OPTIMIZATION AS A TOOL FOR CONSISTENCY MAINTENANCE IN MULTI-RESOLUTION SIMULATION

Darren T. Drewry
Paul F. Reynolds, Jr.

Department of Computer Science
University of Virginia
Charlottesville, Virginia 22904-4740, U.S.A.

William R. Emanuel

Department of Environmental Science
University of Virginia
Charlottesville, Virginia 22904-4123, U.S.A.

ABSTRACT

The need for new approaches to the consistent simulation of related phenomena at multiple levels of resolution is great. While many fields of application would benefit from a complete and approachable solution to this problem, such solutions have proven extremely difficult. We present a multi-resolution simulation methodology that uses numerical optimization as a tool for maintaining external consistency between models of the same phenomena operating at different levels of temporal and/or spatial resolution. Our approach follows from previous work in the disparate fields of inverse modeling and spacetime constraint-based animation. As a case study, our methodology is applied to two environmental models of forest canopy processes that make overlapping predictions under unique sets of operating assumptions, and which execute at different temporal resolutions. Experimental results are presented and future directions are addressed.

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE **2006**	2. REPORT TYPE	3. DATES COVERED **00-00-2006 to 00-00-2006**
4. TITLE AND SUBTITLE **Optimization as a Tool for Consistency Maintenance in Multi-Resolution Simulation**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Virginia,Department of Computer Science,151 Engineer's Way,Cahrlottesville,VA,22094-4740**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**		
13. SUPPLEMENTARY NOTES **The original document contains color images.**		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES **22**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

1 INTRODUCTION

Multi-resolution modeling (MRM) addresses the inconsistencies that arise when multiple simulations of similar phenomena, operating at different levels of temporal and/or spatial resolution, are coupled for a common purpose. These inconsistencies are a product of multiple complexities occurring in the coupled multi-resolution system. In the case of differences in spatial resolution, inconsistencies can occur in the aggregation and disaggregation between the multiple spatial levels. Differences in temporal resolution often lead to high and low frequency representations of the same phenomena, in some cases producing results that do not agree.

Prior research has focused on maintaining internal consistency between model constructs occurring at multiple resolutions. These approaches have met with some success, but are often difficult to implement and in some cases do not solve particular problems related to movement between the various resolution levels. We focus our efforts on the maintenance of external consistency, or external model agreement. Through the proper identification of internal model controls, which may be quite different for the various models comprising a multi-resolution system, we propose methods to guide the models along consistent paths.

Our goal is to identify a valid semi-automated process for effecting multi-resolution modeling. Because there is a strong demand for any capability, we pursue a semi-automated approach in place of any, more elusive, fully automated approach. We explore the application of technology similar to that used in the graphics animation, motion retargeting community, to a pair of environmental models. Using optimization over an objective function and constraints manually chosen by subject matter experts, we have experienced significant success.

This paper is ordered in the following manner: The following section introduces background information on two environmental models which were used as a case study to test our ideas. Also presented in Section 2 is a description of previous work in fields of research that utilize optimization in a manner similar to ours, namely inverse modeling and the animation paradigm of spacetime constraints. This background material leads into a description of our technique which is presented in Section 3. Results of the application of our methodology to two models of environmental processes are described in Section 4. And, finally, we end with a discussion of the results and future directions in Section 5.

2 BACKGROUND

2.1 Related Work

The methodology we present shares similar features with methods developed in two very different fields, both of which utilize simulation and modeling: inverse modeling and spacetime constraint-based animation. We will begin by discussing the field of multi-resolution simulation in general, and then move on to descriptions of the relevant characteristics we have found in the other two fields.

2.1.1 Multi-Resolution Simulation

Multi-resolution consistency maintenance in coupled simulations operating at different levels of temporal and/or spatial resolution has long been recognized as a critical problem in the field of Simulation and Modeling. Inconsistent coupled simulations manifest themselves in the multitude of research areas that utilize simulation as a tool for scientific exploration and understanding. As studies in the natural, social and engineering sciences grow increasingly complex, and seek to incorporate a variety of diverse information, the need to couple models of phenomena at various resolution levels becomes paramount. The complexity of the variety of existing models in any one field, and the range of scales at which relevant phenomena must be represented, present serious multi-resolution simulation challenges.

Several methods have been developed in attempts to solve the problem of multi-resolution simulation inconsistency. We discuss two promising proposals to solve the MRM problem.

Davis and Hillestad (1993) described "variable resolution modeling", in which resolution hierarchies are constructed so as to facilitate the seamless movement between levels of different resolution. This approach relies on the aggregation of entity attributes when changing from high to low resolution, and disaggregation when moving in the opposite direction. Aggregation and disaggregation have been shown to suffer from several problems (Reynolds et al. 1997), including chain disaggregation, transitional latency, and difficulties in mapping between different levels of resolution.

Reynolds et al. (1997) proposed, as an alternative, the concept of multi-resolution entities (MRE). MREs provide a data structure oriented approach to solving the MRM problem. MREs act to maintain information at multiple levels of resolution, as well as map information between the different levels, for all interactions. Internal consistency is maintained through the specification of a core set of attributes which are updated in the MRE for each interaction. While this approach holds promise, to date it has yet to be widely implemented. We believe this is primarily due to the complexities that would be involved in the development of the necessary data structures and the mapping functions between attributes at different resolution levels.

2.1.2 Inverse Modeling

Simulation models typically contain internal parameters whose values affect model predictions. When the phenomena the model is to predict are accurately observed, adjustments to model parameters can be made to ensure correct predictions under similar conditions, assuming model validity. The process of identifying the proper parameter set, and determining the proper values for those parameters, has come to be known as "inverse modeling". This technique has been widely applied in the environmental sciences.

The ability to accurately describe the surface properties of large areas of land is of great importance in certain disciplines of the environmental sciences. Remote sensing data collected for the entire globe by satellite observing platforms have revolutionized the ability of scientists to characterize particular properties of the Earth's surface. Specifically, the amount and type of incoming solar radiation reflected back from the surface is directly related to the properties of the vegetation and soil composing the

region. Intensive investigations of the physics controlling the reflection of radiation incident on plant canopies has led to a set of environmental variables recognized as significant to controlling this process. Models incorporating the physics dependent on these controlling variables have been developed as a means of predicting canopy reflectance for specific vegetation covers in specific locations. The class of models providing this functionality is generally known as canopy reflectance models (Goel and Strebel 1983).

Canopy reflectance models can also be used to identify values of the surface features described by the set of controlling variables discussed above. Through what has been termed "model inversion", an optimization scheme is applied to the canopy reflectance model (Braswell et al. 1996). The optimization process seeks to discover valid values of the controlling variables which produce near identical results of canopy reflectance as those measured by satellite imagery. The optimization scheme consists of an objective function describing the sum of the squared differences between calculated canopy reflectance and those measured. The constraint set over which the optimization is performed consists of a subset of the controlling variables determined to affect the calculation of canopy reflectance in a significant manner. The optimization varies the values of these parameters until a minimum difference between modeled and measured canopy reflectance is achieved.

2.1.3 Spacetime Constraint-Based Animation

The spacetime constraints technique was introduced by Witkin and Kass (1988) as an approach to computer character animation. In the spacetime approach a set of physical constraints for a particular character are developed and utilized as a way of synthesizing physically correct motions which meet user specified goals. The problem is formulated as a nonlinear constrained optimization in which a specified objective function is minimized so as to meet a set of constraints that have been placed on the character and the corresponding motion. The objective function that is minimized typically describes the energy expended by the character in performing the motion. Studies of human biomechanics have shown that nature has crafted the human body to move in an energy efficient manner, lending validity to the minimization of a function describing the energy consumed in completing a particular motion. The constraint set applied to the minimization is developed from the character description, the environment, physical laws that must be obeyed throughout the motion, and the specific goals of the motion. Examples of typical constraints are: limitations on the range of joint angles, constant limb lengths, gravity must be obeyed throughout the simulation, a limb cannot pass through the torso, and a hand must be placed at point (x, y) at time t. The constraint set and the objective function are formulated mathematically and incorporated into a numerical optimization routine which, upon successful convergence, describes the "best" set of character parameters (ie. muscle strengths, joint angles, limb positions, etc.) necessary to perform the desired motion. The optimization routine solves the problem for the entire motion simultaneously, rather than focusing on individual frames. This provides a valid parameter set which meets the specified goals over the entire course of the motion.

As motion capture and other technologies for creating motions developed, spacetime constraints was applied to the problem of adapting preexisting motions to characters other than those that were used to create it. This application has been referred to as "motion

retargeting" (Gleicher 1998) and "motion transformation" (Popovic and Witkin 1999). We will use the term "motion retargeting" in this discussion. The motion retargeting problem differs from the original spacetime constraints approach in a few subtle ways. The first difference is the definition of the objective function used in the optimization procedure. It is not necessary to minimize the energy consumed throughout the course of the motion, as the motion already exists and has been chosen for use because of its inherent qualities, efficient or not. Instead, the objective function in motion retargeting describes a measure of the difference between the original motion and the version adapted to the new character. In this way the optimization procedure can find the motion which satisfies the constraint set and yet deviates from the original motion by a minimal amount. Gleicher (1998) applied spacetime constraints to the retargeting of motions from one character to another with identical structure, but of different size. Popovic and Witkin (1999) solve the problem of mapping captured motion to animated characters, as well as retargeting motion between characters with drastically different kinematic descriptions.

The utility of viewing character motion as one or more discrete time series of joint angles and joint positions has been presented in the literature (Gleicher and Litwinowicz 1998). This view naturally suggests the application of signal processing techniques to modify motion signals so as to preserve the desired qualities of the original while modifying it to meet the goals of the new motion. This viewpoint will provide an analogy to the multi-resolution simulation of canopy carbon assimilation.

2.2 Two Models of Canopy Processes

Our study focuses on two environmental models of vegetative canopy processes, CANOAK and DOLY. While designed for different purposes, these models share a certain degree of overlap in the environmental quantities they forecast. Brief descriptions of the two models are presented below, focusing on key similarities and differences between them, as well as features used as points of focus for this study.

2.2.1 CANOAK

CANOAK (Baldocchi and Harley 1995; Harley and Baldocchi 1995) was designed as a detailed biophysical model of an oak canopy. The goal of this model is to simulate the high temporal resolution water, energy and carbon dioxide (CO_2) fluxes between an oak canopy and the atmosphere. These scalar fluxes are measured using eddy covariance techniques at the growing network of flux tower sites throughout the United States and around the world. CANOAK has been validated at a site in Oak Ridge, Tennessee, and has been shown to have good performance in predicting water, energy, and CO_2 exchange over a spectrum of timescales ranging from diurnal to annual (Baldocchi and Harley 1995; Baldocchi and Wilson 2001).

The model separates a forest canopy into forty horizontal layers, with leaves distributed equally throughout each layer. Both micrometeorological and eco-physiological modules are coupled in an effort to describe the physical and biological aspects of land-atmosphere exchange. The micrometeorological modules compute the detailed biophysical processes of turbulent diffusion, leaf and soil gas and energy exchange, and radiative transfer occurring throughout the detailed microclimate of the multi-layer canopy. Eco-physiological components account for leaf stomatal

conductance, photosynthetic carbon assimilation, respiration and transpiration, all of which are driven by the micrometeorological computations. The photosynthetic assimilation of CO_2 is modeled by the popular formulation presented by Farquhar et al. (1980).

Simulation time proceeds in hourly steps, in which a set of observed meteorological conditions drive the canopy micrometeorological module, producing estimates of solar radiation and sunlit and shaded leaf fractions. In the same time step, the eco-physiological component uses solar radiation information to compute stomatal conductance and fluxes of water, energy and carbon. An iterative algorithm revises these predictions until equilibrium is achieved. Hourly results are output, as are daily averages.

2.2.2 DOLY

DOLY (Woodward et al. 1995) is a model of canopy primary productivity. This model was created to be run at sites across the Earth's surface, simulating global carbon uptake and allocation. It was developed to be incorporated into a coarse-grained dynamic vegetation component coupled to global climate models (GCMs). Focus is given to nutrient and water uptake, and the corresponding impacts of these processes on carbon assimilation rates. Canopy conductance is dependent on temperature and soil water and nutrient content, and acts to control soil water loss and carbon assimilation. Carbon assimilated by photosynthesis is allocated to respiration and annual plant growth. Carbon and moisture budgets constrain leaf area index (LAI), the ratio of leaf area to ground surface area. DOLY divides the vegetation canopy into horizontal layers each corresponding to one unit of LAI. Experiments have favorably demonstrated the ability of the model to simulate global distributions of leaf area index and annual net primary productivity.

DOLY executes at simulated time steps of one month. Each month of simulated time, DOLY takes average meteorological inputs and produces values including canopy conductance, CO_2 assimilation, and water vapor exchange. The calculations are performed by consecutively scaling down from the canopy to leaf to cell, where the biochemical processes actually occur. The calculated values at the cellular level are then scaled up to the canopy as a whole.

Comparing DOLY and CANOAK can be a challenging endeavor given the disparity in design and function of the two models. Several characteristics of these models make them interesting for the purposes of this study. Both models use the Farquhar et al. (1980) biochemical carbon assimilation formulation for the prediction of canopy photosynthetic rates, although differences in the details of the implementations exist. The modeling of canopy processes in general is approached from different philosophical perspectives. CANOAK implements a detailed biophysical approach, while DOLY is designed to be a coarser, more computationally efficient model, approximating detailed structure and other model formulations that cannot be evaluated over large landscapes. Additionally, the wide gap in temporal resolution of the two models makes them a suitable case study for a test of our methodology.

3 THE OPTIMIZATION-BASED APPROACH

We have explored the application of constrained optimization to multi-resolution modeling. An introductory overview of our approach has been presented by Reynolds (2002). Our approach takes advantage of numerical optimization as a method to establish a set of model parameter values which meet specified criteria, or goals. These parameters may be model constants or relationships between model entities and attributes. The correctness of the parameter values generated by the optimization procedure is enhanced by establishing an appropriate objective function and constraint set for the specific problem being addressed. The parameter set produced by the optimization may not be unique in that multiple parameter sets may minimize the objective. The generated parameter set does however represent a solution to the originally inconsistent coupled simulation over the period of simulation time for which the minimization was performed. The appropriate objective function and constraint set for a coupled simulation system is very specific to the subject matter expressed in the models, and therefore depends greatly on the knowledge of a domain expert. We offer general guidelines in developing an objective function and constraint set, present a simple example of the application of the technique, and describe the details of a case study involving two environmental models.

3.1 Model Optimization

Numerical optimization is a tool that has found wide application in many branches of engineering and the sciences. The general optimization problem is formulated with three basic constructs: an objective function, a set of unknown parameters, and a set of constraints. The objective is the function whose value the optimization algorithm seeks to minimize or maximize. The value of the objective function is varied by changing the values of the unknown parameters, each of which has an effect on the value of the objective function. The allowable set of values that each of the parameters may take is defined by the constraint set, in which bounds are placed on the acceptable values of parameters, or relationships between the parameters. We focus on constrained nonlinear minimization using one objective function subject to linear and nonlinear equality and inequality constraints. The minimization problem we seek to solve can be described as:

$$\min OBJ(\mathbf{p}) \text{ subject to } \begin{cases} \mathbf{f}_{eq}(\mathbf{p}) = \mathbf{c}_{eq} \\ \mathbf{f}_{ineq}(\mathbf{p}) \geq \mathbf{c}_{ineq} \\ \mathbf{lb} \leq \mathbf{p} \leq \mathbf{ub} \end{cases} \qquad (1)$$

where $\mathbf{p}$ is the vector containing the parameter set, $OBJ(\mathbf{p})$ is the objective function, $\mathbf{f}(\mathbf{p})$ is the constraint vector, and $\mathbf{lb}$ and $\mathbf{ub}$ refer to the vectors of lower and upper bounds, respectively. The subscripts eq and $ineq$ specify equality and inequality constraints, respectively. The Matlab™ $fmincon$ (The MathWorks, Inc. 2001) routine provides functionality to perform this type of optimization and is used in the experiments described below.

General methods to solve the optimization problem under all circumstances do not exist. Optimization algorithms come in several varieties and the choice of which algorithm to use depends on the requirements of the problem to be solved. Variants can

differ in the number of objective functions, specification of the form of the objective function, existence of constraints, and the specification of the allowable constraints. It should be noted that due to the difficulties associated with optimization, convergence is not guaranteed. There is the possibility of the algorithm becoming trapped in a local minimum, which makes the initial guess of the parameter values crucial to finding the best solution.

3.1.1 The Objective Function

In discrete-event simulation, model predictions can naturally be viewed as discrete time series. Simulations progressing at different temporal resolutions will produce time series of model predictions at different frequencies. A high temporal resolution model of particular phenomena will capture high frequency dynamics that will not be present in the functioning and predictions of a low temporal resolution model of the same phenomena. The disparity in the frequency of model dynamics can quickly lead to inconsistencies in coupled model behavior. The goal of multi-resolution modeling is to minimize this disparity between model behaviors and outputs to within acceptable limits. The objective function formulation must capture this goal. The minimization algorithm can then determine the best set of model parameter values that minimize the difference between model behaviors.

Spacetime motion retargeting utilizes objective functions that represent the difference between the original motion and that of the retargeted motion. The minimization finds a set of parameter values that maintain the characteristics of the original motion to the largest extent possible, while meeting the goals of the new motion which have been captured in the constraint set. Likewise, inverse modeling utilizes objective functions that represent the least squares difference between observations and the values predicted by models. The goal then is to find a set of parameter values, contained in the associated constraint set, which guide the model to produce the observed results. In the multi-resolution simulation approach we advocate a similar formulation of the objective function.

It is necessary to minimize the difference between each of the particular model behaviors and outputs that determine the level of agreement between two models operating at different resolutions. This can be accomplished by incorporating each of these n characteristics into the objective function as a summation of squared differences:

$$OBJ(\mathbf{p}) = \sum_{j=1}^{m} \sum_{i=1}^{n} w_j \left(H_{ij} - L_{ij}(\mathbf{p})\right)^2 \tag{2}$$

where $\mathbf{w}$ is the optional vector of weights between 0 and 1 used to signify the importance of each of the values represented in the objective function (Goel and Strebel 1983; Privette et al. 1996). H and L represent the actual model values that are differenced, with H being the value produced by the high resolution model, and L representing that produced by the low resolution model. The set of Hs and Ls incorporated into the objective function defines the overlapping calculations made by the multi-resolution models that are critical to maintaining consistent behavior between them.

3.1.2 The Constraint Set

The constraint set must be formulated to provide a set of parameters whose final values, once the optimization procedure has completed execution, can be used to execute the coupled multi-resolution system in an externally consistent manner. This requires the choice of constraint parameters that impact the system enough to bring the coupled models into agreement. The constraint set can be viewed as a set of "control knobs" which, when turned the proper amount in combination, guide the coupled simulations into agreement with each other. In order to successfully guide the simulations into agreement, the model(s) must be sufficiently sensitive to variations in the constraint parameters so that changes made to these values have the possibility of moving model behaviors closer together. Absolute model constants, constant values that are seen as unchangeable, are obviously a bad choice, even if the model shows sensitivity to changes in them, because model validity will suffer by adjusting them. This implies that not only must the model display sensitivity to the constraint parameters, those parameters must have an acceptable range of possible values over which model validity is not sacrificed.

Sensitivity analysis provides one useful tool for understanding the degree to which a simulation is affected by specific parameters. Repeated replication of an isolated simulation in an identical manner except for changes to a single model parameter will allow for determination of the magnitude of the effect of that parameter on the simulation. This intensive approach may not be necessary if sensitivity to particular model parameters is already understood.

3.2 A Simple Example

As an initial example we present two "models" that are simple linear combinations of harmonic functions. The optimization procedure we present views the models as black-boxes, regardless of the level of complexity contained within. We define the following two functions:

$$y(t) = A \cdot \sin(m \cdot t) + B \cdot \sin(n \cdot t) - C \cdot \cos(p \cdot t) \tag{3}$$

$$z(t) = G \cdot \cos(q \cdot t) + H \cdot \cos(r \cdot t) \tag{4}$$

The parameter values, chosen at random, are: A=10, B=2, C=5, G=3, H=15, m=0.5, n=2, p=0.1, q=0.3, and r=0.8. Figure 1 displays line plots of the two functions generated using integer values of t from t = 1...25. The correlation coefficient between y(t) and z(t) was calculated to be 0.0724.

A minimization problem was constructed for the function y(t). The objective function to be minimized was the sum of the squares of the difference between the corresponding values of y(t) and z(t) for each value of t:

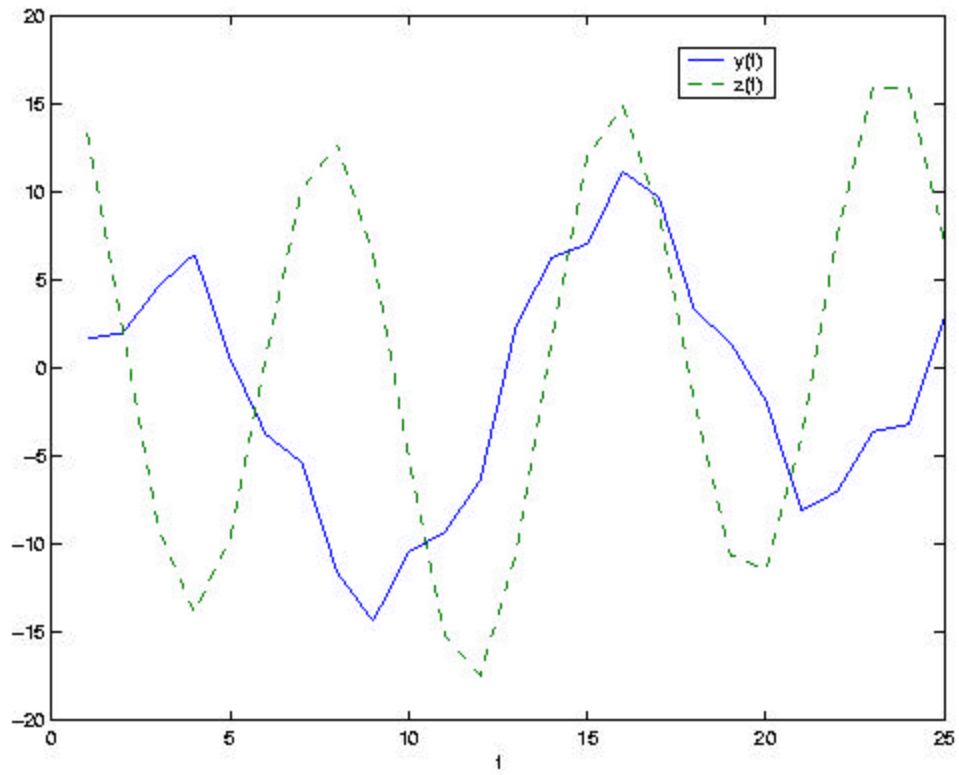$$OBJ(\mathbf{p}) = \sum_{i=1}^{25} \left(z(t_i) - y(\mathbf{p}, t_i) \right)^2 \tag{5}$$

9

Figure 1: Initial plots of y(t) (solid/blue) and z(t) (dashed/green).

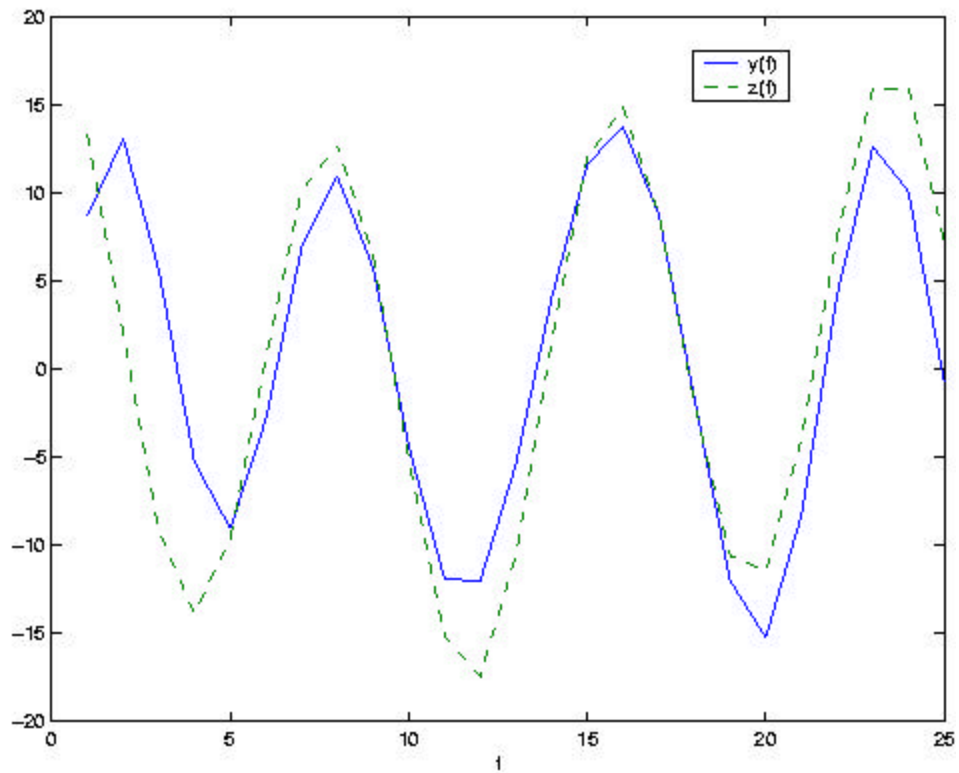Figure 2: Plots of y(t) (solid/blue) and z(t) (dashed/green) after minimization.

All of the parameters in y(t), symbolized by the vector **p**, were incorporated into the constraint set, and allowed to vary between a lower bound of 0 for all parameters, and an upper bound of 20 for A, B and C, and an upper bound of 3 for m, n, and p. The initial guesses for the optimal parameter values were those listed above that were used to generate Figure 1. The minimization produced the following final parameter values: A=12.2, B=4, C=5.46, m=0.9, n=1.1, and p = 0.6. Figure 2 is a plot of y(t) and z(t) with the final parameter values produced by the minimization. The correlation coefficient between y(t) and z(t) using the new parameter values was calculated to be 0.8701.

The optimization procedure was able to "pull" the first "model", y(t), into much closer agreement with z(t), as demonstrated by Figures 1 and 2, as well as the much greater linear correlation between the values of y(t) and z(t). In essence, the two simple models used in this demonstration were brought into much closer agreement through the determination of optimal values of parameters of one of the models.

3.3 Canopy Process Model Case Study

Two models of forest canopy processes, DOLY and CANOAK, are a suitable pair of models for an initial test of our multi-resolution methodology. As discussed above, these two models make similar predictions of environmental phenomena with significant differences in model detail and temporal resolution. In addition, the same underlying biochemical carbon assimilation formulation is used in both models, with differences in representation and implementation.

Meteorological data sets capable of providing the necessary input data for both models were obtained from sites in Oak Ridge, Tennessee and Petersham, Massachusetts. These data sets are approximately one year in length. The data was compiled separately for each of the two models, as input data for each of them is required in different units and at different simulation time intervals. To distinguish between these two sites we will refer to them as "Site 1" (Tennessee) and "Site 2" (Massachusetts).

As the high frequency temporal resolution model, CANOAK is capable of simulating the diurnal variations in water, energy and carbon fluxes of a forest canopy. Detailed biophysical and canopy representations enhance the ability of CANOAK to partition incoming solar radiation and model the magnitudes of radiation incident on leaf surfaces, as well as scalar dispersion within the canopy. DOLY, operating at monthly time steps, accepts monthly averaged values and produces monthly estimates. Coarse representations of radiation transfer, precipitation, and various other canopy measures enhance the computational efficiency of DOLY, while sacrificing resolving power and possibly accuracy. These factors led us to the assumption that under most circumstances, CANOAK will provide the more accurate predictions. For this reason, we perform our case study using CANOAK's predictions as the goal for coupled model agreement. The optimization is applied to DOLY, in an effort to find an acceptable parameter set capable of bringing it into agreement with CANOAK.

In this study we focus only on the carbon assimilation predictions of the two models. The magnitude of carbon assimilation depends on many factors, including incident radiation on the leaf surface, water and nutrient availability, and atmospheric conditions such as relative humidity. This variety of factors controlling carbon assimilation makes it a good gauge of the accuracy of canopy process model predictions. The assessment of predictive accuracy is not the goal of this study, but rather the ability to bring two models of potentially different predictive accuracies into agreement. We identified a challenge

for the test of our methodology by focusing on the point in these two models where predictive accuracy is most strained and corresponding predictions significantly diverge.

The objective function used in this study consisted of the sum of the squared differences of carbon assimilation predictions at each point of comparison. No multiplicative weights were used as only carbon assimilation was used as the determinant of model agreement. The points of comparison were limited to monthly outputs due to the time step of DOLY. CANOAK output was monthly averaged to produce a value that could be directly compared to those produced by DOLY. During winter months CANOAK simulates leaf drop, and carbon assimilation goes to zero during this period. We ignore these periods and instead focus on the points at which both models are actively calculating carbon assimilation.

The formation of the constraint set was guided by a sensitivity analysis performed on DOLY. Several model constants and internal relationships were identified as possibly having a significant impact on carbon assimilation predictions. Sensitivity analysis was performed for the following model constructs: the Beer's Law constant, the cloud attenuation factor, two stomatal conductance parameters, the stomatal conductance response to changes in soil water content, initial soil water content, and the cellular carboxylation rate. The results of two of the sensitivity analysis, those for the Beer's Law constant and the stomatal conductance parameter 'c1', are displayed in Figures 3 and 4, respectively. The large impact of the value of the Beer's Law constant on DOLY carbon assimilation calculations is apparent in Figure 3. In comparison, the 'c1' sensitivity results displayed in Figure 4 show a lack of sensitivity to significant adjustments to the value of this model parameter. These results point to the possible inclusion of the Beer's Law constant, and exclusion of the 'c1' stomatal conductance parameter, in the final constraint set. Several of the possible constraints listed above were individually tested for sensitivity with respect to carbon assimilation calculations. Several of these quantities had a reasonably large impact on carbon assimilation calculations when varied over a large enough range. We conducted several experiments, using subsets of these constructs as constraint sets in our optimization methodology, and present here the results of one such experiment.

The set of quantities added to the constraint set for this experiment was: the effect of soil water content on stomatal conductance, the Beer's Law constant, and the cellular carboxylation rate. It should be noted that the specific ranges over which to vary these quantities, or even whether they should be varied at all, is a complicated issue depending on the particular model in question and the specific assumptions that have gone into the construction of that model. Modification of parameters is a delicate process. Determination of which parameters to vary and the ranges over which they can be varied must involve the opinions of subject matter experts. The choice of parameters manipulated in our experiments was guided by collaboration with several subject matter experts (who at times did not agree). A side benefit of the process we employed is intense consideration of model behavior that we believe can only improve model validity.

The Beer's Law constant (BL) is used to measure radiation attenuation within a forest canopy. Radiation is a necessary component of photosynthesis, and so the value of the Beer's Law constant acts as a direct control on carbon assimilation. Likewise, cellular carboxylation (CX) has a direct impact on carbon uptake. Soil water content (SW) acts to
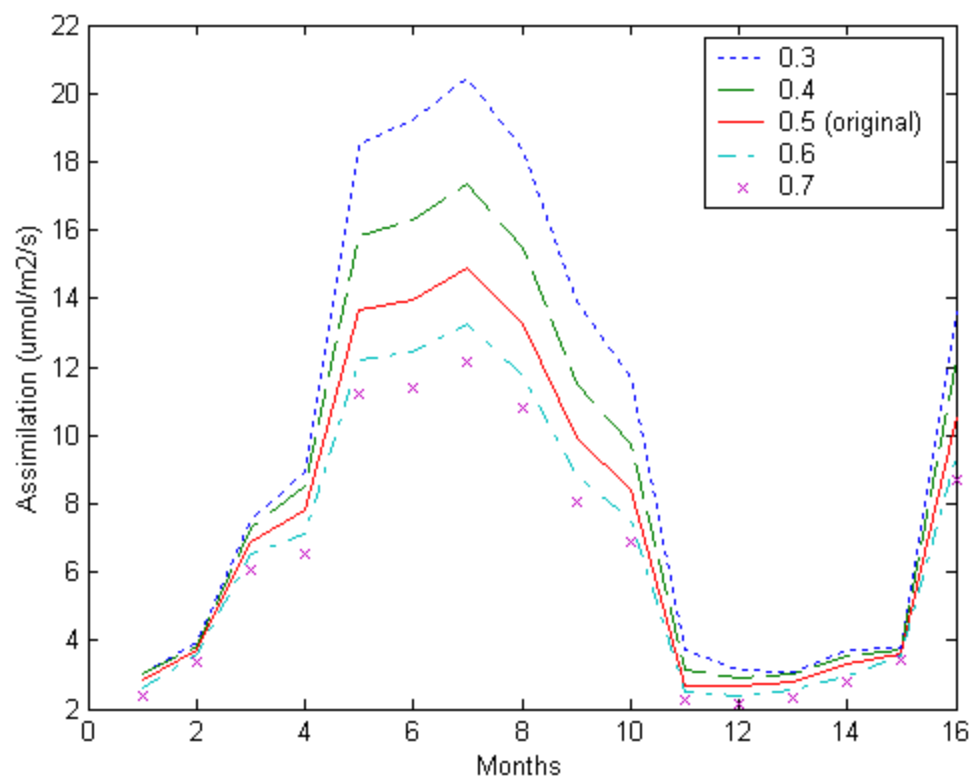
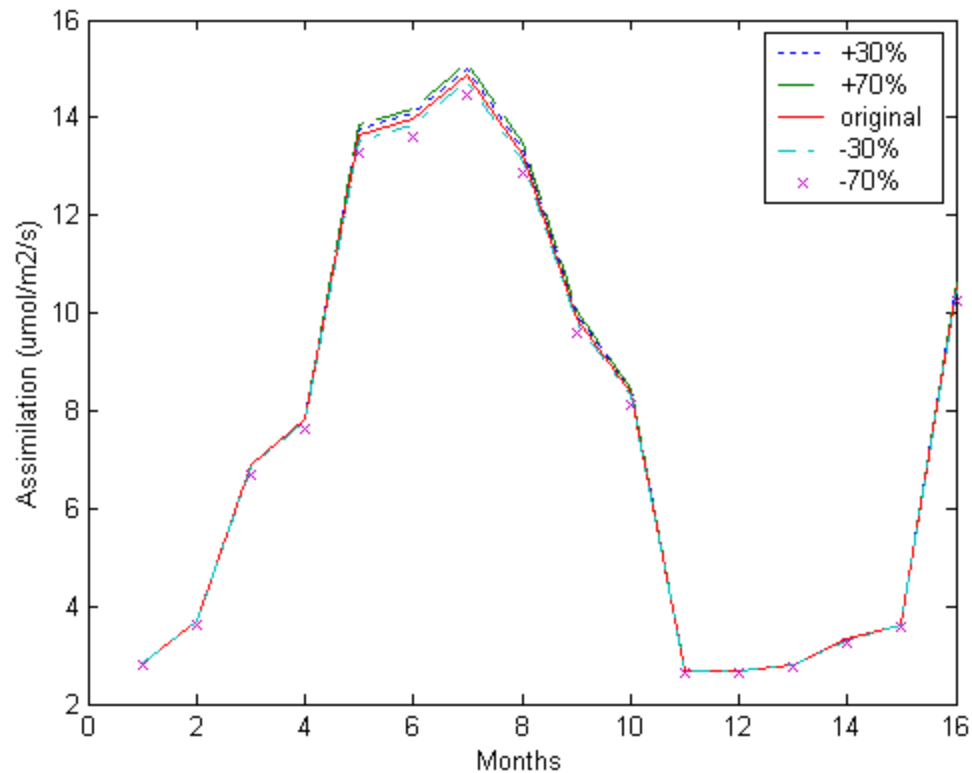Figure 3: Carbon assimilation sensitivity analysis results for the Beer's Law constant.

Figure 4: Carbon assimilation sensitivity analysis results for the stomatal conductance parameter 'c1'.

constrain stomatal conductance, which in turn limits the amount of available carbon at the sites of photosynthesis, providing an indirect control on carbon assimilation.

A flow chart for the multi-resolution system constructed for this case study is presented in Figure 5. The set of initial parameter values is input to DOLY and used in the initial model run. The sum of the squared differences between DOLY and CANOAK carbon assimilation calculations is calculated by the objective function, and a determination is made as to whether the current parameter set satisfies the goal of model agreement. If the level of agreement between the predictions of the two models is satisfactory, then the parameter set values are recorded and execution completes. Otherwise, the minimization routine varies the parameter values in an effort to further minimize the difference between the predictions of the two models. DOLY is iterated over until a satisfactory parameter set is found, or until the user defined number of iterations has been exceeded.
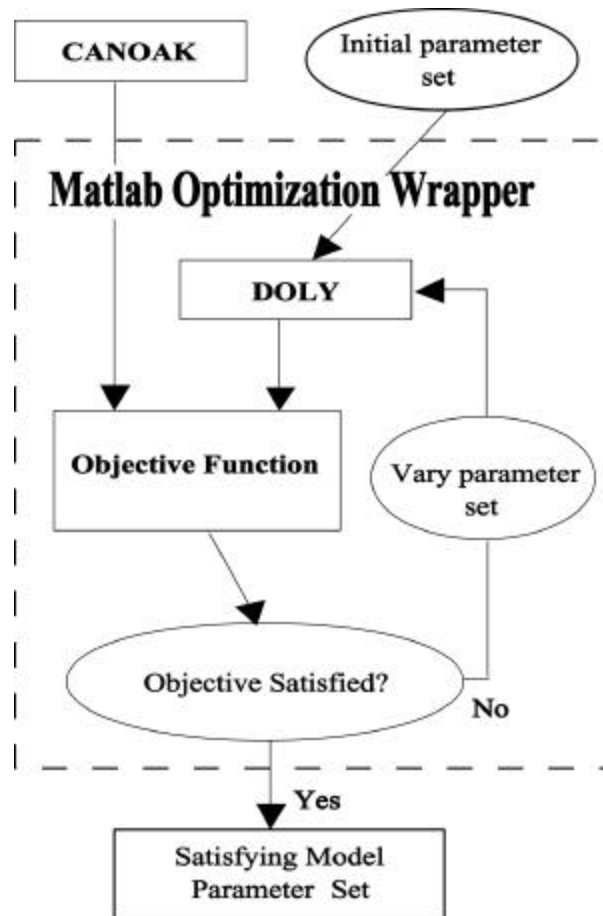

Figure 5: The multi-resolution modeling system constructed for the canopy process model case study.

4 EXPERIMENTAL RESULTS

Comparison of the carbon assimilation predictions of CANOAK and DOLY shows an as yet unexplained significant difference between the two. This can be seen in Figure 6, which displays the results of the initial DOLY and CANOAK runs at Site 1 (Tennessee), labeled "DOLY original run" and "CANOAK monthly avg", respectively. Closing this gap is the goal of applying our multi-resolution methodology to these two canopy process models. Also included in this figure are the results of running DOLY using the optimal parameter values produced by running the optimization procedure at this site. This is labeled "DOLY with Set 1 vals" in Figure 6. It can be seen from the plots that the optimal parameter values for Site 1 produce a very nice fit between DOLY and CANOAK. Figures 7 and 8 display linear correlation plots of the carbon assimilation calculations of DOLY and CANOAK initially, and after the optimization approach was applied, respectively. The straight lines in each plot represent an exact linear fit. A nearly one to one correspondence was created between the predictions of the two models through the application of our methodology.
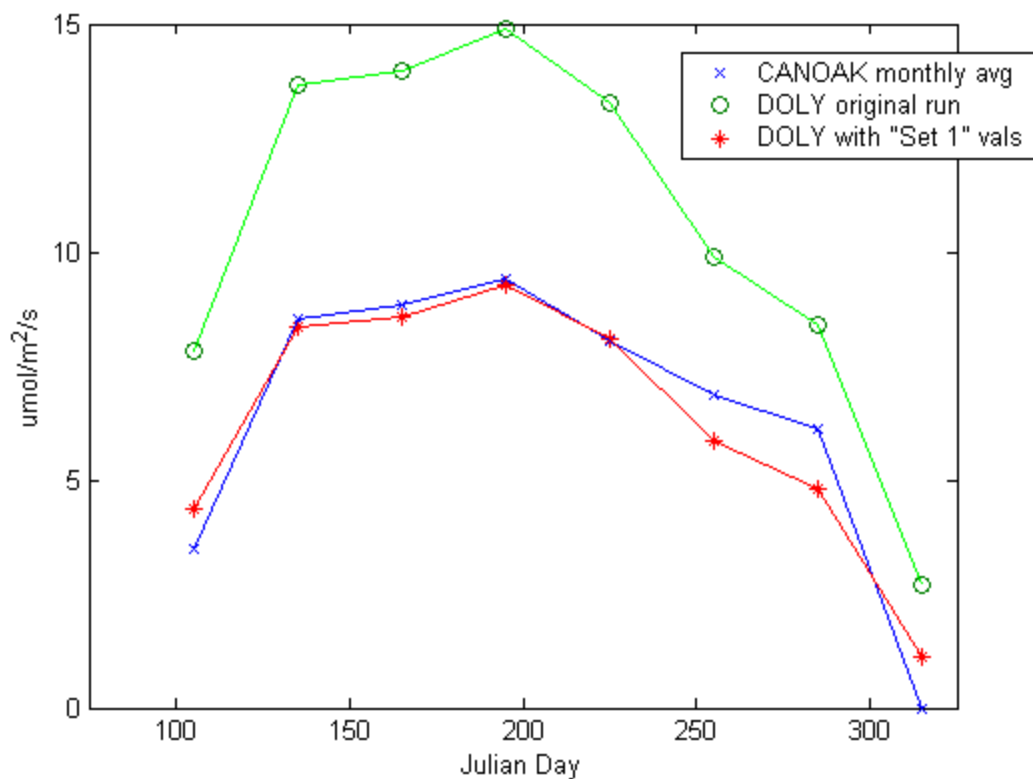

Figure 6: Results of model runs at Site 1, including DOLY results for the original model run and those produced with the optimized parameter values (Set 1).

A metric we used to evaluate the level of consistency, or agreement, between DOLY and CANOAK was the total absolute difference (TAD) between corresponding carbon assimilation calculations. The original difference between the two models was calculated
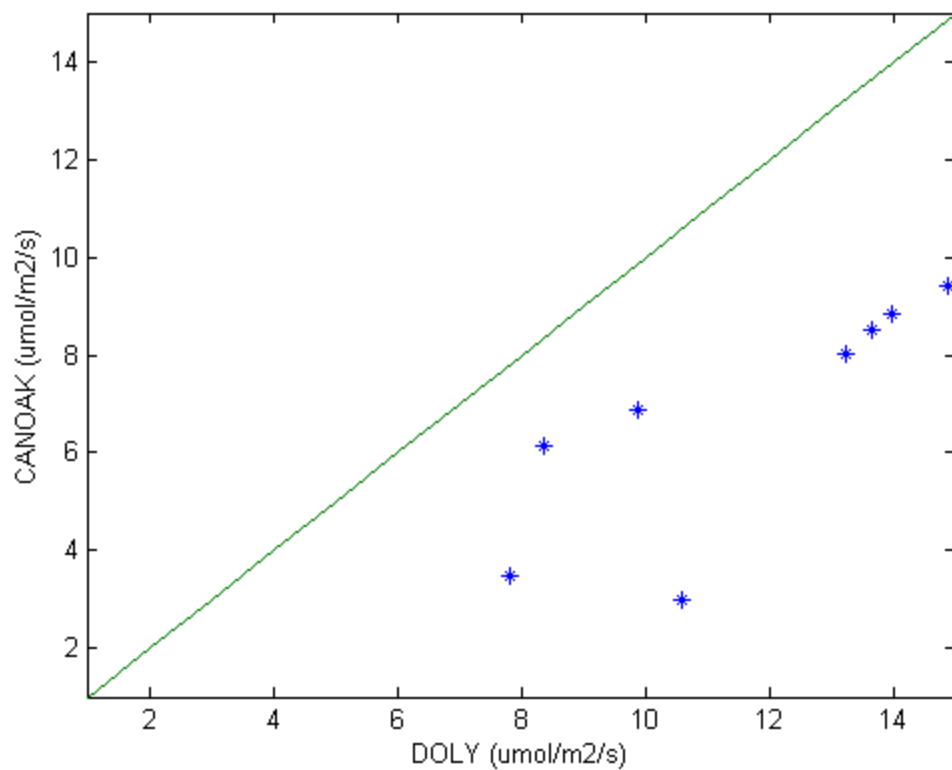
15

Figure 7: Linear correlation plot of canopy carbon assimilation before application of the optimization methodology.
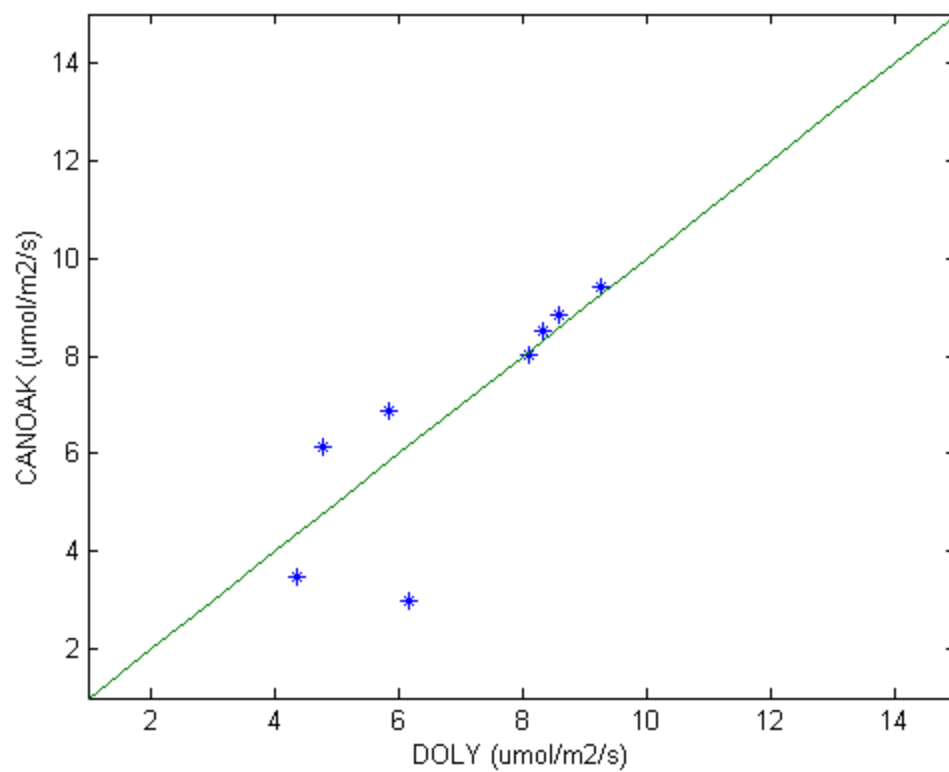

Figure 8: Linear correlation plot of canopy carbon assimilation after application of the optimization methodology.

to be 33.3 at Site 1. Upon inclusion of the optimal parameter values for Site 1, the TAD was reduced to 5.00, a clear demonstration of the ability of this process to produce a parameter set capable of bringing the two models into agreement at a single site.

The optimization at Site 1 produced the following DOLY parameter values: BL=0.5109, CX=-0.3 and SW=0.0642. This parameter set will be referred to as "Set 1". These values display a minor change in the value of BL, with significant effects produced by an increase in SW and a decrease in CX. This set of parameter values is one of many that may produce the desired effect demonstrated here. Nevertheless, these values indicate the importance of SW and CX in guiding carbon assimilation calculations in the formulation used in DOLY.

In order to make an initial test of the degree of robustness of the parameter set produced above, we ran both models under a second scenario. This second scenario was based on the field data taken from Site 2 (Massachusetts), which is approximately 10 latitudinal degrees north of Site 1. This latitudinal gradient, and the different environmental conditions experienced at Site 2, provided the opportunity to test the robustness of the Set 1 parameter values under different conditions than those under which the optimization produced Set 1.

CANOAK was run for Site 2 to produce the assimilation values which would act as the goal for DOLY to achieve. Three different DOLY runs were then made. The model was first run with its original parameter values. This output is labeled "DOLY original run" in Figure 9. The second DOLY run incorporated the optimized parameter values produced in the first set of experiments, Set 1. The results of this model run are labeled "DOLY with Set 1 vals" in Figure 9. In this plot it can be seen that the Set 1 values did not sufficiently generalize to the Site 2 conditions. DOLY significantly undershot the goal, producing a TAD of 14.10. This is in comparison to a TAD of 13.14 for the original DOLY run at Site 2. From Figure 9 it can be seen that the DOLY run with the Site 1 parameter values undershoots the CANOAK goal by approximately the same amount as the original DOLY run misses the CANOAK values. The only small improvement occurred during the month of maximum carbon assimilation, in which the difference between the DOLY run with Site 1 values and CANOAK was less than the difference with the original DOLY run. The difference between the goal and DOLY seems to be spread out more evenly when using the Site 1 values.

These results demonstrate the difficulty in developing a particular set of parameter values that are valid over a wide range of execution conditions. The resolution of this issue is a very difficult, and possibly application-specific, problem meriting more research. Here we present a step toward a possible solution to this problem.

In order to determine a more appropriate set of values at Site 2 for the same parameter set (BL, CX and SW) we optimized over DOLY model runs at Site 2. This optimization produced the following set of values: BL=0.55, CX=-0.1079 and SW=-0.1689. This set of parameter values will be referred to as "Set 2". Plots of the original DOLY run at Site 2, the CANOAK run at Site 2, and the DOLY run with the Set 2 parameter values are given in Figure 10. Once again, we see the ability of the optimization procedure to produce a successful set of parameter values for a specific site. Here the TAD was reduced from 13.14, the original value at Site 2, to 6.14.
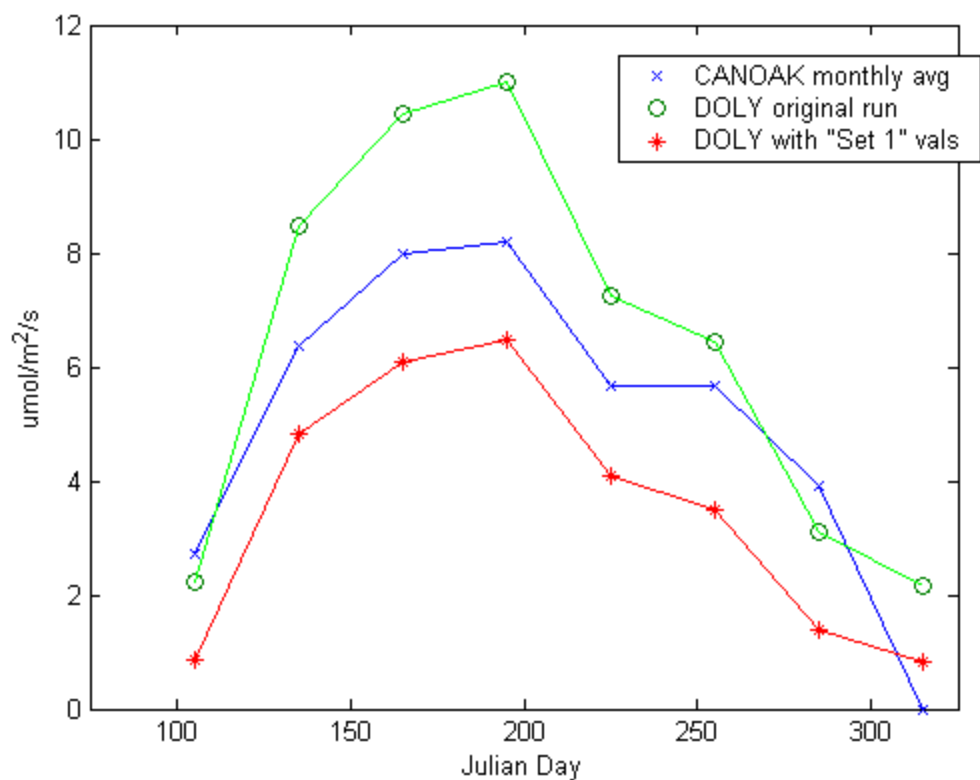
Figure 9: Results of model runs at Site 2, including DOLY results for the original model run and those with the optimized parameter values from Set 1.
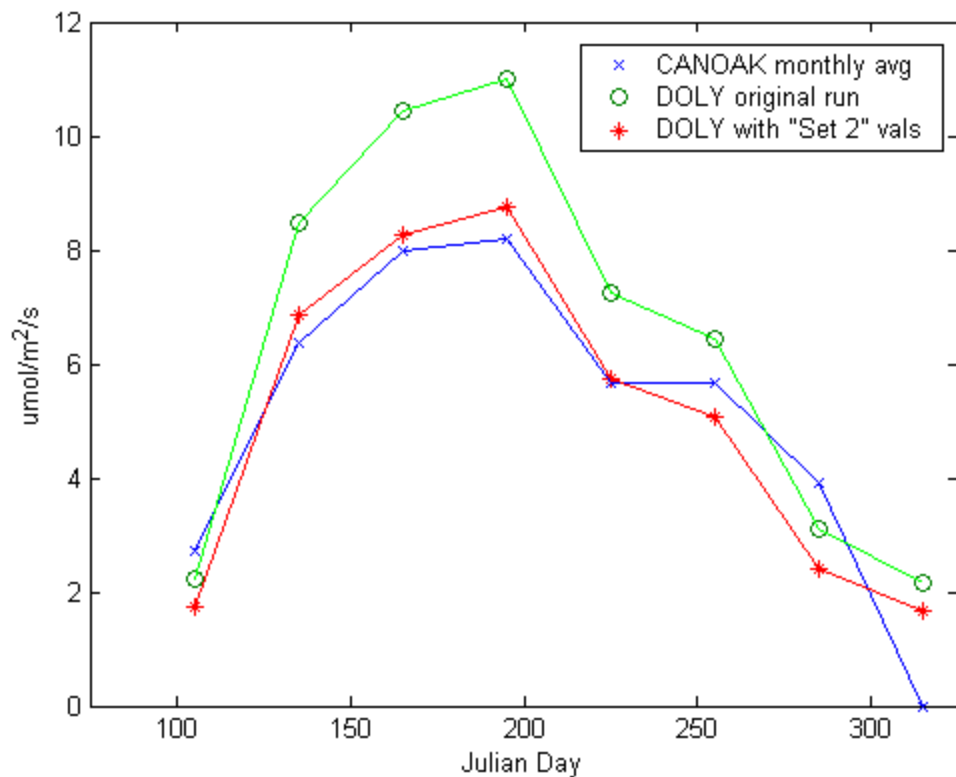

Figure 10: Results of model runs at Site 2, including DOLY results for the original model run and those produced with the optimized parameter values (Set 2).

18

In an attempt to find a single set of parameter values that could be used to simulate both scenarios effectively, we created a third parameter set (Set 3) as a simple average of the values contained in Sets 1 and 2. This set of average parameter values produced the results presented in Figures 11 and 12. For the Site 1 case (Figure 11), the Set 3 parameter values produced a TAD of 8.65. This value is not very different from the 5.00 TAD produced using the Set 1 parameter values, which were created specifically for this scenario. For the Site 2 case, a TAD of 8.59 was produced. Once again, this does not differ by much from the 6.14 TAD value produced by the specific set of parameter values produced for this scenario. While the averaged parameter values of Set 3 are not as precise as the site-specific parameter values in guiding DOLY's predictions into agreement with those of CANOAK, they are clearly a large improvement when compared to the results obtained in the original model runs for both Sites 1 and 2. While there certainly are examples which would cause this averaging procedure to fail, these results point the way to the exploration of statistical methods that could be used to create an acceptable, general parameter set from several scenario-specific sets created from a variety of different scenarios.
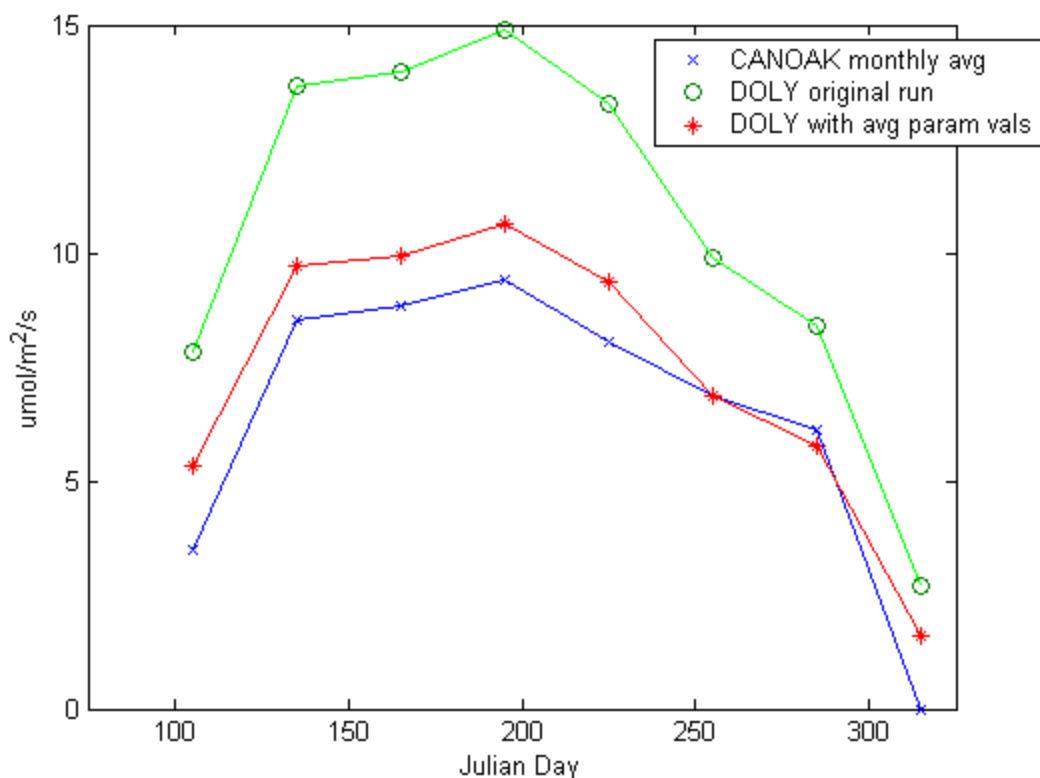

Figure 11: Results of model runs at Site 1, including DOLY results for the original model run and those produced with the averaged optimized parameter values (Set 3).
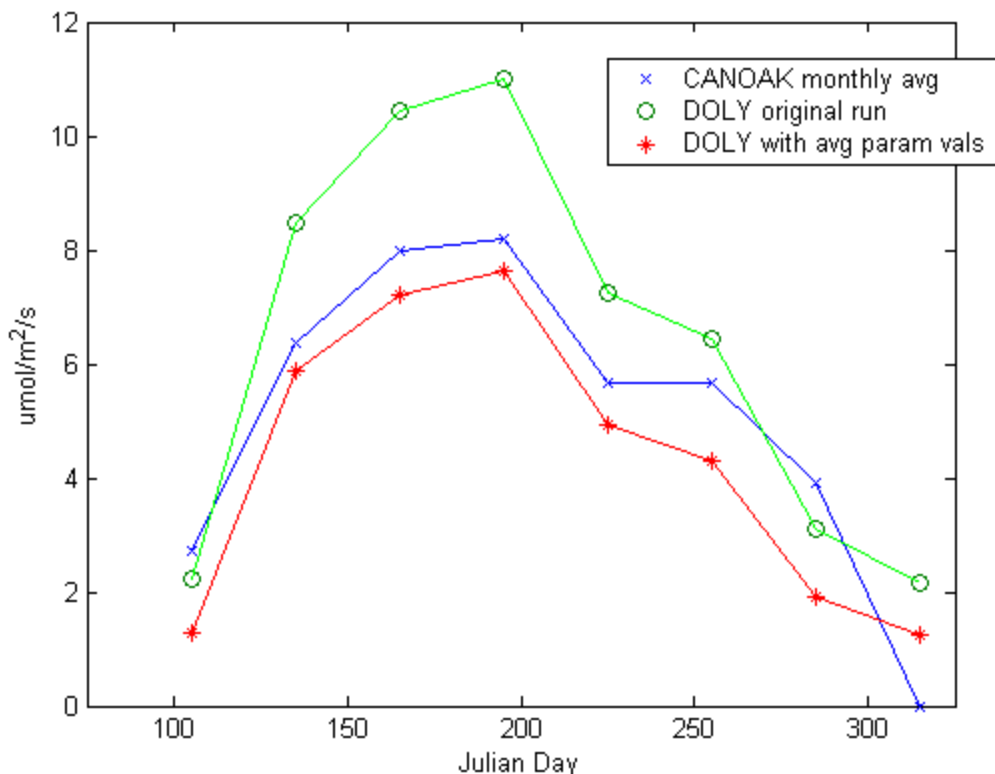
Figure 12: Results of model runs at Site 2, including DOLY results for the original model
run and those produced with the averaged optimized parameter values (Set 3).

5   DISCUSSION

The results presented above demonstrate the promise of the application of
optimization to problems in multi-resolution simulation. The parameterization of DOLY
found above represents one of many possible parameterizations. The choice of the
parameter set is significant as it represents a set of model constructs that have an impact
on the model behavior of interest, carbon assimilation calculations.

There is much work to be done in discovering the degree of utility that this
methodology will have in maintaining multi-resolution consistency in the variety of areas
that need a solution to this problem. Work is necessary to understand the degree to which
constraint characteristics are general or specific to the particular modeling study at hand.
The ability to automate the constraint capture process would be extremely useful and
merits further study. Spacetime constraint systems have been developed which allow
animators to graphically set and vary constraints. This type of system may provide a good
starting point in the search for methods to automatically generate appropriate constraints.

The results presented here represent experiments at specific locations, under specific
sets of circumstances and over specific periods of time. In order to apply this technology
to a broader array of problems with greater confidence, it will be necessary to explore the
degree to which a satisfying parameter set can be transferred to different simulation
circumstances. This question is directly linked to the set of constraints chosen for the
problem, and so the suggestions made previously regarding the study of constraint set
construction apply here as well. As well, exploration of the use of advanced statistical

20

techniques to combine the results derived from optimizations for multiple scenarios may provide positive results.

ACKNOWLEDGMENTS

REFERENCES

Baldocchi, D.D., and P.C. Harley. 1995. Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest. II. Model testing and application. *Plant, Cell and Environment* 18:1157-1173.

Baldocchi, D.D., and K.B. Wilson. 2001. Modeling CO_2 and water vapor exchange of a temperate broadleaved forest across hourly to decadal time scales. *Ecological Modelling* 142:155-184.

Braswell, B.H., D.S. Schimel, J.L. Privette, B. Moore III, W.J. Emery, E.W. Sulzman, and A.T. Hudak. 1996. Extracting ecological and biophysical information from AVHRR optical data: An integrated algorithm based on inverse modeling. *Journal of Geophysical Research* 101 (18): 23,335-23,348.

Davis, P.K., and R. Hillestad. 1993. Families of models that cross levels of resolution: issues for design, calibration, and management. In *Proceedings of the 1993 Winter Simulation Conference*, ed. G.W. Evans, M. Mollaghasemi, E.C. Russell, and W.E. Biles, 1003-1012.

Farquhar, G.D., S. von Caemmerer, and J.A. Berry 1980. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* 149:78-90.

Gleicher, M. 1998. Retargetting motion to new characters. In *Proceedings of SIGGRAPH 98.* 32-42.

Gleicher, M. and P. Litwinowicz. 1998. Constraint-based motion adaptation. *The Journal of Visualization and Computer Animation.* 9 (2):65-94.

Goel, N.S., and D.E. Strebel. 1983. Inversion of vegetation canopy reflectance models for estimating agronomic variables. I. Problem definition and initial results using the Suits model. *Remote Sensing of Environment* 13:487-507.

Harley, P.C., and D. D. Baldocchi. 1995. Scaling carbon dioxide and water vapour exchange from leaf to canopy in a deciduous forest. I. Leaf model parametrization. *Plant, Cell and Environment* 18:1146-1156.

The MathWorks, Inc. 2001. *MATLAB The Technical Language of Computing, Using MATLAB, Version 6.* Available online via <http://www.mathworks.com/access/helpdesk/help/pdf_doc/matlab/using_ml.pdf> [accessed on January 18, 2002]

Privette, J.L., W.J. Emery, and D.S. Schimel. 1996. Inversion of a vegetation reflectance model with NOAA AVHRR data. *Remote Sensing of Environment* 58:187-200.

Reynolds, P.F. Jr. 2002. Using space-time constraints to guide model interoperability. In *Proceedings of the 2002 Spring Simulation Interoperability Workshop*.

Reynolds, P.F. Jr., A. Natrajan, and S. Srinivasan. 1997. Consistency maintenance in multiresolution simulations. *ACM Transactions on Modeling and Computer Simulation* 7(3):368-392.

Witkin A., and Z. Popovic. 1988. Spacetime constraints. In *Proceedings of SIGGRAPH 88*. 22:159-168.

Woodward, F.I., T.M. Smith, and W.R. Emanuel. 1995. A global land primary productivity and phytogeography model. *Global Biogeochemical Cycles* 9(4):471-490.