ELSEVIER

Fuzzy Sets and Systems 113 (2000) 53–67

FUZZY sets and systems

www.elsevier.com/locate/fss

A multiple criteria decision support system for testing integrated environmental models

D. Scott Mackay[a,*], Vincent B. Robinson[b]

[a] *Department of Forest Ecology and Management and Institute for Environmental Studies, University of Wisconsin-Madison, Madison, WI 53706, USA*

[b] *Department of Geography, University of Toronto, Erindale Campus, Mississauga, Ont., Canada L5L 1C6*

Received November 1998

Abstract

Spatial models of ecological and hydrological processes are widely used tools for studying natural systems over large areas. However, these models lack specific mechanisms for reporting output uncertainty contributed by model structure, and so testing their suitability for studying a large range of problems is difficult. This paper describes a method of evaluating the uncertainty contributed by underlying assumptions used in constructing integrated environmental models from two or more sub-models that were developed for different purposes. Integrated environmental models are typically constructed from many individual process-based models. Conflicting assumptions between these sub-models, e.g., spatial scale differences, are easily overlooked during model development and application. This “semantic error” cannot be predicted prior to simulation, as it may only emerge through the interaction of sub-models applied to a particular set of data used to drive a simulation. Model agreement is proposed and demonstrated as a way to detect problems of model integration at the state variable level within an integrated ecosystem model. This model agreement is then propagated to model response variables using multiple criteria to examine their sensitivity, predictability, and synchronicity to the measured uncertainty in state variables. These three properties are combined under fuzzy logic in order to provide decision support on where, for a given time during simulation the sub-models agree on a particular response variable. This paper describes the details of the approach and its application using an existing integrated environmental model. The results show that, for a given set of model inputs and application, integrated environmental models may have spatially variable levels of agreement at the sub-model level. The results using RHESSysD, a spatially integrated ecosystem hydrology model, indicate that semantic error in estimates of plant available soil moisture are consistent with observations of the need for resetting events, such as flooding, to initialize the model to a point where further simulation results can be trusted. These results suggest that a dynamic selection of sub-models may be warranted given a reasonable method of determining sub-model disagreement during simulation. Fuzzy set theory may be a useful tool in arriving at such a model selection process as it allows for a relatively straightforward synthesis of numerous model evaluation criteria with a large quantity of output from the model.

Keywords: Artificial intelligence; Control theory; Decision making; Environmental modeling; Operators; Multicriteria analysis; Spatial scale

* Corresponding author. Tel.: +1-608-262-1669; fax: +1-608-262-9922.

E-mail addresses: dsmackay@facstaff.wisc.edu (D.S. Mackay); vbr@geophagia.erin.utoronto.ca (V.B. Robinson)

0165-0114/00/$ - see front matter
PII: S0165-0114(99)00012-3

1. Introduction

Concern over the role of human activity on our environment has increased the demand for integrated, spatially-distributed, environmental models that address the interactions of human activity, the terrestrial biosphere, climate, and hydrology. Furthermore, the widespread availability of geographical information systems (GIS) to support spatial data processing and analysis is making spatial models accessible to a larger audience that includes policy makers. As a result, there has never been a greater need for decision support tools to help in evaluating the applicability of complex environmental models to a given problem. One of many possible requirements for such models is that their underlying assumptions be consistent with a potential application of the model. Models abstract from reality to varying degrees by assuming that certain properties of the system represented can be ignored. Some of these simplifying assumptions will be known to a model end-user, while others will not. The more deeply rooted assumptions, such as those based on spatial aggregation, are often poorly documented and thus less well understood to model end-users. If explicit knowledge of the assumptions underlying a process-based model is unknown, then conflicting assumptions resulting from a combination of such models will undoubtedly produce unexpected, and often difficult to evaluate, output. Conflicting assumptions in geographical information systems have previously been described as semantic errors [27,35], since they are usually attributable to differences in interpretation of a modeled reality and usually contribute to the overall error. Semantic error in the context of integrated models also contributes to error in the form of over-prediction or under-prediction of model output. While there are many sources of error in environmental models it is this semantic error contributed by model structure that we address here.

For the purposes of this paper integrated models are composed of many individual process-based models (or sub-models) that describe different processes. Each sub-model represents an interpretation of system behavior, intended for the purpose of (1) elucidating complex system behavior, or (2) predicting system behavior [15]. A model developed as a null hypothesis may be used for making predictions if it demonstrates a convincing level of validity. Such adaptation is a potential source of uncertainty for environmental models, which are not subject to formal proof and so are neither unconditionally validated nor invalidated [28]. Since the conditional requirements for using a model are either not reported or are subject to interpretation, there is a danger that the perceived "...theoretical rigor that underlies the models will engender uncritical belief in their prediction" [20].

This paper presents a methodology and example prototype software system for testing combinations of process models intended to be integrated into an environmental model and applied to a particular problem. A new method of managing semantic errors caused by the structure of integrated environmental models is developed to evaluate model suitability for answering specific questions. Here it is suggested that a decision about sub-model suitability as part of an integrated model can be made given a knowledge of how model response variables are affected by semantic error determined by a divergence from some expected system behavior. The decision-making framework upon which this decision is made is referred to as *model agreement*. This framework is described in the next section. A method of measuring semantic error in the context of a particular integrated model is then described, and finally we will step through an example to illustrate the application of the system.

2. Decision making framework

2.1. Model agreement

Hayes-Roth and Jacobstein [12] suggest that knowledge bases give application programs increased flexibility by lending them reasoning ability. Previous work in knowledge-based simulation modeling has demonstrated improved control over model development, analysis of model results, and model use [4,14]. Knowledge-based modeling has improved the rigor and provability of some earth science models. For instance, decision-support systems have aided in the application of both hydrological and ecosystem models [7,10,30,34,36]. Decision-support systems help in choosing, from a repository of models, a collection of models that is suitable for a given problem. While the model selection process remains largely human-guided a number of model description languages have

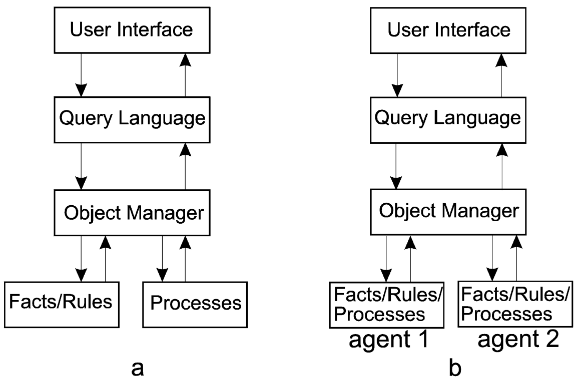

Fig. 1. Model-based management can be approached from the perspective of the model as (a) black-box, or (b) an integral part of the software system in which all components, e.g., sub-models, must be viewed as acting together towards achieving a common goal.

emerged to formalize some of the human decision processes and information management associated with model development [3,5,6,19,33,38].

Model information management can be approached at different levels (Fig. 1). In the model information management strategy that is most commonly seen within the GIS community (Fig. 1a), all decision rules are held separate from the process models and decisions are made centrally [10,21,36]. This is considered a coarse level of integration in which semantic errors within the process models cannot be detected, as they depend on the interactions between sub-models during simulation. A finer level suggested here (Fig. 1b) allows sub-models to measure semantic error that emerges through model interaction. The degree to which semantic errors in the linkages between models affect model output is here referred to as model agreement.

Model agreement follows from model theoretic ideas, which are applied to databases and knowledge bases [18]. A model theory clearly states the semantics of a knowledge base, *KB*, and provides the software system an unambiguous understanding about which statements added to *KB* agree with *KB* (Fig. 2a). The statements based on model theory are either true or false in the context of the existing database or knowledge base. Here we must deal with the fact that process models do not lend themselves to mathematical proof [28], and so, by definition, semantic equivalence of process models must follow a different paradigm that is not based on a Boolean set. Instead, a continuous measure of agreement between models is suggested (Fig. 2b). Two models are in perfect agreement if they give the same system response under the same conditions. Integrated models that are designed top-down, in which all sub-models are designed together for a specific goal, will have sub-models that are compatible in their underlying space, time, and attribute assumptions. However, large integrated models are more likely to be constructed in a bottom-up fashion, in which sub-models that were originally designed for different goals are combined to solve a new set of goals. Thus, integrated models are most likely to have conflicting assumptions between sub-models, and so the sub-models may diverge from perfect agreement. Instead of being true or false, the combined models do not totally agree on all variables of interest to the end-user of the model.

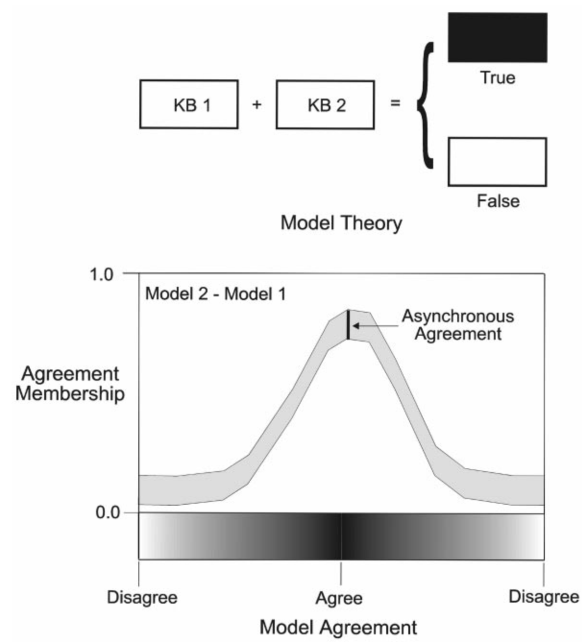

Fig. 2. Model theory requires perfect agreement between knowledge bases KB 1 and KB 2. Model agreement measures a shift from a central concept representing perfect agreement between models.

While it is obvious that a detected semantic error can be corrected given a better model, we suggest that a given integrated model structure may perform admirably over a wide range of applications and then easily trusted for all applications. However, there is as yet little theory on how to determine when an integrated model is misbehaving at a semantic level. Here it is suggested that semantic errors can be detected for a larger number of variables than more traditionally tested variables that are corroborated against measurements. Many modeled variables have no field-measured values for which a rigorous model corroboration can be made. For instance, the spatial variability of soil moisture is difficult to measure over large areas, and yet it has a profound influence on surface runoff, which is routinely measured. For the model described here we found that semantic errors could be detected in storage, or state, variables such as water table depth. The propagation of this detectable error to other model variables, such as soil moisture, give the modeler insight on how and when to change model structure. Of course, even the "best" integrated models may have undetectable semantic errors, for which the methodology presented here cannot help. However, the methodology relies on the modeling being able to specify all simplifying assumptions used to translate what is currently known about the physical system into a model structure. The discussion that follows assumes that semantic error is directly related to violation of one or more underlying assumptions of the models and that this error is propagated to variables of interest to the model developer. A multiple criteria method of transferring model agreement from measured semantic error to output variables is described in the next section.

2.2. Transfer of model agreement

Fig. 3 shows the overall flow of processes within the decision support system. If we assume that the response variable is defined in some query to the integrated modeling system, then such a query could have the form, $\langle \alpha, \theta, \tau, \mu(\varepsilon_\alpha) \rangle$, where α is the goal of the query (or response variable of interest), θ and τ respectively define the spatial and temporal domains within which α is defined, and $\mu(\varepsilon_\alpha)$ is a membership function that describes the extent to which α is a member of the fuzzy set [23] describing model agreement for the goal. We assume that a membership of 1.0 will be used to denote that α was computed with perfect agreement between sub-models, and a membership of 0.0 will denote perfect disagreement.

The simulation model provides agreement amounts in both the response and state variables. Agreement in response variables and state variables are distinguished in that agreement in a response variable, ε_α, is said to be functionally related to semantic error in a state variable, ε_ϕ:

$$\varepsilon_\alpha = f\langle \varepsilon_\phi, M \rangle, \tag{1}$$

where the function, f, combines one or more criteria, and M represents the integrated model itself. The actual computation of ε_α is made by parallel simulation in which an output variable α is calculated twice: (1) α_1 with ε_ϕ uncorrected and (2) α_2 with ε_ϕ corrected. The difference between α_1 and α_2 is ε_α, which can then be compared to ε_ϕ to assess how semantic error is transferred to variables of interest. A one-to-one relationship between ε_α and ε_ϕ is not necessary to determine the actual transfer of agreement, as the final agreement is reported as a membership regardless of the absolute sign of the relationship between ε_α and ε_ϕ. For instance, there could be a form of hysteresis in which ε_α may take on completely different values at different times while ε_ϕ remains constant. Hysteresis is partially accounted for by incorporating a synchronicity criterion as described below.

We identified three criteria for measuring the transferral of semantic error to goal variables. The criteria are: (1) sensitivity, (2) predictability, and (3) synchronicity. These criteria were selected to be consistent with indicators used in sensitivity analysis to assess model performance. Sensitivity describes the total amount of variability in a response variable that is attributable to a given amount of model disagreement measured for a state variable. The membership function for sensitivity is computed as

$$\mu\langle \sigma_\alpha \rangle = \min \left\{ 1 - \left[\frac{\varepsilon_\alpha(i,t) - \min(\varepsilon_\alpha)}{\max(\varepsilon_\alpha) - \min(\varepsilon_\alpha)} \right] \right\} \tag{2}$$

for all i patches and time intervals, t, within a defined spatial domain. A spatial domain is represented as a collection of one or more patches. Patches may be considered as a group in order to reduce the number of end-member response functions and facilitate easier presentation to a model designer. The membership function that results from the transformation in

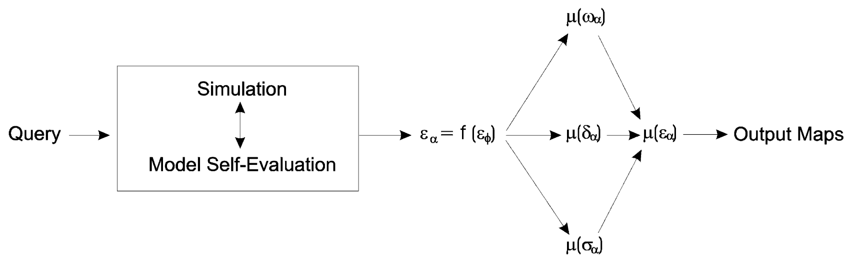

Fig. 3. Flow of processes within the decision support system. RHESSysD is contained within the box, while all components to the right of the box are considered part of the decision support system.

Eq. (2) describes the worst case sensitivity of the response variable to measured disagreement. The shape of this function indicates how sensitive the response variable is to model disagreement.

Predictability describes the rate of change of ε_α with respect to ε_ϕ. Slowly or smoothly varying responses are predictable over a small range of variation in semantic error. Sharp transitions or discontinuities make it difficult to predict response variable behavior for even small amounts of model disagreement. A response variable whose behavior is linearly related to disagreement in a state variable would yield a predictability membership function with a value of 1.0. The membership function for predictability is computed as

$$\mu(\delta_\alpha) = \min\left\{1 - \left[\frac{\delta_\alpha(i,t) - \min(\delta_\alpha)}{\max(\delta_\alpha) - \min(\delta_\alpha)}\right]\right\}, \tag{3}$$

where $\delta_\alpha = \partial\varepsilon_\alpha/\partial\varepsilon_\phi$ is the rate of change of the response variable with respect to measured model disagreement.

Synchronicity describes the degree to which agreement in a response variable follows state variable agreement. In many cases a response variable disagreement may peak at a later time during a simulation than the peak disagreement in a state variable. This lag is attributable to a storage effect in many physical systems. For instance, some water stored in the soil by a rising water table during a spring snow melt often remains in the soil well into the summer months at which time it may be depleted by transpiring vegetation. An error in calculating the position of the water table can result in an error in soil water storage. Although the error in the water table position may occur over only a short time period, e.g., the time it takes to melt a snow pack, the error propagated to soil moisture storage remains for a longer period of time. This storage effect results in an asynchronous response, possibly in the form of some hysteresis, as illustrated in Fig. 2b. The actual width of the grey band is variable and is conceptually defined as the difference in relative residual semantic error between the response and state variables, for a given spatial location, calculated over the full time domain of the simulation. It is computed as follows:

$$\mu(\omega_\alpha(i)) = 1 - |\gamma(i) - v(i)|, \tag{4}$$

where

$$v(i) = \frac{\Delta\varepsilon_\alpha(i) - \min(\Delta\varepsilon_\alpha)}{\max(\Delta\varepsilon_\alpha) - \min(\Delta\varepsilon_\alpha)}, \tag{5}$$

$$\gamma(i) = 1 - \frac{|\varepsilon_\phi(i)| - \min(|\varepsilon_\phi|)}{\max(|\varepsilon_\phi|) - \min(|\varepsilon_\phi|)}, \tag{6}$$

respectively, describe relative magnitude of the range of residual error in a response variable for a given amount of state variable disagreement, and relative magnitude of the state variable disagreement. The range of residual error in the response variable is given by

$$\Delta\varepsilon_\alpha(i) = \max(\varepsilon_\alpha(i)) - \min(\varepsilon_\alpha(i)). \tag{7}$$

Finally, an overall membership function for model agreement with respect to a particular goal variable is given by

$$\mu(\varepsilon_\alpha(i)) = \min[\mu(\sigma_\alpha(i)),\ \mu(\delta_\alpha(i)),\ \mu(\omega_\alpha(i))], \tag{8}$$

where $\mu(\varepsilon_\alpha)$ gives the membership of response variable, α, in the model agreement set for that variable.

We choose to use the minimum function on the grounds that low membership in any one of sensitivity, predictability, and synchronicity should produce undesirable integrated model response. How these individual criteria interact with each other is otherwise unknown. Furthermore, it is assumed that disagreement is given by the complement of agreement, $\neg\mu(\alpha)$, such that a group of sub-models are in agreement with respect to α if $\mu(\varepsilon_\alpha) \geqslant 0.5$.

The preceding discussion is easily extended to multi-variate goals by recognizing that each goal variable in a query can have its own membership function and decision-making process. However, specific rules for combining the individual membership functions to assess overall agreement of the multi-variate query are beyond the scope of the present work. In the next section we describe a particular environment model and a method of measuring semantic error within it.

3. Measuring model semantic error

To demonstrate our approach we used the Regional HydroEcological Simulation System-Dynamic (RHESSysD). RHESSysD is a spatial information processing and modeling environment for simulating water, carbon, and nutrient fluxes (Fig. 4). The model incorporates a hydrology model, TOPMODEL [22,26], to link a collection of landscape patches via ground water. Vertical water balance is maintained within each patch using a collection of models to represent evaporation, transpiration, soil drainage, capillary rise, snow melt, and storage of water. Details of the mechanics of the integrated model are given in [8,25]. Some details of the hydrology model that determines the position of the water table at each time step during simulation are described here. The position of the water table in turn affects soil water storage, which is used by vegetation and determines the spatial extent of saturated areas that can generate surface runoff. The position of the water table is also used to control the rate of ground water flow, which sustains stream flow. In TOPMODEL the depth to a water table in thin, well-drained soils on moderate to steep slopes is represented by

$$z_i = \langle z \rangle - \frac{1}{f}(\lambda - TSI), \tag{9}$$

where z_i is the water table depth at a particular land surface area (or patch), $\langle z \rangle$ is a mean water table depth, f is a soil hydraulic parameter, λ is the mean topography-soils index (TSI) [29], where TSI accounts for horizontal water flow convergence in concave areas, water retention in areas with relatively low soil transmissivity, flow divergence in convex areas, and rapid drainage in areas of high soil transmissivity [26]. Patches with high TSI will tend to have small z values while patches with low TSI will have large z values. Eq. (9) allows for a redistribution of water based on a probability distribution of TSI with no regard for spatial position other than its connected to a single simulation unit. This allows TOPMODEL to effectively move water laterally without knowing explicit connections between patches. This lack of a spatially explicit hydrological model poses a challenge for constructing spatially distributed models such as RHESSysD, yet these simpler models remain in the mainstream because of their applicability to addressing watershed management problems for which there may be a paucity of detailed spatial information [16,20].

Patches are represented as groups of cells within a raster GIS. Patches are in turn organized into subcatchments over which $\langle z \rangle$ and λ are computed. Patches within a given subcatchment that have the same *TSI* are considered hydrologically equivalent. Each patch is assigned vegetation, soil, and other attributes using GIS analysis. *TSI* assumes that subsurface recharge is spatially uniform in order to allow the water table to rise and fall uniformly with the subcatchment average water recharge. A net positive recharge raises the water table the same amount everywhere, while a net water loss lowers the water table the same everywhere. The model generates runoff from any patch in which $z_i < 0$. The uniform rise and fall of the water table is only considered a reasonable representation under conditions of spatially uniform water recharge to the water table. Otherwise, transient localized water table mounding occurs for short periods of time in areas of greater than average recharge. Spatial variations in snow melt rates, evaporation and transpiration rates, and drainage determine the actual spatial distribution of water recharge in RHESSysD, which is not required to be uniform. The result is a spatial scale mismatch between TOPMODEL and the other models within RHESSysD. Mackay [9] tested the validity of

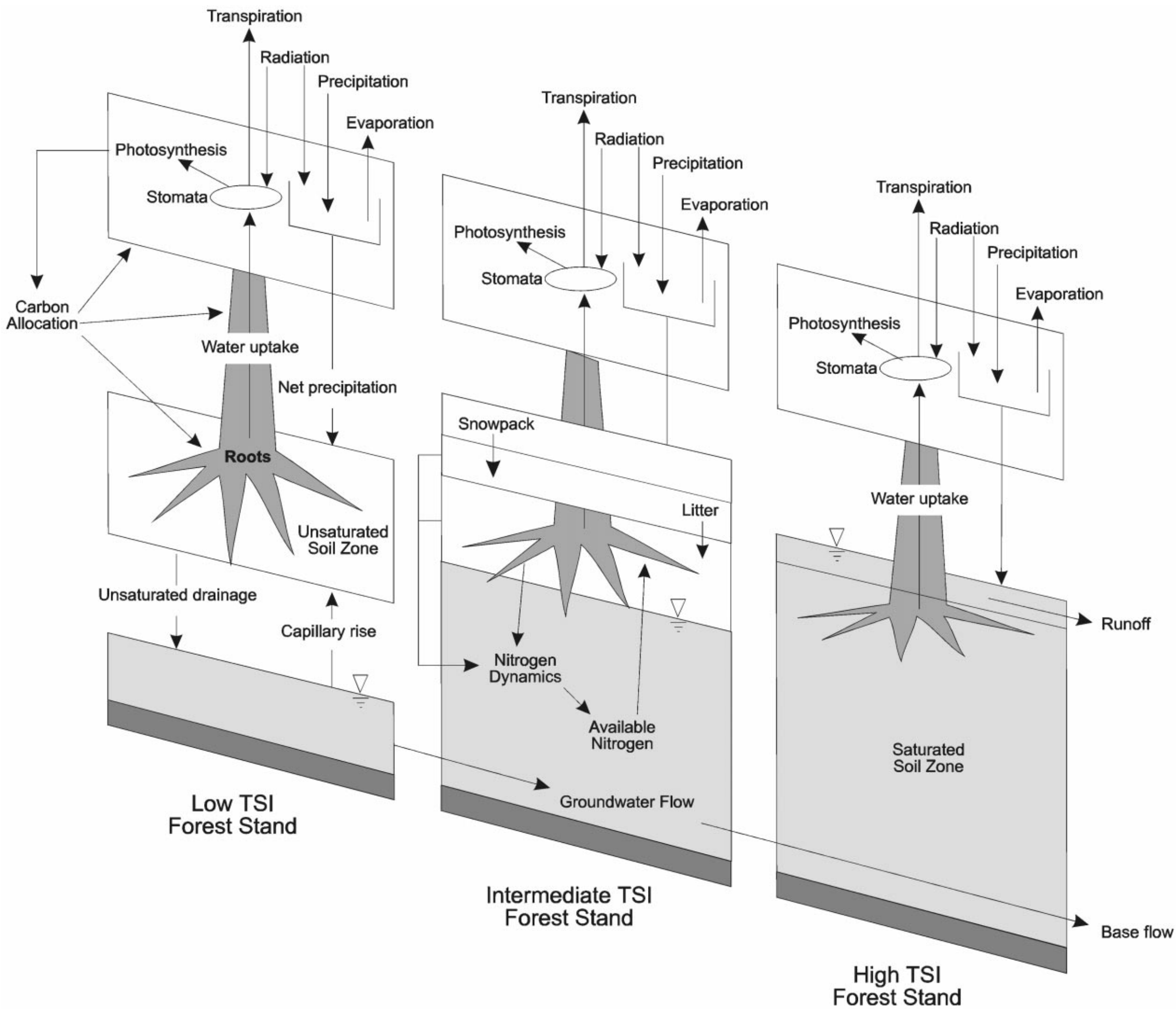

Fig. 4. Physical processes represented by RHESSysD. Shown are three patches (e.g., forest stands) with vertical water fluxes and horizontal water fluxes.

Eq. (9) given spatial variation in surface-subsurface water flux processes, by incorporating a model self-evaluation tool that differentiates between expected subsurface flow and actual subsurface flow as represented by TOPMODEL.

Fig. 5 describes the process of differentiating between expected subsurface flow and semantic error due to variable recharge. Patches are spatially arranged into groups by elevation, allowing for water redistribution using Eq. (9) within each elevation group, but restricting between-group water redistribution to occur only from higher to lower elevation. This restriction assumes that the partitioning of the landscape into a series of hillslope facets attached to streams captures the dominant landscape dissection and that major pathways of surface runoff are not contained within hillslopes. It is recognized that this assumption may not always hold and it may be possible that runoff generating areas higher on the slope will occur. However, it is the redistribution of excess recharge that occurs in runoff generating areas to source areas that constitutes a semantic error here. By restricting water

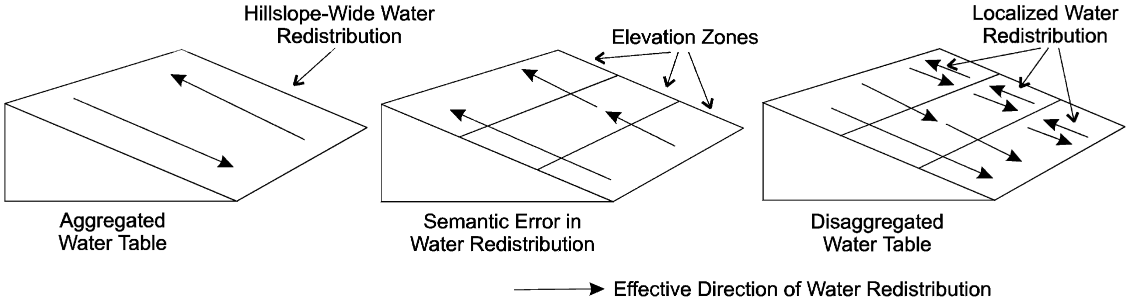

Fig. 5. Conceptual representation of semantic error in the redistribution of water to compute water table depths.

redistribution we hope to approximate the behavior of explicit routing models that move water in only a down slope direction without creating an explicit routing model that would be difficult to integrate into the existing model structure. The unrestricted water redistribution of TOPMODEL allows water to effectively be moved in an up slope direction under certain conditions in which higher than average recharge occurs in the lower elevation areas within the subcatchment. For example, snow melt rates are usually higher at lower elevation sites than at high elevation sites due to the higher temperatures, and so snow melt processes can violate the uniform recharge assumption. The difference between unrestricted water redistribution and restricted water redistribution gives the patch level semantic error.

Following Moore and Thompson [32] Eq. (9) is solved for patch i using the current water table position at patch j. Solving for z_j gives an estimate for z_j as

$$\hat{z}_{j_i} = z_i + \frac{1}{f}(TSI_i - TSI_j). \tag{10}$$

A residual between stored water table depth, which was computed in a previous iteration of the model using the correction described below in Eq. (13), derived from Rules 1 and 2 (below), and the estimated depth is computed as

$$\Delta R_{j_i} = z_j - \hat{z}_{j_i}, \tag{11}$$

where z_j is explicitly stored from the previous time step. Eq. (11) describes the effect of redistributing water from (to) i to (from) j, but we still need to determine how this potentially contributes to semantic error. Since we would like to preserve the down-slope movement of water that is produced by explicit routing models, semantic error can be quantified to determine if lower elevation patches contribute water to higher elevation patches. The following rules define up-slope and down-slope components of the water redistribution:

Rule 1 (*Water incorrectly redistributed to an up-slope patch from patch i*):

$$\forall j, \in \textit{sub-area} \mid \textit{elevation}(j) > \textit{elevation}(i),$$
$$i \in \textit{sub-area}$$
$$\Rightarrow \varepsilon_{i_{\text{upslope}}} = \text{MAX}\left(0, \sum_{j=1}^{k} \Delta R_{j_i} w_j\right).$$

Rule 2 (*Water incorrectly distributed from a down-slope patch to patch i*):

$$\forall j, \in \textit{sub-area} \mid \textit{elevation}(j) < \textit{elevation}(i),$$
$$i \in \textit{sub-area}$$
$$\Rightarrow \varepsilon_{i_{\text{downslope}}} = \text{MIN}\left(0, \sum_{j=1}^{k} \Delta R_{j_i} w_j\right),$$

where w_j is weighting of sub-area j by its relative area. Total semantic error, which is error in the position of the water table due to violation of the uniform recharge assumption of TOPMODEL, at patch i is

$$\varepsilon_{\phi}(i) = \varepsilon_{i_{\text{upslope}}} + \varepsilon_{i_{\text{downslope}}} \tag{12}$$

where $\varepsilon_{\phi}(i)$ is positive when the predicted water table is too deep and negative when the modeled water table is too near the surface. Given a knowledge of semantic error a "true" patch water table is computed as

$$z_i = (\langle z \rangle - \varepsilon_\phi(i)) - \frac{1}{f}(TSI_i - \lambda), \tag{13}$$

where $\varepsilon_\phi(i)$ is used as a correction to the mean water table depth used in redistribution. This corrected model (Eq. (13)) is then applied iteratively with Eq. (11) in the next simulation time step. By using Eqs. (9) and (13) in parallel two sets of output variables are generated representing, respectively, uncorrected and corrected sets. The differences between respective corrected and uncorrected variables, $\varepsilon_\alpha(i)$, that arise using the different sets of state variable calculations are then applied in Eq. (8) to obtain model agreement (or disagreement) about α at patch i.

4. Results

We tested the decision support system using data sets obtained for Onion Creek, a 13 km^2 watershed in the Central Sierra Nevada in California. Relief in this watershed ranges from about 1600 m a.s.l to 2600 m a.s.l. Annual precipitation averages about 1300 mm and occurs primarily as snowfall. In this region hydrological processes are dominated by snow pack storage and snow melt. We partitioned the watershed [24] into subcatchments, elevation bands, and TSI [29] patches (Fig. 6). We then ran the decision support system with the following query:

Where do the sub-models agree on soil moisture predictions at the start of the growing season?

In this query the goal is soil moisture, the spatial domain is defined by the spatial layers shown in Fig. 6, and the temporal domain is the start of the growing season, which corresponds with the end of the snow melt period. Fig. 7 shows six zones defined by spatially aggregating TSI patches. This aggregation is made to improve presentation clarity and reduce computational requirements. The actual number of spatially aggregated areas can be defined by the end-user. Fig. 8 shows the corresponding membership functions for (a) sensitivity, (b) predictability, and (c) synchronicity of soil water variability as functions of semantic error in the water table position for the six zones. In these relations the x-axis represents semantic error as computed by Eq. (12), with positive values corresponding to locations where the water table depth is over-predicted by the model and negative values corresponding to where the water table is predicted to be too close to the surface. The plots were generated by fitting a weighted, localized robust line [39] through the points representing the lower envelope of the scatter of points representing the respective relationships. The lower envelope is presented as the most conservative of memberships. In general, the memberships tend to be larger at more positive levels of semantic error. This is expected to occur in higher recharge areas, which are typically saturated during the snow melt period regardless of the semantic error and hence do not influence plant available soil water greatly. This is clearly shown in the final membership functions for agreement on soil moisture for each of the six aggregated spatial areas (Fig. 9). The final membership functions were plotted using the same fitting method as described above [39]. The highest memberships tend to occur in areas around streams, where soil moisture is near saturation with or without the semantic error removed. Fig. 8a shows that drier (lowest TSI) areas tend to be more sensitive to small amounts of error in predicting the water table position. This suggests that the integrated model provides semantically reasonable plant available soil moisture prediction in valley bottom areas and poor prediction of soil moisture in drier areas. This is not surprising given that TOPMODEL was designed for predicting stream flow response, not providing reasonable estimates of plant available soil water.

Mackay [9] showed that semantic error in the water table position approaches zero everywhere at the end of snow melt. This period of time corresponds with the beginning of the vegetation growing season. Thus, an evaluation of the memberships for each aggregated area when semantic error of the water table position is zero was used to evaluate the results of the query. Fig. 10a shows the fuzzy memberships in spatial set *agreed soil water*, which were produced by mapping the fuzzy memberships from Fig. 9 at a point where semantic error in the water table position (ε_ϕ) is zero, for each of the six respective zones. Areas of low soil moisture agreement membership occur where soils should normally have low water content, but have increased water content due to error in the water table depth redistribution. Also shown (Fig. 10b) is the crisp

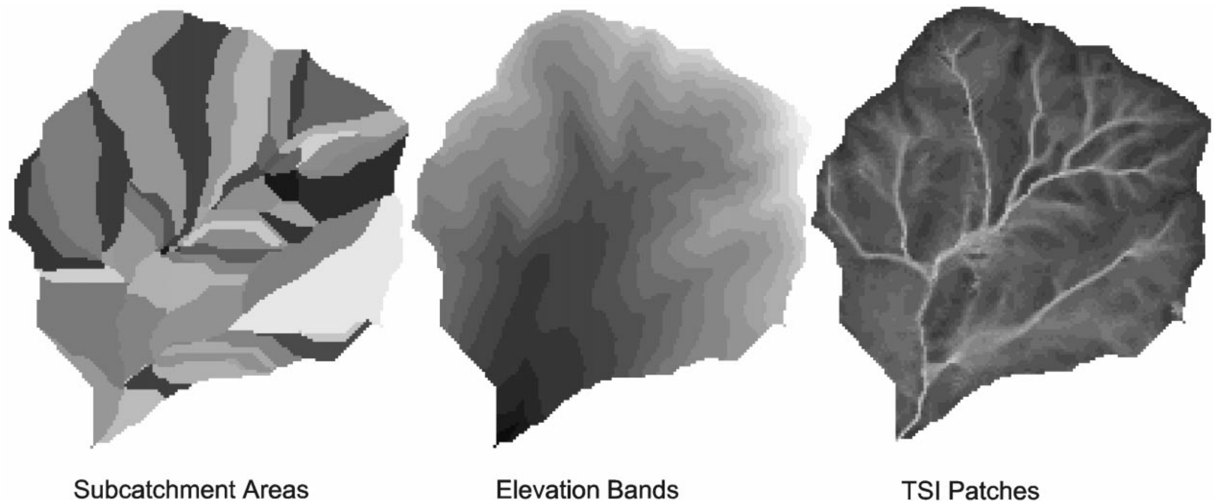

Fig. 6. Partitioning of Onion Creek into subcatchment areas, elevation bands or zones, and TSI patches for spatial representation of input to RHESSysD.

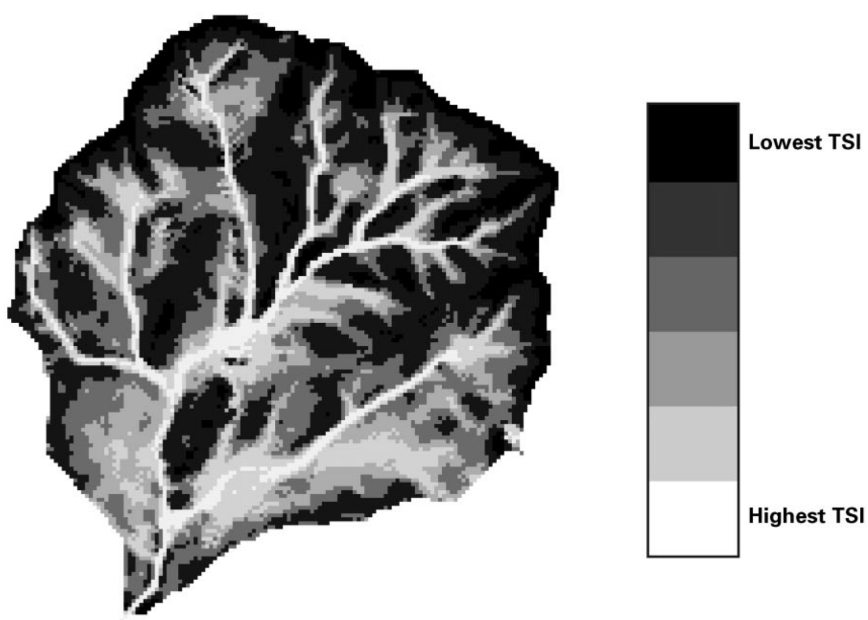

Fig. 7. Spatially aggregated TSI zones used for model agreement analysis and presentation of final memberships.

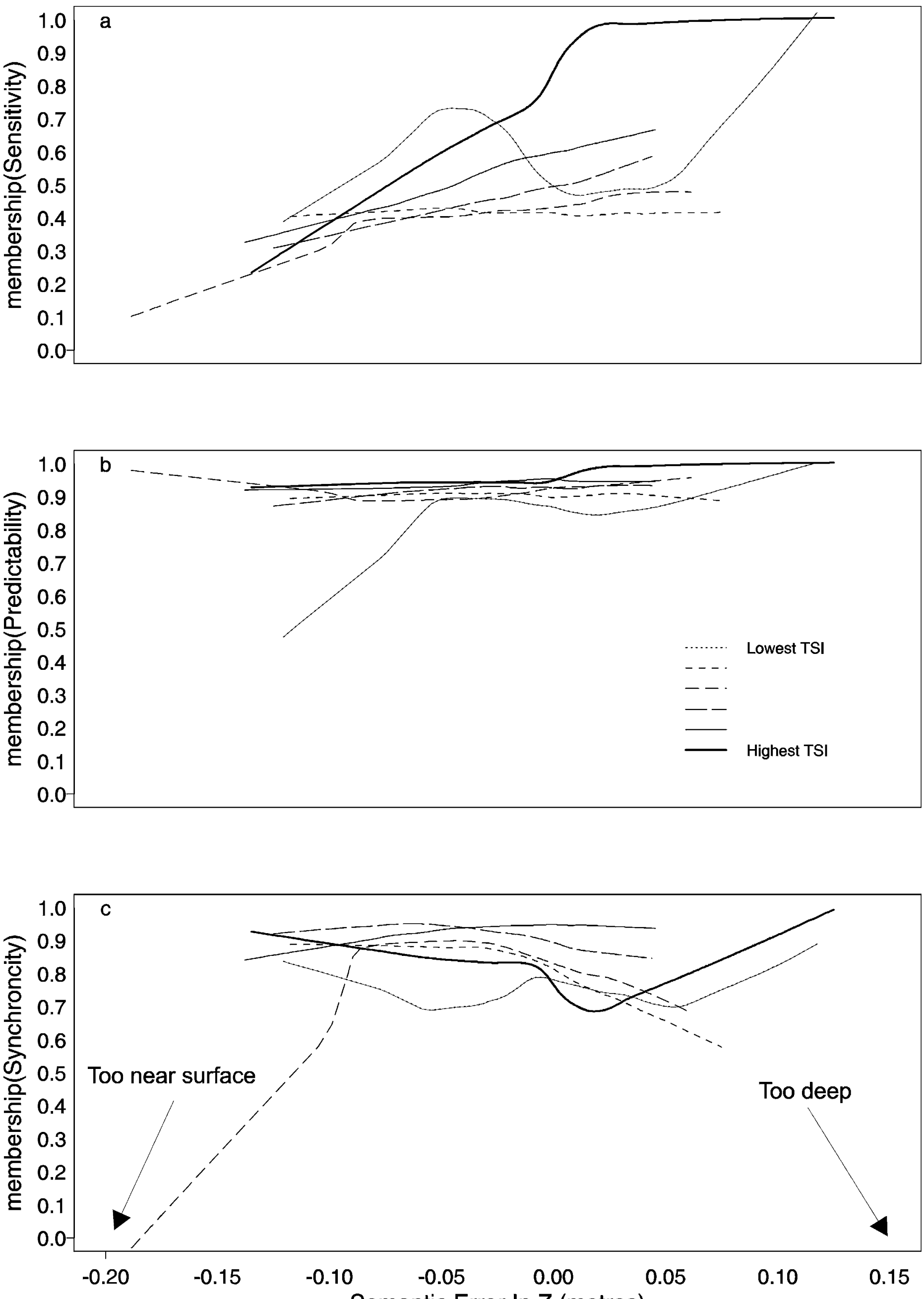

Fig. 8. Plots of (a) sensitivity memberships, (b) predictability memberships, and (c) synchronicity memberships for soil moisture versus semantic error in the water table, for each of the respective aggregate TSI zones.

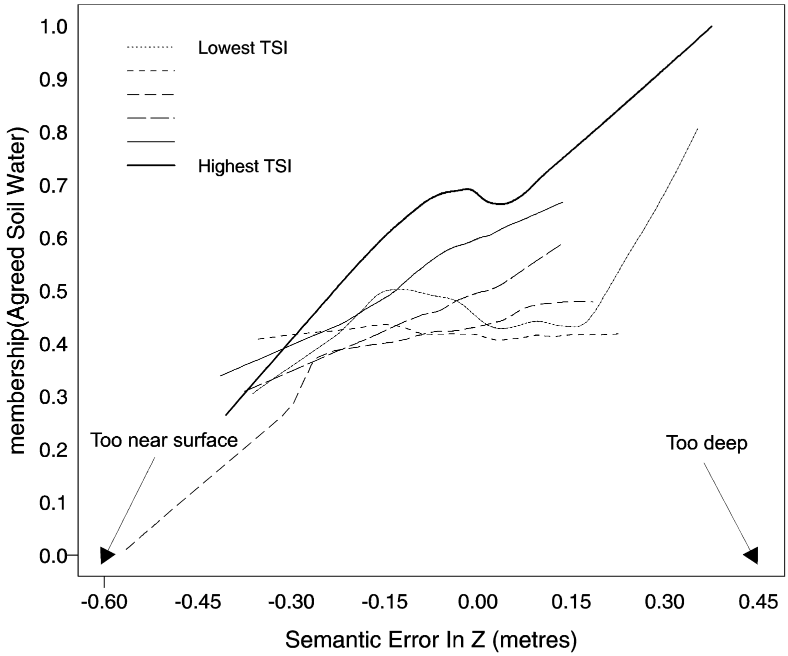

Fig. 9. Overall agreed soil water memberships for the 6 aggregate TSI zones.

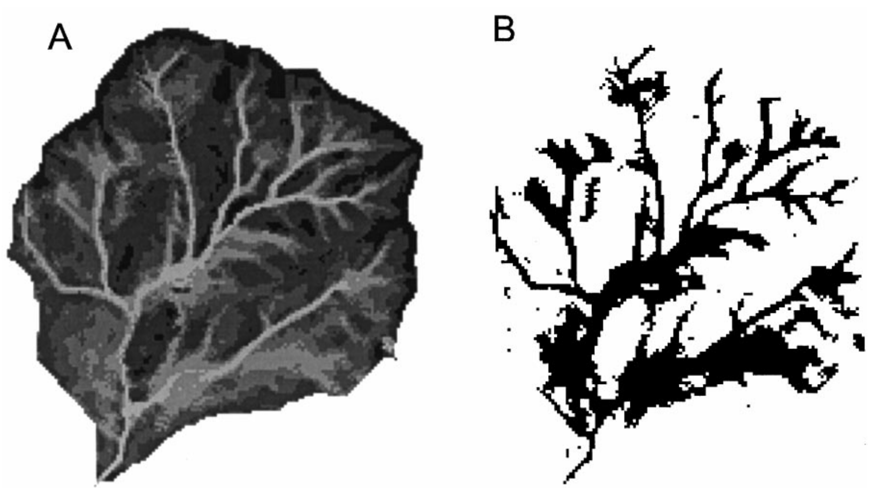

Fig. 10. Final (a) fuzzy and (b) crisp maps showing where the sub-models within RHESSysD agree/disagree on soil moisture at the start of the growing season. In (a) brighter areas have higher fuzzy memberships, while in (b) the black areas represent where the models agree.

spatial set derived from the fuzzy spatial set by mapping only the areas that have fuzzy memberships of at least 0.5. The areas mapped in black correspond to areas of greatest likelihood of soil saturation and hence runoff generation. The results are consistent with the need for a resetting event, such as a soil surface saturation, to establish a reasonable set of initial conditions for simulation [8,25]. For the purposes of decision making the areas mapped in black are where the integrated model successfully answers the query.

These are areas where the integrated model gives soil moisture estimates that are not adversely affected by model structure.

5. Discussion

Simulation models typically generate large quantities of output, which usually have to be analyzed by the model end-user in order to assess model quality. For instance, simulations of the watershed used here have 1450 TSI patches, which are simulated for each of 365 days. This represents over a half-million soil moisture output values, which may or may not be trustworthy given the structure of the integrated model and the characteristics of the input data. Model self-evaluation provides an initial basis for evaluating the model by measuring model semantic error in computing the position of the water table. How goal variables respond to this error provides deeper insight on model agreement. The significance of transient disagreements between sub-models is that for certain queries to the integrated model (e.g., soil moisture), sub-model incompatibility may or may not preclude the use of these models. The semantic analysis presented here allows for the identification of areas within the spatial and temporal domains where sub-models provide agreed upon answers to specific queries. This approach differs from more commonly used model testing tools such as sensitivity analysis (e.g., Monte Carlo analysis in which many realizations of equal likelihood are run) for a model treated as a black box [1,17,13]. Sensitivity analysis cannot distinguish between uncertainty due to data quality and uncertainty due to model structure [17], whereas the approach presented here focuses on model structure problems rather than comparison of model output to data.

Here fuzzy sets are used as a method of combining several criteria for evaluating model performance. This differs from other applications of fuzzy sets in the GIS community. For instance, systems such as SRAS [37] and SOLIM [2] use fuzzy set theory to linguistically express spatial or co-spatial relations using experience of expert subjects to produce membership functions. Here, membership functions are generated by a simulation model using a fixed set of criteria for evaluation, for the purpose of controlling the selection of sub-models to be combined. Once a decision is made on model agreement for a given query, a specific set of model alternates can be selected. For instance, a model that represents the water table in a mathematically more explicit way, but is computationally intractable for large spatial areas [31], may be preferred for certain types of questions over small spatial areas. However, for many queries aggregating to a level appropriate for capturing the dominant spatial variability of processes avoids the need for extremely accurate, fine resolution data sets [11]. Since most data cannot always be obtained to support detailed physically-based models, more conceptually-based models, which are based on simplified relationships that use more readily available data, will remain in widespread use [6,16]. Thus, model self-evaluation is an important step towards providing general tools for determining the range of acceptable queries for integrated models constructed from conceptual models.

6. Conclusions

The most significant result of this work is the introduction of self-evaluation into spatially distributed process models, and its application as a decision-making tool for selecting or rejecting a combination of sub-models for a given application. The model, RHESSysD, incorporates this self-evaluation in the sense that it identifies and keeps track of semantic error in the water table position for every modeled patch. In essence, two parallel simulations, one using Eq. (9) (uncorrected) and the other using Eq. (13) (corrected), are maintained by the model. All other aspects of the model remained unchanged between the parallel versions, making the comparative analysis between the two simulated results feasible. The results of these simulations are then passed to the decision support system for comparative analysis using the three criteria. The ability to manage the interactions between sub-models during simulation allows for the identification of model semantic inconsistencies that could be used to trigger the replacement of particular sub-models as the simulation model executes. The use of fuzzy set theory allows the decision to be based on a combination of criteria allowing for different indicators of uncertainty in the model response variables to be considered. Fuzzy sets theory thus facilitates the analysis of a large quantity of simulation output and its synthesis more decision making. The current software system is only a prototype and is geared towards analysis of output from RHESSysD. Future work should consider how alternative models can be selected and used based on the decision processes described in this paper. In addition, future work with the method must consider how queries involving multiple response variables can be addressed, and an integration of memberships made across a suite of variables.

Acknowledgements

This work was partially supported by a start-up grant from the Graduate School of the University of Wisconsin-Madison to the first author. Data for Onion Creek were provided by the U.S. Forest Service and by NTSG, University of Montana. The authors gratefully acknowledge Pauline van Gaans and an anonymous reviewer for their careful and thoughtful review of this manuscript.

References

[1] A.M. Binley, K.J. Beven, A. Calver, L.G. Watts, Changing responses in hydrology: assessing the uncertainty in physically based model predictions, Water Resources Res. 27 (6) (1991) 1253–1261.

[2] A.X. Zhu, L.E. Band, B. Dutton, T. Nimlos, Automated soil inference under fuzzy logic, Ecol. Modeling 90 (1996) 123–145.

[3] B.P. Zeigler, Object-Oriented Simulation with Hierarchical, Modular Models: Intelligent Agents and Endomorphic Systems, Academic Press, San Diego, CA, 1990.

[4] C. Loehle, Applying artificial intelligence techniques to Ecol. Modelling 38 (1987) 191–212.

[5] C.G. Wesseling, D. Karssenberg, P.A. Burrough, W.P.A. van Deursen, Integrating dynamic environmental models in GIS: the development of a dynamic modelling language, Trans. GIS 1 (1996) 40–48.

[6] D.A. Bennett, M.P. Armstrong, F. Weirich, An object-oriented modelbase management system for environmental simulation, in: M.F. Goodchild, B.O. Parks, L.T. Steyaert (Eds.), GIS and Environmental Modelling: Progress and Research Issues, GIS World Books, Fort Collins, CO, 1996, pp. 439–444.

[7] D.H. Jamieson, K. Fedra, The ‘WaterWare’ decision-support system for river-basin planning. 1. conceptual design, J. Hydrol. 177 (1996) 163–175.

[8] D.S. Mackay, L.E. Band, Forest ecosystem processes at the watershed scale: dynamic coupling of distributed hydrology and canopy growth, Hydrol. Processes 11 (1997) 1197–1218.

[9] D.S. Mackay, Integrating self-evaluating hydrological and ecological models of different spatial scales, in: Proc. GIS/LIS’97, American Society of Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Falls Church, VA, 1997, pp. 486–498.

[10] D.S. Mackay, V.B. Robinson, L.E. Band, A knowledge-based approach to the management of geographic information systems for simulation of forested ecosystems, in: W.K. Michener, J.W. Brunt, S.G. Stafford (Eds.), Environmental Information Management and Analysis: Ecosystems to Global Scales, Taylor & Francis, London, 1984, pp. 515–538.

[11] E.F. Wood, M. Sivapalan, K. Beven, L. Band, Effects of spatial variability and scale with implications to hydrologic modeling, J. Hydrol. 102 (1988) 29–47.

[12] F. Hayes-Roth, N. Jacobstein, The state of knowledge-based systems, Comm. ACM 37 (3) (1994) 27–39.

[13] G.B.M. Heuvelink, Error propagation in quantitative spatial modelling: applications in geographical information systems, Thesis, University of Utrecht, 1993.

[14] H. Abelson, M. Eisenberg, M. Halfant, J. Katzenelson, E. Sacks, G.J. Sussman, J. Wisdom, K. Yip, Intelligence in scientific computing, Comm. ACM 32 (5) (1989) 546–562.

[15] H. Caswell, The validation problem, in: B.C. Patten (Ed.), Systems Analysis and Simulation in Ecology, Academic Press, New York, 1976, pp. 313–325.

[16] I.D. Moore, T.W. Norton, J.E. Williams, Modelling environmental heterogeneity in forested landscapes, J. Hydrol. 150 (1993) 717–747.

[17] J. Freer, K. Beven, B. Ambroise, Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach, Water Resources Res. 32 (7) (1996) 2161–2173.

[18] J.D. Ullman, Principles of Database and Knowledge-Base Systems, Computer Science Press, Rockville, MD, 1988.

[19] J.W. Rozenblit, P.L. Janowski, An integrated framework for knowledge-based modeling and simulation of natural systems, Simulation 57 (1991) 152–165.

[20] K. Beven, Changing ideas in hydrology – the case of physically-based models, J. Hydrol. 105 (1989) 157–172.

[21] K. Fedra, Distributed models and embedded GIS: strategies and case studies of integration, in: M.F. Goodchild, B.O. Parks, L.T. Steyaert (Eds.), GIS and Environmental Modeling: Progress and Research Issues, GIS World Books, Fort Collins, CO, 1996, pp. 413–418.

[22] K.J. Beven, M.J. Kirkby, A physically based, variable contributing area model of basin hydrology, Hydrol. Sci. Bull. 24 (1) (1979) 43–69.

[23] L. Zadeh, Fuzzy sets, Inform. and Control 8 (1965) 339–353.

[24] L.E. Band, A terrain-based watershed information system, Hydrol. Processes 3 (1989) 151–162.

[25] L.E. Band, P. Patterson, R. Nemani, S.W. Running, Forest ecosystem processes at the watershed scale: incorporating hillslope hydrology, Agri. Forest Meteorol. 63 (1993) 93–126.

[26] M. Sivapalan, K. Beven, E.F. Wood, On hydrologic similarity 2. A scaled model of storm runoff production, Water Resources Res. 30 (1987) 1665–1679.

[27] M.F. Worboys, S.M. Deen, Semantic heterogeneity in distributed geographic databases, ACM SIGMOD Record 20 (4) (1991) 30–34.

[28] N. Oreskes, K. Shrader-Frechette, K. Belitz, Verification, validation, and confirmation of numerical models in the earth sciences, Science 263 (1994) 641–646.

[29] P. Quinn, K. Beven, P. Chevalier, O. Planchon, The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models, Hydrol. Processes 5 (1991) 59–79.

[30] R. Ford, S. Running, R. Nemani, A modular system for scalable ecological modeling, Comput. Sci. Eng. (Fall) (1994) 32–44.

[31] R.B. Grayson, I.D. Moore, T.A. McMahon, Physically based hydrologic modeling 2. Is the concept realistic?, Water Resources Res. 28 (10) (1992) 2659–2666.

[32] R.D. Moore, J.C. Thompson, Are water table variations in a shallow forest soil consistent with the TOPMODEL concept?, Water Resources Res. 32 (3) (1996) 663–669.

[33] R.M. Keller, M. Rimin, A. Das, A knowledge-based prototyping environment for construction of scientific modeling software, Autom. Software Eng. 1 (1994) 79–128.

[34] S.M. Dunn, R. Mackay, R. Adams, D.R. Oglethorpe, The hydrological component of the NELUP decision-support system: an appraisal, J. Hydrol. 177 (1996) 213–235.

[35] V.B. Robinson, A.U. Frank, About different kinds of uncertainty in collections of spatial data, Proc. Auto-Carto 7, American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, Falls Church, VA, 1985, pp. 440–449.

[36] V.B. Robinson, D.S. Mackay, Semantic modeling for the integration of geographic information and regional hydroecological simulation management, Comput. Environ. and Urban Systems 19 (1996) 321–339.

[37] V.B. Robinson, Interactive machine acquisition of a fuzzy spatial relation, Comput. Geosci. 16 (1990) 857–872.

[38] W.P.A. van Deursen, Geographical information systems and dynamic models, Thesis, University of Ultrecht, 1995.

[39] W.S. Cleveland, Robust locally weighted regression and smoothing scatterplots, J. Amer. Statist. Soc. Assoc. 74 (1979) 829–836.