RESEARCH **Open Access**

A spatially distributed, deterministic approach to modeling *Typha domingensis* (cattail) in an Everglades wetland

Gareth Lagerwall[1], Gregory Kiker[1*], Rafael Muñoz-Carpena[1], Matteo Convertino[1], Andrew James[2] and Naiming Wang[3]

Abstract

Introduction: The emergent wetland species *Typha domingensis* (cattail) is a native Florida Everglades monocotyledonous macrophyte. It has become invasive due to anthropogenic disturbances and is out-competing other vegetation in the region, especially in areas historically dominated by *Cladium jamaicense* (sawgrass). There is a need for a quantitative, deterministic model in order to accurately simulate the regional-scale cattail dynamics in the Everglades.

Methods: The Regional Simulation Model (RSM), combined with the Transport and Reaction Simulation Engine (TARSE), was adapted to simulate ecology. This provides a framework for user-defineable equations and relationships and enables multiple theories with different levels of complexity to be tested simultaneously. Five models, or levels, of increasing complexity were used to simulate cattail dynamics across Water Conservation Area 2A (WCA2A), which is located just south of Lake Okeechobee, in Florida, USA. These levels of complexity were formulated to correspond with five hypotheses regarding the growth and spread of cattail. The first level of complexity assumed a logistic growth pattern to test whether cattail growth is density dependent. The second level of complexity built on the first and included a Habitat Suitability Index (HSI) factor influenced by water depth to test whether this might be an important factor for cattail expansion. The third level of complexity built on the second and included an HSI factor influenced by soil phosphorus concentration to test whether this is a contributing factor for cattail expansion. The fourth level of complexity built on the third and included an HSI factor influenced by (a level 1–simulated) sawgrass density to determine whether sawgrass density impacted the rate of cattail expansion. The fifth level of complexity built on the fourth and included a feedback mechanism whereby the cattail densities influenced the sawgrass densities to determine the impact of inter-species interactions on the cattail dynamics.

Results: All the simulation results from the different levels of complexity were compared to observed data for the years 1995 and 2003. Their performance was analyzed using a number of different statistics that each represent a different perspective on the ecological dynamics of the system. These statistics include box-plots, abundance-area curves, Moran's *I*, and classified difference. The statistics were summarized using the Nash-Sutcliffe coefficient. The results from all of these comparisons indicate that the more complex level 4 and level 5 models were able to simulate the observed data with a reasonable degree of accuracy.

(Continued on next page)

* Correspondence: gkiker@ufl.edu
[1]Frazier Rogers Hall, University of Florida, PO Box 110570, Gainesville, FL 32611-0570, USA
Full list of author information is available at the end of the article

(Continued from previous page)

Conclusions: A user-defineable, quantitative, deterministic modeling framework was introduced and tested against various hypotheses. It was determined that the more complex models (levels 4 and 5) were able to adequately simulate the observed patterns of cattail densities within the WCA2A region. These models require testing for uncertainty and sensitivity of their various parameters in order to better understand them but could eventually be used to provide insight for management decisions concerning the WCA2A region and the Everglades in general.

Keywords: Typha, Modeling, Ecology, Dynamics, Model complexity, Water conservation area 2A, Transport and reaction simulation engine, Regional simulation model

Introduction

The Everglades, commonly known as the "River Of Grass" Douglas (1947), in southern Florida, USA, once covered some 28,500 km^2. This wetland ecosystem was sustained by the Kissimmee River, flowing through Lake Okeechobee and southwards as a shallow, slow-moving sheet of water flowing freely to the estuaries of Biscayne Bay, Ten Thousand Islands, and Florida Bay. The channelization of the Everglades around 1948 caused the reduction of the original wetland areas by up to 50%, with related declines in dependent wildlife. In addition to the changes in hydrology, continuous mining, agriculture, and urbanization activities have resulted in invasive and exotic plants becoming established in place of the original vegetation, altering habitats and often forming mono-crop stands (single species environments) (Odum et al. 2000).

The Comprehensive Everglades Restoration Plan (CERP) was implemented in 2000 (USACE, S.F.R.O 2010a) with the explicit goal of restoring some of the Everglades' former extent and ecosystem functioning. The main focus of CERP has been on improved management of water quantity and water quality with the assumption that if the water quantity and quality are adequate, the ecology will follow suit. There is, however, an increasing focus on the ecological impacts of various management decisions, and these efforts center on improving species diversity and protecting existing habitats (USACE, S.F.R.O 2010b). In an effort to achieve these goals, stormwater treatment areas (STA) were constructed just south of the Everglades agricultural area (EAA) to filter out phosphorus from the water before releasing it into the water conservation areas (WCA). The WCAs act as impoundments for water storage and flood control as well as serving as wildlife habitat. Water flows from these WCAs into the Everglades National Park (Guardo et al. 1995).

Typha domingensis as an invasive species

The emergent wetland species *Typha domingensis* (cattail) is a native Everglades monocotyledonous macrophyte, typically occurring as a sparse complement alongside *Cladium jamaicense* (sawgrass) stands. These two species have significantly different morphology, growth, and life history characteristics (Miao and Sklar 1998), and this has enabled the cattail to expand prolifically under the altered conditions in the Everglades. In the 1980s, the area covered by cattail stands in WCA2A doubled, expanding southward into the sawgrass marshes (Willard 2010). Cattail has hence been labeled as an indicator species, or species of concern, and its distribution is used to determine the effectiveness of various water management decisions. Cattail expansion has been studied extensively (Miao 2004; Wu et al. 1997; Newman et al. 1998), and it has been determined that there are four main external factors that affect its growth and aid in cattail's dominance over sawgrass. These factors include water depth, hydroperiod, soil phosphorus concentration, and disturbance (Newman et al. 1998). It was determined that the optimum water depth at which cattail grows is between 24 and 95 cm (Grace 1989), with a hydroperiod of 180–280 days (Wetzel 2001). In terms of soil phosphorus concentration, cattail has been found to be invading the natural sawgrass habitats of WCA2A along a soil phosphorus gradient running from the northwest (high concentrations) to the southeast (low concentrations). Urban et al. (1993) mention that, given an adequate water depth, soil phosphorus concentration is the next most important factor in determining cattail expansion/invasion. In creating their water quality model for simulating soil phosphorus concentrations downstream of the Everglades STAs, Walker and Kadlec (1996) determined that the lower bound soil phosphorus concentration for the optimum growth of cattail was 540 mg/kg. Fires and other disturbances such as hurricanes were also found to affect the colonization of areas by cattail by altering local topography and nutrient concentrations (Newman et al. 1998).

Ecological model designs to address everglades systems

In order to assess these various influences on cattail and other ecological components, a variety of computation models were designed and implemented. These models aid our understanding of complex systems and allow scientists and managers to evaluate different ecological outcomes of decisions before the more costly task of their implementation (Fitz et al. 2011). To ensure numerical efficiency, most spatially distributed models have their

equations, laws, and assumptions "hard-coded" into their programming code. This creates a "fixed-form" model, with changes in the functioning coming through extensive code re-writes and careful redesign around logical structures. Dynamic "free-form" simulation models, such as STELLA (Costanza and Voinov 2001), QnD (Kiker and Linkov 2006; Kiker et al. 2006), and the Kepler system (Ludascher et al. 2006) are generally written using an object-oriented programming (OOP) language such as C++ (Stroustrup 2000) or Java (Arnold and Gosling 1998), as opposed to a linear language such as FORTRAN (Cary et al. 1998). When interacting with free-form models and their algorithms, designers do not interact directly with the program code. Rather, they influence objects through placing data, storage, and logical structures into either a graphical user interface (STELLA, Kepler) or within a meta-code structure such as the eXtensible Markup Language (XML) (Harold 1998).

There are a number of fixed-form ecological models currently in use across the Everglades region. Of these, the Across Trophic Level System Simulation (ATLSS) (Gross 1996) and the Everglades Landscape Model (ELM) (Fitz and Trimble 2006b) are probably the most well-known. These and most other models available for modeling cattail in the Everglades are entirely qualitative, that is, they involve switching between one species and another. The majority of these current ecological models are also stochastic, that is, based on probabilities and a degree of randomness and uncertainty. They generally run as post-process models, using hydrological data output by other models such as the South Florida Water Management Model (SFWMM) (Fitz et al. 2011).

The ATLSS vegetation succession model is used to determine the succession of one habitat type to another (e.g., sawgrass to cattail). The ATLSS model simulates with an annual time step on square 500 m cells and uses a stochastic cellular automata model to switch between vegetation types. Currently there is no way to determine vegetation densities within vegetation types (Duke-Sylvester 2005).

The ELM model uses a counter to switch between species by accumulating days of water level and soil phosphorus concentration above certain limits. The model then switches between species based on their preferred hydroperiod and historical soil phosphorus concentrations (Fitz and Trimble 2006a). The ELM model is the only currently available simulation tool for evaluating water quality across the Everglades landscape and does not simulate detailed ecological features (Fitz et al. 2011).

Another modeling effort by Wu et al. (1997) used Markov chain probabilities to switch between *Cladium* and *Typha* species. This model was in fact used to inform the ATLSS nutrient and fire disturbance model (Wetzel 2003). Again, this is a stochastic, species-specific, presence/absence-type model.

A modeling effort by Tarboton et al. (2004) developed a set of habitat suitability indices (HSI) for evaluating water management alternatives. These HSIs provided a range of probabilities for a particular species occurring across the landscape and were based predominantly on local hydrological conditions such as depth (maximum, minimum, and mean), hydroperiod, velocity, and flow direction.

Given that water quantity (depth) and quality (soil phosphorus concentration) affect cattail (and other plants) growth and distribution, there is a need to integrate these components to determine the more detailed biological outcomes of an Everglades ecological model. There is also a need for a quantitative model to provide continuous density values for specific vegetation rather than simply presence/absence information. Given that the Everglades restoration includes a large and ongoing research effort, there is a need to efficiently test and explore potentially useful algorithms in an adaptable, ecological modeling engine.

The RSM/TARSE ecological model

A combined effort of the University of Florida, the South Florida Water Management District (SFWMD), and the US Geological Survey created the Transport and Reaction Simulation Engine (TARSE) (Jawitz et al. 2008), which was originally designed to run in line with the SFWMD-developed Regional Simulation Model (RSM) (SFWMD 2005c) to simulate soil phosphorus dynamics in the Everglades system. The OOP structure of this coupled hydrologic/water quality model, along with the user-definable inputs and interactions, allowed for the extension of this model beyond its original purpose into ecological processes and features. The coupled RSM/TARSE (henceforth referred to as RTE) model, implemented with the goal of modeling ecological features within the southern Florida landscape and presented in this paper, is a spatially distributed, free-form model simulating cattail biomass distribution and dynamics across WCA2A. Using the RTE model to couple vegetation dynamics with phosphorus dynamics has been alluded to by Jawitz et al. (2008), Muller (2010), and Perez-Ovilla (2010) during their respective TARSE-influenced, WQ simulations. Zajac (2010) used vegetation types to influence Manning's *n* and evapotranspiration coefficients. These parameters were informed by initial vegetation types and not by changing vegetation distribution and density over time.

There is therefore a definite need for the RTE model, which allows one to model a vegetation species quantitatively and ultimately determine the ecological impact of various management scenarios falling under the CERP initiative. This new engine would accommodate different algorithms or new species as available data or new

knowledge becomes available. It would allow for interactions and feedback effects within species as well as among different species and with other environmental factors.

Objectives and hypotheses

The primary objective of this paper is to test and apply a new spatially distributed, deterministic, free-form (user-definable), quantitative ecological model of cattail dynamics. A significant advantage of this free-form modeling approach is that multiple ecological algorithms of differing complexity can be quickly implemented and tested simultaneously, instead of through time-consuming code additions. As a first step of our objective, we tested the influence of increasing cattail model complexity on reducing uncertainty in simulated output (Lindenschmidt 2006). Five levels of increasing complexity were selected to model the cattail densities. These five levels of complexity were chosen to correspond with various hypotheses regarding the growth and spread of cattail in the Everglades, namely:

1. Whether cattail growth is density dependent.
2. Whether water depth is an important factor for cattail expansion.
3. Whether soil phosphorous is a contributing factor for cattail expansion.
4. Whether sawgrass density impacts the rate of cattail expansion.
5. Whether inter-species interactions between cattail and sawgrass contribute to the observed cattail dynamics.

Following the methodology used by Jawitz et al. (2008), a simple logistic function (Keen and Spain 1992) formed the base of the complexities with water depth and soil phosphorus concentration [the two most important factors influencing cattail growth according to Newman et al. (1998)] and sawgrass interaction influencing the higher levels of complexity. A second step in our objective was to use an existing ecosystem and its monitoring data to analyze performance of our five candidate models. The entire WCA2A vegetation dataset (1991, 1995, and 2003), obtained from Rutchey et al. (2008), was chronologically divided into model training and testing sections. Training of the model was conducted for the years 1991–1995, where the growth factor (found in Equation 3) was fitted to the level 1 complexity. As a third step in our objective, model testing was conducted on the two time periods of 1991–2003 (testing 1) and 1995–2003 (testing 2), respectively, with the testing 2 time period being equivalent to a blind test (due to different initial conditions). The 1991 and 1995 vegetation maps were used to initialize the training, testing 1, and testing 2 simulations, respectively. Model output from the training, testing 1, and testing 2 simulations was compared with the 1995 and 2003 vegetation maps. Model output was compared to observed patterns, and the most accurate level of complexity thus determined.

Methods

In order to reproduce the observed cattail patterns, both hydrological and water quality data were used as inputs for the ecological model. To this end, it was decided to use the Regional Simulation Model (RSM), which was developed by the South Florida Water Management District (SFWMD) to replace the popular SFWMM, coupled with the Transport and Reaction Simulation Engine (TARSE) to provide the base structure for modeling cattail dynamics across the test site.

The Regional Simulation Model (RSM)

Developed by SFWMD, the RSM simulates hydrology over the South Florida region. It is often thought of as the successor to the successful SFWMM, referred to as the "2-by-2" model for its 2 mile resolution (SFWMD 2005a). The RSM operates over a variable triangular mesh grid, in contrast to the 3.22 km (2 mile) square grid of the SFWMM; this enables higher resolution in areas of concern as well as the ability to delineate canals (SFWMD 2005c). The RSM uses a weighted, implicit, finite volume method to simulate two-dimensional diffusional flow and hence implicitly simulates groundwater flow and surface water flow (SFWMD 2005c). The OOP design structure of RSM allows for the abstraction and modularity of various components (SFWMD 2005b). A result of this is that there are two engines that comprise the RSM, namely the Hydrologic Simulation Engine (HSE) and the Management Simulation Engine (MSE). The HSE simulates all the hydrological processes, while the MSE simulates various management or control regimes. These two engines interact at runtime to provide an accurate representation of the hydrodynamics of the region (SFWMD 2005c).

Simulating transport and reactions using TARSE

The TARSE was recently developed to simulate water quality (WQ) components within the RSM model for areas in the Everglades system (Jawitz et al. 2008). The TARSE model was designed to be as generic as possible, to allow multiple water quality components to be simulated with a simple change in the input file. It was first implemented as another engine to be incorporated within the RSM framework, along with the HSE and MSE, called the Water Quality Engine (WQE). Due to its structure, the WQE does not simulate hydrology and requires a hydrologic driver to feed it values of flow and depth at every time step (SFWMD 2008b). TARSE has since been decoupled from RSM and implemented with other hydrologic drivers such as Flow and Transport in a Linked Overland-Aquifer Density Dependent System (FTLOADDS) (Wang et al. 2007; Muller 2010) and VFSMOD (Muñoz-Carpena et al.

1999; Perez-Ovilla 2010). TARSE solves the advection-dispersion-reaction equations (ADRE) over an unstructured triangular mesh (James and Jawitz 2007). The ADRE is represented by Equation 1, and every term is a function of a two-dimensional spatial coordinate x, with components (x_1, x_2), and time, t.

$$\frac{d(\Phi hc)}{dt} + \nabla(chu - hD^* \cdot \nabla c) + hf_2 c = hf_1 c_1 \qquad (1)$$

Where t is time [T], $c(x,t)$ is the concentration [M/L^3], and $\Phi(x,t)$ is the porosity of the medium (which may be 1 for surface water) [L^3/L^3]. $h(x,t)$ is the water depth [L] or thickness of the saturated zone in groundwater flow, $u(x,t)$ is the specific discharge [L/T] of water (either surface or groundwater), and $D^* = D^*(u(x,t))$ is the dispersion tensor (a function of u). $f_1(x,t)$ is a source rate [M/L$^{3\cdot}$T] with associated concentration c_1, and $f_2(x,t)$ is a first-order decay rate [M/L$^{3\cdot}$T]. The density [M/L^3] of the water is assumed to be constant.

The basis of TARSE involves transfers (e.g., settling, diffusion, growth) between various stores, such as soil water column solutes, pore water solutes, macrophytes, and suspended solids. The specifics of these stores, and the transfers among them are user-definable in the XML input file (Jawitz et al. 2008). TARSE equations are composed of pre-equations, equations, and post-equations. Pre- and post-equations are used for implementing conditional ("if-then-else") statements as part of pre- and post-processing after the main processing in the equations. For example, pre-processing could be used to determine if the current water depth [m] is above the threshold for cattail optimum growth and thus reduce the depth influence factor accordingly. If the depth is less than the optimum growing depth, then the influence factor decreases accordingly. The logic just described is represented by Equation 2, as described by Grace (1989), where cattail optimum depth is 70 cm.

$$\begin{aligned} &If\,(depth > cattail\_optimum\_depth) \\ &Then \\ &depth_H SI = 1 - \left(\frac{depth - cattail\_optimum\_depth}{109}\right) \\ &Else \\ &depth_H SI = 1 - \left(\frac{cattail\_optimum\_depth - depth}{112}\right) \end{aligned}$$
$$(2)$$

The main equations are structured as ordinary differential equations (ODE) (SFWMD 2008a).

The RSM/TARSE coupling represents possibly the first time that a free-form dynamic system model has been integrated with a fixed-form, spatially distributed, hydrologic model (Muller 2010). This unique coupling, with user-defined interactions operating across a spatially distributed domain, lends itself to simulating

ecological behaviors (growth, death, movement, and feeding) as well as the original WQ interactions. The model can currently only solve ADRE movement and as such is insufficient for ecological/animal movement. Attempts to include some form of Lagrangian-type movement in this model are discussed by Lagerwall (2011).

Model application

In order to test the influence of increasing complexity on reducing uncertainty in model output (Lindenschmidt 2006), five levels of increasing complexity were selected to model the cattail densities. Following the methodology used by Jawitz et al. (2008), a logistic function (Keen and Spain 1992) was used for the most basic, level 1 complexity, due to its density dependent growth and rapid (exponential) early stages of growth. The logistic function is represented in Equation 3.

$$\frac{dP}{dt} = GF \times P \times \left(1 - \frac{P}{K}\right) \qquad (3)$$

Where P is the population density [M/L^2], t is time [T], GF is the constant growth rate [T^{-1}], and K is the carrying capacity or maximum population density [M/L^2].

Level 2 is a water-depth-influenced level 1 complexity. A water depth factor (habitat suitability index) ranging from 0 to 1 is multiplied by the carrying capacity in the logistic function. The depth factor decreases linearly from 1 as the current depth either rises above or drops below the optimum (70 cm) growing depth. This depth factor can be seen in Equation 4.

$$\frac{dP}{dt} = GF \times P \times \left(1 - \frac{P}{K \times DepthF}\right) \qquad (4)$$

Where P is the population density [M/L^2], t is time [T], GF is a constant growth rate [T^{-1}], $DepthF$ is the water depth factor [L/L], K is the carrying capacity or maximum population density [M/L^2].

Level 3 is a soil-phosphorus-influenced level 2 complexity, with the soil phosphorus factor being incorporated in a similar fashion to the depth factor and can be seen in Equation 5.

$$\frac{dP}{dt} = GF \times P \times \left(1 - \frac{P}{K \times (DepthF + phosphorusF)/2}\right) \quad (5)$$

Where P is the population density [M/L^2], t is time [T], GF is a constant growth rate [T^{-1}], $DepthF$ is the water depth factor [L/L], $phosphorusF$ is the soil phosphorus factor [M/L^3/M/L^3], and K is the carrying capacity or maximum population density [M/L^2]. The soil phosphorus factor behaves like a logistic function, increasing from 0 to 1 as soil phosphorus concentration increases to 1,800 from 200

mg/kg, as described by Walker and Kadlec (1996), and can be seen in Equation 6.

$$\text{phosphorusF} = \left(1 + e^{-\frac{\text{phosphorus} - 1034}{144}}\right)^{-1} \quad (6)$$

Where *phosphorusF* is the soil phosphorus HSI, ranging from 0 to 1, and *phosphorus* is the current soil phosphorus concentration (mg/kg).

Level 4 builds on a level 3 complexity with an added sawgrass interaction factor, much like the soil phosphorus and depth factors. It decreases linearly from 1 to 0.16 as sawgrass densities increase to 1,958 from 0 g/m^2 (Doren et al. 1999), which is their reported maximum density (Miao and Sklar 1998). The sawgrass is set to grow according to a level 1 complexity as in Equation 4, thus the level 4 complexity is represented by Equation 7.

$$\frac{dP}{dt} = GF \times P \times \left(1 - \frac{P}{K(DepthF + phosphorusF + sawgrassF)/3}\right) \quad (7)$$

Where *P* is the population density [M/L^2], *t* is time [T], *GF* is a constant growth rate [T^{-1}], *DepthF* is the water depth factor [L/L], *phosphorusF* is the soil phosphorus factor [M/L^3/M/L^3], *sawgrassF* is the sawgrass influence factor [M/L^2/M/L^2], and *K* is the carrying capacity or maximum population density [M/L^2]. The sawgrass factor varies according to Equation 8.

$$\text{sawgrassF} = 1 + (-0.84 \times (\text{sawgrass}/K_{SAW})) \quad (8)$$

Where *sawgrassF* is the sawgrass HSI ranging from 0 to 1, *sawgrass* is the current sawgrass density, and K_{SAW} is the sawgrass carrying capacity.

The level 5 complexity is the same as level 4, but with a density-dependent influence on the level 1 sawgrass model, which is represented by Equations 9 and 10, respectively.

$$\frac{dP}{dt} = GF \times P \times \left(1 - \frac{P}{K \times cattailF}\right) \quad (9)$$

Where *P* is the population density [M/L^2], *t* is time [T], *GF* is a constant growth rate [T^{-1}], *cattailF* is the cattail factor ranging from 0 to 1, and *K* is the carrying capacity or maximum population density [M/L^2].

$$\text{cattailF} = 1 + (-0.84 \times (\text{cattail}/K_{CAT})) \quad (10)$$

Where *cattailF* is the cattail HSI ranging from 0 to 1, *cattail* is the current cattail density, and K_{CAT} is the cattail carrying capacity.

The depth, soil phosphorus, and sawgrass interaction factors are all calculated using the pre-equations, similar to that presented in Equation 2. These factors are then incorporated into the main growth equations, presented in Equations 4, 5, 7 and 9 representing levels of complexity 2 through 5, respectively.

In TARSE, components are listed as either mobile or stabile. Mobile components are moved in the water using the ADRE equations, while the stabile components do not move and only undergo the reaction part of the ADRE. Given the complexities associated with simulating wind-borne or water-borne transportation of seeds and rhizome expansion—which is another mode of expansion noted by Miao (2004)—all mesh elements were initialized (seeded) with cattail, with areas originally not containing cattail being seeded with the minimum value of 10 g(dry weight)/m^2. This assumption represents the presence of a seed bank, providing cattail the opportunity to colonize an area as soon as conditions become favorable. Vegetation then is modeled as a stabile component, with no means for dispersal, or in another way we assume "infinite dispersal." The latter assumption is supported by very high values of dispersal for seeds in the Everglades, enhanced by the diffused presence of biotic (animals) and abiotic (water, wind) dispersal vectors (Miao and Sklar 1998). Also, as a result of this current inability for modeled dispersal, the maximum influence that the aforementioned factors such as phosphorusF, sawgrassF, and cattailF can have has been limited so that they reduce the cattail population to 1% of its maximum density.

Test site

The test site used for ecological model development and testing was the WCA2A (Figure 1). WCA2A is a 547 km^2 managed wetland just south of Lake Okeechobee, FL, and accounts for about 6.5% of the total area of the Everglades. It came into existence in 1961 with the construction of the L35-B canal and receives inflow from the Stormwater Treatment Areas (STAs), before discharging into downstream water conservation areas, and eventually into the Everglades National Park (Urban et al. 1993). According to Rivero et al. (2007b), the region has an average annual temperature of 20°C, and precipitation between 1,175 and 1,550 mm. The elevation range in WCA2A is between 2.0 and 3.6 m above sea level, which generates a slow sheet flow from the northwest to the southwest of the region. The hydrology is controlled by the SFWMD at a number of inlet and outlet structures (green squares in Figure 1) along the surrounding canals (blue lines in Figure 1). The landscape is composed of dominant sawgrass marshes, shrub and tree island communities, and invasive cattail communities (van der Valk and Rosburg 1997). WCA2A has
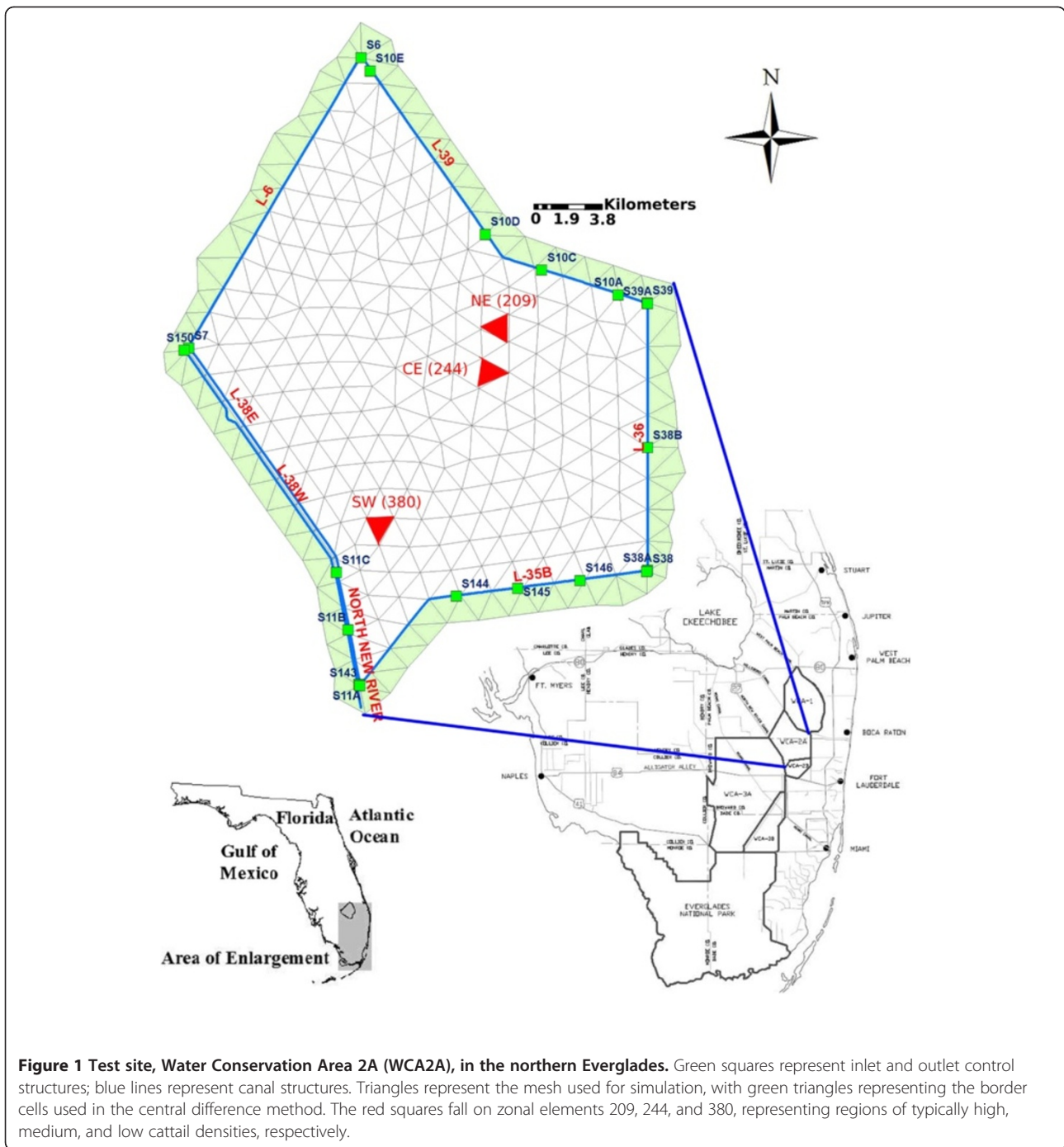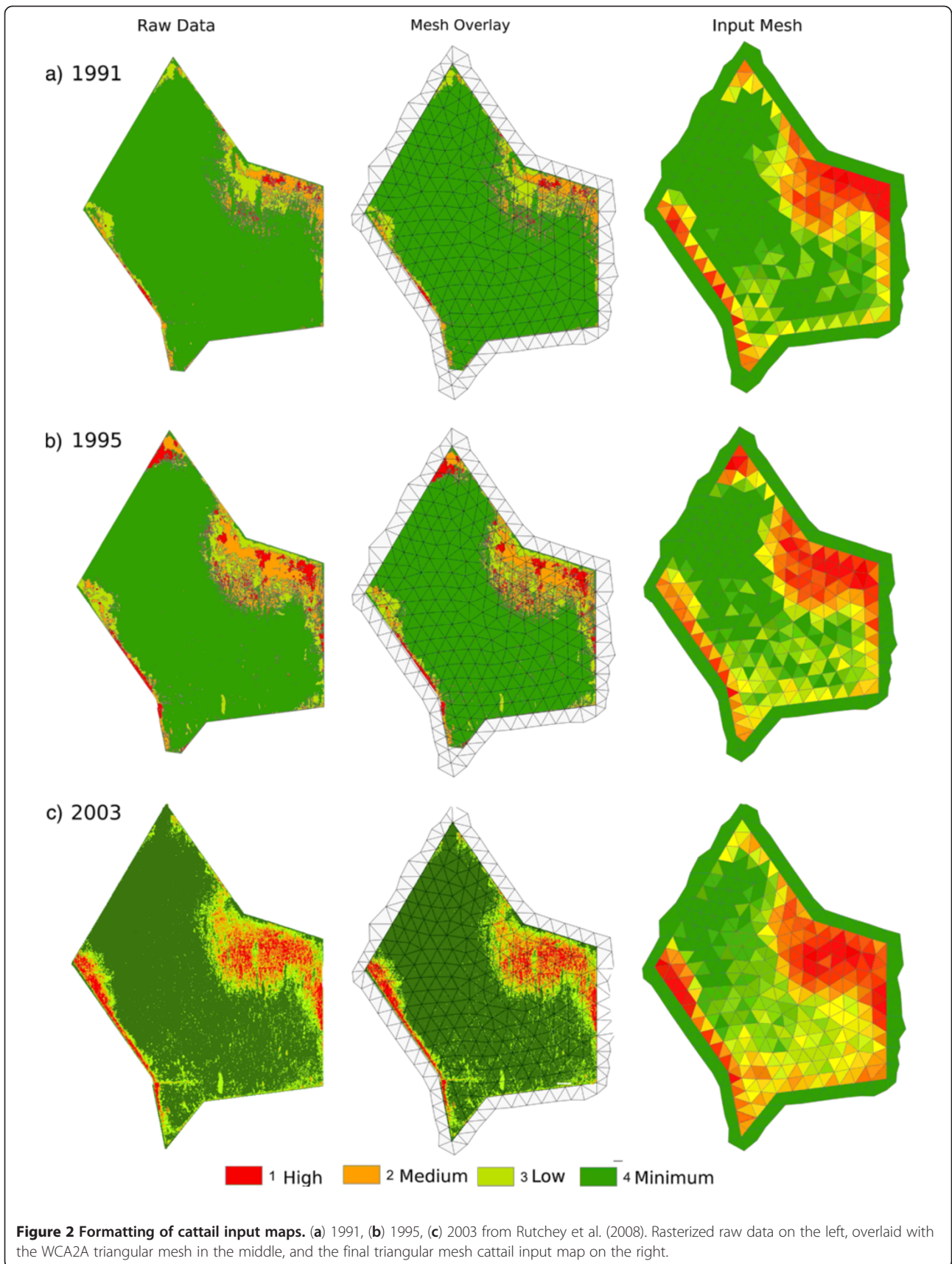
Figure 1 Test site, Water Conservation Area 2A (WCA2A), in the northern Everglades. Green squares represent inlet and outlet control structures; blue lines represent canal structures. Triangles represent the mesh used for simulation, with green triangles representing the border cells used in the central difference method. The red squares fall on zonal elements 209, 244, and 380, representing regions of typically high, medium, and low cattail densities, respectively.

been used extensively as a research site by the SFWMD, with extensive trial and monitoring programs for a number of biogeochemical components, especially soil phosphorus and vegetative structure (Rivero et al. 2007a). The triangular mesh grid used for simulation is also displayed in Figure 1, with the green border cells used for numerical stability of the hydrological RSM component. An overview of the HSE setup for WCA2A, which provides the

hydrological operating conditions, can be found in SFWMD (2008c).

Initial conditions, boundary conditions, and time series data

Cattail vegetation maps (Figure 2) are used for the initial conditions as well as for comparing model output with measured data. Hydrological time series are used for initial and boundary conditions along the surrounding canals.

Figure 2 Formatting of cattail input maps. (**a**) 1991, (**b**) 1995, (**c**) 2003 from Rutchey et al. (2008). Rasterized raw data on the left, overlaid with the WCA2A triangular mesh in the middle, and the final triangular mesh cattail input map on the right.

Using RSM, the hydrological boundary conditions are converted into depth values across the domain, which are then used as inputs in the level 2 complexity algorithm. Soil phosphorus concentration maps provide initial conditions and an influence factor for the level 3 complexity algorithm. Sawgrass vegetation maps are used as initial conditions for the level 1 complexity sawgrass model, which serves as an influence factor for the level 4 and level 5 complexity cattail algorithms. The following sections provide additional detail on these model inputs.

Hydrological time series

The hydrology of WCA2A is controlled primarily by the operation of control points along the S10 and L35-B canals. The hydrology data were obtained from the SFWMD, which uses the WCA2A site as a test site for the RSM. The average depth for the region ranges from 60 to 90 cm (SFWMD, 2008c). The input dataset consisted of a daily time series of hydraulic head values (m) at the inlet and outlet control structures of WCA2A (represented by the green squares in Figure 1) for the years 1979–2000 (Wang 2009). The time series have since been updated to 2008 for all control structures using data collected from the DBHYDRO website (SFWMD 2009).

Soil phosphorus

A gradient of soil phosphorus exists along WCA2A, with a high concentration near the inlets at the north, and a low concentration at the outlets in the south. This soil phosphorus gradient has been widely documented and studied (DeBusk et al. 1994; Grunwald et al. 2004, 2008; Rivero et al. 2007a,b; Grunwald 2010). Given the unavailability of spatial soil phosphorus data beyond map classifications (Grunwald 2010), soil phosphorus input maps were created by overlaying the WCA2A mesh on the existing maps obtained from Grunwald et al. (2004, 2008). The soil phosphorus map of 1990 was used for the model training period of 1991–1995, while the soil phosphorus map of 2003 was used for both the testing 1 (1991–2003) and testing 2 (1995–2003) simulation periods. Due to the poor quality of these soil phosphorus input maps and the inability of TARSE to adequately simulate phosphorus dynamics in the WCA2A region (as it is still in development), the soil phosphorus concentration itself was not simulated, i.e., the static soil phosphorus concentration provided by the input maps was used to inform the model throughout the simulation period.

Cattail and sawgrass

Vegetation maps for WCA2A were obtained for the years 1991, 1995 (Rutchey 2011), and 2003 (Wang 2009), which were all used in Rutchey et al. (2008). These maps provided density (g/m^2) distributions across the test site for cattail. The negative correlation between sawgrass and cattail has been reported by Doren et al. (1999) and Richardson et al. (2008), and various other vegetation maps of the area, namely 1991 (Jensen et al. 1995), 1995 (SFWMD 1995), 1999 (SFWMD 1999), and 2003 (Wang 2009), confirm this negative correlation. Although sawgrass density is related to more environmental factors than only cattail density (Miao and Sklar 1998), a simple negative correlation with the cattail maps was used in order to assign densities to the sawgrass maps. For example, high sawgrass density values (1,600 g/m^2) were assigned to regions with typically low cattail density values, and low sawgrass density values (600 g/m^2) were assigned to regions with high cattail density values.

The program ArcMap (ESRI Environmental Systems Resource Institute 2010) was used to create a uniform raster map from the original images which had a minimum mapping unit of 50 m^2 (Rutchey et al. 2008). The vegetation class values were converted to density values according to Table 1, with vegetation class 4 (other) relating to the absolute minimum (residual) cattail density, representing the seed bank. The input file was created by overlaying the mesh grid of 385 triangles (510 triangles total—which includes a row of triangles along the border) on the rasterized vegetation map and calculating the mean value of all raster cell density values within each triangular element. This new aggregated map was used to create the input file. A graphical overview of this process for the data maps can be seen in Figure 2.

The final sawgrass maps are viewable in Figure 3. The maximum densities of 1,240 g/m^2 for cattail and 1,958 g/m^2 for sawgrass were reported by Miao and Sklar (1998). An overview of the parameter descriptions for the increasing levels of complexity can be found in Table 2.

Statistical analysis of simulated and monitored biomass

Besides a side-by-side visual comparison of the model output, there were three sets of statistical analysis techniques that were used to compare the model results and the raw data. These metrics, commonly used in literature for comparing both single- and multi-species patterns (Fortin and Dale 2005; Muneepeerakul et al. 2008; Convertino et al. 2009), analyzed the local, global, and autocorrelation structure of observed and modeled vegetation patterns. All metrics were accompanied by a

Table 1 Cattail class and density values for formatting data maps

Vegetation class	Cattail density value (g/m^2)	Sawgrass density value (g/m^2)
1 - High density cattail	1,000	10
2 - Medium density cattail	600	600
3 - Low density cattail	200	1,000
4 - Other	10	1,600
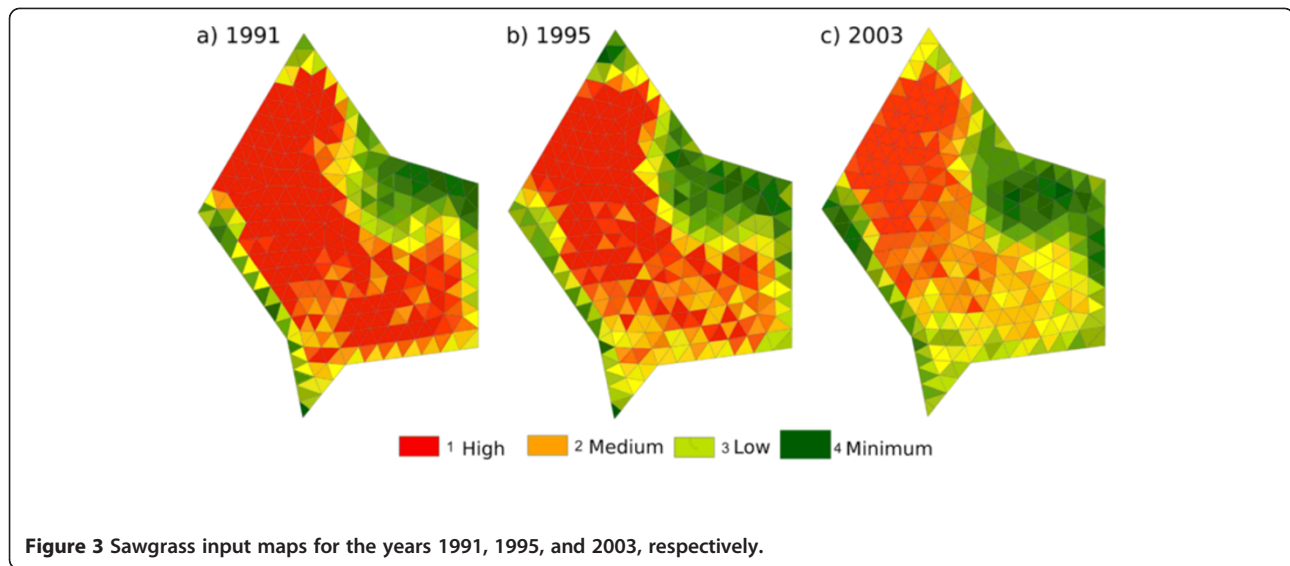
Figure 3 Sawgrass input maps for the years 1991, 1995, and 2003, respectively.

Nash-Sutcliffe coefficient (McCuen et al. 2006), represented by Equation 11, which provides a singular number for the comparison of the model statistics and how they compare to the observed data. The coefficient is a comparison of model results vwith the mean of the data.

$$E_f = 1 - \frac{\sum_{i=0}^{n}(y_i - \hat{y})^2}{\sum_{i=0}^{n}(y_i - \bar{y})^2} \qquad (11)$$

Where E_f is the Nash-Sutcliffe coefficient, $\hat{y}$ is the predicted variable, y_i is the observed variable, $\bar{y}$ is the mean of the observed variable, and n is the sample size. A Nash-Sutcliffe value of 1 means that the model completely matches the data, while a value of 0 means that the model performs no better than the mean of the data. Any value less than 0 is interpreted as a poor representation of the data.

A direct comparison between model output and the data was performed with the use of a classified difference technique (Kiker 1998). Since the data maps were initialized with a minimum density of 10 g/m^2 to account for movement between triangular elements that is not simulated in this model application, a difference between model output and the data value falling within 20 g/m^2

was considered a "perfect" match. This is loosely based on the fact that Miao and Sklar (1998) reported a roughly 10% error in measurement of the maximum density of 1,240 g/m^2. So, for example, if the data value was 10 g/m^2 (representing a typical non-cattail region), and the model output was 12 g/m^2, with a difference of 2 g/m^2 (falling within the 20 g/m^2 range), then this would be considered a "perfect" match. The next class of differences lies within the 200 g/m^2 range, which is the value assigned to the low cattail density class during the formatting and creation of the input data maps. This 200 g/m^2 range is also half the range between the successively higher cattail density classes. The third class of differences lies within 400 g/m^2, which can be thought of as a data class difference (e.g., between low and medium densities) or also as being within 40% of the maximum possible difference (the maximum data density is set as 1,000 g/m^2). Finally, any difference above the 400 g/m^2 threshold is placed in the fourth class of differences and represents a significant misrepresentation of the data by the model.

A box and whiskers plot (Ott and Longnecker 2004) was created with all model element values compared with their corresponding data element values. The desired figure is a plot with the means and ranges

Table 2 Parameter description for the increasing levels of complexity studied

Parameter	Parameter description	Levels influenced	Affected variables	Parameter equation/logic
Cattail	Cattail density	1,2,3,4,5	Cattail	Population density
CATGF	Cattail growth rate	1,2,3,4,5	Cattail	Rate of increase of population
DepthF	Water depth influence	2,3,4,5	Cattail carrying capacity, Cattail	Equation 2
phosphorusF	Soil phosphorus concentration influence	3,4,5	Cattail carrying capacity, cattail	Equation 6
Sawgrass	Sawgrass density	4,5	Sawgrass, cattail carrying capacity, cattail	Population density
SAWGF	Sawgrass growth rate	4,5	Sawgrass	Rate of increase of population

corresponding to the associated data ranges. The box and whiskers plots cover the entire range of possible values from 0 to 1,240 g/m^2.

Moran's *I* statistic (Cliff and Ord 1970; Paradis 2010) was used to determine the spatial autocorrelation between cells separated by an increasing distance. Moran's *I* is represented by Equation 12.
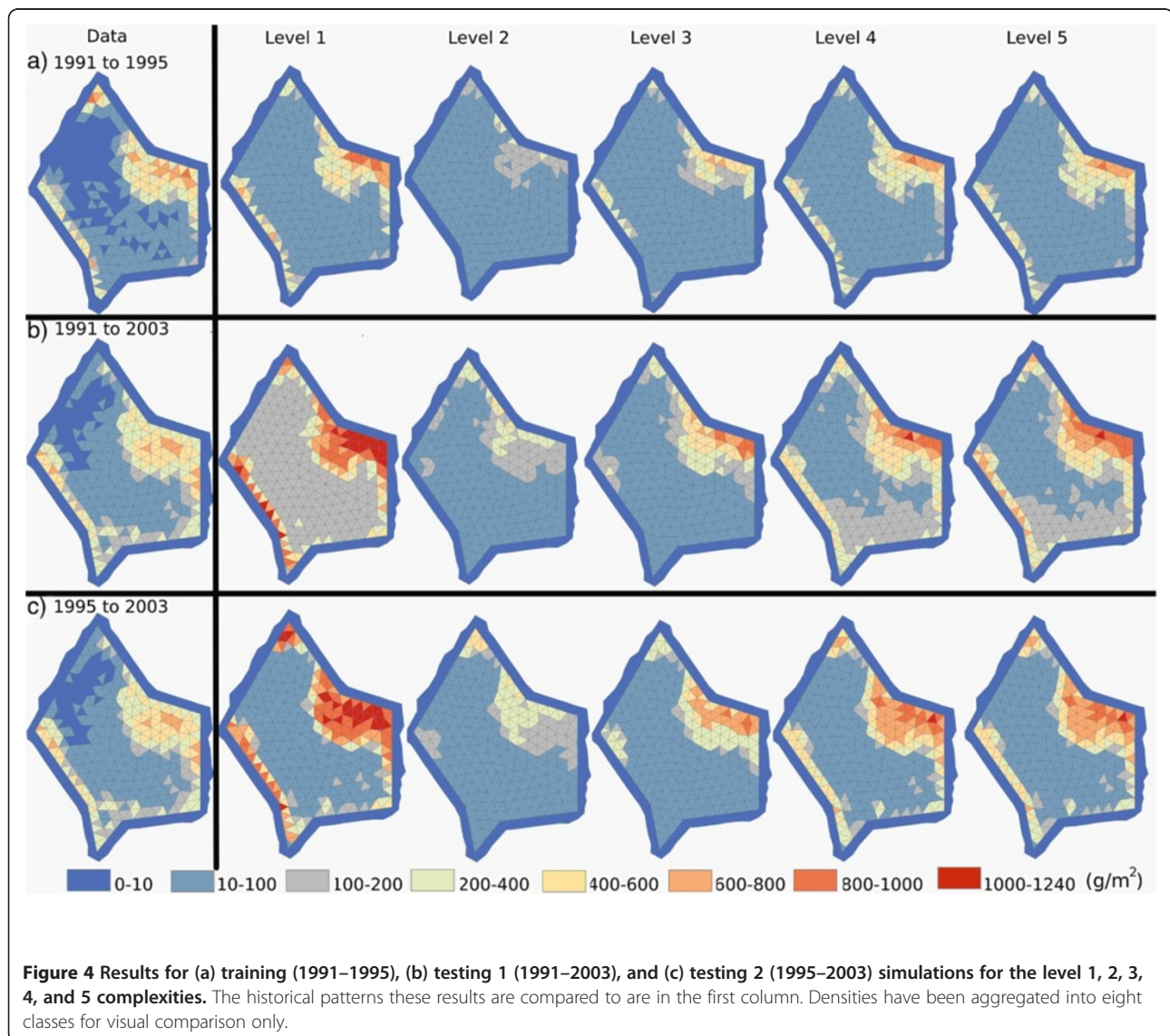
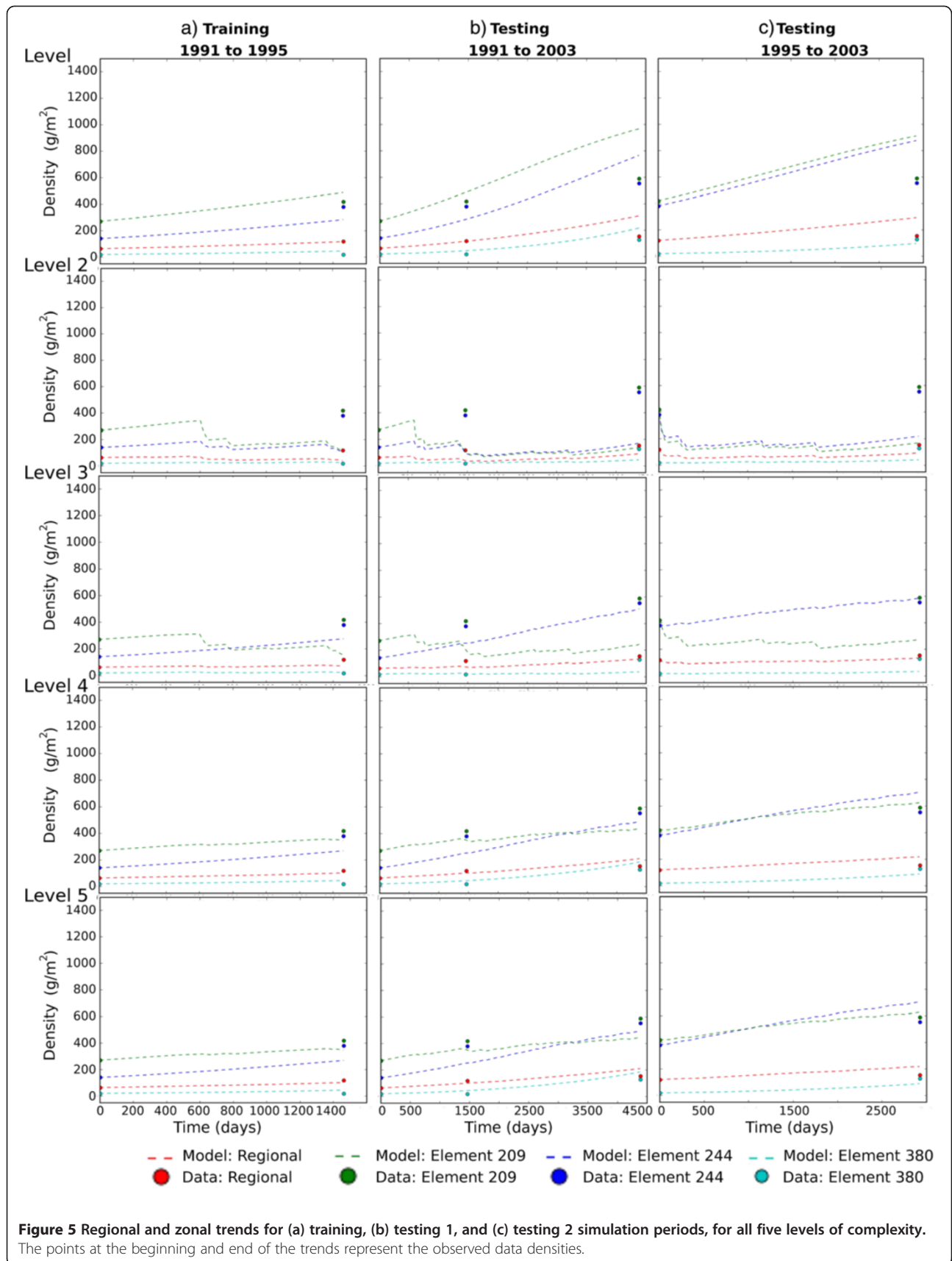$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x)(x_j - x)}{W \sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{12}$$

Where x_i is the current cell value, x_j is the value of the cell separated by a given distance, x(bar) is the mean, and W is the number of cells surrounding the current one and found within the given distance. These values are plotted against an increasing cell-pairwise distance,

as in Marani et al. (2006), to determine the trend in spatial autocorrelation across the entire region.

A landscape-scale abundance-area plot (Martin 1980; Michalski and Peres 2007) was used to measure the average change in density across the test site. One hundred randomly distributed cells are used as base cells. From these, the densities of all cells falling within a given radius are summed. This total is then divided by the number of base cells and plotted against the area of circles with an increasing radius as in Martin (1980).

A trend in the regional mean density was plotted with a daily timestep for a visual comparison of the trends between the different levels of complexity. This was repeated for the individual levels of complexity and selected zones (elements) within the region, for a more detailed view of the effect of external parameters on different areas of the region. Elements 209, 244, and 380,

Figure 4 Results for (a) training (1991–1995), (b) testing 1 (1991–2003), and (c) testing 2 (1995–2003) simulations for the level 1, 2, 3, 4, and 5 complexities. The historical patterns these results are compared to are in the first column. Densities have been aggregated into eight classes for visual comparison only.

Figure 5 Regional and zonal trends for (a) training, (b) testing 1, and (c) testing 2 simulation periods, for all five levels of complexity.
The points at the beginning and end of the trends represent the observed data densities.

located in the northeast, central, and southwest, were selected as representative elements for typically high, medium, and low cattail densities, respectively. These elements are marked by red squares in Figure 1 and are useful for evaluating local vegetation indicators.
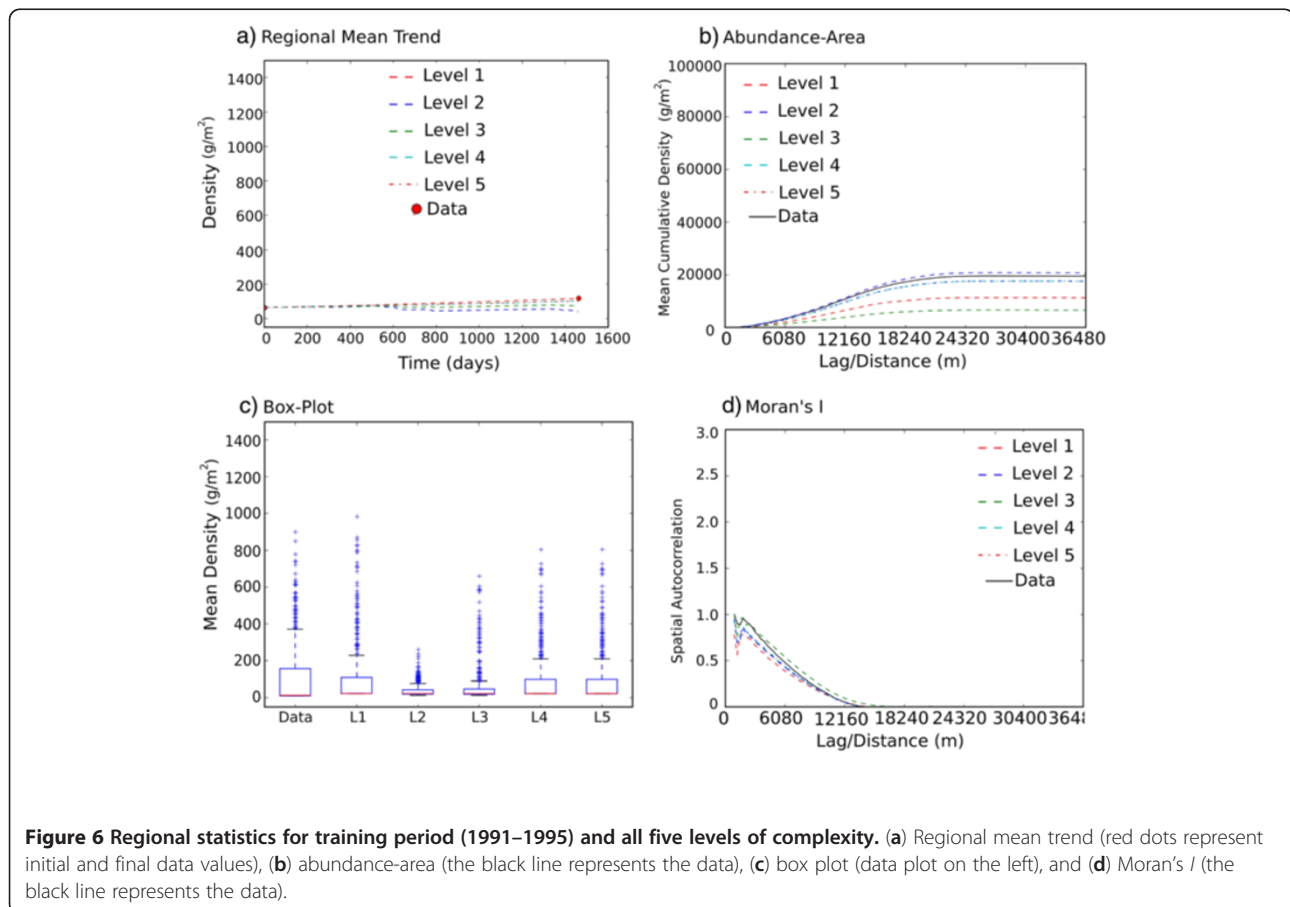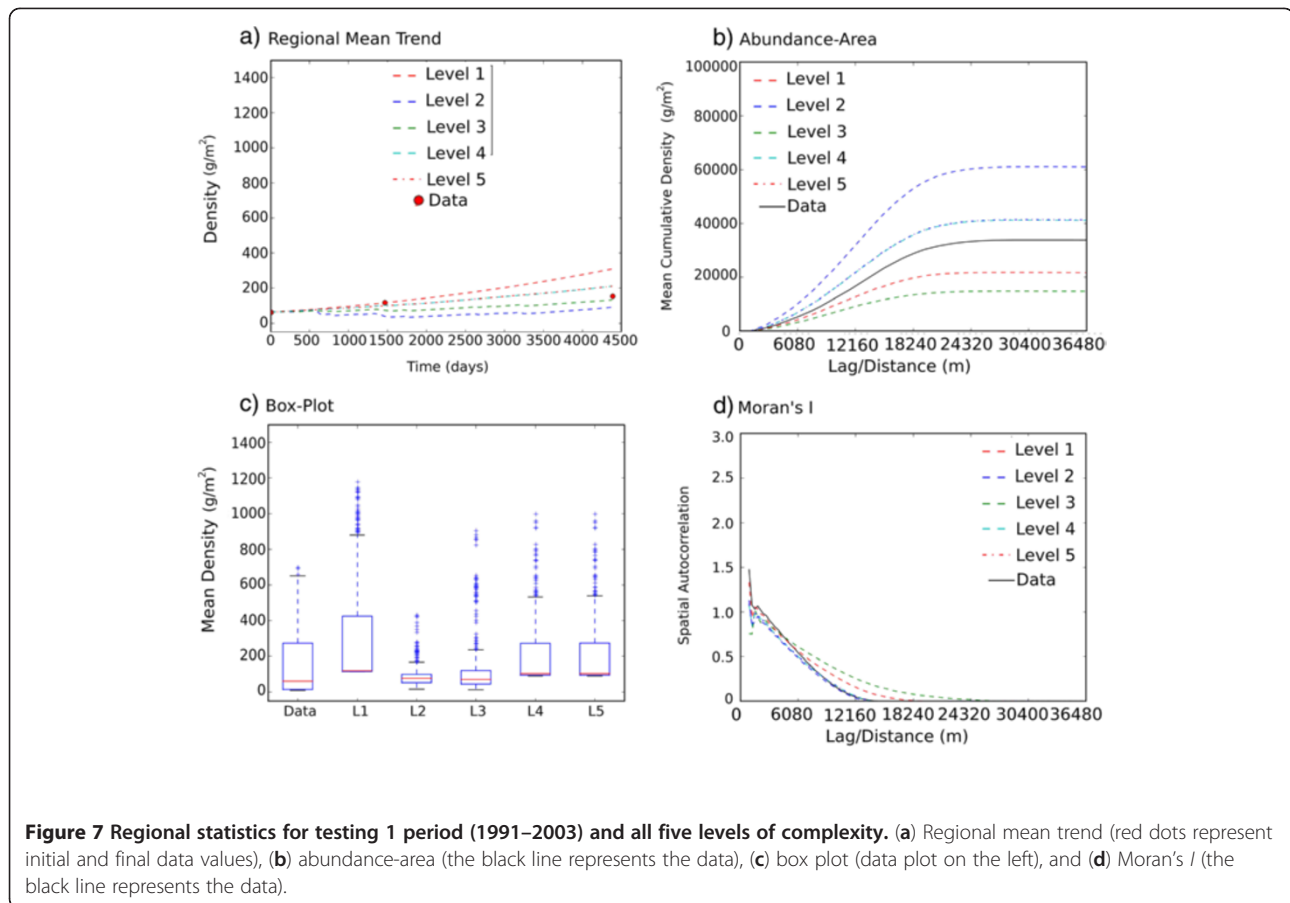
Model training and testing

There were three time periods over which the model was simulated using the available data maps of 1991, 1995, and 2003. Training was performed for the time period 1991–1995 using the level 1 complexity to establish the growth rate (6.7×10^{-9} g/g·s), and results from the other levels will be due solely to the effect of their included external parameters. It is therefore expected that the results of the other levels of complexity will not be as accurate as the level 1 complexity for this time period. Testing of the model was performed for the time period 1991–2003. This provides an extended forecast based on the original calibration time period and initial data. Finally the 1995–2003 time period was used as a blind test of the model, using different initial conditions and determining its ability to accurately predict the density distribution of the 2003 cattail map.

Results and discussion

From the cattail maps of Figure 2 and those in Rutchey et al. (2008), a trend in cattail distribution over the years is observable. It appears that cattail density and distribution increased from 1991 to 1995. From 1995 to 2003 the general distribution continued to increase but with more dispersed patches of high-density cattail. This may be related to a reduction in the overall dispersal or to an increased local speciation. Through the use of best management practices, the total phosphorus load entering WCA2A for the period 1995–2004 was reduced by roughly 36% (Richardson et al. 2008), which may have also had a role in the dispersal noted above.

The results of the simulations and analyses are displayed in Figures 4, 5, 6, 7, and 8. Figure 4 shows the model output maps for the different simulation periods, and all five levels of complexity, compared to the final data maps. These density maps have had their values aggregated into eight classes for visual comparison only. A better depiction of these trends is found in the classified difference maps of Figure 9 below. Figure 5 shows a time series plot for the five levels of complexity across all three simulation periods. It provides added insight into the trends of the model, without relying purely on
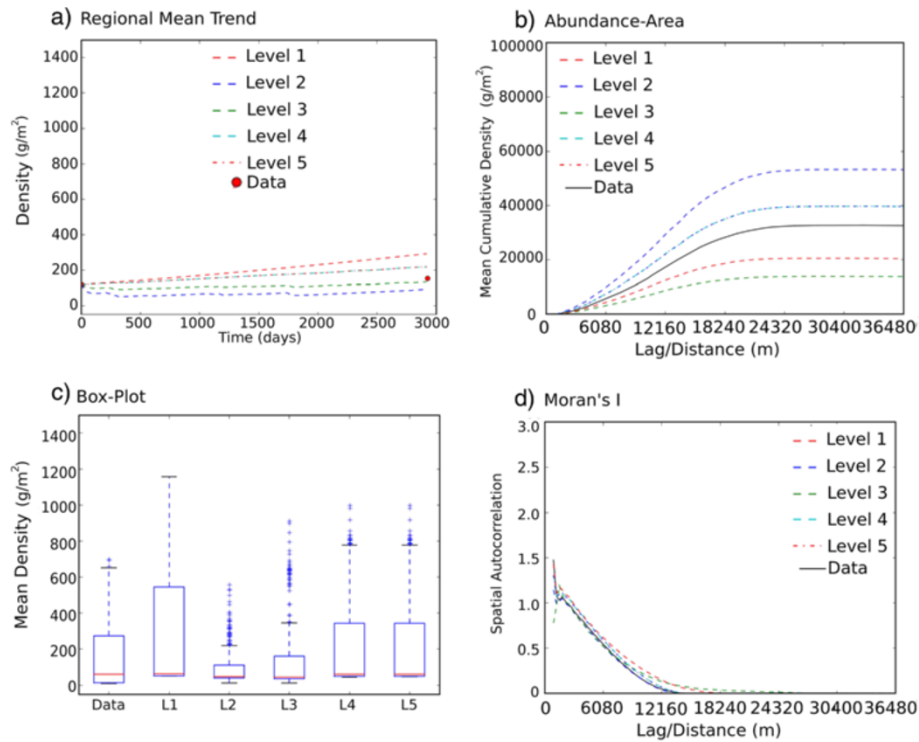

Figure 6 Regional statistics for training period (1991–1995) and all five levels of complexity. (**a**) Regional mean trend (red dots represent initial and final data values), (**b**) abundance-area (the black line represents the data), (**c**) box plot (data plot on the left), and (**d**) Moran's *I* (the black line represents the data).

Figure 7 Regional statistics for testing 1 period (1991–2003) and all five levels of complexity. (**a**) Regional mean trend (red dots represent initial and final data values), (**b**) abundance-area (the black line represents the data), (**c**) box plot (data plot on the left), and (**d**) Moran's *I* (the black line represents the data).

the end points. The plots are for the regional mean density (R), in red, and elements 209 (blue), 244 (green), and 380 (cyan). The three statistics and comparison time series for the calibration period 1991–1995 can be found in Figure 6. The regional mean time series plot for all five levels of complexity can be found in Figure 6a, the abundance-area plot in Figure 6b, the boxplot in Figure 6c enables a comparison of the spread of model densities with that of the observed data, and the Moran's *I* plot is found in Figure 6d. Figures 7 and 8 display the same three statistics and regional mean density trends as in Figure 6 for the other two simulation periods, namely 1991–2003 and 1995–2003.

When considering the first hypothesis, or level of complexity, that cattail growth is density dependent, we note the following points. For the training (1991–1995) time period, the level 1 complexity's spatial density distribution (Figures 4 and 9) is the most similar to the observed 1995 data. The density trend (Figure 5) is smooth and slowly increasing for all observed points (red dots). The regional trend ends directly on the data density. The southwest (element 380) and central (element 244) trends over-predict the data points. The

abundance-area statistic (Figure 6b) follows the data trend (black line) the closest. The mean and distribution of densities (Figure 6c) are relatively close to the data. The Moran's *I* statistic follows the data (black line) trend relatively closely (Figure 6d). All of these results from the training period are expected because this level of complexity was used for calibration over this time period. For the two testing simulation periods, the level 1 complexity clearly overestimates the historical data (Figures 4 and 9). The density trend (Figure 5b,c) remains smooth but overestimates the observed data, except for element 380 in Figure 5c which remains low, possibly due to the low initial starting density and relatively short time period. The abundance-area statistic (Figures 7b and 8b) shows significant over-prediction of the data trend (black line). The mean density is still low, but the distribution is significantly skewed toward the higher densities (Figures 7c and 8c). This is evidence that a spatial distribution of densities is more informative than simply using the mean for the area or a presence/absence type model. Moran's *I* statistic follows the data (black line) trend relatively closely (Figures 7d and 8d). The results of these analyses confirm that although
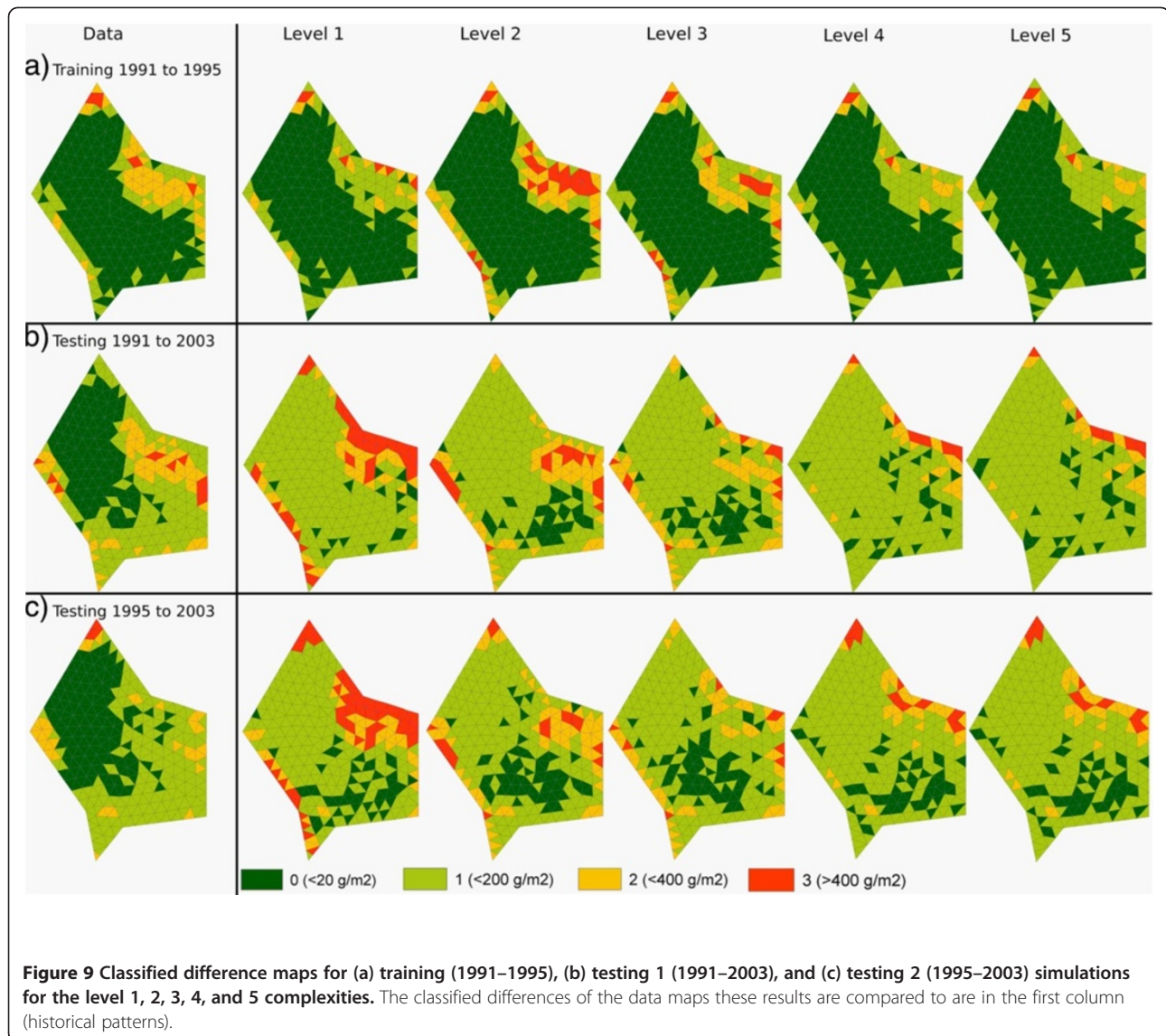
Figure 8 Regional statistics for testing 2 period (1995–2003) and all five levels of complexity. (**a**) Regional mean trend (red dots represent initial and final data values), (**b**) abundance-area (the black line represents the data), (**c**) box plot (data plot on the left), and (**d**) Moran's *I* (the black line represents the data).

cattail may indeed have a density-dependent/logistic growth pattern as we are able to simulate observed data during the training period, our inability to simulate observed data for the two training periods indicates that there are certainly other parameters affecting the growth and distribution of this species.

When considering the second hypothesis, or level of complexity, that cattail growth/expansion is dependent on water depth, we note the following points. For all time periods (training, testing 1, and testing 2), the level 2 complexity's spatial density distribution (Figures 4 and 9) is consistently lower than the observed values. This is confirmed in the trend analysis (Figure 5a,b,c), where all the observed elements (209, 244, and 380) and the regional trend are consistently below the observed values. The only exception is element 380 in Figure 5a, where there is hardly any change in the element's density, and this is possibly due to the low initial density value of that element. The abundance-area statistic for all time periods (Figures 6b, 7b, 8b) is significantly lower than the observed trend. Similarly, the distribution of densities for all time periods (Figures 6c, 7c, 8c) is much reduced. For the Moran's *I* statistic, the model is relatively close to the

data trend but consistently has a longer (the longest) tail. This implies that cells further away have an observable impact on the density of any other cell. This would be due to the fact that the water depth in every cell has an effect/influence on every other cell in the region. We know that water depth is an influential factor in cattail growth (Newman et al. 1998; Miao and Sklar 1998), however the results of these analyses indicate that the current model (level 2 complexity) is overly influenced by this parameter. It is expected that the influence of this parameter will be reduced as it is "diluted" with other parameters in the higher complexity models.

When considering the third hypothesis, or level of complexity, that cattail growth/expansion is dependent on soil phosphorus concentration, we note the following points. The spatial density distribution (Figures 4 and 9) for level 3 lies somewhat in-between that for level 1 and level 2. Except for the training period, which slightly under-predicts the observed values, the two testing periods appear to more accurately predict the observed density distribution. This is confirmed with the trend analysis (Figure 5a,b,c), where at least the regional trend is at or relatively close to the observed values. As with

Figure 9 Classified difference maps for (a) training (1991–1995), (b) testing 1 (1991–2003), and (c) testing 2 (1995–2003) simulations for the level 1, 2, 3, 4, and 5 complexities. The classified differences of the data maps these results are compared to are in the first column (historical patterns).

the level 2 complexity, element 380 tends to under-predict the observed value. However, element 209 tends to predict the observed value better than either of the previous two levels of complexity. The abundance-area statistic (Figures 6b, 7b, 8b) shows consistent under-prediction of the observed trend, but also shows consist-ently higher values than the level 2 trend and is closer to the data than the level 1 trend. The distribution of dens-ities for all time periods (Figures 6c, 7c, 8c), although greater than the level 2 complexity, is still significantly lower than the observed distribution. The Moran's *I* trend is followed closely for all time periods (Figures 6d, 7d, 8d). The results of these analyses confirm that soil phosphorus is a significant influencing factor in the dis-tribution of cattail, although the water depth parameter remains highly influential. The level 3 complexity is

better able to predict cattail in areas of typically high phosphorus or of high cattail density than the previous two levels of complexity.

When considering the fourth hypothesis, or level of complexity, that sawgrass density may impact the rate of cattail expansion, we note the following points. The spatial density distribution (Figures 4 and 9) is closer to the observed values than the previous levels of complexity. This is confirmed in the trend analysis (Figure 5 a,b,c), where most notably all of the ele-ments tend to better predict the observed values, except for element 244 in Figure 5c, which over-predicts the observed density and in turn raises the regional trend above the observed value as well. The abundance-area statistic only slightly under-predicts the observed trend during the training time period

(Figure 6b). During the two testing time periods, the statistic indicates a slight over-prediction of the observed trend, but results show better predictions than any of the previous levels of complexity. The density distribution (Figures 6c, 7c, 8c) is significantly higher than the level 2 and level 3 complexities, and equal to (Figure 6c; training) or less than (Figures 7c, 8c; testing) the level 1 complexity. This means that the level 4 complexity consistently approximates the observed densities for the region better than the other levels of complexity for all time periods, albeit with slightly elevated minimum densities. The Moran's *I* statistic (Figures 6d, 7d, 8d) follows the observed trend relatively well for all time periods. Although the

level 4 complexity tends to have slightly elevated minimum densities, like the level 1 complexity, the general result from these analyses is that the level 4 complexity is able to simulate the cattail densities through the region consistently better than any of the previous levels of complexity. We can thus conclude that including a simulated sawgrass density does indeed impact the rate of cattail expansion and improve simulation results.

When considering the fifth hypothesis, or level of complexity, that inter-species interactions between cattail and sawgrass contribute to the observed cattail dynamics, we find the following: The spatial density distribution (Figures 4 and 9) does not predict the
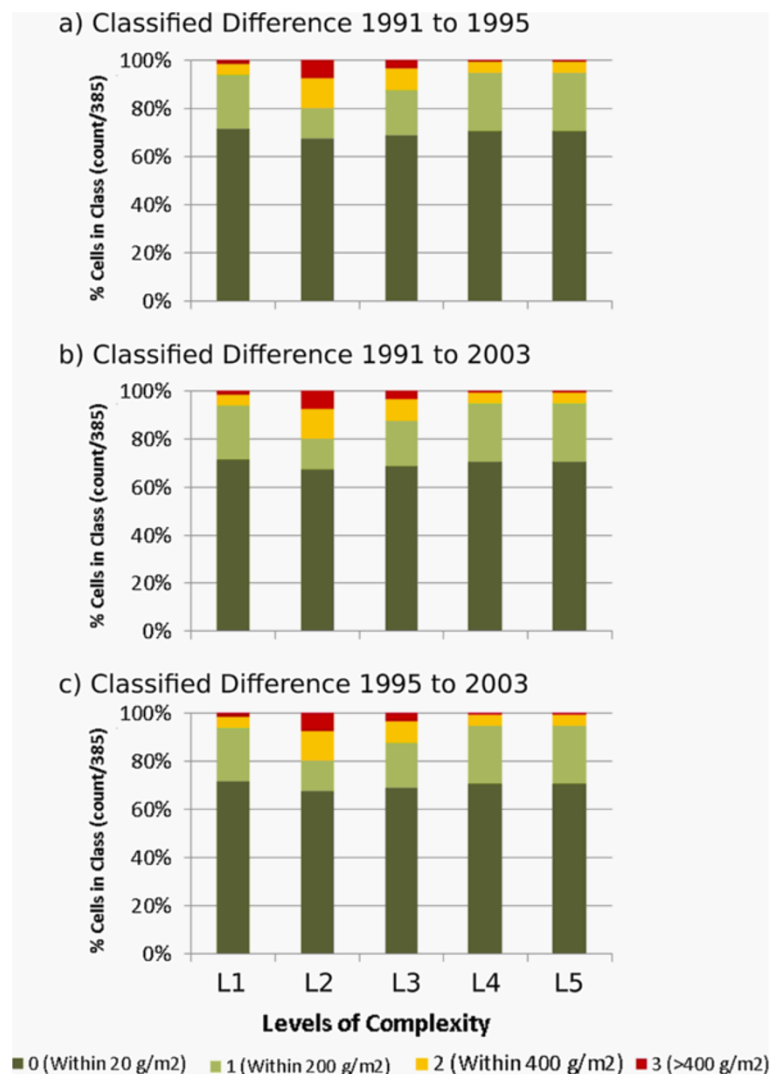
Figure 10 Classified difference summary. Percentage of cells occurring within each class, for all levels of complexity and time periods (**a**) training (1991–1995), (**b**) testing 1 (1991–2003), and (**c**) testing 2 (1995–2003).

observed values significantly better than the level 4 complexity. The trend analysis (Figure 5a,b,c) is almost identical to that of the level 4 complexity in every respect. All of the statistical analyses and distributions for all time periods (Figures 6b,c,d; 7b,c,d; 8b, c,d) are almost identical to those of the level 4 complexity. The result of these analyses is that the level 5 complexity does not predict the observed values with greater success than the level 4 complexity. While inter-species interactions might well have an effect with a different model structure, the current modeling arrangement has shown the beginning of diminishing returns with respect to model complexity and predictive capability.

With regard to the Moran's *I* statistic, all the complexity levels followed the same basic trend as the data (represented by the black line) and were all 0 by around the 18,240 m mark. This distance corresponds approximately to the width of the region, while the total distance of 36,480 m in the plot corresponds to the longest north–south distance of the region. It is believed that the statistic drops to 0 by the 18,240 m mark due to overlapping and boundary effects and that this elevates the Nash-Sutcliffe coefficient for all levels of complexity in this statistic.

A summary of the Figure 9 classified difference maps can be found in the bar chart of Figure 10, which shows the percentage of triangular elements falling within each class for all five levels of complexity and simulation periods. Upon further inspection of these plots, the level 4 and level 5 complexities consistently outperform the other levels of complexity, with either the highest percentage of combined classes 0 (< 20 g/m^2) and 1 (< 200 g/m^2), or the lowest percentage of combined classes 2 (< 400 g/m^2) and 3 (> 400 g/m^2).

A summary of the three statistics found in Figures 6b,c,d; 7b,c,d; and 8b,c,d is provided by the Nash-Sutcliffe coefficients in Table 3 and can be visually compared in Figure 11, with the box plots (or 1-to-1 comparisons) located in Figure 11a, abundance-area in Figure 11b, and Moran's *I* in Figure 11c. From Figure 11 it can be noted that the level 4 and 5 complexities, which include depth, soil phosphorus, and sawgrass interactions, consistently perform better than the other levels of complexity. A point to note regarding the level 5 complexity is that despite the fact that it does not offer a significant improvement in predictive capability over the level 4 complexity, it does not predict the observed values any worse than the level 4 complexity either.

Conclusions

The methods of modeling cattail for ecological models currently in use were compared, their similarities and differences were noted, and a knowledge gap identified: there doesn't yet exist a method of quantitatively and deterministically determining the spatial distribution of cattail in the

Table 3 Summary of Nash-Sutcliffe values comparing model and observed data for box plot, Moran's *I*, and abundance-area statistics (represented by Figures 6, 7, and 8, respectively) for level 1, level 2, level 3, level 4, level 5, training (199–1995), testing 1 (1991–2003), and testing 2 (1995–2003) simulations

Year	level	1-to-1 Box plot	Moran's *I*	Abundance
1991-1995	1	0.74	0.98	0.98
	2	0.13	0.99	−0.94
	3	0.49	0.95	0.23
	4	0.74	0.98	0.96
	5	0.74	0.98	0.96
1991-2003	1	−0.75	0.97	−1.89
	2	0.02	0.86	−0.35
	3	0.23	0.98	0.44
	4	0.49	0.98	0.77
	5	0.49	0.98	0.76
1995-2003	1	−0.95	0.99	−0.80
	2	0.14	0.94	−0.29
	3	0.36	0.97	0.51
	4	0.39	0.99	0.77
	5	0.39	0.99	0.77

Everglades. A coupled free-form/fixed-form model was introduced to solve this problem. An added benefit of the free-form nature of the RSM/TARSE coupled model is the user-definable equations of interaction, which can be modified as data and/or new theories become available. This new ecological implementation of the model (RTE) was successfully applied towards modeling cattail dynamics across the WCA2A test site for training (1991–1995), testing (1991–2003), and blind test (1995–2003) simulation periods. Five algorithms, with increasing complexity, were used to match the historical data. Upon analysis of the performance of these different levels, it can be concluded that the level 4 and 5 complexities, which include depth, soil phosphorus, and sawgrass interaction parameters, are the most suitable models for matching the historical data. The Nash-Sutcliffe coefficient was used to distinguish the success of different models.

Both local and landscape-scale indicators were used to perform the comparison between historical and modeled cattail patterns. The average local cattail density was estimated with a box-plot analysis; the pairwise-cell comparison of local cattail densities was analyzed with Moran's *I*; and, the regional increase with area of the local cattail density was estimated through the abundance-area relationship. The box-plot and the abundance-area were the most meaningful patterns to discriminate models in terms of their ability to represent the observed patterns.
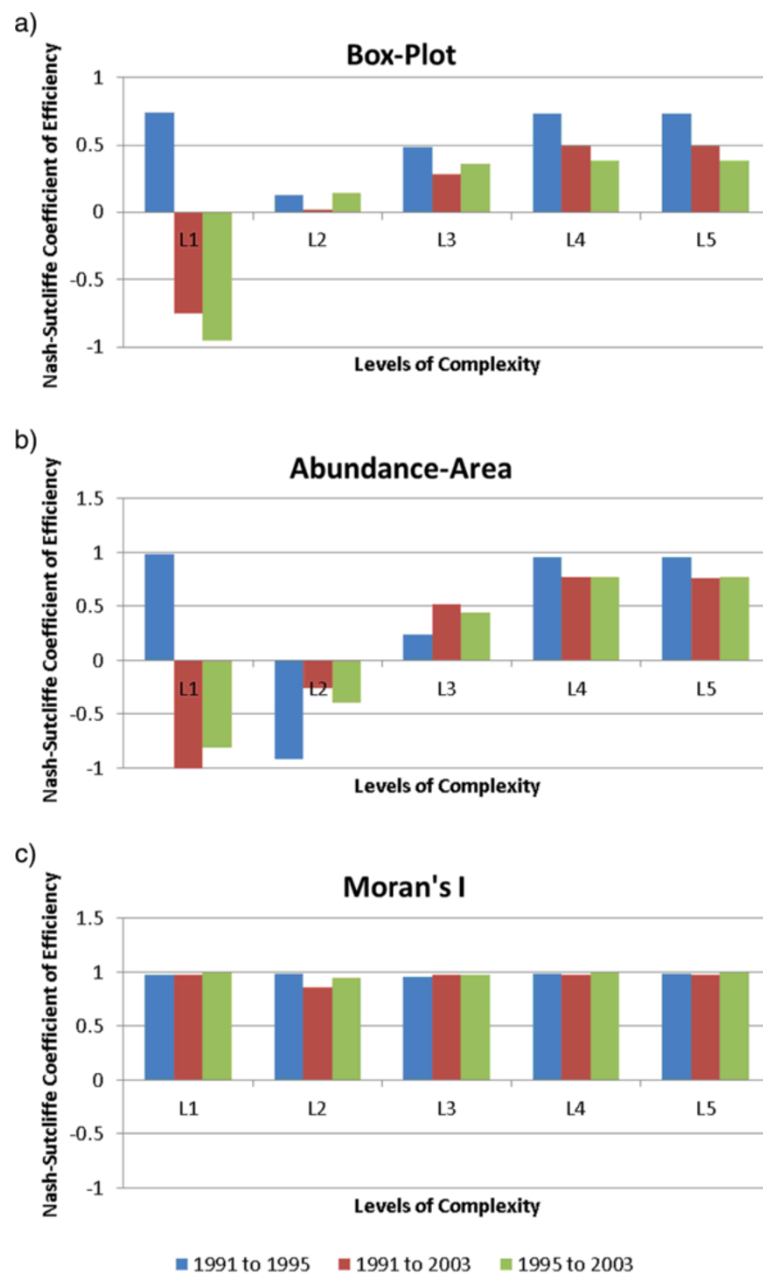
Figure 11 Nash-Sutcliffe summary of statistics. A graphical representation of Table 3. The level 4 and 5 complexity models perform consistently well in comparison with all the other models.

The autocorrelation structure of the cattail patterns were well represented by all the models at each complexity level. This is possibly due to the fact that through overlapping and boundary effects, cattail densities leveled off after roughly half the distance (top to bottom) that was used to calculate the statistic. It may be more representative if future calculations considered only half this maximum distance, where the variations would carry a greater weighting.

Our simulation results would be in agreement with the studies of Newman et al. (1998) and Miao and Sklar (1998), in which water depth and soil phosphorus concentration were the most important factors aiding in cattail expansion. Our results also include an interaction parameter with sawgrass, which is of interest in the region. Thus, we confirm the importance of considering species dependencies or interactions in reproducing the cattail patterns even

in water-controlled areas in which the anthropic-driven variables would be expected to dominate the species processes and the resulting patterns.

Limitations of our current modeling approach may include the element/triangle size, with a range of 0.5–1.7 km^2 (Wang 2009). This constraint was dictated by the choice of the RSM that simulates hydrological processes. Although the imposed grid-unit has a relatively coarse size in which there is still considerable heterogeneity of the environmental features (Zajac 2010), RTE has proven to be capable of reproducing the dynamics of cattail and sawgrass at the landscape scale using the level 4 and level 5 complexities. This makes it a valuable tool for exploring potential management scenarios in water conservation areas in the Everglades and possibly in other water-controlled wetlands.

Further investigations would consider the quantification of the importance of water-controlled drivers and species traits (dispersal) for vegetation patterns, the stability/instability states of species under varying stressors, the prediction of future management scenarios, and the comparison with neutral-based models.

In terms of further model development and added complexity, efforts have been made towards more accurate representation of fauna movement through the use of Eulerian–Lagrangian (grid-independent) particle movement (Lagerwall 2011), as well as using vegetation types/densities to influence the hydrology with a dynamically linked Manning's *n* parameter (Zajac 2010). While creating more dynamically linked parameters is an ongoing task, these linkages remain a challenge to implement due to the difficulties associated with parameterizing (training) a model with feedback effects. This feedback relationship between ecological and hydrological model components may be quite important to the function and resilience of these ecosystems and is certainly a subject of further research.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

GL conducted the majority of the research, model adaptation for ecology, and writing of the paper. GK provided ecological modeling expertise, general guidance, help in developing the five levels of complexity, paper writing, and review contributions. RMC provided statistical insights, provided critical review on model design, and ensured that the general logic of the paper was maintained. MC provided expertise in the ecological statistics and contributed to paper writing, formatting, and review. AJ provided RSM/TARSE model expertise. NW provided RSM and WCA2A expertise, supplied raw vegetation maps, and provided critical review on model design. All authors read and approved the final manuscript.

Author details

[1]Frazier Rogers Hall, University of Florida, PO Box 110570, Gainesville, FL 32611-0570, USA. [2]Soil and Water Engineering Technology, Inc., 3960 Magnolia Leaf L, SuwaneeGA 30024, USA. [3]Hydrologic and Environmental Systems Modeling, South Florida Water Management District, 3301 Gun Club Rd, West Palm Beach, FL 33406, USA.

References

Arnold K, Gosling J (1998) The Java programming language, 2nd edn. Prentice Hall, Upper Saddle River, NJ

Cary JR, Shasharina SG, Cummings JC, Reynders JVW, Hinker PJ (1998) Comparison of C++ and Fortran 90 for object-oriented scientific programming. Comp Phys Comm 105:20–36

Cliff AD, Ord K (1970) Spatial autocorrelation: a review of existing and new measures with applications. Econ Geography 46:269–292

Convertino M, Muneepeerakul R, Azaele S, Bertuzzo E, Rinaldo A, Rodriguez-Iturbe I (2009) On neutral metacommunity patterns of river basins at different scales of aggregation. Water Resour Res 45:W08424

Costanza R, Voinov A (2001) Modeling ecological and economic systems with STELLA: part III. Ecol Model 143:1–7

DeBusk WF, Reddy KR, Koch MS, Wang Y (1994) Spatial distribution of soil nutrients in a northern-Everglades marsh: Water Conservation Area 2A. Soil Soc Am 58:543–552

Doren RF, Armentano Thomas V, Whiteaker Louis D, Jones Ronald D (1999) Marsh vegetation patterns and soil phosphorus gradients in the Everglades ecosystem. Aqua Bot 56:145–163

Douglas MS (1947) The Everglades: river of grass. Rinehart, New York

Duke-Sylvester S (2005) Initial performance measures and information related to the ATLSS vegetation succession model. http://atlss.org/VSMod. Accessed 31 July 2010

ESRI (Environmental Systems Resource Institute) (2010) ArcMap 10.0. ESRI, Redlands, CA

Fitz CH, Trimble B (2006a) Documentation of the Everglades Landscape Model: ELM v2.5. South Florida Water Management District, West Palm Beach, FL

Fitz CH, Trimble B (2006b) Everglades Landscape Model (ELM). http://my.sfwmd.gov/portal/page/portal/xweb%20-%20release%202/elm. Accessed 31 July 2010

Fitz HC, Kiker GA, Kim JB (2011) Integrated ecological modeling and decision analysis within the Everglades landscape. Crit Rev Environ Sci Technol 41 (S1):517–547

Fortin MJ, Dale MRT (2005) Spatial analysis, a guide for ecologists. Cambridge University Press, Cambridge

Grace JBL (1989) Effects of water depth on *Typha latifolia* and *Typha domingensis*. Am J Bot 76:762–768

Gross LJ (1996) ATLSS home page. http://atlss.org/. Accessed 31 July 2010

Grunwald S (2010) Phosphorus data for WCA2A. Personal Communication. University of Florida, Gainesville

Grunwald S, Reddy KR, Newman S, DeBusk WF (2004) Spatial variability, distribution and uncertainty assessment of soil phosphorus in a South Florida wetland. Environmetrics 15:811–825

Grunwald S, Ozborne TZ, Reddy KR (2008) Temporal trajectories of phosphorus and pedo-patterns mapped in Water Conservation Area 2, Everglades, Florida, USA. Geoderma 146:1–13

Guardo M, Fink L, Fontaine Thomas D, Newman S, Chimney M, Bearzotti R, Goforth G (1995) Large-scale constructed wetlands for nutrient removal from stormwater runoff: an Everglades restoration project. Environ Manage 19 (6):879–889

Harold ER (1998) XML: Extensible Markup Language, 1st edn. IDG, Foster City

James AI, Jawitz JW (2007) Modeling two-dimensional reactive transport using a Godunov-mixed finite element method. J Hydrol 338:28–41

Jawitz JW, Muñoz-Carpena R, Muller S, Grace KA, James AI (2008) Development, testing, and sensitivity and uncertainty analyses of a Transport and Reaction Simulation Engine (TaRSE) for spatially distributed modeling of phosphorus in South Florida peat marsh wetlands. Scientific Investigations Report 2008–5029. United States Geological Survey, Reston, VA

Jensen JR, Rutchey K, Koch MS, Narumalani S (1995) Inland wetland change detection in the Everglades Water Conservation Area 2A using a time series of remotely sensed data. Photogramm Eng Rem Sens 61(2):199–209

Keen RE, Spain JD (1992) Computer simulation in biology. Wiley-Liss, New York

Kiker GA (1998) Development and comparison of savanna ecosystem models to explore the concept of carrying capacity. PhD Dissertation. Cornell University, Ithaca

Kiker, G. A. & Linkov, I. 2006. The QnD Model/Game System: Integrating Questions and Decisions for Multiple Stressors. pp. 203-225 in Arapis, G., Goncharova, N. & Baveye, P. Ecotoxicology, Ecological Risk Assessment and Multiple Stressors. Netherlands: Springer. (1-4020-4475-5)

Kiker, G. A., Rivers-Moore, N. A., Kiker, M. K. & Linkov, I. 2006. QnD: A modeling game system for integrating environmental processes and practical management decisions. pp. 151-185 in Morel, B. & Linkov, I. Environmental Security and Environmental Management: The Role of Risk Assessment. Netherlands: Springer. (1-4020-3892-5)

Lagerwall GL (2011) Modeling *Typha domingensis* in an Everglades wetland. Dissertation. University of Florida, Gainesville

Lindenschmidt KE (2006) The effect of complexity on parameter sensitivity and model uncertainty in river water quality modeling. Ecol Model 190:72–86

Ludascher B, Altintas I, Berkley C, Higgins D, Jaeger E, Jones M, Lee Edward A, Tao J, Zhao Y (2006) Scientific workflow management and the Kepler system. Concurr Comp Pract Exper 18:1039–1065

Marani M, Tommaso Z, Belluco E, Silvestri S, Maritan A (2006) Non-neutral vegetation dynamics. PLoS One 1(1):e78

Martin TE (1980) Diversity and abundance of spring migratory birds using habitat islands on the Great Plains. Cooper Ornithol Soc 82:430–439

McCuen RH, Knight Z, Cutter AG (2006) Evaluation of the Nash-Sutcliffe Efficiency Index. Hydrol Eng 11:597–602

Miao S (2004) Rhizome growth and nutrient resorption: mechanisms underlying the replacement of two clonal species in Florida Everglades. Aquat Bot 78:55–66

Miao SL, Sklar FH (1998) Biomass and nutrient allocation of sawgrass and cattail along a nutrient gradient in the Florida Everglades. Wetlands Ecol Manage 5:245–264

Michalski F, Peres CA (2007) Disturbance-mediated mammal persistence and abundance-area relationships in Amazonian forest fragments. Conserv Biol 21:1626–1640

Muller S (2010) Adaptive spatially-distributed water-quality modeling: an application to mechanistically simulate phosphorus conditions in the variable-density surface-waters of coastal Everglades wetlands. PhD Dissertation. University of Florida, Gainesville

Muneepeerakul R, Bertuzzo E, Lynch HJ, Fagan WF, Rinaldo A, Rodriguez-Iturbe I (2008) Neutral metacommunity models predict fish disversity patterns in Mississippi-Missouri basin. Nature 453:220–222

Muñoz-Carpena R, Parsons JE, Gilliam JW (1999) Modeling hydrology and sediment transport in vegetative filter strips. J Hydrol 214:111–129

Newman S, Schutte J, Grace J, Rutchey K, Fontaine T, Reddy K, Pietrucha M (1998) Factors influencing cattail abundance in the northern Everglades. Aquat Bot 60:265–280

Odum HT, Odum EC, Brown MT (2000) Wetlands management. In: Environment and society in Florida. CRC Press, Boca Raton

Ott RL, Longnecker MT (2004) A first course in statistical methods. Curt Hinrichs, Belmont, CA

Paradis E (2010) Moran's autocorrelation coefficient in comparative methods. http://cran.r-project.org/web/packages/ape/vignettes/MoranI.pdf. Accessed 7 August 2010

Perez-Ovilla O (2010) Modeling runoff pollutant dynamics through vegetative filter strips: a flexible numerical approach. PhD Dissertation. University of Florida, Gainesville

Richardson CJ, King Ryan S, Vymazal J, Romanowicz Edwin A, Pahl James W (2008) Macrophyte community responses in the Everglades with an emphasis on cattail (*Typha domingensis*) and sawgrass (*Cladium jamaicense*) interactions along a gradient of long-term nutrient additions, altered hydroperiod, and fire. Ecol Stud 201:215–260

Rivero RG, Grunwald S, Bruland GL (2007a) Incorporation of spectral data into multivariate geostatistical models to map soil phosphorus variability in a Florida wetland. Geoderma 140:428–443

Rivero RG, Grunwald S, Osborne TZ, Reddy KR, Newman S (2007b) Characterization of the spatial distribution of soil properties in Water Conservation Area 2A, Everglades, Florida. Soil Sci 172:149–166

Rutchey K (2011) *Typha domingensis* maps of WCA2A for the years 1991 and 1995. Personal communication. South Florida Water Management District, West Palm Beach

Rutchey K, Schall T, Sklar F (2008) Development of vegetation maps for assessing Everglades restoration progress. Wetlands 172(2):806–816

SFWMD (1995) Land cover land use 1995. http://my.sfwmd.gov/gisapps/sfwmdxwebdc/dataview.asp?query=unq_id=297. Accessed 11 November 2009

SFWMD (1999) Land cover land use 1999. http://my.sfwmd.gov/gisapps/sfwmdxwebdc/dataview.asp?query=unq_id=1593. Accessed 11 November 2009

SFWMD (2005a) Documentation of the South Florida Water Management Model version 5.5. South Florida Water Management District, West Palm Beach, FL

SFWMD (2005b) Regional Simulation Model (RSM) Hydrologic Simulation Engine (HSE) user's manual. South Florida Water Management District, West Palm Beach, FL

SFWMD (2005c) Regional Simulation Model (RSM) theory manual. South Florida Water Management District, West Palm Beach, FL

SFWMD (2008a) RSM water quality user manual (draft). South Florida Water Management District, West Palm Beach, FL

SFWMD (2008b) RSMWQE theory manual (draft). South Florida Water Management District, West Palm Beach, FL

SFWMD (2008c) WCA2A HSE setup. South Florida Water Management District, West Palm Beach, FL

SFWMD (2009) DBHYDRO. http://my.sfwmd.gov/dbhydroplsql/show_dbkey_info.main_menu. Accessed 04 August 2010

Stroustrup B (2000) The C++ programming language, specialth edn. Addison-Wesley, Westford, MA

Tarboton KC, Irizarry-Ortiz MM, Loucks DP, Davis SM, Obeysekera JT (2004) Habitat suitability indices for evaluating water management alternatives. South Florida Water Management District, West Palm Beach, FL

Urban NH, Davis SM, Aumen NG (1993) Fluctuations in sawgrass and cattail densities in Everglades Water Conservation Area 2A under varying nutrient, hydrologic, and fire regimes. Aquat Bot 46:203–223

USACE, S.F.R.O (2010a) CERP: The plan in depth - part 1. http://www.evergladesplan.org/about/rest_plan_pt_01.aspx. Accessed 3 August 2010

USACE, S.F.R.O (2010b) CERP: The plan in depth - part 2. http://www.evergladesplan.org/about/rest_plan_pt_02.aspx. Accessed 3 August 2010

van der Valk AG, Rosburg TR (1997) Seed bank composition along a phosphorus gradient in the northern Florida Everglades. Wetlands 17(2):228–236

Walker WW, Kadlec RH (1996) A model for simulating phosphorus concentrations in waters and soils downstream of Everglades stormwater treatment areas. Draft. US Department of the Interior Everglades National Park, Homestead, FL, http://publicfiles.dep.state.fl.us/DEAR/GoldAdministrativeRecord/Item%2027/018752.PDF

Wang N (2009) 2003 Vegetation map; dss hydrology input files. Personal communication. South Florida Water Management District, West Palm Beach, FL

Wang JD, Swain ED, Wolfert MA, Langevin CD, James DE, Telis PA (2007) Application of FTLOADDS to simulate flow, salinity, and surface-water stage in the southern Everglades, Florida. Scientific Investigations Report 2007–2010. United States Geological Survey, Florida

Wetzel PR (2001) Plant community parameter estimates and documentation for the Across Trophic Level System Simulation (ATLSS). East Tennessee State University, Johnson City

Wetzel PR (2003) Nutrient and fire disturbance and model evaluation documentation for the Actoss Trophic level System Simulation (ATLSS). East Tennessee State University, Johnson City

Willard DA (2010) SOFIA - FS-146-96. http://sofia.usgs.gov/publications/fs/146-96/. Accessed 3 August 2010

Wu Y, Sklar FH, Rutchey K (1997) Analysis and simulation of fragmentation patterns in the Everglades. Ecol Appl 7(1):268–276

Zajac ZB (2010) Global sensitivity and uncertainty analysis of spatially distributed watershed models. PhD Dissertation. University of Florida, Gainesville