UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
NATIONAL HEALTH AND ENVIROMENTAL EFFECTS
RESEARCH LABORATORY
WESTERN ECOLOGY DIVISION
200 S.W. 35TH STREET • CORVALLIS, OREGON 97333

Ecohydrology Model Review
Reconciliation Memo

November 5, 2010

EPA asked four scientists highly experienced in the development and application of ecohydrology models to provide comprehensive, independent reviews of EPA-WED's ecohydrology modeling framework, and to comment on its utility, ease of use, reliability, and predictive capability. The reviewers were selected to represent a range of backgrounds and potential end users in academia (Dr. Lawrence Band, University of North Carolina, and Dr. Irena Creed, University of Western Ontario), industry (Dr. George Ice, National Council for Air and Stream Improvement), and government (Dr. Steven McCutcheon, EPA).

The four reviewers were asked to address 7 charge questions based on their experience with the VELMA ecohydrology model (formerly known as GTHM-PSM). Prior to an initial 1-day meeting at the Corvallis EPA laboratory on December 10, 2009, each reviewer received a model information package containing program code, a user's manual and tutorial, example simulations and validation results for three LTER sites (H.J. Andrews, Konza Prairie sand Hubbard Brook), and other supporting information. At the initial meeting the reviewers were invited to perform model tests over the next several months using field site data of their own choosing to assess model performance with respect to the charge questions. Drs. Band, Creed and Ice applied VELMA to sites at which they have worked extensively: the Coweeta LTER site in North Carolina (Dr. Band); the Turkey Lakes Watershed in Ontario (Dr. Creed); and the Alsea Watershed in Oregon (Dr. Ice). I communicated with the reviewers during the review concerning their model tests, and provided calibration assistance when requested. On March 30, 2010, the reviewers, the review coordinator and I met a final time at the EPA facility in Research Triangle Park, NC, to discuss the model tests and address any final questions before the reviewers prepared their written responses to the charge questions.

VELMA is a relatively new model that continues to undergo significant development. Although the model development team (Dr. Marc Stieglitz' Georgia Tech laboratory and me) had previously performed extensive tests for the H.J. Andrews, Konza Prairie, and Hubbard Brook LTER sites, this review was the first opportunity for anyone outside of our team to perform model tests. The tests reported by the reviewers for their applications of VELMA are proving extremely valuable, providing many new and helpful insights about the applicability and utility of the model, while also confirming many of our own findings.

In this reconciliation memo, I summarize and discuss the reviewers' assessments of the December 2009 version of the modeling framework, focusing on strengths and limitations that were identified in response to the 7 charge questions listed in ***bold italics***, below. Many issues were identified by more than one reviewer, and I have consolidated comments that came up in multiple reviews. I have summarized the reviewers' major recommendations as bullets in blue text in the following narrative, and again in a more concise manner in Table 1 in the Summary section at the end of this memo.

I want to express my thanks to the reviewers for their insightful and constructive reviews. Their comments are already proving extremely useful in our efforts to improve the model to make it more useful to EPA and other users. I also want to express my gratitude to Virginia Houk, Review Coordinator for EPA-NHEERL, for her expert guidance throughout the entire review process. The efforts of all involved have made this a very positive and productive experience.

Finally, I want to thank my VELMA colleagues, Marc Stieglitz, Feifei Pan and Alex Abdelnour, for their valuable advice during the course of this review. Alex graciously provided timely calibration assistance for the Turkey Lakes and Alsea applications. It has been a pleasure working with all of them on the VELMA project. While I've made every effort here to accurately address the reviewers' comments, any errors are my own.

Robert McKane, Ph.D.
Lead PI for Ecohydrology Model Development
US Environmental Protection Agency

Charge Question 1:

Do GTHM-PSM (VELMA) and MEL provide a valid and useful framework for assessing the integrated responses of multiple ecosystem services to interacting stressors? Do the models address the appropriate hydrological and biogeochemical processes, and interactions among these processes, for supporting ecosystem service assessments? Are the model's predictions supported by analyses and results? How does this framework compare with similar models?

Strengths

Regarding this first and perhaps most important charge question, the consensus was that a major strength of the VELMA model is its relative simplicity, and that despite this simplicity, the model can simulate catchment-scale water, carbon and nutrient budgets with accuracy similar to more complex models – e.g., "It is impressive that the catchment wide annual water and carbon budgets appear to be simulated well with a minimum amount of calibration." It was noted that with minimal training, new users were able to implement VELMA for the Turkey Lakes and Coweeta sites in less than 6 weeks, compared to about 6 months for similar models (experienced VELMA users require less than 2 weeks to calibrate the model for a new site). Furthermore, the consensus was that VELMA provides a valid framework for assessing tradeoffs among multiple ecosystem services, e.g., regulation of water quality and quantity, carbon sequestration, greenhouse gases, soil productivity, and sinks/sources of reactive nitrogen. In this regard, it was noted that "the model performed remarkably well in the HJ Andrews experimental watershed", the site for which the model development team has most intensively calibrated and tested the model (Abdelnour et al. in review).

However, the reviewers were also in agreement that there are a number of hydrological and biogeochemical processes that are currently not represented in VELMA that will need to be added to improve its applicability to a greater range of ecohydrological regions and management practices. They made a number of very helpful recommendations, described below, to address these limitations. The reviewers were very cognizant of the tradeoffs – simplicity and speed versus explanatory power – associated with adding more complexity to VELMA. In my view, the main recommendations that follow are essential and can be incorporated without significantly compromising the model's simplicity.

Limitations/Recommendations

- Include multiple land cover types to address complex landscapes

A chief recommendation was that the current single land-cover version of VELMA (e.g. all forest, or all grassland) needs to incorporate multiple land cover types to address more complex landscapes having a mixture of agriculture, grasslands, wetlands, forests, etc. This recommendation is being addressed now under an EPA contract with Georgia Tech. The new multiple land cover version of VELMA is currently being tested for the Flint Hills in Kansas and the Willamette Basin in Oregon – two contrasting regional landscapes characterized by a complex mosaic of land cover types and land uses. Obviously, spatial variability in soil properties will also have a major effect on ecohydrological processes. Therefore, our tests of the new multiple land cover version will also take advantage of VELMA's existing capability to represent spatial variability in soils, defined by differences in texture, carbon and nitrogen concentrations, and drainage characteristics. As the reviewers recommended, a key goal is to simulate how spatial variability in vegetation and soils affects the cycling and transport of water and nutrients within real landscapes, so that best management practices (BMPs) can be better identified. For example, how can riparian buffers best be designed with regard to cover type, width, and flow pathways to reduce surface and subsurface runoff of agricultural fertilizers to streams? Similarly, to what extent can natural or constructed wetlands reduce stream nutrient loads in agricultural and forest areas? In which hydrogeomorphic settings are wetlands and riparian buffers most effective? Our ultimate objective is to quantify how spatial and temporal variability in vegetation and soils influence the ecohydrological processes that ultimately regulate water quality and quantity, greenhouse gases, food and fiber production, and other important ecosystem services.

The Turkey Lakes Watershed application nicely illustrates this need. VELMA accurately simulated the increase in stream DIN levels for the first year after this forest watershed was clearcut, but then overestimated subsequent (years 2-7) post-harvest DIN levels, presumably because the single land cover version of VELMA could not simulate observed reductions in DIN associated with denitrification as the stream passed through a wetland above the outlet. This is a good example of a model result suggesting that something is missing, in this case a landscape feature that provides important nitrogen regulating processes. Thus, addressing this problem will require an accurate representation of the wetland-forest mosaic in the Turkey Lake Watershed. This will be a good test of the multiple land cover/soils version of VELMA.

In summary, the multiple land cover/soils version of VELMA is our highest priority, given that it will directly support several ongoing projects: a Regional Applied Research Effort (RARE) with EPA Region 7 to assess the effects of rangeland prescribed fires on ecosystem services in the Flint Hills; the Willamette Ecosystem Services Project (WESP); and a project supporting EPA's national Biofuel Research Initiative to assess the effects of biofuel feedstock production on ecosystem services.

- Include phenological controls on evapotranspiration

It was noted that VELMA overestimated wintertime evapotranspiration (ET) for Coweeta, a site with a predominantly deciduous forest in a climate with relatively warm winter temperatures. Consequently, actual ET during winter months is limited less by temperature than by a lack of leaf area after trees shed their leaves. This situation is quite different than the other sites for which VELMA had been previously tested – HJ Andrews, Konza, and Hubbard Brook – where low winter temperatures alone are sufficient to limit ET, regardless of leaf phenology. Thus, VELMA's applicability to sites like Coweeta is limited because the simple ET submodel in VELMA is temperature-driven and does not currently include phenology or leaf area. However, phenological triggers for leaf flush (ET "on") and leaf fall (ET "off") will be easy to implement, requiring just a few lines of code. Triggers based on degree-days, rather than Julian days, will be more effective for simulating these phenological events and how these may shift in time in response to climate change. This change will be incorporated into the new multiple land use version of the model currently under construction.

- Include nitrogen fixation by plants and soil organisms

As the Alsea application so clearly highlighted, incorporating nitrogen fixation in VELMA will be essential for simulating carbon and nitrogen dynamics in systems where N-fixing species are an important component of the landscape. Nitrogen fixation will be modeled as simply as possible, borrowing methods from existing models such as MEL (Rastetter et al. 2001) and RHESSys (Tague and Band 2004), and incorporating these into the multiple land-cover version of VELMA. This enhancement is a priority for EPA's Willamette Ecosystem Services Project, given the abundance of N-fixing red alder in the Oregon Coast Range and its impact on stream nitrate levels (Compton et al. 2003).

- Improve the model calibration procedure, GUI interface, and user manual

The reviewers' recommendations on these interrelated points have been very helpful. In the current version of VELMA, users must hunt through several subroutines in the program code to make calibration adjustments – a frustrating procedure. A number of improvements to VELMA's calibration procedure are being incorporated into the multiple land-cover version. First, the current graphical user interface (GUI) will be redesigned to allow adjustment of all hydrological and biogeochemical calibration parameter values in one place.

Second, the new GUI will provide an information link for each parameter that provides its definition, associated model equations, its range of calibration values for existing sites, and a list of recommended visualization tools for assessing model performance. VELMA already includes a wide range visualization tools under the "Display" tab of the current GUI, but new users will benefit from some guidance on which tools to use to assess how well the simulated water, carbon and nutrient budgets are matching available validation data.

Finally, we are exploring various options for automating the calibration procedure – e.g., using "Monte Carlo simulation to help get you in the right ballpark", as one reviewer suggested. This could reduce model implementation time even more, and would allow a more exhaustive exploration of the range of possible parameter solutions. This is an important point given the problem of "equifinality" – i.e., is a particular solution giving the right answers for the right reasons (Kirchner 2006)? This issue was mentioned by several reviewers, e.g., "Many different models are capable of getting the hydrograph right, but not necessarily for the right reasons. VELMA seems to be getting the hydrograph right at a lower level of effort than many other models. It is very valuable that VELMA has outputs such as nutrients and soil moisture status (for separate layers) that allow us to test and see if the model is performing well for the right reasons."

In addition to Monte Carlo methods, I have been discussing other automated calibration options for VELMA with colleagues at the Corvallis USDA-ARS lab who are using a multi-objective evolutionary algorithm and Pareto optimization to calibrate the SWAT model for watershed-scale applications (Confessor and Whittaker 2007). This approach has major advantages for quickly and accurately identifying an optimal set of parameter values that satisfies multiple validation objectives, e.g., spatial and temporal data describing stream water quality and quantity, plant and soil C and N stocks and fluxes, and so on. This becomes increasingly important as land use and land cover complexity increases. Because incorporation of an optimization algorithm will require a fair amount of work, this will be addressed as resources allow after the other improvements discussed in this section are completed.

- Replace VELMA's logistic function for modeling rainfall-runoff processes

In regard to VELMA's logistic function for controlling runoff, one reviewer noted that "The calibration parameters related to a logistic approximation cannot have physical significance because of the purely empiric nature of the function… Lack of hydrologic physical significance in the short-term prevents reliable forecasting… The field of hydrology is dominated by what turned out to be oversimplification in the form of the Soil Conservation Service curve number method (NRCS 2001, McCutcheon 2003…)."

While I agree with this reviewer that VELMA is a simplification of rainfall-runoff processes represented in much more complex models (e.g., HSPF), the view that VELMA's logistic curve approach does not have a physical basis is incorrect. VELMA's hydrological submodel consists of a quantifiable set of physically-based parameters (Abdelnour et al. 2010; Abdelnour et al. in review). These physical parameters include the saturated hydraulic conductivity (*ks*) and its change with depth both vertically and laterally (*ksv* and *ksl*), and texture-specific constants for soil porosity (Φ), soil moisture field capacity (θ_{fc}), and maximum soil water storage (s_m). The logistic function uses 2 texture-specific parameters to predict vertical drainage (D_i) based on the degree of saturation (s/s_m) per soil layer i. The logistic function has been generically calibrated to capture the breakthrough characteristics of soil water movement for defined soil texture classes (Buol et al. 2003), and has been demonstrated to capture the fast "switching" to higher flow when soil moisture approaches field capacity (Abdelnour et al. in review).

Thus, with VELMA's comparatively few "free" parameters – 3 total, including *ks, ksv* and *ksl* – it has been possible to simulate hydrologic processes such as vertical and lateral flow, changes in soil water storage, and terrestrial flow pathways by which water (and nutrients) are transported within hillslopes and watersheds to streams. *A purely empirical approach would be incapable of simulating these spatially-distributed hydrologic processes*. For example, the Soil Conservation Service's curve number method is limited to predicting the amount of runoff at a particular point (outlet) within a watershed, without providing any information about the physical processes and pathways by which the water arrived there. Therefore, VELMA and the empirically-based SCS curve number approach should not be confused.

Furthermore, VELMA's simplified approach to modeling hydrologic processes has considerable advantages for quantifying uncertainty and identifying process-level controls. The remarks of Kirchner (2006) are especially relevant here: "Jakeman and Hornberger (1993) showed that typical rainfall-runoff data only contain enough information to constrain simple hydrological models with up to four free parameters. By contrast, many catchment models have dozens of free parameters. Each additional parameter represents a whole new dimension of parameter space, so the overparameterization problem grows rapidly and nonlinearly with the number of free parameters… These considerations imply that in order to know whether we are getting the right answers for the right reasons, we will need to develop reduced-form models with very few free parameters." Recommendations concerning quantification of uncertainty are further addressed under Charge Question 7.

- Check the simulated water balance for VELMA's spin-up subroutine

One reviewer (Dr. McCutcheon) stated that another reviewer (Dr. Band) "seems to have found a water balance issue from repeating sequences of meteorological data that does not bode well for establishing a scientific basis for VELMA (L. Band, email, April 12, 2010)." This issue was promptly resolved when it came up in April 2010 and had to do with a bug in VELMA's "spin-up" routine. The immediate solution was to create a single long (e.g., 200-year) meteorological input file, rather than require the spin-up routine to generate that by repeatedly rerunning a shorter input file, e.g., rerunning a 20-year record of observed data ten times (B. McKane email to L. Band, April 13, 2010). I apologize for forgetting to notify the other reviewers about this solution.

Charge Question 2:

One of the proposed benefits of this framework is that it was designed to assess more than one process in a watershed. Is this model capable of performing the functions of more than one of the single-purpose models currently available? With regard to EPA's programmatic objectives, what are the advantages and disadvantages of this approach?

Strengths

As the following statements indicate, the reviewers were in general agreement that a major strength of the model is its ability to assess multiple processes and their interactions within a watershed:

> "VELMA directly links water cycling, nutrient cycling and carbon cycling. This is an advance over approaches that use separate, unlinked models for each of these purposes. There are existing models that also do this (e.g., RHESSys). An advantage to VELMA is the relative simplicity of linking together these processes while maintaining the important feedbacks between them, and linking together patches along hydrologic flow paths."
>
> "The model is capable of performing multiple functions, in that it can simulate both hydrological and biogeochemical processes under different land use/land cover and climate change scenarios."
>
> "The treatment of multiple ecosystem services is very valuable. VELMA addresses many key ecosystem services but it misses some others that are very important to the forest products industry such as sediment and temperature."

The important point here is that the integration of hydrological and biogeochemical processes within VELMA provides a means for defining policy and management strategies for entire ecosystems, not simply individual components of the ecosystem. While recognizing VELMA's current capabilities in this regard, the reviewers identified several ways to improve and extend the model's representation of multiple processes and ecosystem services, as follows.

Limitations/Recommendations

- Test modeled processes for wider range of management conditions

The reviewers noted in several contexts, including this charge question, the difficulties inherent to using research-oriented process models for decision support purposes. For example, "existing models…are designed to research process controls on patterns observed and as such are not easily accessible or transferable to decision makers. While the VELMA model is useful in assessing a number of hydrological and biogeochemical processes in line with the EPA goal…further model testing and development is required under various management conditions to see its validity to the EPA." I agree completely and will discuss how we are addressing this recommendation under Charge Question 7, which focuses on decision support aspects.

- Include a means for partitioning total aboveground biomass into component tissues

It was noted that while VELMA can accurately model ecosystem carbon budgets, the plant biomass simulation is simple and does not partition carbon into stem or foliar stores. This limits VELMA's ability to represent a number of important processes, including canopy interception of precipitation, and leaf area (LAI) regulation of plant productivity in response to variability in solar radiation. As described under Charge Question 1, it will be straightforward to incorporate a simple algorithm in VELMA for modeling stem and leaf biomass (and LAI) as an allometric function of any particular land cover type's total plant biomass. Root biomass and its depth distribution are already simulated in a similar way. This approach is consistent with our overarching goal to keep the model as simple and computationally efficient as possible. More mechanistic approaches for partitioning biomass are available (e.g., Rastetter et al. 1997), but these would severely limit VELMA's applicability to large basins and landscapes.

Charge Question 3:

Do the models address the appropriate stressors for supporting ecosystem service assessments? What improvements could be made?

Strengths

The results reported by the reviewers for their test watersheds illustrate the ability of VELMA to capture the interactive effects of climate (temperature and precipitation), land use (harvest, fire, etc.), topography, and soil physical and chemical properties on key ecohydrological processes, including the transport and discharge of water and nutrients to streams, and the accumulation and turnover of carbon and nutrients in vegetation and soils. These ecohydrological processes directly support a wide range of ecosystem services, including regulation of water quality and quantity, provisioning of food and fiber, regulation of greenhouse gases (CO_2, N_2O and NOx), and regulation of sources and sinks of reactive nitrogen within watersheds. The reviewers' written comments provide detailed test results in this regard. For example, it was noted that VELMA "can simulate both hydrological and biogeochemical processes under different land use/land cover and climate change scenarios" and is "capable of evaluating ecosystem services (e.g., tradeoffs between climate change, resource extraction and water resources)."

However, as discussed below, the reviewers were also in agreement that some improvements need to be made in the way that some forcing variables are represented in VELMA. Their recommendations specifically aim to make the model more broadly applicable across a wider range of environments.

Limitations/Recommendations

- Include microclimate spatial variation

VELMA currently represents climatic drivers homogenously across a watershed, so addressing this recommendation is one of our highest priorities. We are collaborating on this with Dr. Chris Daly of Oregon State University. Dr. Daly is the developer of PRISM, a state-of-the-art climate mapping system that is in wide use globally. A new high-resolution (50 meter) version of PRISM was a product of our initial collaboration with Dr. Daly and includes slope/aspect radiation effects on temperature and precipitation among other features (Daly et al. 2007). He has since improved on this version of PRISM to account for topographically driven cold air pooling and consequent atmospheric decoupling in mountain valleys (Daly et al. 2009). We are currently working with Dr. Daly and other Oregon State University colleagues to apply this version of PRISM to VELMA to the HJ Andrews LTER site as a proof-of-concept demonstration. The general principles underlying PRISM are applicable to any environment.

- Include solar radiation effects on snow dynamics, ET and NPP

This recommendation is closely related to the preceding and, therefore, also a high priority. VELMA does not currently include solar radiation as a driver, so cannot capture the effects of spatial (slope and aspect) and temporal (seasonal) variations in radiation on snow melt, ET and NPP. The inclusion of topographic controls on microclimate variation (see preceding paragraph) will make it possible to address these limitations. Although it was noted that the current version of VELMA can closely approximate annual catchment water and carbon budgets, I agree that including radiation effects will be necessary to simulate the interactive effects of multiple climate drivers (radiation, temperature and precipitation) on ecosystem processes in the correct manner in all biomes. For example, using the current version, it has been difficult to consistently capture the dynamics of snow accumulation and melt in mountainous areas like the HJ Andrews, Turkey Lakes, and Hubbard Brook. To incorporate radiation effects in VELMA, we will add a simple algorithm that adjusts total incoming radiation according to slope, aspect and LAI, where LAI will be modeled as an allometric function of each cover type's total biomass and phenological stage. The effectiveness of these changes for simulating spatial and temporal variations in snowmelt, ET and NPP will be tested initially for the HJ Andrews LTER, where appropriate validation data exist at the plot, hillslope and catchment scales.

- Include effects of roads and streamside BMPs on stream runoff, sediments and temperature

Several reviewers noted that VELMA currently does not represent the effects of roads on discharge of water, nutrients and sediments. I completely agree with their assessment that incorporation of these capabilities is important and feasible, and that this will be more in the range of a moderate to major new development. Therefore, in terms of priorities, this enhancement will be addressed as soon as possible after the preceding modifications are made. As was noted, the methods for modeling road effects on hydrology and stream nutrient and sediment loads are available in other ecohydrological models, including TOPOG (Vertessy et al 1993) and RHESSys (Tague and Band, 2000), and DHSVM (Wigmosta et al 1994). Modeling the combined effects of roads and streamside BMPs will also be facilitated by the multiple land cover version of VELMA now under development. In this context it is worth noting that even the reviewed version of VELMA has been used recently to simulate effects of streamside BMPs involving spatial patterns (cut blocks) of forest harvest, e.g., indicating that near-stream harvesting results in disproportionate increases in stream flow and nutrient losses compared to upslope harvesting (Abdelnour et al. in review).

Charge Question 4:

Are the data required to implement and apply the models reasonable? In what kinds of terrain, soils, and vegetation are the models most precise? What improvements could be made?

<u>Strengths</u>

The consensus was that the data requirements are reasonable. For example:

> "Data requirements for VELMA are fairly moderate. This is an advantage, making it possible to apply the model in a relatively short amount of time."
>
> "With suitable calibration, a good approximation to catchment scale runoff and above ground biomass production can be developed…and reasonable timing and magnitude of dissolved nutrient concentrations are achieved. This can be accomplished with a manageable amount of calibration and data requirements in these sites."

As discussed below, an excellent recommendation was made with respect to preparation and analysis of digital elevation model (DEM) data. Several other recommendations suggested under this charge question were also raised, and addressed, under previous charge questions – e.g., effects of roads, streamside BMPs, and multiple lands cover.

Limitations/Recommendations

- Improve the preprocessing of digital elevation model (DEM) data

It was noted that the method used to preprocess DEM data for input into VELMA is cumbersome, especially with regard to selection of the outlet for a watershed. Although this can be addressed in part by using the "Magnifier" application in Windows, this is not an optimal solution. I also agree with the reviewers that it will be important to compare the functionality of VELMA's DEM processor against existing packages such as the Whitebox Geospatial Analysis System. It will be important to determine which flow routing algorithm performs best, especially in relatively flat terrain that has proven notoriously difficult to model using existing models. The Calapooia Watershed in the Willamette Valley will be a good test ground in this regard. We are not tied to any particular DEM processor and will go with whatever proves to works best.

Charge Question 5:

Are the model output variables and associated visualization tools useful for understanding and communicating the effects of interacting stressors on eco-hydrologic processes and ecosystem services? What improvements could be made?

Strengths

The consensus was that "The model output variables and visualization tools are good. The visualization tools are a major strength…" and the "movies (animations), which show the real time fit between the simulated and observed values, are effective." The combination of line graphs and spatial data (maps) describing responses of hydrological and biogeochemical variables is intended to provide model users and decision makers with a variety of ways to better understand the model output. A main motivation for developing VELMA in a visual programming environment (http://Processing.org/) was to improve communication among scientists, policy makers and resource managers, specifically in regard to the application of scientific research results to environmental decision making. This emphasis is described in our recent paper "Enabling the Dialog – Scientist <> Resource Manager <> Stakeholder: Visual Analytics as Boundary Objects" (Cushing et al. 2009).

In the context of purely scientific applications, one reviewer noted that visualization tools alone "tend to obscure the difficulty in simultaneously simulating peak flows and base flows reliably." I agree, and emphasize that VELMA's visualization tools should be used as just a first step in evaluating model output, mainly to gain a quick intuitive grasp of overall model performance. For example, is the current calibration "in the ballpark" with respect to seasonal and annual water, carbon and nutrient budgets? Calculation of standard goodness-of-fit statistics can certainly provide a more rigorous assessment of model performance. However, statistics alone are less effective, and often incapable, of identifying precisely when and where a mismatch between simulated and observed results may be occurring. Obviously, both sets of tools – visual and statistical – are necessary for a complete understanding and assessment of model performance.

Limitations/Recommendations

- Improve control of output variables for specific grid cells

It was noted that "It would be useful to be able to select specific grid cells to develop output time series for selected variables, both for diagnosis and as primary model needs." This excellent recommendation could be accomplished in the current version by modifying code, but this would be very cumbersome. Instead, we will enable this feature through the GUI, a much more user-friendly approach that can be accomplished fairly easily. This will be incorporated in the multiple land cover version now under development.

- Include flexible data output summaries

A recommendation was made to allow users more flexibility in selecting options for generating data summaries for daily, monthly or annual time steps, and for switching from calendar year to water year. This will be straightforward to incorporate in the new version.

- Include a wider range of statistics for assessing model performance

One reviewer incorrectly assumed that we rely solely on the Nash-Sutcliffe statistic for comparing simulated and observed streamflow, perhaps because I focused on this useful statistic during my discussions with the review panel. Rather, as illustrated by the set of statistics we generated for the HJ Andrews (Abdelnour et al. in review) and for an initial Turkey Lakes analysis, we typically calculate the correlation coefficient, root mean square error, Nash-Sutcliffe statistic, baseline adjusted first degree efficiency, and baseline adjusted modified index of agreement. Currently, these statistics are calculated in Excel using the model's standard output files. In this regard, the reviewer's comment does raise a good point, i.e., that it would be very useful to automatically generate these statistics as part of the output generated at the end of a run. This can be accomplished fairly easily by modifying the program code.

Charge Question 6:

Are the model's spatial and temporal scales of application useful for ecosystem service assessments? Does the research plan for scaling up GTHM-PSM from 5th order watersheds to much larger landscapes and basins adequately address the main computational and practical issues? What improvements could be made?

Strengths

While the reviewers recognized that VELMA's simplicity provides a good foundation for scaling up to larger watersheds, they also recognized that the computationally intensive nature of spatially-distributed ecohydrological models makes them difficult to scale up to much larger basins and regional landscapes. They offered several excellent recommendations in this regard.

Limitations/Recommendations

- Scale up VELMA to larger basins and regional landscapes

As the reviewers noted in several contexts, two types of limitations – scientific and technological – must be overcome to scale up spatially-distributed ecohydrological models like VELMA. Addressing the scientific aspects is an essential first step for any scaling effort. A specific recommendation in this regard was to "determine the limits of model scaling from low to higher order catchments, watersheds, or basins." To date, we have demonstrated that VELMA accurately scales up from 1st order catchments to 5th order catchments *without recalibration*. For example, we initially calibrated VELMA for an intensively studied 0.1 km^2 watershed (WS-10) at the H.J. Andrews (HJA) site in Oregon, and verified that this same calibration accurately simulated water, carbon and nutrient budgets for progressively larger watersheds, up to the HJA's 64 km^2 Lookout Creek Basin – a jump in spatial scale of nearly 3 orders of magnitude. These initial scaling tests lend credence to the model's scientific underpinnings for simulating ecohydrological processes across multiple spatial scales.

However, such tests do not address practical technological (hardware and software) limitations to scaling up VELMA. Regardless of accuracy, it clearly would be computationally impractical to apply the current version of the model at the scale of the 30,000 km^2 Willamette Basin. Therefore, we are pursuing several potential hardware and software strategies to address this need.

The first and simplest strategy is to optimize the VELMA program code, which was not done for the version under review. This is now being addressed by Kevin Djang, a professional programmer and contractor with Computer Sciences Corporation, and Dr. Allen Brookes, a computer scientist/systems modeler who joined EPA-WED as a Principal Investigator in October 2010. We do not expect optimization of the VELMA code to entirely solve the scaling problem, but it is already helping. For example, after just one week of “profiling” bottlenecks within VELMA, the contractor was able to increase simulation speeds by almost an order of magnitude by modifying how the program code handles memory and the visualization subroutines.

A second scaling strategy is to run VELMA on a multi-processor computer, whereby individual processors running in parallel are assigned to discrete sub-basins. VELMA output from individual sub-basins would then be fed into a stream network model (e.g., Liu and Weller 2008) that integrates discharge of water and nutrients for hydrological connected basins via a routing algorithm that accounts for transit times and nutrient losses within the stream network (see below). In this strategy, the scale of application would be limited by the number of available processors. My proposal to purchase a Cray CX-1 “desktop supercomputer” (64 processors and 4 terabytes of internal storage) was ranked #2 nationally in the annual competition for EPA capital equipment purchases. Contingent on final approval and funding, the CX-1 will arrive in 2012. In the meantime, we can run VELMA on Dr. John Bolte’s Cray CX-1 at Oregon State University. Dr. Bolte is an EPA Expert Hire and Interim Lead PI on the Willamette Ecosystem Services project.

A third scaling strategy is to employ a coarser simulation grid. For example, simulation times could be decreased by an order of magnitude by going from a 30-m to a 100-m grid. Developing the capability to use a variable grid size would be very useful, so that high resolution is only maintained where required – e.g., riparian buffers and other hydrological and biogeochemical “hotspots”.

A fourth scaling strategy is to reprogram VELMA to perform its calculations using Graphical Processing Units (GPUs) rather than Computer Processing Units (CPUs). All computers have both processors, but whereas most scientific models rely on CPU technology, GPUs are specifically designed to rapidly perform many computing tasks in parallel. GPUs are the engine behind computationally intensive video games (Xbox, etc.) and the making of animated Hollywood movies, but have found limited usage in scientific applications. However, Dr. Stieglitz’ Georgia Tech lab recently used CUDA, a parallel programming architecture, to adapt GPU technology to run an Everglades ecohydrological model. Compared to the CPU-based version of that model, the GPU version runs about 60 to 100 times faster (depending on the computer) while reproducing the same results. Although VELMA is more complex than the Everglades model, the technical principles and feasibility for converting VELMA to a GPU-based model are similar.

We are confident that a combination of the preceding strategies will enable us to apply VELMA accurately and at reasonable speeds for the kinds of regional-scale applications that our EPA clients and stakeholders require, e.g., the Willamette Ecosystem Services Project in Region 10, and the Flint Hills Regional Applied Research Effort in Region 7. For example, an upper estimate of the combined increase in simulation speed from implementing just the first, second and fourth scaling strategies (code optimization = x10, multi-processor computer = x30, and GPU technology = x60) would be about 18,000 times faster than presently possible. Given that the version of VELMA reviewed here required almost 1 hour to simulate the 64 km^2 HJ Andrews Basin, and that the Willamette Basin is about 500 times larger (30,000 km^2), simulations at the Willamette Basin scale could potentially require less than 2 minutes per simulation year. This assumes that simulations at the Willamette Basin scale would be set up to efficiently manage computer memory and disk storage, e.g., in similar fashion to other model applications now being conducted for the Willamette Basin via the ENVISION decision support system (see Charge Question 7 for details). Even if hardware and software inefficiencies reduce this potential gain in simulation speed by one or two orders of magnitude, the final gains are still substantial.

- Include in-stream processes affecting nutrient concentrations

I agree that addressing this recommendation will become increasingly important as VELMA is scaled up to larger landscapes and basins (per the preceding recommendation). The simplest approach will be to run VELMA in conjunction with a stream network model that includes an empirically-based regression model describing the effects of in-stream processes on the downstream attenuation of nutrients (e.g., Mulholland et al. 2008, Wolheim et al. 2006). Thus, VELMA will provide spatially-explicit terrestrial inputs of nutrients (NO_3, NH_4, DON and DOC) to the stream network model, which will in turn predict how these inputs are attenuated downstream, both as a function of biological processing and increases in flow. This general approach has been used successfully by Liu and Weller (2007) in the Chesapeake Basin.

- Convert model output format from single vector to binary arrays

It was noted that "At present the model makes use of single vector ascii files for all inputs and preprocessing routines, including elevation data, meteorological drivers, etc. While this is an initial method that provides the ability to easily inspect and read through datasets, it is inefficient and cumbersome, and will become a barrier as the system is expanded to run on larger areas. Simple solutions would involve using standard, binary data structures." I agree. This recommendation will be straightforward to implement as part of the multiple land cover version of VELMA.

Charge Question 7:

Will the linkage of GTHM-PSM (VELMA) to the ENVISION decision support tool provide a useful means for policymakers and land managers to explore the consequences of alternative decisions on tradeoffs among multiple ecosystem services? What improvements could be made?

Strengths

The review team's exposure to the ENVISON decision support tool was limited to the one presentation given by the developer, Dr. John Bolte, during the introductory meeting in Corvallis, Oregon. In addition, the development of the linkage of VELMA and ENVISION is in an early stage. Thus, the general feeling was that "More information would be necessary to understand how ENVISION interacts with client models in order to provide feedback on the utility of their interactions." Nonetheless, based on their prior experiences with decision support tools, the reviewers offered a number of supportive comments on the potential usefulness of linking VELMA and ENVISION. For example, "Environment model coupling and linkage is receiving substantial attention at present, as the community develops interdisciplinary models for earth systems simulations" and "I have seen the power of such decision support tools in bringing together government, industry and community groups together to explore the consequences of alternative management strategies on ecosystem services and can testify to the importance of such tools. The linkage of GTHM-PSM (VELMA) to ENVISION will provide a useful means for policy makers and land managers, as long as the inputs/outputs are clear, and feedbacks between the two are effectively represented."

The reviewers also made the following recommendations that will prove very helpful as we develop the linkage of VELMA and ENVISION.

Limitations/Recommendations

- Clarify EPA's need for the ENVISION decision support tool and it linkage to VELMA

A number of very useful decision support tools are available, some of which are already being applied within EPA, or potentially could be. The main difference between ENVISION (http://envision.bioe.orst.edu/) and other decision support tools, such as ReVA (http://www.epa.gov/reva/) or InVEST (http://www.naturalcapitalproject.org/InVEST.html), is its use of flexible "decision rules" rather than predetermined, inflexible scenarios. ENVISION's decision rules are user-defined and designed to *dynamically* simulate spatial and temporal changes in ecosystem services in response to alternative policy and land use decisions. That is, the decision rules dynamically determine when and where changes in land use and land cover (LULC) occur within a landscape.

For example, at the beginning of a simulation, ENVISION passes initial LULC spatial (GIS) information to ecological process models (such as VELMA) that, in turn, simulate how the initial LULC conditions affect the hydrological, biogeochemical and population processes that control ecosystem services of interest. Continuing this loop, the ecological models pass their output to ENVISION at the end of a specified time interval (e.g., 1 year). ENVISION then evaluates the modeled changes in ecosystem services in biophysical and economic terms. If specified criteria are not met for this updated evaluation of ecosystem service status and trends, ENVISION's user-defined decision rules will then automatically trigger the implementation of remedial policies affecting land management decisions for the next time step. For example, population growth could trigger expansion of an urban growth boundary. Or, exceedance of a stream nutrient TMDL could trigger a rule requiring wider riparian buffers and/or less fertilizer inputs for a particular kind of crop type. ENVISION users have the option of stopping a simulation at any point to reassess and intervene with a modified set of decision rules.

ENVISION also provides users with a powerful set of visualization tools. These include graphs and maps of status and trends in biophysical processes and final ecosystem services. The visualizations are designed with clients and stakeholders in mind, to provide them an intuitive grasp of how alternative decisions could affect tradeoffs among multiple ecosystem services across the landscape. For example the following link illustrates one type of ENVISION visualization, a virtual "flyover" of the Puget Sound region showing land cover changes in response to alternative land use plans for managing population growth (http://www.youtube.com/watch?v=XiLgh9r1GM8&feature=related).

No other similar framework of which we are aware has such a fully developed combination of decision support capabilities. ENVISION is a mature framework at this point, having been applied to multiple locations in collaboration with numerous stakeholder groups (http://envision.bioe.orst.edu/caseStudies.htm). ENVISION has been peer-reviewed via the scientific literature (http://envision.bioe.orst.edu/Publications.htm) and major granting agencies, including four successful National Science Foundation grants since 2001 totaling more than $7M. These include grants funded by NSF's Dynamics of Coupled Natural and Human Systems Program, and the Water Sustainability and Climate Program (grant details can be found at http://www.nsf.gov/awardsearch/piSearch.do?SearchType=piSearch&page=1&QueryText=&PIFirstName=John&PILastName=Bolte&IncludeCOPI=true&PIInstitution=&PIState=&PIZip=&PICountry=&Restriction=0&Search=Search#results).

The linkage of VELMA to ENVISION will be an important step forward for assessing ecosystem service tradeoffs in a decision support environment. To date, the models linked to ENVISION generally have been fairly simple single-purpose models that individually simulate changes in hydrology or biogeochemistry or other ecological processes. ENVISION's shared landscape GIS provides a means for loosely coupling the various single-purpose process models, but not in the fully integrated fashion for which VELMA was designed. Thus, the linkage of VELMA to ENVISION aims to enable scientists, clients and stakeholders to dynamically capture important feedbacks among a large number of interacting processes and "final ecosystem services" (*sensu* Boyd and Banzhaf 2007). We believe that capturing these feedbacks is critical for accurately representing potential tradeoffs among bundled ecosystem services in response to interacting stressors (Bolte et al. 2010).

- Develop ENVISION-VELMA linkage based on continuous communication/feedback among clients, stakeholders, modelers and decision support developers

Several reviewers recommended that communication between model developers and decision makers be emphasized to optimize the development of VELMA and the ENVISION-VELMA framework for decision making purposes. I agree. My colleagues and I have been working with a variety of clients and stakeholders from the outset of all the projects to which this decision support framework will be applied. The clients and stakeholders involved with the Willamette project are many, and are listed in the project's Implementation Plan (Bolte et al. 2010). Similarly, my work on the Flint Hills Regional Applied Research Project has been planned and carried out in complete collaboration with our EPA Region 7 clients. Most recently, I organized and held a workshop in September 2010 to gather client and stakeholder input for a new EPA-ORD Biofuels Research Initiative project I'm leading. The goal of the Biofuels project is to assess the unintended effects of biofuel feedstock production on multiple ecosystem services in the Willamette Valley and Flints Hills. This workshop was well attended by clients and stakeholders from EPA Regions 7 and 10, the ORD's Biofuel Research Initiative Program, the Ecosystem Services Research Program, the Oregon Department of Agriculture, the Washington State Department of Agriculture, and the USDA-ARS.

For each project, we have solicited input from our clients/stakeholders, specifically to assist in the development of realistic decision rules for implementing decision support applications. The early involvement of clients/stakeholders in this process is helping us to more effectively tailor the decision support framework to their needs. This should also make it easier to engage them later in exercising the completed framework. The focus will be on exploring the consequences of alternative policies and land management decisions that they are considering. We expect that continual engagement of our clients/stakeholders throughout the lifespans of these projects will lead to better environmental decision making.

In summary, our intent from the beginning has been to produce a broadly applicable framework capable of addressing multiple ecological endpoints and services of interest to a wide range of decision makers. I think our clients would agree that the development of VELMA and its linkage to ENVISION has not been an ad hoc process, as one reviewer suggested, but quite the opposite, i.e., one planned with considerable forethought to serving multiple purposes and clients.

- Examine other decision support tools such as WARMF, RTI-Fire and Fuel Extension, and CEEOT to learn how these models address decision-making for forest industry issues and concerns

We will pursue this excellent recommendation by (1) reviewing the literature describing these decision support tools, and (2) contacting the model developers and persons in the forest industry who are using these tools to make management decisions. The goal will be to look for useful features in these models that could be incorporated into ENVISION-VELMA to better address key forest industry issues and concerns – e.g., assessing the effectiveness of forest practice rules and BMPs in meeting water quality and other environmental objects; assessing sustainability of forest practices; providing a measure of benefits and costs of alternative management practices; meeting certification requirements; and assessing watershed impacts of forest bioenergy management practices. All of these issues are highly relevant to the Willamette Ecosystem Services Project, given the predominance (>60%) of forest cover in the Basin.

- Quantify uncertainty for the ENVISION-VELMA modeling and decision support framework

This is one of the more difficult and important challenges facing any modeling and decision support effort. This is especially true of EPA's Ecosystem Services Research Program (ESRP) and similar efforts that aim to address long-term (years to decades) effects of interacting stressors in real landscapes subject to expanding human influences, and all the uncertainties inherent to that. Because the topic of model uncertainty has come up repeatedly in various contexts besides this review, including the All Scientists ESRP annual meeting in October 2010, I've saved this topic for last so that I can address it in more detail.

There is a large body of literature on uncertainty analysis (e.g., Cacuci 2003; Cacuci et al. 2005; Saltelli et al. 2004) and references therein. Our approach to uncertainty analysis will be guided in large part by the Willamette Ecosystem Services Project (WESP) Implementation Plan (Botle et al. 2010). The WESP plan discusses some of the challenges and approaches that we are pursuing for quantifying model uncertainty:

> "Within WESP we recognized the need to address uncertainty at two fundamental levels: 1) uncertainty associated with model calibration and validation, and 2) uncertainty associated with aggregate analysis within an alternative futures context. We address model uncertainty first, followed by uncertainty associated with integrated assessments within the DSP (ENVISION and associated models)."
>
> "Uncertainty analysis is an important component of the overall modeling strategy. For statistical models, uncertainty analysis should be straightforward, as all error terms are usually generated in the model fitting (calibration) process. For mechanistic (process-oriented) models, complete uncertainty analysis is a major challenge. It may be possible to estimate univariate distributions for uncertain parameters, but the error covariance between parameters (which can be quite important in error propagation) is often quite difficult to determine. Further, estimation of error in the model equations is an even greater challenge. Thus, it is important, when presenting the results of an uncertainty analysis, to clearly indicate how the error terms were estimated, and what error terms were not included in the analysis (Johnston et al. 2008)."

“Model validation involves comparing predictions and observations for an application of a model on data that were not used in calibration. Given the limited amount of data in many modeling studies, the modeler may be forced to calibrate with one year's data and validate on the data collected in the following year. It is unlikely that two consecutive years will be vastly different, so the validation exercise may not too meaningful. Ideally, validation should indicate that the model will perform as the actual ecosystem responds when management actions change some of the forcing variables (e.g., nutrient loading) substantially. Unfortunately, having data that allow this degree of validation is likely to be quite rare. Thus an option to consider is to rate the rigor of the validation exercise by how much the calibration data distributions differ from the validation data distributions (Johnston et al. 2008).”

“Although predictions of complex models (statistical or process-based) cannot be validated in a strict sense (Rastetter 1996), comparisons against a variety of observations and criteria are essential for characterizing model uncertainty and building confidence in a model. In part, such confidence will be proportional to the range of environmental conditions across which a model accurately portrays responses. Similarly, validation methods that employ multiple criteria for assessing model performance provide the most rigorous means of establishing confidence in a model. For example, Reynolds and Ford (1999) describe a multiple-criteria model assessment methodology for characterizing uncertainty associated with ecological theory, model structure, and assessment (validation) data. We will investigate such methodologies for testing WESP models. The success of this will depend on the availability of high quality data sets for the purposes of model calibration and validation. The long history of scientific research and environmental monitoring in the Willamette Basin and surrounding region are providing a data-rich foundation for this effort, as evidenced by our model tests for the HJ Andrews LTER and Alsea Watershed studies.”

“Sensitivity analysis is another aspect that will be developed in the WESP Modeling Strategy. Traditional sensitivity analyses determine the rate of change of model output as a single input is varied by a small amount, while holding the other inputs constant. These rates of change are often converted into relative measures to facilitate comparisons among different inputs. Sensitivity analyses may be used for several purposes, including (a) determining which inputs contribute most to output variability, (b) identifying parameters that are insignificant and may be eliminated from the model, and (c) identifying parameters that require better estimation in order to reduce output uncertainty. McKane et al. (1997) present an example of a sensitivity analysis for a process-based biogeochemistry model that served these purposes.”

“Uncertainty analysis within the WESP decision platform [ENVSION] provides additional challenges, primarily related to the complex interactions of multiple parameters, models, and decision processes that are not based on numerical algorithms. The WESP decision platform integrates qualitative descriptions of decision variables and non-algorithmic decision processes with traditional numerical models. The paradigm for modeling in these synoptic alternative futures assessments is more one of comparative analysis than parameter estimation, i.e., comparing alternative scenarios for their differential outcomes for the suite of services. The first order differences across the suite (with bundling diagrams, for example) are powerful communication results from these kinds of assessments, where the goal is to communicate not a prediction of a specific result, but rather to generate a portfolio of possible outcomes for the alternatives. We contrast the traditional predict-then-act paradigm, which pairs models of rational decision making with methods for treating uncertainty derived largely from the sciences and engineering (Raiffa 1968; Lempert et al. 2003), with an explore-then-test paradigm which is emerging as a viable approach for more complex, decision-oriented assessments. The preferred course of action in predict-then-act assessments is the one that performs “best” given some (typically small) set of assumptions about the likelihood of various futures and the landscape processes that will be sustained if these assumptions prove true. Such assessments are strongly tied to the validity of their assumptions. These approaches are fraught with challenges, especially when applied over the spatial and temporal extents at which important long-term environmental processes operate and when the ecosystem services that people rely on are taken into consideration (Holling 2001). In contrast, the explore-then-test approach seeks actions that are shown to perform well, i.e., are robust across a large number of plausible future alternatives. We define robust decisions as those that result in resilient system behaviors more likely to achieve outcome goals under a variety of possible future trajectories, where aspects of these trajectories are more or less certain. By encompassing a broad range of future possibilities and uncertainties, e.g., local manifestations of climate change, these approaches offer greater potential to be responsive to opportunities and adaptive to problems. By virtue of their exploration of broad sets of contingencies, they also have the potential to serve as constructive means for forging consensus among diverse groups of citizens and policy makers (Lempert et al. 2003)."

"Also important in decision making is the analysis of the tradeoffs among alternative courses of action and the need to address uncertainty. Multi-criteria assessment (MCA) methods are oriented to the multi-dimensional character of many natural resource management problems (Hajkowicz 2007; Kiker et al. 2005). They are designed to overcome the problems of multiple objectives, incompatible units, the need to consider both qualitative and quantitative data, and the need to incorporate stakeholder knowledge and preferences (Chee 2004). These tools are inherently capable of integrating biological, social and economic data, and are ideal for assisting evaluation in data-poor situations."

"Envision employs the use of Monte-Carlo simulation, using statistical descriptors of decision processes and, optionally, scenario-specific statistical descriptors of model inputs, to allow a given scenario to be run multiple times to produce a distribution of possible outcomes. This approach, coupled with a multi-criteria assessment reflecting user-specified weighting of importance of various ecosystem services, is well suited to the explore-then-test paradigm we will emphasize in the decision platform. Further, this approach allows the assessment of landscape vulnerability, i.e. the likelihood of variant or invariant landscape properties to exist under multiple scenario instantiations and multiple policy sets, another useful aspect of uncertainty analysis."

We will also investigate the use of emerging GIS-based strategies for assessing uncertainty associated with effects of changes in land-cover/land use change for environmental planning efforts. For example, Semmens et al. (in press) describe how alternative future scenarios can be used as input to hydrologic models and compared with existing conditions to evaluate potential environmental impacts as part of this process. However, model error can be significant and potentially compounded when projecting future land-cover/use change and management conditions. To address this problem, Semmens et al. used repeat observations of land cover/land use as a proxy for projected future conditions. They then used a systematic analysis of model efficiency during simulations based on observed land-cover/use change to quantify error associated with simulations for a series of known "future" landscape conditions over a 24-year period. Calibrated and uncalibrated assessments of relative change over different lengths of time were also presented to determine the types of information that can reliably be used in planning efforts for which calibration is not possible.

Thus, as VELMA and ENVISION-VELMA applications mature and increase in spatial and temporal scale, we will adopt methods from a variety of sources that prove useful for quantifying uncertainty for our modeling and decision support framework.

Summary

Table 1, below, summarizes the reviewers' main recommendations pertaining to each of the seven charge questions. Their recommendations are organized in Table 1 with regard to priorities, as I see them, for further developing VELMA and its linkage to ENVISION, specifically to address EPA client and stakeholder needs. Addressing all the recommendations would certainly improve VELMA for this purpose. The prioritization of tasks in Table l provides a reasonable path forward for making these improvements, to the extent that available resources allow. It will also be necessary to continually assess the extent to which each change will improve the model's accuracy and explanatory power versus decreasing its computational efficiency and applicability to larger spatial and temporal scales. This is an important tradeoff to consider, given our clients' interests in regional-scale risk assessments and their need to unambiguously link effects to causes.

It is worth noting again that the version of VELMA reviewed here was an initial attempt at a parsimonious solution to this tradeoff. Despite the limitations/opportunities highlighted by this review, this initial version of VELMA provided a relatively simple, spatially-distributed framework for assessing the effects of changes in climate and land use on ecohydrological processes that support a number of vital ecosystem services. The reviewers' tests of the model for their study sites, and our own tests for the HJ Andrews, Konza and Hubbard Brook sites, illustrated the model's potential for predicting these effects with reasonable effort and accuracy.

Finally, I want to once again thank the reviewers for their insightful and constructive comments. It was a pleasure working with them and I greatly appreciate the time and effort they put into learning VELMA. Their recommendations have already proven extremely valuable.

Table 1. Summary of reviewer recommendations and priorities for improving EPA-WED's ecohydrological modeling and decision support framework.

Charge Question	1st Priorities	2nd Priorities	3rd Priorities
1	• Multiple land cover types	• Phenological control of ET • N fixation	• Improve model calibration procedure, GUI and manual • Automate optimization for multiple processes and services
2	• Test processes for wider range of LULC conditions	• Partition plant biomass into leaves (LAI), stems, roots	
3	• Microclimate spatial variation • Radiation effects on snow, ET, NPP	• Effects of roads & streamside BMPs	
4			• Improve DEM processing
5		• Improve control of output for specific grid cells	• Flexible data output summaries
6	• Scale up to larger watersheds, basins	• Instream nutrient processes • Binary output format	
7	• Clarify EPA's need for ENVISION • Develop ENVISION-VELMA linkage with client/stakeholder input	• Quantify model uncertainty	• Learn from other decision support tools to address forest industry decision support needs

References

Abdelnour, A., M. Stieglitz, F. Pan and R. McKane. Catchment hydrological responses to forest harvest intensity and spatial pattern. In review.

Abdelnour, A., F. Pan, M. Stieglitz, and R. McKane. 2010. VELMA User's Manual. EPA-WED internal document.

Bolte, J., R. McKane, D. Phillips, N. Schumaker and D. White. 2010. Community based demonstration projects: Willamette Ecosystem Services Project (WESP) Implementation Plan. USEPA Report: EPA/600/R-10/073.

Boyd, J. and S.Banzhaf. 2007. What are ecosystem services? The need for standardized environmental accounting units. Ecological Economics 63: 616-626.

Buol, S.W., Southard, R.J., Graham, R.C., and McDaniel, P.A. (2003). Soil Genesis and Classification, 5th Edition. Iowa State Press - Blackwell, Ames, IA.

Cacuci, D. G. 2003. Sensitivity and uncertainty analysis, Volume I: theory. CRC Press.

Cacuci, D. G., M. Ionescu-Bujor, and I. M. Navon. 2005. Sensitivity and uncertainty analysis, Volume II: applications to large-scale systems. CRC Press.

Chee, Y.E. 2004. An ecological perspective on the valuation of ecosystem services. Biological Conservation 120(4): 549-565.

Confesor, R.B., Whittaker, G.W. 2007. Automatic Calibration of Hydrologic Models with Multi-Objective Evolutionary Algorithm and Pareto Optimization. Journal of the American Water Resources Association.43:1-9.

Compton, J.E., M.R. Church, S.T. Larned and W.E. Hogsett. 2003. Nitrogen Export from Forested Watersheds in the Oregon Coast Range: The Role of N 2-fixing Red Alder

Cushing, J.B., N. Kopytko, N. Stevenson-Molnar, L. Zeman, S. Stafford, J. Bolte, B. Bond, D. Lach, and R. McKane. 2009. Enabling the Dialog – Scientist <> ResourceManager <> Stakeholder: Visual Analytics as Boundary Objects, IEEE Intelligent Systems, Special Issue on AI, E-Government and Politics 2.0 (Hsinchun Chen, ed.), Sept-Oct. 2009.

Daly, C., J.W. Smith, J.I. Smith and R.B. McKane. 2007. High-Resolution Spatial Modeling of Daily Weather Elements for a Catchment in the Oregon Cascade Mountains, United States. J. Appl. Meteor. Climatol., 46, 1565–1586.

Daly, C., D.R. Conklin and M.H. Unsworth. 2009. Local atmospheric decoupling in complex topography alters climate change impacts. International Journal of Climatology. International Journal of Climatology

Hajkowicz, S. 2007. Allocating scarce financial resources across regions environmental management in Queensland, Australia. Ecological Economics 61(2) 208-216.

Holling, C.S. 2001. Understanding the complexity of economic, ecological and social systems. Ecosystems 4(5): 390-405.

Jakeman, A. J., and G. M. Hornberger. 1993. How much complexity is warranted in a rainfall-runoff model? Water Resources Research, 29, 2637–2649.

Johnston, J., R. McKane, B. Rashleigh, S. Raimondo, S. Yee, R. Kreis, K. Reckhow, M. White, R. Lassiter, G. Laniak, and D. White. 2008. Modeling Strategy for ESRP. Unpublished internal document, US EPA.

Kiker, GA, TS Bridges, A. Varghese, T.P. Seager, and I. Linkov. 2005. Application of Multicriteria decision analysis in environmental decision making. Integrated Environmental Assessment and Management 1(2): 95-108.

Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42

Lempert, R.J., S.W. Popper, and S.C. Bankes. 2003. Shaping the next one hundred years: New methods for quantitative, long-term policy analysis. RAND Corporation, Santa Monica, CA, 187 pp.

Liu, Z., and D.E. Weller. 2008. A stream network model for integrated watershed modeling. Environmental Modeling and Assessment 13: 291-303.

McKane, R., E. Rastetter, G. Shaver, K. Nadelhoffer, A. Giblin, J. Laundre and F. Chapin. 1997. Climatic effects on tundra carbon storage inferred from experimental data and a model. Ecology 78:1170-1187.

Mulholland, P.J., et al. 2008. Stream denitrification across biomes and its response to anthropogenic nitrate loading. Nature 452: 202-205 (13 March 2008).

Raiffa, H. 1968. Decision analysis: introductory lectures on choices under uncertainty. Oxford, England: Addison-Wesley.

Rastetter, E.B., P.M. Vitousek, C. Field, G.R. Shaver and D. Herbert. 2001. Resource Optimization and Symbiotic Nitrogen Fixation. Ecosystems, 2001, Volume 4, Number 4, Pages 369-388.

Rastetter, E. B. 1996. Validating models of ecosystem response to global change. BioScience 46(3):190-198.

Reynolds, J.H., and E.D. Ford. 1999. Multi-Criteria Assessment of Ecological Process Models. Ecology 80: 538-553.

Saltelli, A., S. Tarantola, F. Campolongo, and M. Ratto. 2004. Sensitivity analysis in practice: a guide to assessing scientific models. Wiley, New York.

Semmens, D.J., M. Hernandez, D.C. Goodrich, and W.G. Kepner. A retrospective analysis for forecasting hydrologic change. In press.

Waichler, S., et al. (2005), Simulation of water balance and forest treatment effects at the HJ Andrews Experimental Forest, Hydrological Processes, 19(16), 3177-3199.

Warren, J., et al. (2005), Vertical stratification of soil water storage and release dynamics in Pacific Northwest coniferous forests, Agricultural and Forest Meteorology, 130(1-2), 39-58.

Wigmosta, M., et al. 1994. A distributed hydrology-vegetation model for complex terrain, Water Resources Research, 30(6), 1665-1680.

Wolheim, W.M., et al. 2006. Relationship between river size and nutrient removal. Geophysical Research Letters, Vol. 33, L06410.