakes. As with the previous example, our ability to reliably simulate high latitude hydrology, and to make current models available to the community, is limited. Other examples of new science are those related to emergent properties, as spatial and temporal scales change, at the interface of hydrologic zones (e.g. land-biosphere-atmosphere or surface-groundwater), or resulting from interactions across disciplines (e.g. hydrology and ecology, biogeochemistry or climate). A community effort that greatly enhances the scope of the hydrologic and related Earth system processes that can be modeled, and that makes these model advances available to the community, is a critical step in the evolution of hydrology as a distinct geoscience.

We can do it today

The advances in the computer hardware and software technology have resulted in numerous advances in hydrological modeling. Hydrological models can now run in real time at very high temporal and spatial resolution for global spatial and many decades (50-100 years) of temporal coverage.

*Advances in computer power***.** In terms of computer power we have seen tremendous increase over the past 2 decades. To put in plain terms, a typical laptop today when compared to a University mainframe from 1975 is 800 times cheaper, 10,000 times physically smaller, 750 and 3000 times more memory and hard disk respectively and 3000 times faster. In the late 1980s, computing was carried out using mainframe computers for larger and more computationally intensive algorithms whereas the less intensive algorithms were run on desktop workstations and few on personal computers. In the 1990s, there was less dependence on workstations, large codes (such as general circulation models) still used mainframes and/or supercomputer centers but the personal computers became faster and its use was much greater in a wider range of hydrology. In this decade, with the exception of very large simulations, personal computers have been the mainstay of hydrological computing. For very large problems, there is still a use of the IBM supercomputer at NCAR or the supercomputer at Oak Ridge National Laboratory. In essence, the simulations that were carried out on workstations in the 1980s can be done many times faster on personal computers today. The computational power of large computers (those used for GCM simulations) has increased from 170 GFlops (National Aeronautical Laboratory, Japan, 1995) to 1 PetaFlop (Sandia RoadRunner).

Advances in software engineering. In terms of software, the computer science and engineering communities have proposed many paradigms for designing complex codes that are more robust, while also having the pragmatic benefit of being less expensive to build and maintain. These paradigms include object-oriented programming, component-based architectures, and service-oriented architectures, among others. Many scientific communities, from the genomic sciences, to astronomy, to the atmospheric sciences, are leveraging these advanced software engineering approaches to advance scientific knowledge within their own domains [Foster, 2005; Hey and Trefethen, 2005]. Hydrology has also seen a growth the availability of digital data through the proliferation of *in situ* sensors, remote sensing devises, and global-scale climate models. There is, therefore, a pressing need in the hydrologic sciences to leverage advanced software engineering practices to create a new generation of hydrology models and modelers able to make use of the growing data sets in their efforts to advanced hydrologic understanding. Adopting advanced software engineering approaches are the key to achieving this goal.

Remote sensing and in situ *sensor networks are ripe for assimilation*. Data assimilation in earth sciences and hydrological modeling has come a long way since the use of assimilation in atmospheric modeling [Kalnay, 2003]. The advent of the TERRA and AQUA satellite sensor systems launched by NASA in 1999 and 2002 respectively with a number of sensors on each platform [LaMoreaux, 2001] has contributed to the availability of data for land and atmosphere for soil moisture, temperature, precipitation, vegetation, atmospheric temperature and water vapor to name a few variables. These data along with ground observations from either established *in situ* networks such as the Oklahoma Mesonet [Basara and Crawford, 2000] or opportunity based experiments such as the Southern Great Plains 1997 experiment (SGP97) [Jackson et al., 2002] have been assimilated in hydrological models [McLaughlin, 2002]. There has been a large body of work in the past decades that utilize satellite data [Lakshmi and Susskind, 2001; Boni et al., 2001; Walker and Houser, 2001; Lakshmi, 2000].

Modeling is mature enough and more importantly, the community is ready. There are many research groups who are very active in large modeling exercises with inter-disciplinary and multi-institution teams. These applications are in the areas of coupled surface water-subsurface and groundwater system which has a very heavy computational requirement for solving coupled atmosphere-land surface-subsurface equations [Kollet and Maxwell, 2006; Maxwell and Miller, 2005] and the area of coupled land-atmosphere interactions where model spatial resolution is paramount to achieving computational accuracy [Chen et al., in press; Trier et al., 2004] and connections between nutrient cycling and the atmosphere [Xu, 2007]. Another area of high computational requirement that serves as a general category for all of hydrological simulations is model calibration and evaluation of uncertainty [Moradkhani et al., 2005; Mugunthan et al., 2005]. Land information system serves as a general modeling framework for all hydrological modeling [Kumar et al., 2006]. It is apparent from such a diverse range of hydrological science questions presented in this subsection and the extent of the collaborations between individuals, institutions and research areas that hydrological modeling in all its aspects is mature and the community is ready.

We can build upon existing elements and capabilities

CHyMP will truly be something new, but it will be built from a vast wealth of existing elements and capabilities. The ability to leverage existing capabilities will markedly reduce the level of investment required to initiate and bring together a working implementation of CHyMP. The added value that CHyMP will bring to these existing capabilities includes expanding the integration and coupling of processes, as well as facilitating accessibility to the community.

Models that integrate ground water and surface water processes will play an important role in CHyMP because they already couple together significant hydrologic processes. Interest in these types of models has increased lately and many are currently available. Integrated models that were discussed at the workshop include: HydroGeoSphere [Sudicky], PARFlow [Maxwell, 2005], and PIHM [Qu and Duffy, 2007].

Integrated ground water-surface water models typically fall short of adequately representing surface water transients, or 3-D circulation in surface water, but a variety of other models have been developed to fill this need. ADCIRC and CWR-ELCOM are examples of hydrodynamic models that could be included in CHyMP.

Modeling platforms are also currently available in the hydrologic community. Several federal agencies have developed their own version of a modeling platform, with some examples including USDA's OMS, NASA's LIS, USGS's GWFLOW, EPA's FRAMES, and NWS's CHIPS. The DoD has a platform that has been released commercially as three separate packages: WMS, SMS, GMS. GEOtop is another platform developed by Rigon, and CSDMS is a platform under development by Peckham. Several of these existing platforms were represented at the workshop.

Current vision for the platform

The goal of the CHyMP effort is to significantly accelerate the development of advanced hydrological modeling capabilities in order to address complex water issues of the highest priority at national and international levels. Key components of the effort are to:

- Provide a single interface for accessing a platform of modular components that can be linked together to form integrated water cycle models across of range of space-time scales;
- Development of National Water Model that includes all major components of the natural and managed hydrologic cycle;
- Maintain close integration with other CUAHSI activities, in particular Water Data Services (HIS), and other community efforts including CSDMS, NCAR CCSM/CLM, NOAA, and USGS;
- Develop the ability to link to models from other disciplines, including climate, ecology, and biogeochemistry;
- Maintain HPC compatibility/scalability and access to high performance computing;
- Engage the community through working groups and annual meetings.

More specifically, the CHyMP should:

- Represent physics associated with the flow of all terrestrial water: ground water, vadose, streams, lakes, estuaries, glaciers, snow, other;
- Include flexibility to represent many processes;
- Allow for the inclusion of transport of solutes and sediment, chemical and biogeochemical reactions, multiphase flow, porous media deformation, as

well as processes from biology, ecology, environmental engineering, geomorphology, economics, and other fields;

- Accommodate parameters and physics over a wide range of scales;
- Include techniques for adjusting forms of equations to accommodate scales from pores to continents; methods to up-scale and down-scale parameters;
- Accommodate coupling with ocean and atmospheric models;
- Estimate model parameters and uncertainty from large data sets;
- Represent stochastic processes, e.g. parameter distributions, transition probabilities, Monte Carlo, geostatistics, and other stochastic processes;
- Enable visualization of data and output to maximize insights;
- Be easy to use, learn, and teach.

Packages on the model platform could include:

- *Model data and geometry package* to represent the geometry of the simulation region (includes mesh or grid generation, with arrays of constitutive parameters, state variables, and forcing terms);
- *Forward Package* with general capabilities to represent distributed processes;
- *Inverse Package* with capabilities to identify model parameters and uncertainty from large data sets, evaluate management strategies, integrate data from sensor networks, and related;
- *Stochastic Package* to conduct general geostatistical analyses, generate parameter distributions, Markov chain, Monte Carlo analyses, transition probabilities, up-scaling and down-scaling methods, and other techniques;
- *Spatial Analysis Package* to conduct general analyses of distributed parameters;
- *Visualization Package* with general graphical and data display capabilities;
- *Data Package* to read data from, or write data to CUAHSI data services;

The eventual conceptual model for the CHyMP will be fully developed following the series of community workshops described below.

Links to other CUAHSI activities

A unique attribute of the hydrologic sciences community when compared with other surface Earth processes communities such as geochemistry and geomorphology is the routine use of extensive data collected by government agencies in hydrologic research in addition to data collected as part of academic projects. To meet this need, CUAHSI Water Data Services have been developed by the Hydrologic Information Systems project [Maidment, 2008], that index and provide access to the data collected by agencies such as the U.S. Geological Survey, the National Climate Data Center, the Natural Resources Conservation Service and others. In addition, the CUAHSI Water Data Federation uses the same services for the publication of field data.

A key feature of CHyMP will be the tools to use CUAHSI Water Data Services and to publish output of models using the same Water Data Services. The technology for publication has been developed either as part of WDS (for model output at a point) or as part of Unidata's Thematic Realtime Environmental Distributed Data Services (THREDDS) (for gridded output).

The development of CHyMP will also provide important capabilities in the design of observatory networks and of field campaigns by allowing the marginal value of additional data in different locations to be determined. CHyMP would provide the ability to construct alternate simulation models (representing specific objectives of the network) using the library of modules, as well as tools, described above, for setting up the geometry of the problem, ingesting data, and visualizing output. Currently, the resources required for such analyses have been significant barriers to the objective design of observation networks.

Synergies with other community modeling activities

CHyMP development can benefit from a number of ongoing community modeling activities. Some efforts can provide a starting conceptual model, while others can supply a template for software architecture, and still others can provide a community engagement model. Notable examples include:

- The NCAR Community Land Model (CLM);
- The Global and National Land Data Assimilation System models (GLDAS and NLDAS);
- The NASA Land Information System (LIS);

- The Community Surface Dynamics Modeling System;
- The Chesapeake Community Modeling Program;
- The NOAA CHPS.

The CHyMP community is actively reaching out and interfacing with these and other efforts to encourage collaboration and to ensure synergies. For example, the CHyMP and CSDMS communities are now working closely through the CSDMS Hydrology Focus Research Group.

Perhaps the best-known community modeling effort in the geosciences is based at the National Center for Atmospheric Research (NCAR). The NCAR Community Land Model (CLM) [*Dickinson et al.*, 2006] is a component of the NCAR Community Climate System Model [*Boville and Gent*, 1998], and has evolved over the years from its origins as a biogeophysical model that emphasizes vertical rather than horizontal water movement [e.g. *Dickinson et al.,* 1993]. While the CLM effort is making tremendous progress towards its mission of improving global climate modeling, it understandably does not include the resources for detailed understanding and enhanced parameterization at the smaller hillslope- to catchment-scale that are the focus of much contemporary hydrology. Hence, the CLM is an example of a successful and important community model, but it is designed for climate modeling and most hydrologists find limited utility in a model like the CLM. CHyMP could play an important role in CLM development by providing detailed hydrologic process code, and by providing a high-resolution reference model like the National Water Model, to which CLM simulations could be compared. Conversely, the NCAR effort can provide a successful model for working group structure and community engagement.

Another example comes from the Global Land Data Assimilation System (GLDAS) [*Rodell et al.*, 2004], the smaller-scale, higher-resolution (1/8^{o}) National Land Data Assimilation System (NLDAS) and the very high resolution (1 km) Land Information System (LIS) [*Kumar et al*., 2006]. The GLDAS, NLDAS and LIS are modeling *systems* or platforms, described below. In GLDAS for example, multiple land surface models (e.g. CLM, NOAH [*Ek et al.* 2003], VIC [*Liang et al.*, 1994]) can be 'plugged in' to a grid-based modeling environment, and driven with atmospheric forcing to produce output datasets of model states and fluxes at 1/4^{o} to 1^{o} spatial resolution, with output archived at 3-hourly intervals. As in the case of the CLM, the GLDAS, NLDAS and LIS projects are all highly successful, but their sponsoring-agency missions (enhanced prediction through assimilation of remotely-sensed atmospheric and hydrologic data) do not allow for significant community input to model development or for substantial technical support for distributed code. However, model code and output are freely available. As a result, they are examples of models where the source code can be accessed by the community, although the opportunity for input from the community is limited. The CHyMP could clearly build upon the software architecture and assimilation framework of the LIS; and conversely, the LIS could benefit from the community engagement aspects of the CHyMP.

The surface processes community is currently developing the above-mentioned CSDMS [*CSDMS Working Group*, 2004a-c, *http://instaar.colorado.edu/deltaforce/workshop/csdms.html*]. The CSDMS is yet another example of a community modeling platform that will provide a simulation environment in which a community-built, freely-available suite of surface processes components could be integrated to predict erosion, sediment fluxes and landscape evolution across a broad range of spatial and temporal scales. The CSDMS effort provides another important potential template for the development of a community modeling platform in hydrology. The first meeting of the CSDMS Hydrology Focus Research Group revealed the potential for great synergy between the CSDMS and CHyMP efforts. As in the case of the LIS/CHyMP interactions, CSDMS can benefit from the hydrologic expertise and the CUAHSI community upon which CHyMP is based, while the CHyMP could readily adopt the CSDMS software architecture.

Challenges

We recognize two types of obstacles to development of CHyMP, one related to technical challenges that limit its capabilities and another to community attitudes that limit its use. In the following we identify various challenges and briefly explain how we think they could be addressed.

Technological Challenges

Can't get codes to run Legacy codes will play an important role on the platform. We expect it will be feasible to get all legacy codes to run on the platform using a single processor, but it may be beyond the scope of the project to get some codes to run in a parallel computing environment. Parallel versions of these codes developed by the community could be included in this case.

Coupling between models or packages greatly reduces speed Coupling multiple, stand-alone codes will slow the execution speed compared to running the codes individually and as the size of the models becomes larger the execution time will eventually be prohibitive. Multiple methods for coupling models will be supported to provide the flexibility to balance the capabilities and size of a model with run time.

License restrictions prevent best codes from being represented Use of the best codes in some branches of hydrology may be restricted to users who have paid for an appropriate license. We expect that most of the content of CHyMP will be open source, but we will also evaluate the feasibility of including commercial codes.

Community Attitudes

"I'm already doing this. " Community modeling is a concept that has emerged in various forms in nearly all the mission-oriented government agencies, among various academic groups, and in private industry. We recognize that this may cause some groups to feel that the CHyMP project is unnecessary because it is redundant with their efforts.

"Not needed for what I do." Some investigators will not benefit immediately from CHyMP because either their research does not make use of models, or they currently have ready access to the models that they need. CHyMP will be designed to grow through contributions from the community, so investigators that do not initially embrace CHyMP may find that the capabilities improve to the point where they are valuable.

"Other stand-alone codes work better. " The intent will be for CHyMP to include the best codes available, and there will be a mechanism for the community to include codes that they want to be supported.

"Unfairly competes with codes I am selling." CHyMP will be developed in collaboration with private industry and specifications will be available to all commercial software vendors. We expect the market to respond with add-ons and improved functionality. Some commercial codes may be included on CHyMP for users with appropriate licenses.

"Can't tell which codes to use." CHyMP will include benchmark scenarios for comparing performance, as well as evaluations and suggestions from users. This information will help new users decide about code usage.

"Too difficult to use/don't have time to learn." CHyMP will include a Windows interface with on-line help and tutorials to facilitate self-learning. Training workshops and technical support through CUAHSI will also be proposed to maximize benefit to the community.

"Contributing is not worthwhile." A Peer-review process will be considered to ensure the quality and value of functionality and data published on CHyMP. Authorship of contributions will be readily available and citation of authors will be strongly encouraged. The intent is that these measures will make contributions to CHyMP worthwhile for all scientists.

"Accessing high performance computing capabilities is too cumbersome." CHyMP will make use of technology that will community facilitate access to HPC.

Opportunities

A community hydrologic modeling effort will create a number of opportunities for community engagement, for new, interdisciplinary science, and for contributing to environmental policy, water management and decision making. Similarly, new opportunities for funding the effort must be explored.

A number of possible science questions enabled by the CHyMP effort have been mentioned previously. However, it should be noted that community modeling efforts in other fields have also been successful at building vibrant, collaborative developer/user communities. For example, the NCAR Community Climate System Model has a well-established working group structure for model development and testing, and has annual, open meetings for community input. We envision a similar structure for the CHyMP, whereby working groups are established for coordinating model development and testing. Additionally, we expect that selected simulation activities, such as an annual assessment of regional and national water budgets and water availability, will form a focal point for community engagement in the CHyMP.

An integrated, national-scale model such as the National Water Model, because of its potential for comprehensive simulation of natural hydrology and anthropogenic manipulation, could play an important role in water management within regions and at the national scale. The opportunity to bridge the research-policy gap is compelling, and offers an important pathway for bringing advanced simulation tools to bear on national water issues of mounting significance.

We expect that several funding agencies will be interested in supporting the development of the CHyMP. At the January 2009 CUAHSI Board of Directors meeting, several agency representatives from NSF, NASA, EPA, DOE, USGS, NOAA, and the NWS all expressed their support for the CHyMP. Furthermore, important opportunities exist to complement several ongoing community modeling efforts, for example, the NASA LIS, NOAA CHIPS, CSDMS, MODFLOW, etc.

Moving forward

The CHyMP effort will move forward via a series of two additional workshops on science and implementation issues. The 2nd Workshop on a Community Hydrologic Modeling Platform will be held on March 31 and April 1, 2009, in Memphis, TN. This workshop will focus on the key community science questions that a CHyMP can enable, and will begin defining CHyMP needs and requirements. A second report, the CHyMP Science Plan, will result from the workshop. The third workshop will be held within the 18 months that follow the Memphis workshop, and will focus on hardware and software issues. These two workshops will result in a third report, the CHyMP implementation plan. The workshops will also lead to the formation of CHyMP working groups.

Near-term goals of the CHyMP effort are to deliver the above-mentioned reports to the community and to the NSF; to develop a conceptual model for the platform itself; to develop the working group and advisory committee structures; to refine mechanisms for community input; to initiate prototype platform and National Water Models; and to secure funding for long-term development and support.

A positive outcome of the workshops will be an open solicitation for the development of the CHyMP. Additionally, other agencies may contribute to CHyMP development by funding proposals that plan to contribute to or use the CHyMP.

References

Basara J., and K. Crawford (2000) Improved installation procedures for deep-layer soil moisture measurements, Journal of Atmospheric and Oceanic Technology, 17: 879-884.

Boni G., D. Entekhabi and F. Castelli (2001) Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation, Water Resources Research 37: 1713-1722.

Boville, B.A. and P.R. Gent (1998) The NCAR Climate System Model, version 1, Journal of Climate, 11,6: 1115-1130.

Chen, F., K.W. Manning, M.A. LeMone, S.B. Trier, J.G. Alfieri, R. Roberts, M. Tewari, D. Niyogi, T.W. Horst, S.P. Oncley, J.B. Basara, and P.D. Blanken (2006) Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System, Journal of Applied Meteorology and Climate, in press.

Dickinson, R.E., K.W. Oleson and G. Bonan, et al. (2006) The Community Land Model and its climate statistics as a component of the Community Climate System Model, Journal of Climate, 19: 2302-2324.

Ek, M.B., K.E. Mitchell and Y. Lin et al. (2003) Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, Journal of Geophysical Research-Atmosphers, 108: 8851.

Foster, I. (2005) Service-oriented science, Science, 308(5723): 814-817.

Giles, J. (2006) Water crisis happening now, Nature, Published online 21 August 2006, doi:10.1038/news060821-1.

Heffernan, O. (2008) They say they want a revolution, Nature, 453: 268-269.

Hey, T. and A.E. Trefethen (2005) Cyberinfrastructure for e-science, Science, 308(5723): 817-821.

IPCC (2007) Climate Change 2007: The Physical Science Basis. Contribution of Working Group 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom, and New York, NY, USA.

Jackson, T.J., A.Y. Hsu, A.M. Shutko, Y. Tishchenko, B. Petrenko, B. Kutuza, and N. Armand (2002) N. Priroda microwave radiometer observations in the SGP97 hydrology experiment, International Journal of Remote Sensing, 23: 231-248.

Kalnay, E. (2003) Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, ISBN 0521796296, 9780521796293, 341 pp.

Kollet, S.J. and R.M. Maxwell (2006) Integrated surface-groundwater modeling: A free surface overland flow boundary condition in a parallel groundwater flow model, Advances in Water Resources, 29: 945-958.

Kumar, S. V. and C.D. Peters-Lidard et al. (2006) Land information System: An interoperable framework for high resolution land surface modeling, Environmental Modeling and Software, 21: 1402-1415.

Lakshmi, V. (2000) A simple surface temperature assimilation scheme for use in land surface models, Water Resources Research, 36: 3687-3700.

Lakshmi, V., and J. Susskind (2001) Utilization of Satellite Data in Land Surface Hydrology: Sensitivity and Assimilation, Hydrological Processes, 15: 877-892.

LaMoreaux, P.E. (2001) EOS Terra Mission, Environmental Geology, 40: 1299-1300.

Liang, X., D.P. Lettenmaier and E.F. Wood et al. (1994) A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models, Journal of Geophysical Research-Atmospheres, 99: 14415-14428.

Maidment, D.R. (Ed.) (2008) CUAHSI Hydrologic Information System: Overview of Version 1.1., Consortium of Universities for the Advancement of Hydrologic Science Inc., Washington, D.C., 92pp. Available from: <http://his.cuahsi.org/documents/HISOverview.pdf>.

Maxwell, R.M. and N.L. Miller (2005) Development of a coupled land surface and groundwater model, Journal of Hydrometeorology, 6(3): 233-247.

McLaughlin, D. (2002) An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering, Advances in Water Resources, 25(8-12): 1275-1286.

Moradkhani H., K.L. Hsu, H. Gupta, S. Sorooshian (2005) Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resources Research, 41: W05012.

Mugunthan, P., C.A. Shoemaker, R.G. Regis (2005) Comparison of Function Approximation, Heuristic and Derivative-based Methods for Automatic

Calibration of Computationally Expensive Groundwater Bioremediation Models, Water Resources Research, 41: W11427.

Qu, Y.Z. and C.J. Duffy (2007) A semidiscrete finite volume formulation for multiprocess watershed simulation, Water Resources Research, 43: W08419.

Rodell, M., P.R. Houser and U. Jambor et al. (2004) The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85(3).

Sudicky, E.A., J.P. Jones and Y.J. Park et al. (2008) Simulating complex flow and transport dynamics in an integrated surface-subsurface modeling framework, Geosciences Journal, 12: 107-122.

Trier, S., F. Chen, and K. Manning, 2004: A Study of convection initiation in a mesoscale model using high-resolution land surface initiation conditions. Mon. Wea. Rev., 132, 2954-2976.

Walker JP, P Houser (2001), A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations, Journal of Geophysical Research (Atmospheres) 106: 11761-11774

Xu, T., Incorporation of Aqueous Reaction Kinetics and Biodegradation into TOUGHREACT: Application of a Multi-region Model to HydrobiogeoChemical Transport of Denitrification and Sulfate Reduction, *Vadose Zone Journal*, In press, 2007.