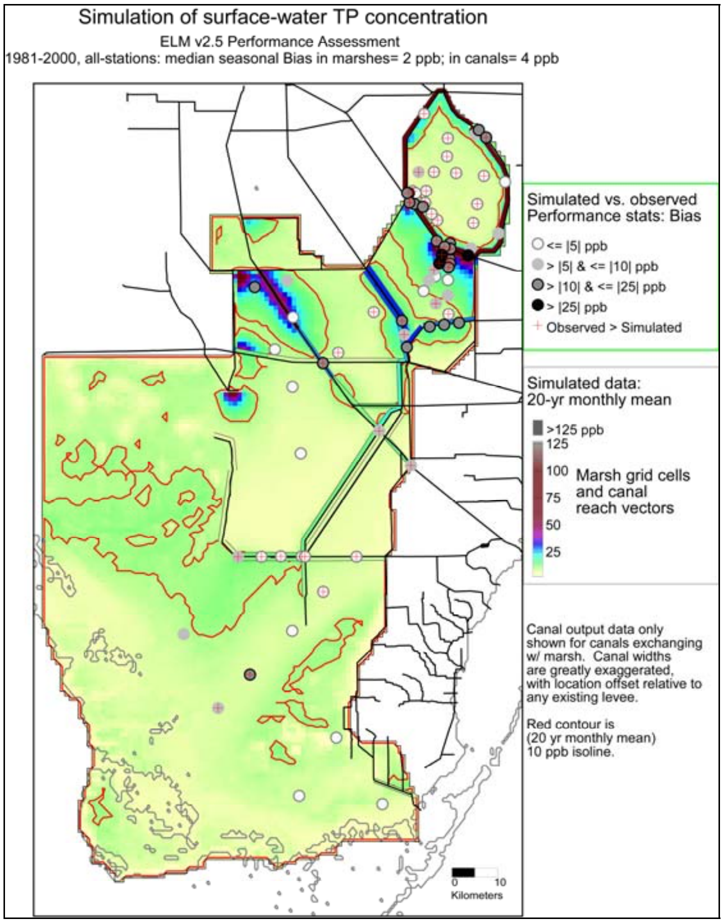
Documentation of the Everglades Landscape Model: ELM v2.5

Chapter 6: Model Performance



Simulation of surface-water TP concentration

ELM v2.5 Performance Assessment

1981-2000, all-stations: median seasonal Bias in marshes= 2 ppb; in canals= 4 ppb

Canal output data only shown for canals exchanging w/ marsh. Canal widths are greatly exaggerated, with location offset relative to any existing levee.

Red contour is (20 yr monthly mean) 10 ppb isoline.

http://my.sfwmd.gov/elm

July 10, 2006

Chapter 6: Model Performance

Chapter 6: Model Performance 6-1
- 6.1 Overview 6-2
- 6.2 Performance expectations 6-3
  - 6.2.1 Model application niches 6-3
  - 6.2.2 ELM v2.5 application niche 6-3
  - 6.2.3 Establishing performance expectations 6-3
- 6.3 Performance evaluation methods 6-5
  - 6.3.1 Calibration process 6-5
  - 6.3.2 Validation process 6-10
  - 6.3.3 Performance evaluation methods 6-11
- 6.4 Model updates 6-14
- 6.5 Model configuration 6-15
- 6.6 Performance results 6-15
  - 6.6.1 Ecological performance 6-15
  - 6.6.2 Hydrologic performance 6-29
  - 6.6.3 Ecological consistency 6-45
  - 6.6.4 Validation 6-48
- 6.7 Discussion 6-51
  - 6.7.1 Model performance summary 6-51
  - 6.7.2 Uncertainty & expectations 6-52
  - 6.7.3 Performance refinements 6-52
  - 6.7.4 Conclusions 6-53
- 6.8 Literature cited 6-54
- 6.9 Appendix A: Computational methods for statistics 6-56
- 6.10 Appendix B: Time series & CFDs: TP (separate pdf) 6-58
- 6.11 Appendix C: Time series & CFDs: stage (separate pdf) 6-137
- 6.12 Appendix D: Water budgets, ELM & SFWMM 6-220
- 6.13 Appendix E: Time series & CFDs: CL (separate pdf) 6-231

6.1 Overview

As described in the Introduction Chapter of this documentation, an overarching Goal of the ELM is to understand and predict ecological dynamics across the greater Everglades landscape. For the current ELM v2.5, the specific Objectives of the model application are those of the two ecological Performance Measures that involve the "water quality" aspect of ecosystem dynamics across the landscape: 1) surface water phosphorus concentration, and 2) accumulation (net load) of phosphorus in the ecosystem.

The overall approach of (developing and) calibrating the ELM was to start by simplifying the complex Everglades ecosystems by processes and by space. Generally, this involved first considering the most important ecosystem "drivers" within a simplified spatial domain. The hydrology and water quality drivers were evaluated using a variety of statistical and visualization methods. Hydrologic performance was generally evaluated and calibrated first, followed by water quality and its associated ecosystem dynamics. A stepwise, hierarchical process followed, evaluating each module of the total system behavior. In this context of the fully integrated ELM, specific aspects of water column phosphorus calibration are required to be associated with reasonable behavior in other ecosystem properties. The best model parameter set becomes that which provides acceptable performance of the primary model application Performance Measures, while maintaining other ecosystem dynamics that are, at minimum, consistent with our best understanding of the Everglades.

In its regional (~10,000 km^2) application at 1 km^2 grid resolution, the current ELM version 2.5 is available to assess relative differences in ecological performance of Everglades water management plans - at decadal time scales. Hydrologic performance of the ELM is comparable to the South Florida Water Management Model within the Everglades. While consistency with that primary tool for Everglades water management is important, the focus of ELM is on the associated ecological assessment. Extensive data are available for calibrating-validating surface water phosphorus (P) concentrations; during a 2-decade period, the model had a 2 ug L^{-1} (ppb) median bias in predictions of that Performance Measure within the marshes and canals. Predicted P accumulation along a multiple- decade eutrophication gradient showed a high degree of concordance with P accumulation estimates from radionuclide markers. With other predicted ecological attributes and rates being consistent with available observations, there is cumulative, strong evidence of model skill in predicting phosphorus trends in the regional Everglades landscape at the relevant decadal time scales.

6.2 Performance expectations

6.2.1 Model application niches

For model users and stakeholders, a fundamental concern is simply: how well does the model work? To be useful, it is critical that model goals and objectives are clearly stated, and that the design and performance of the model is shown to meet those goals. Towards this end, it is critical that a model is understood within the context of its "application niche" (as discussed by D.P. Loucks[1]). The application niche should be a juxtaposition of A) the real or perceived needs of the "users" and B) the realistic capabilities portrayed by the model developers. The intersection of A & B is the intended target of the model application – a basic point that is sometimes lost in practice as a result of inadequate communication.

6.2.2 ELM v2.5 application niche

The ELM application niche is broadly defined in the Introduction Chapter of this documentation, is specified in detail in the Model Application Chapter, and demonstrated in practice in this Model Performance Chapter. The model Performance Measures to be used in comparing relative benefits of alternative management plans define the specific Objectives of the model, including the spatio-temporal scale of application. While there are requests (and expectations) for ELM to address a larger suite of ecological questions, the relatively narrower subset of *current* model Objectives defined by the Model Developers should be considered to be the *current* application niche of the ELM. It is this application niche that is to be considered when evaluating the ELM.

As described in the Introduction Chapter of this documentation, an overarching Goal of the ELM is to understand and predict ecological dynamics across the greater Everglades landscape. For the current ELM v2.5, we emphasize that the available ecological Performance Measures are those involving the "water quality" aspect of ecosystem dynamics across the landscape: 1) surface water phosphorus concentration, and 2) net accumulation of phosphorus in the ecosystem.

6.2.3 Establishing performance expectations

6.2.3.1 ELM

The expectations of hydrologic simulations in the Everglades are reasonably well-understood by most users. Perhaps this is largely due to the context of hydrologic modeling in south Florida, which has a multi-decadal history of applications, with a relatively well monitored system in which the physics are reasonably well understood.

---

[1] D.P. Loucks of Cornell University made a variety of recommendations on modeling and peer review to the South Florida Water Management District in: Loucks, D. P. 2003. Modeling and Peer Review Protocols for Use in HSM (OOM) and IMC for CERP and RECOVER. Report to SFWMD, West Palm Beach.

There is less of a common understanding of the expected performance of regional Everglades models that simulate ecological (including water quality) dynamics. Nutrients are subject to many more processes (such as uptake by plants, release by soils, etc) than are water depths. Moreover, there is about an order of magnitude fewer observed data available relative to hydrologic data (in the Everglades): the quantity of water flowing into a basin may be reasonably well-known on a daily basis, but the associated nutrients are generally sampled less than 5 - 10% of that time (see the Data and the Uncertainty Chapters). Observations in the marsh, used to compare to the model output, can be even less frequent than those input data. This combination of very infrequent data collections in the Everglades, along with highly-variable, random processes, necessitates the more complex assumptions for any water quality or ecological model relative to those involving physical hydrology.

6.2.3.2 Other Everglades hydrologic models

The South Florida Water Management Model (SFWMM, sometimes referred to as the "2x2") is the primary tool used to evaluate managed hydrology in the south Florida landscape, including the greater Everglades region. This model was used to evaluate relative hydrologic benefits under different water management alternatives for the Comprehensive Everglades Restoration Plan (USACE and SFWMD 1999), in addition to a wide variety of other planning applications. The two-mile by two-mile square (~10.4 km^2) grid of the SFWMM has a relative accuracy in predicting stage that has been well-accepted for evaluating water management alternatives for the greater Everglades and much of south Florida in general. The documentation for the SFWMM v5.5 is available at:

http://www.sfwmd.gov/org/pld/sfwmm_doc/menu.htm

which includes statistical evaluations of the model performance in predicting stage in the greater Everglades. For the 82 marsh stage monitoring locations common to the ELM domain, the statistical comparisons of SFWMM daily output data to daily observed data indicated very good performance, as indicated by the median values for each statistic: $R^2 = 0.81$, Nash-Sutcliffe Efficiency = 0.67, Root Mean Square Error = 0.12 m, and Bias = 0.0 m. The computational methods used in these statistics are the same as those defined later in this chapter.

As a "second generation" simulator of managed hydrology in south Florida, the South Florida Regional Simulation Model (SFRSM,

http://www.sfwmd.gov/site/index.php?id=342)

is designed to have significantly increased flexibility and model performance relative to the current SFWMM. While portions of the SFRSM are still under development, its advanced design, and the very good performance of early prototypes, indicate that it will provide significant improvements as a replacement for the SFWMM in the future.

6.2.3.3 Other Everglades water quality models

A modeling effort that was accepted to evaluate water quality throughout most of the Everglades region is the Everglades Water Quality Model (EWQM). The EWQM was used in evaluating phosphorus surface water quality under different water management alternatives for the Comprehensive Everglades Restoration Plan (CERP) (USACE and SFWMD 1999). Raghunathan et al. (2001) presented evidence that the model was reasonably well calibrated relative to its objectives of predicting phosphorus transport and fate under different strategies of reducing phosphorus inputs across this large region. Specific statistics were provided in a referenced report (Limno-Tech 1997), which showed that (during the 1979 to 1989 simulation period) the mean observed vs. predicted phosphorus concentrations within most of the hydrologic basins differed by 6 – 23 ug l^{-1}, while one basin (WCA-1) exhibited differences >100 ug l^{-1}. The presented range of spatial and temporal variations in modeled phosphorus accumulation rates within WCA-2A usually overlapped the point estimates of measured phosphorus accumulation rates. As a tool for making relative comparisons of project alternatives within most Everglades basins, the model was judged acceptable for CERP planning purposes. However, refinement of the model was discontinued, and it is no longer available.

6.3 Performance evaluation methods

The methods of evaluating and improving the performance of a distributed, integrated ecological model are wide ranging, usually involving both analytic tools and science-based judgments. Ultimately one seeks to communicate the cumulative evidence of how well the model meets its objectives: an evaluation of the model performance in history-matching is a fundamental component of that communication. Here we attempt to summarize the methods that we used in evaluating the ELM performance.

6.3.1 Calibration process

Definitions abound, but a reasonably concise definition of the calibration of distributed simulation models is “the adjustment of model parameters in order to improve the match between simulated and observed spatio-temporal dynamics”. Improving this history-match for a model, however, involves much more than parameter adjustments. Model performance is the net result of multiple model development & refinement decisions, including the selection of algorithms and their aggregation, the influence of initial & dynamic boundary conditions, and the understanding and accounting for the wide range of other uncertainties associated with models (e.g., see Uncertainty Chapter). In this methodological summary, we do not attempt to characterize the past decade of ELM refinement and calibration, with performance improvements as our understanding (i.e., data) of the landscape advanced. Rather, we generically summarize how to take advantage of the basic design of the model to evaluate the model performance, and improve the history-match via selective adjustment of the most sensitive, or important, parameters.

Thus, this methodological section does not explicitly describe the interplay between research and modeling, nor the decisions made in improving algorithms or in data synthesis. The rationale for, and results of, those critical modeling decisions are described in the Data, the Model Structure, and the Uncertainty Chapters (each including references to associated publications). Given an “acceptable” assemblage of model code and boundary condition data, the basic steps in parameter adjustment to best meet the ELM goals and objectives are summarized here.

6.3.1.1 Parameter optimization

Parameter optimization is optionally part of the process of calibrating models. Towards this end, automated parameter optimization procedures are rapidly becoming an integral component of calibrating simulation models, most notably for physically based groundwater and other hydrologic models. Although we have recently explored methods for parameter optimization in integrated ecological models (Villa et al. 2004), we have not yet utilized formal, automated optimization methods. One conceptual constraint has been the development of objective functions (for the targeted behaviors) that incorporate the non-linear spatial and temporal interactions among multiple variables. Nevertheless, for optimizing specific (e.g., hydrologic) variables in a model such as the ELM, there may be increasing feasibility in using newer parameter optimization methods to improve model performance. At this point, in lieu of automated calibration procedures, we employ "manual" calibration methods in a hierarchical, or stepwise, process of increasing complexity associated with the modeled processes and spatio-temporal scales.

6.3.1.2 Calibrating integrated ecosystem models

The ELM simulation involves the dynamic spatio-temporal interaction among a suite of fundamental ecosystem variables and processes. As discussed in the earlier Chapter on "Ecological Models: Wetlands", the number of interacting model processes increases the complexity of this modeling approach. An integrated ecosystem design, however, can lessen the degree to which the model is dependent upon historical correlations, increasing the degree to which the model responds mechanistically to (previously unobserved) input forcing data. An integrated model that explicitly considers such responses can potentially be applied across a broader range of input conditions than a more statistically-derived model that is restricted to envelopes of past observations.

Another important aspect of this integrated design is that each of the whole- ecosystem components (or modules) are explicitly evaluated in space and time, enforcing the need to verify that each component of the ecosystem behaves realistically. Our modeling process does not "allow" for final performance evaluations to be restricted to an isolated component of the system; the dynamics of each fundamental component are explicitly considered to some level.

Achieving integrated and balanced cycles of elements in models of complex ecosystems requires a significant investment of effort in system understanding and synthesis. The cybernetic nature of ecosystems has evolved over millennia, and it is unlikely that its actual complexity can be captured by computer simulation anytime in the near future. However, synthesizing the fundamental drivers and emergent properties of basic ecosystem interactions is a feasible goal – as outlined in this ELM documentation report.

The ELM described in this documentation, with its core General Ecosystem Model (Fitz et al. 1996), simulates a simple yet complete carbon cycle of an ecosystem: atmospheric carbon is fixed by living plants, incorporated into dead organic matter, and lost from the system via oxidation. Likewise, a comprehensive phosphorus cycle is incorporated, including dynamic stoichiometry associated with the flows among the fundamental "live" and "dead" phosphorus storages. The hydrologic cycle is also complete, considering surface and subsurface storages and flows. A calibration of one ecosystem component in ELM must be achieved in tandem with realistic behavior in the rest of the ecosystem components. This is not the case in simpler models of (an) isolated ecosystem component(s), in which the behavior of the remaining ecosystem components is not considered.

Thus, the calibration goals of ELM extend beyond the specific Performance Measure to be used in model applications. In integrating a simple representation of a complex ecosystem, one ELM calibration goal is to obtain output of principal ecosystem properties that not only are mass-balanced[2], but that exhibit realistic dynamics across space and time. The definition of this realism is dependent on the spatial and temporal quality of available data that are specific to the Everglades, as presented in the results of this Chapter. More specific calibration goals involve the scrutiny of formal Performance Measures that are specific to the intended applications.

For the current ELM v2.5, our intended applications target phosphorus "water quality" Performance Measures (see Model Application Chapter). In this context of the fully integrated ELM, specific aspects of water column phosphorus calibration are required to be associated with reasonable behavior in other ecosystem properties. For example, in early development efforts we observed model parameter sets that exhibited statistically-acceptable water column P concentrations, but which were suboptimal because they also were associated with less-realistic rates of processes such as soil accretion or periphyton growth. <u>The best parameter set becomes that which provides acceptable performance of the primary model application Performance Measures, while maintaining other ecosystem dynamics that are, at minimum, consistent with our best understanding of the Everglades.</u>

There is no mathematical "guarantee" that the current parameter set is unique and optimal. However, the tightly interactive nature of the algorithms highly constrains the range of parameter values that result in acceptable whole-ecosystem dynamics. These "final" results (for any particular model version) are intended to demonstrate realistic ecosystem behaviors across a heterogeneous, regional landscape within decadal time scales of ecological relevance. Thus, the methods of evaluating the general performance, and the more specific application Performance Measures, are intended to demonstrate a reasonable degree of confidence in the application of the ELM under widely varying environmental inputs.

6.3.1.3 Processes and scales

The overall approach of (developing and) calibrating the ELM was to start by simplifying the complex Everglades ecosystems by processes, by space, and to some extent by time. Generally, this involved first considering the most important ecosystem drivers within a simplified spatial domain. The calibration procedure paralleled that used in our stepwise, hierarchical sensitivity analysis (see Uncertainty Chapter). The intensively studied and spatially simple domain of Water Conservation Area 2A (WCA-2A) was used as an important test bed for improving our understanding of simulated and observed behaviors. Hydrologic and nutrient transport/fate were considered important ecosystem drivers, and their dynamics were scrutinized in the subregional application. This model testing and parameter refinement process was iterated until the performance of the targeted

---

[2] Mass balance is ensured by the code design, and is verified in detailed budget outputs at multiple spatial and temporal scales. See the User's Guide Chapter for further details.

variable(s) was deemed suitable for interim calibration purposes. That iterative sequence then expanded in scope, evaluating a broader suite of model ecosystem components along with those important ecosystem drivers. Where appropriate, the lessons-learned from this intensively studied area were subsequently applied at the larger spatial domain of the regional Everglades landscape.

From the perspective of numerical solutions, the ELM was designed to be scaleable, in that the same source code, parameters, and (where appropriate) boundary conditions are used in model applications at different grid scales and domains. For example, in the case of processes which are usually scale-dependent, such as horizontal dispersion of surface water constituents, the algorithms were designed to ensure consistency of results across a range of grid scales, as described in the Model Structure Chapter. Of course, if raw data support higher resolution variables such as initial land surface elevation, processes such as water flows will potentially respond differently to fine vs. coarse scale spatial data. However, if coarse-resolution (e.g., 1000 m) input map data are simply resampled into finer grid resolutions, the results across scales are very similar. Depending on the application, some differences can still exist when using such resampled data because of the influence of scale-dependent implementations of other boundary condition grid data, and scale-dependent raster-vector topology of water management features (i.e., canals and water control structures). While of interest for landscape pattern and other analyses, such scaling considerations are not explored in detail in this documentation, which primarily focuses on the regional (greater Everglades) 1000 m grid scale application.

While the subregional model applications can be used to address specific questions that involve processes and patterns at fine spatial resolution, these applications were developed largely as a learning tool in order to improve the performance of the regional ELM. Relative to the greater Everglades region, there are substantially fewer habitat types and less complex water management features in a basin such as WCA-2A. Additionally, finer-scaled subregional applications aided our understanding of the influence of boundary conditions, and helped determine optimal ways to represent fine-scaled features at the 1000 m regional grid scale. For example, the 500 m grid scale subregional application was used to explore finer scaled spatial patterns and flows in WCA-2A, relative to the 1000 m subregional application for that domain, and relative to the 1000 m grid regional (greater Everglades) application. Similarly, a 200 m grid subregional application in Water Conservation Area 1 (WCA-1, or A.R.M. Loxahatchee National Wildlife Refuge) provided useful insights into the complexity of the topographic relationships in the marsh-canal (raster-vector) hydrologic exchanges along the uninterrupted “perimeter” canal bordering that entire basin’s domain.

Because even the regional simulation run-times are short[3], most simulations included the entire 1981 – 2000 period of record, with post-processing evaluations made either on the initial 1981-1995 calibration period, the 1996-2000 validation period, or the entire simulation period. However, the model can simulate any user-selected time period for which initial and dynamic boundary condition data are available. As indicated in the

---

[3] See User’s Guide Chapter; a modern PC executes a 20-year simulation of the regional ELM in slightly over one hour.

Model Application Chapter, meteorological (but not water control structure flow) boundary condition data are available for the period from 1965 – 2000.

6.3.1.4 Hydrologic calibration

The first step in the ELM calibration process is to "get the water right", as the physics of the Everglades are a primary driver of the other ecological dynamics across the landscape. The user can edit the ELM runtime configuration file to select the desired combination of vertical and horizontal solution modules. By simply "turning off" all vertical modules except those of hydrology and a tracer, the ELM can be run as a stand-alone hydrologic model, without any dynamic feedbacks from time-varying vegetation or soils.

The model sensitivity analysis in the Uncertainty Chapter provides a summary of the relative sensitivity of the global (GP_*) and habitat-specific (HP_*) parameters, which are fully defined in the Data Chapter. The following are the principal parameters that were adjusted in hydrologic calibration:

- Evapotranspiration (GP_calibET, HP_MAC_MAXLAI)
- Surface roughness (HP_MAC_MAXROUGH, and to some extent, HP_MAC_MINROUGH)
- Groundwater flows & storage (GP_calibGWat, and to some extent, HP_HYD_POROSITY)
- Levee seepage & (spatially rare) canal berm/lip-roughness (Seep, edgeMann)

Depending on the status of the calibration process (i.e., seeking preliminary ball-park accuracy, or more accurate near-final history-matching), a variety of comparisons were made between output and target data. Some targets were "soft" performance indicators, such as basin-wide flow budgets from the SFWMM that included groundwater and levee seepage flows. The primary calibration targets were more rigorous comparisons of simulated and observed stage elevations at monitoring sites distributed throughout the landscape. While short-term (ca. hours/days) overland flow velocities were not explicitly calibrated (due to lack of data), spatial and temporal distributions of a longer-term chloride "natural" tracer were evaluated after fundamental within-basin budget characteristics were deemed reasonable.

When the objectives of the current iteration of the calibration process were completed, the remaining (non-hydrologic) ecological modules were invoked in the configuration file, and the performance re-evaluated and refined if needed. Generally at this point, the calibration process moved into phosphorus water quality calibration, with its associated ecosystem dynamics.

6.3.1.5 Ecological calibration

The next major step in the ELM v2.5 calibration process was refinement of the phosphorus water quality performance characteristics. Because of the tightly-coupled code among soils, floc, macrophytes, periphyton, and surface/ground- water phosphorus, all (or none) of those modules must be executed during ecological simulations, i.e., selected in the runtime configuration file[4]. While the primary application goal for this ELM v2.5 is related to “water quality”, we emphasize that water column phosphorus and its associated model performance evaluation is coupled to multiple ecosystem processes, and the demarcations among “water quality” and the rest of the ecosystem are somewhat blurred in the process of model calibration.

The model sensitivity analysis in the Uncertainty Chapter provides a summary of the relative sensitivity of the global (GP_*) and habitat-specific (HP_*) parameters, which are fully defined in the Data Chapter. Without repeating those that also significantly affect hydrologic performance, the following are the principal parameters that were adjusted in water quality (and associated ecological) calibration:

- Periphyton (GP_ALG_RC_MORT, GP_ALG_RC_PROD, GP_C_ALG_KS_P)
- Soils (GP_DOM_DECOMP_POPT, GP_DOM_RCDECOMP, GP_TP_K_SLOPE)
- Water column P (GP_TPpart_thresh)

Other parameters, such as the net production and the mortality rate of macrophytes (HP_PHBIO_RCMORT , HP_PHBIO_RCMORT) were adjusted primarily in the context of improving performance characteristics of other components of the ecosystem. In that context, the primary calibration parameters in the list above were not necessarily always adjusted for water column phosphorus performance goals, but for capturing other ecosystem dynamic characteristics: soils, in particular, were a truly fundamental integrator of the model ecosystem dynamics. The spatial and temporal relationships among 1) the production and mortality of plants with 2) the concomitant rates of soil accretion, in 3) response to wetting/drying and phosphorus inflows, determined the degree to which the model captured the basic dynamics of the Everglades wetlands.

6.3.2 Validation process

More so than in the case of calibration, there are many interpretations of the definition of model “validation”. As discussed in the Uncertainty Chapter, whether “classical validation” can be effectively used in the practice of model applications is questionable. A model may be claimed to be validated in the classical sense when the period of simulation is extended somewhat in time with previously- unused input data, even when the important drivers (e.g., rainfall, nutrient loads) in the new period of simulation are effectively similar to those observed during the calibration period. Importantly, after a “classical” validation, any change to model code or parameters requires that the new model version be validated again. Most desirable for confidence in model utility is the demonstration of useful model performance across as large a range of system drivers as possible. Thus, without attempting to subjectively define “validation” requirements, the confidence in the model utility can advance as knowledge of the system behavior

---

[4] During initial development and refinement of the ELM (prior to v2.1), the algorithms’ code supported the ability to turn off (not execute) any combination of the vertical solution modules, maintaining the associated variables at constant values throughout the simulation. Subsequent development has encoded an even tighter integration among non-hydrologic modules, with some state variables being updated in multiple modules. Thus, all of the non-hydrologic modules need to be executed during an ecological simulation; otherwise, phosphorus mass balance violations will be shown in the budget outputs. In order to facilitate the initial testing of new modules, such as nitrogen biogeochemistry, the ELM code will be revised to once again provide that option for running a simulation with static variables in any of the modules.

advances, with concomitant advances in model refinement. The objective is thus to increase the confidence in the model capabilities.

Despite the difficulties in attempting to define and adhere to validation paradigms, we “classically” validated the ELM with the update from ELM v2.1 to the interim ELM v2.2. We had previously demonstrated the ELM calibration performance during the 17-year period from January 1979 – December 1995 (ELM_Team 2002). Because the behavior of the entire regional domain of ELM during those years had been used in the calibration, the “classical” validation of the model involved the period of simulation update from January 1996 – December 2000. This interim update to ELM v2.2 was used to demonstrate the “classical” validation of the model in predicting water stage and surface water phosphorus concentrations.

As described in the Data Chapter and another section of this Chapter, important forcing data within the calibration period were modified as a result of quality assurance processes at the South Florida Water Management District. Time constraints prevented us from formally recalibrating the ELM during the previously-used 1979-1995 period for the interim v2.2, and instead we evaluated the model performance when using all of the newly available (and theoretically improved) data. For purposes of validating the algorithms and parameters used in the ELM v2.1, the ELM v2.2 had no changes to dynamic calculations in the equations, nor were there effectively changes[5] to model parameters. Statistical evaluations of the differences in observed vs. simulated water stage and surface water phosphorus concentrations were used to evaluate the (1981-1995) calibration and (1996-2000) validation performance of ELM v2.2, in addition to comparing ELM v2.2 and v2.1 during their common period of simulation. As noted in another section of this Chapter, some model refinements were subsequently made to take advantage of enhanced Everglades understanding (data), leading to the current release of ELM v2.5.

6.3.3 Performance evaluation methods

6.3.3.1 Statistical metrics

Simulated data were compared with observations obtained from the South Florida Water Management District’s databases (see Chapter on Data Description). For statistical evaluations of the hydrologic performance, at each monitoring site distributed throughout the region we compared daily predicted and observed stages using calculations of the correlation coefficient (R^2), Bias, root-mean-square-error (RMSE), and Nash-Sutcliffe Efficiency (Eff). These statistical metrics are the same as those used for the SFWMM

---

[5] While the intent was to leave all parameters identical to those in ELMv2.1, two parameters were modified due to the use of potential evapotranspiration (pET) input data, in lieu of internal calculations of that potential from raw meteorological data that was input to ELM v2.1. In that version, part of the method of determining pET involved calculating plant canopy transpiration in response to the calculated saturation vapor pressure deficit. In v2.2 (& higher), the plant-contribution to actual ET required adjustment through the maximum Leaf Area Index in some habitats (that had relatively high maximum values) in order to approximately match actual ET in ELMv2.1 and v2.2. Specifically, the maximum Leaf Area Index parameter for several habitats required a reduction to a value of no greater than 3.5 in any habitat, and the global (across the domain) pET calibration mulitplier parameter was modified slightly (from 1.05 to 0.90). These parameters were modified prior to viewing output from the 1996-2000 extension of the simulation.

(and other hydrologic models), and are well- supported by the spatial and temporal scales and quality of the input data.

For evaluating the water quality performance, we compared temporal aggregations of predicted total phosphorus (TP) concentration in surface water, using metrics of the Bias and RMSE (ug l^{-1}). For these evaluations, the simulated and observed TP concentration data were aggregated into “bins” of arithmetic means within wet (May 1 – September 30) and dry (October 1 – April 30) seasons within each water year of the simulation period. The input data do not support useful time series comparisons for these water quality evaluations (see the earlier section in this Chapter, and supporting data analyses in the Uncertainty Chapter). Moreover, the application Performance Measures are targeted to long term eutrophication trends. For these reasons, the statistical metrics of model performance did not include time series goodness of fit measures, the dynamics of which are subject to the data uncertainties discussed elsewhere. Rather, we determined the magnitude of offsets between observed and simulated data at the monitoring sites, in order to evaluate how well the model captured the long term, spatially distributed (gradients of) eutrophication of the ecosystems across the greater Everglades spatial domain.

See the Appendix A of this Chapter for computational methods for these statistics.

6.3.3.2 Graphical indicators

In order to further evaluate the model performance, we used a variety of quantitative graphical methods that are useful relative indicators of performance through space and time. Stage hydrographs of simulated and observed data (shown relative to the dynamic land surface elevation) at each monitoring site provide insight into any specific periods of time when the simulated stage departs from corresponding observed data. These graphical comparisons are shown at several levels of temporal aggregation: none (daily), seasonal, and water-year, including the 95% Confidence Intervals of data for the temporally-aggregated data. Cumulative Frequency Distributions (and 95% Confidence Intervals) of simulated and observed stages are provided for each location, providing a rapidly- visualized period-of-simulation performance summary within and among monitoring sites. Similarly, time series and Cumulative Frequency Distributions are provided for comparing observed and simulated TP concentrations in surface water at each monitoring site. To minimize the potential for users to “erroneously” infer instantaneous point comparisons at each monitoring site, we only present the temporally-aggregated data, with their associated 95% Confidence Intervals.

An important component of determining the performance of this model involves an evaluation of eutrophication gradients in the Everglades. The most intensively studied area (with respect to length of time and number of processes/variables) is the strong eutrophication gradient in Water Conservation Area 2A (WCA-2A). Two research and monitoring transects downstream of inflow water control structures have been used to document and understand phosphorus eutrophication in the Everglades (multiple references, with many summarized in (McCormick et al. 2002)). Comparisons of simulated and observed data on water column phosphorus concentration, net accumulation of phosphorus in the ecosystem, and other ecosystem attributes are shown relative to the distance from the upstream source of the water and nutrient loading.

6.3.3.3 Indicators of consistency

The above statistical and graphical comparisons of simulated and observed data are a fundamental component of evaluating the "model skill" in capturing the specific Performance Measures and related ecosystem dynamics. Beyond those comparisons, there are other indicators of how well the model performs, including indicators of its consistency relative to other models and relative to other, less rigorously quantified, ecological patterns and trends. These indicators of consistency may involve varying degrees of numerical analyses, but their presentation is intended to increase the cumulative weight of evidence that the model realistically portrays the landscape dynamics.

Hydrologic flows

One useful hydrologic flow indicator is the relative comparison of the basin-wide hydrologic budgets of the ELM and the SFWMM. The SFWMM is currently accepted for management applications, and is used to provide output data on managed water control structure flows to other models such as the ELM when simulating future scenarios. It is therefore useful to provide another measure of its consistency with the SFWMM, beyond the two models' stage calibration statistics. We make these budget-comparisons through a quantitative graphical comparison for each of the principal flows constituting the managed hydrologic budget for each year in the simulation. For a finer scaled comparison in space, we also present side by side summary maps of the long-term hydroperiod in the greater Everglades domain that is common to both models.

Another indicator of the relative accuracy of water flows is an evaluation of the simulated vs. observed data on chloride concentration in surface waters. As discussed in the Data Chapter, chloride is assumed to be a conservative tracer of flows, although the available spatial and temporal sampling constrains its use to that of relatively coarse indicator of relative water flow regimes. In the freshwater Everglades, the chloride input concentrations are sampled at the same frequency (with similar missing data constraints), at most of the same input water control structures, as phosphorus. Thus, the same temporal data quality constraints apply to chloride model inputs, and the associated analyses of model performance are simply presented as the percent difference in the mean simulated and observed values, relative to the observed values[6]. As with the surface water phosphorus graphical analyses, the aggregated time series and Cumulative Frequency Distributions are provided for comparing observed and simulated chloride concentrations in surface water at each monitoring site distributed throughout the greater Everglades.

Landscape patterns

The spatial patterns of ecosystem dynamics are integral to the overall goals this landscape model. In the above method descriptions, we summarize a rigorous suite of analyses of the spatial and temporal trends in model and observed data that relate to the phosphorus "water quality" Performance Measures intended for ELM v2.5 application. In particular,

---

[6] This simple relative index is generally more useful for chloride than for phosphorus, as the latter is commonly found in background (or unimpacted-region) concentrations that are extremely low (<10 ug l-1, close to the detection limit of 4 ug l-1). Thus, at a site whose mean is 8 ug l-1, a 4 ug l-1 difference between simulated and observed data is well within the margin of data uncertainty and appropriate modeling expectations, yet would exhibit a high relative error of 50%.

the spatial distribution of these measures of performance are an important consideration in evaluating the ELM. Beyond the gradients of those spatially distributed "point" measures, we present summary output maps as general indicators of the model consistency with spatial patterns of eutrophication gradients. These multi-decadal summaries of variables related to phosphorus eutrophication are shown for visualization of the spatial trends in variables that include soil phosphorus concentrations and cattail succession. These spatial summaries are not part of the intended model application Performance Measures, and are thus provided only as indicators of the degree to which the regional landscape trends are captured in the simulation.

There are existing observed data that can be used to generate landscape maps of soil attributes and habitat types, and we have made spatial comparisons of simulated and observed patterns in earlier subregional versions of ELM (Fitz and Sklar 1999). Those types of comparisons will be extended in spatial domain, and expanded with respect to their evaluation methods. Moreover, we have initiated potential collaborations[7] to investigate the application of multivariate geographic clustering applications (Hargrove and Hoffman 2005, Hoffman et al. 2005) to synthesize the multiple outputs of ELM into aggregate habitat types involving more than vegetation type alone. We anticipate that the next release, ELM v3.0, will be used to evaluate more of the spatial and temporal patterns of ecosystem variables distributed across the greater Everglades landscape.

6.4 Model updates

As described in other Chapters, the current release[8] ELM v2.5 has a number of improvements over the last release, ELM v2.1. However, the principal dynamic algorithms and most of the associated parameters used in ELM v2.5 are largely the same as those in the prior v2.1. Some of the primary differences among versions are associated with updated data used for boundary conditions, including some initial conditions (primarily land surface elevation). As discussed in an earlier section of this Chapter, prior to adjusting most parameters or source code, we evaluated the model performance using those improved data sets, including an extended period-of-simulation that encompassed the years 1981-2000 (vs. through-1995 in v2.1). That first interim data-driven update (v2.2) was used to "classically" validate the response of the model to new data forcing data.

In updating from the interim ELM v2.2 to the current ELM v2.5, the primary modifications that influenced model calculations involved the inclusion of dynamic stage input data along the edges of the domain boundary. This included daily stage along freshwater (generally urban and agricultural) lands, and monthly tidal fluctuations along the Florida Bay and Gulf of Mexico boundaries. The calculated slope of canal reach vectors was modified to be constant from beginning to ending points (instead of following land surface contours), and a canal parameter was added to allow the incorporation of a "lip" or berm along the side of a canal that does not include a levee.

---

[7] Personal communication, W. Hargrove, Environmental Sciences Division, Oak Ridge National Laboratory

[8] For simplicity, any full public release version is denoted only by the primary and secondary version attributes (see Model Refinement Chapter). The tertiary version attribute of this July 10, 2006 model release is ELM v2.5.2. Any subsequent public model release will be denoted by v2.6 or higher.

Related modifications were made to canal segmentation in Water Conservation Area 1 (A.R.M. Loxahatchee National Wildlife Refuge), improving the flow regimes between the continuous "perimeter" canal and adjacent marshes, including subsequent outflows from the S-10 structures that flow into Water Conservation Area 2A. Summaries of the data and code modifications since ELM v2.1 are found in the Model Refinements Chapter. Full descriptions of the current algorithms and data are found in the Data and Model Structure Chapters.

6.5 Model configuration

In ELM v2.5, the model was configured to simulate historical conditions inclusive of the years 1981 – 2000. The domain was that of the regional ELM, employing a 1 km^2 grid mesh encompassing all of the Water Conservation Areas, Holey Land, Rotenberger Tract, parts of the Model Lands near the C-111 canal region, and most of Everglades National Park and Big Cypress National Preserve. The vector topology of the canal/levee network and the point locations of water control structures were constant during the simulation period. The habitat succession module was operating, as were all other ecological modules, providing dynamic feedbacks among the physics, chemistry, and biology of the mosaic of ecosystems in the landscape. Dynamic boundary conditions included daily data on rainfall, potential evapotranspiration, managed water control structure flows with associated constituent concentrations, and stage (along the borders of the domain, including annually-recurring, monthly mean tidal amplitudes). Full descriptions of the requisite data and the functionality of the algorithms and source code are provided in other Chapters of this documentation.

6.6 Performance results

6.6.1 Ecological performance

6.6.1.1 Surface water P concentration: statistical metrics

The marsh and canal TP concentration monitoring locations used in evaluating the model performance are shown in Figure 6.1. Table 6.1 shows the statistical performance metrics for the simulated vs. observed total phosphorus concentration data at each location during the 1981-2000 simulation period. The median Bias of all predicted TP concentrations in the marsh for the 1981-2000 period of record was 2 ug l^{-1} (ppb), and slightly higher (4 ug l^{-1}) in canal predictions. The spatial distribution of the long-term mean surface water concentration (Figure 6.2) indicates strong gradients of eutrophication in northern WCA-2A, the Miami Canal inputs to northern WCA-3A, and a localized band encircling the interior perimeter of WCA-1. Biases lower than 5 ppb do not appear in any spatial trend, but higher variability associated with high mean concentrations resulted in higher biases in and immediately adjacent to canals.

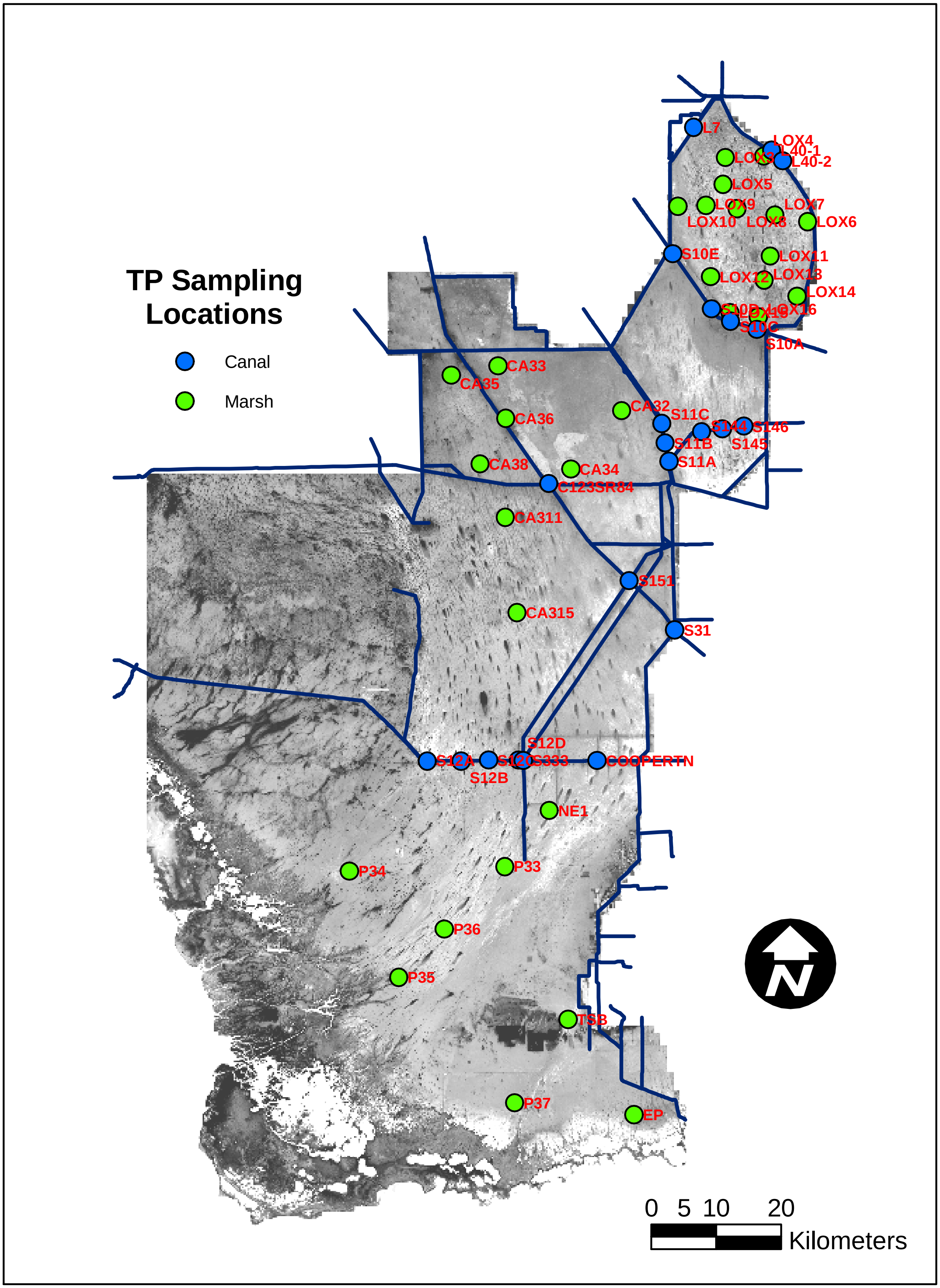


Figure 6.1 Map of most TP and CL monitoring site locations (see also Figure 6.1b).

Figure 6.1b. Map of water quality monitoring locations in WCA-1 and WCA-2A. Note that the scale of the grid-cell interactions with canal vectors results in effectively zero-distance from the canals for a number of the monitoring sites, particularly in WCA-1 (A.R.M. Loxahatchee National Wildlife Refuge).

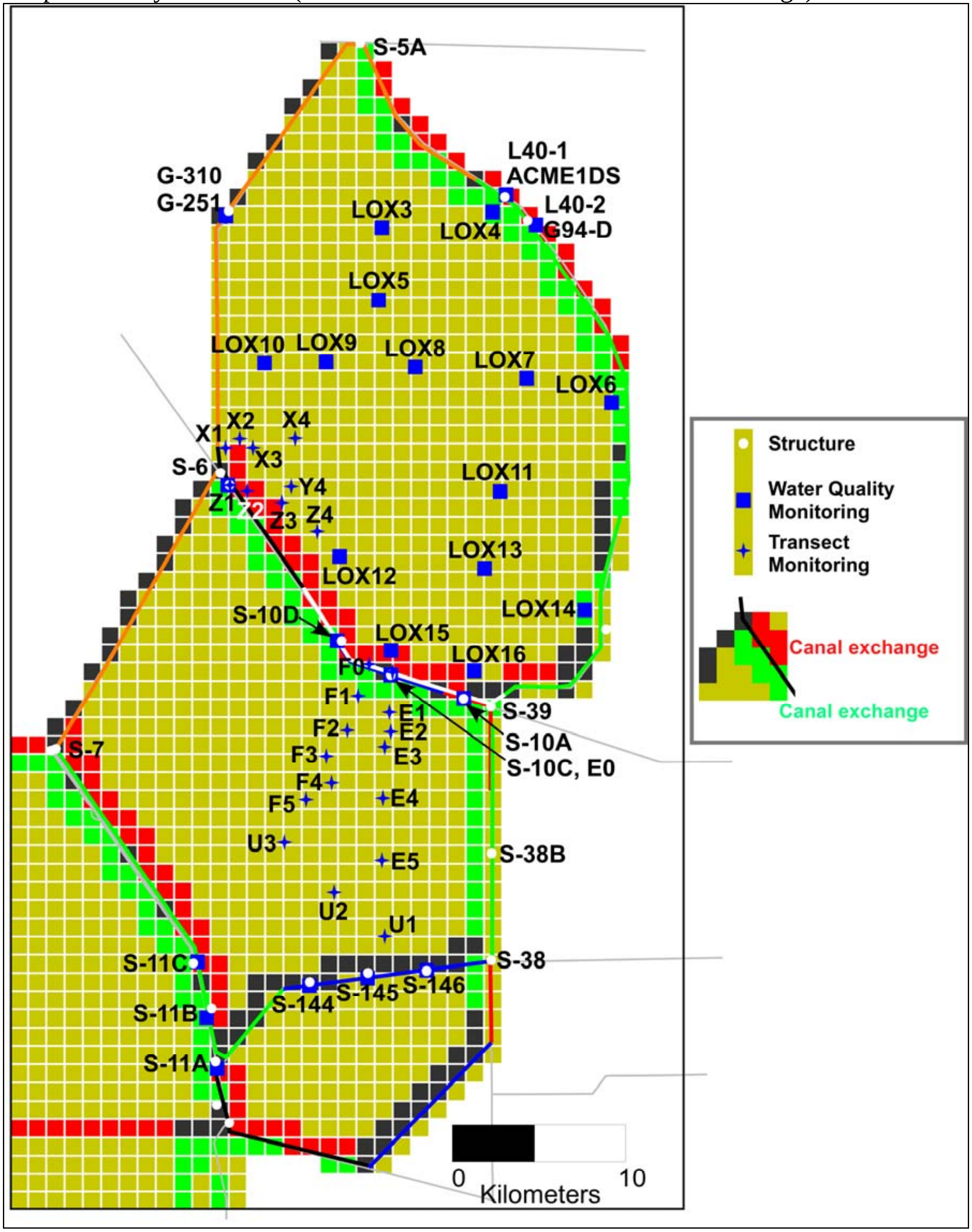
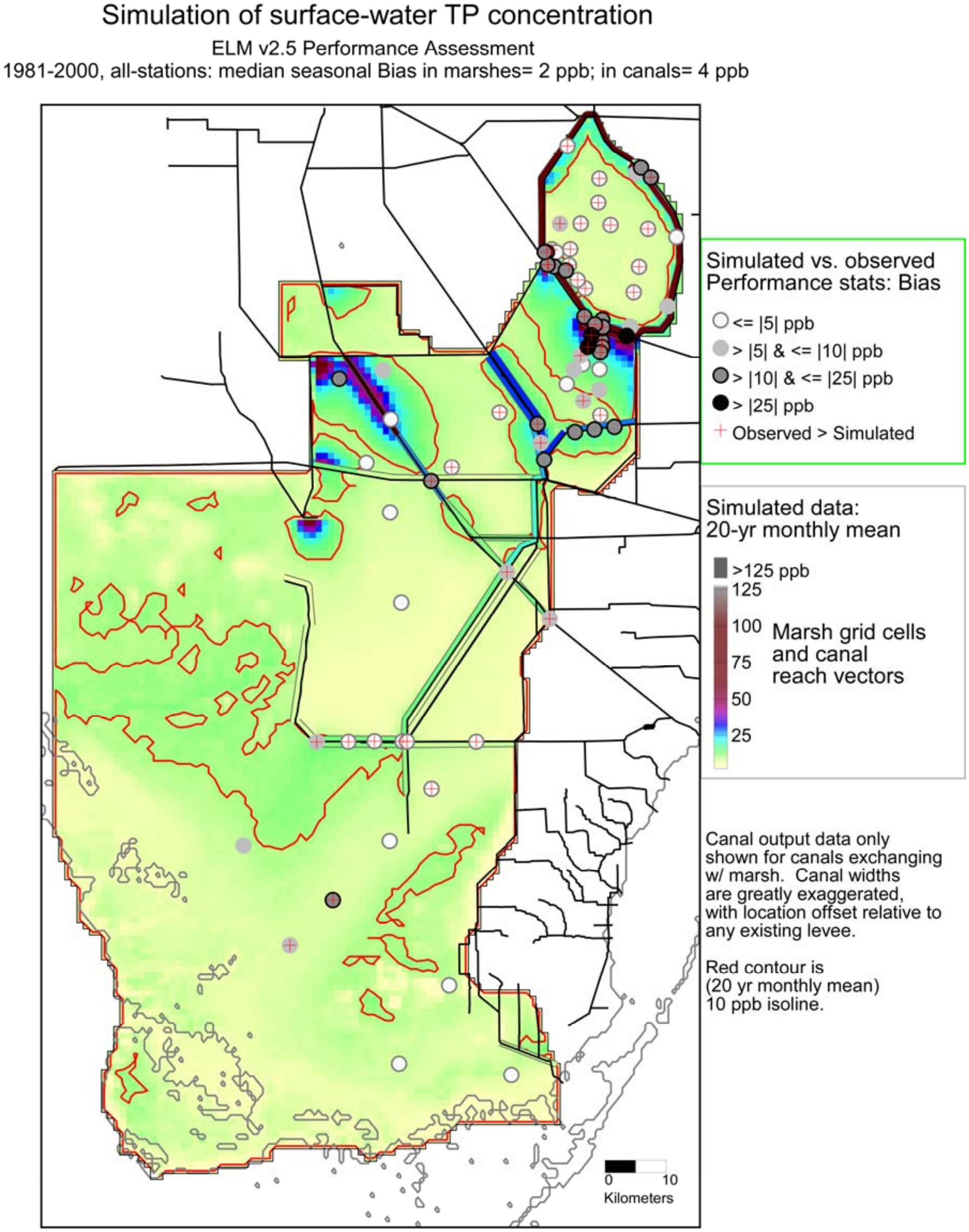


Figure 6.2 Map of statistical bias in model predictions of observed total phosphorus (TP) concentrations in marsh and canal locations. Background map is the simulated mean monthly TP concentration during 1981-2000. Statistics are detailed in Table 6.1.



Canal output data only shown for canals exchanging w/ marsh. Canal widths are greatly exaggerated, with location offset relative to any existing levee.

Red contour is (20 yr monthly mean) 10 ppb isoline.

Table 6.1. Statistical evaluation of simulated vs. observed surface water phosphorus concentration, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are ug l^{-1} (ppb).

					1981-2000		
Site	Basin	Site type	N	ObsMean	RelBias	Bias	RMSE
LOX4	WCA1	Marsh	12	10	-0.92	-9	11
LOX3	WCA1	Marsh	11	11	0.43	5	7
LOX5	WCA1	Marsh	13	10	0.32	3	5
LOX9	WCA1	Marsh	13	9	0.44	4	5
LOX10	WCA1	Marsh	12	10	0.53	5	6
LOX8	WCA1	Marsh	14	9	0.31	3	4
LOX7	WCA1	Marsh	14	8	0.32	3	3
LOX6	WCA1	Marsh	14	8	-0.43	-3	5
LOX11	WCA1	Marsh	14	9	0.46	4	5
LOX12	WCA1	Marsh	14	8	0.32	2	3
LOX13	WCA1	Marsh	14	9	0.45	4	5
LOX14	WCA1	Marsh	14	8	-1.22	-10	11
LOX15	WCA1	Marsh	14	8	-1.87	-14	16
LOX16	WCA1	Marsh	14	9	-0.70	-6	7
CA33	WCA3A	Marsh	14	13	-0.46	-6	8
CA35	WCA3A	Marsh	14	12	-1.74	-21	22
CA32	WCA3A	Marsh	14	8	0.13	1	2
CA36	WCA3A	Marsh	14	30	-0.13	-4	10
CA38	WCA3A	Marsh	14	9	-0.15	-1	4
CA34	WCA3A	Marsh	14	10	0.21	2	4
CA311	WCA3A	Marsh	14	6	-0.66	-4	5
CA315	WCA3A	Marsh	14	6	-0.11	-1	2
NE1	ENP	Marsh	29	10	0.43	4	7
P33	ENP	Marsh	30	8	-0.03	0	3
P34	ENP	Marsh	26	6	-0.91	-6	6
P36	ENP	Marsh	30	17	0.64	11	24
P35	ENP	Marsh	29	13	0.57	8	16
TSB	ENP	Marsh	30	8	-0.53	-4	6
P37	ENP	Marsh	28	6	-0.66	-4	5
EP	ENP	Marsh	27	6	-0.22	-1	3
X1	WCA1	Mar. Trans.	10	40	0.58	23	33
X2	WCA1	Mar. Trans.	10	16	0.22	3	7
X3	WCA1	Mar. Trans.	10	11	-0.40	-5	10
X4	WCA1	Mar. Trans.	9	10	0.44	5	5
Y4	WCA1	Mar. Trans.	10	12	0.31	4	13
Z1	WCA1	Mar. Trans.	10	42	0.07	3	14
Z2	WCA1	Mar. Trans.	9	14	-1.35	-19	23
Z3	WCA1	Mar. Trans.	10	10	-1.73	-17	19
Z4	WCA1	Mar. Trans.	10	9	0.34	3	6
E1	WCA2A	Mar. Trans.	13	65	0.24	15	30
E2	WCA2A	Mar. Trans.	12	58	0.33	19	29
E3	WCA2A	Mar. Trans.	12	39	0.28	11	21
E4	WCA2A	Mar. Trans.	13	15	-0.28	-4	7
E5	WCA2A	Mar. Trans.	13	9	-0.76	-6	8
F1	WCA2A	Mar. Trans.	14	120	0.27	32	72
F2	WCA2A	Mar. Trans.	13	67	0.49	33	47
F3	WCA2A	Mar. Trans.	13	29	0.30	9	13
F4	WCA2A	Mar. Trans.	13	19	-0.01	0	5
F5	WCA2A	Mar. Trans.	13	11	-0.51	-6	8
U1	WCA2A	Mar. Trans.	13	11	0.00	0	8
U2	WCA2A	Mar. Trans.	13	14	0.41	6	29
U3	WCA2A	Mar. Trans.	14	9	-0.45	-4	7

Table continued on next page...

Table 6.1 continued. Statistical evaluation of simulated vs. observed surface water phosphorus concentration, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are ug l^{-1} (ppb).

			1981-2000 (continued)				
Site	Basin	Site type	N	ObsMean	RelBias	Bias	RMSE
L7	WCA1	Canal	8	118	0.04	4	54
L40-1	WCA1	Canal	20	62	-0.16	-10	34
L40-2	WCA1	Canal	20	84	0.16	13	30
S10A	WCA1	Canal	25	54	-0.79	-43	60
S10C	WCA1	Canal	26	81	-0.21	-17	41
S10D	WCA1	Canal	39	99	0.11	11	37
S10E	WCA1	Canal	23	88	0.17	15	40
X0	WCA1	Can. Trans.	8	53	-0.26	-14	26
Z0	WCA1	Can. Trans.	8	60	-0.10	-6	19
E0	WCA1	Can. Trans.	13	86	0.20	17	36
F0	WCA2A	Can. Trans.	12	93	0.23	22	35
S144	WCA2A	Canal	29	19	-0.56	-11	19
S145	WCA2A	Canal	35	16	-0.77	-13	19
S146	WCA2A	Canal	29	16	-0.78	-13	20
S11A	WCA2A	Canal	33	27	-0.49	-13	26
S11B	WCA2A	Canal	32	44	0.13	6	23
S11C	WCA2A	Canal	39	55	0.43	23	32
C123SR84	WCA2A	Canal	26	46	0.48	22	27
S151	WCA3A	Canal	40	27	0.29	8	19
S12A	WCA3A	Canal	39	16	0.33	5	20
S12B	WCA3A	Canal	39	14	0.19	3	14
S12C	WCA3A	Canal	40	14	0.09	1	7
S12D	WCA3A	Canal	40	14	0.14	2	6
S333	WCA3A	Canal	39	15	0.22	3	8
COOPERTN	WCA3A	Canal	20	11	0.35	4	5
S31	WCA3B	Canal	26	21	0.38	8	17
		Median All:	14	14	0.13	2	11
		Median Canal:	28	45	0.13	4	24
		Median Marsh:	14	10	0.10	2	7

6.6.1.2 Surface water P concentration: graphical indicators

These visualizations of the temporal trends in simulated and observed data are an important component of understanding the model performance, particularly with respect to recognizing any unique aspects of the data dynamics at a particular site. Figure 6.3a shows an example of the time series of seasonally-averaged phosphorus concentrations in canals. The model effectively captured the spatial differences between northern Everglades canals with relatively high (ca 70 ppb) mean concentrations, down to canals in the central/southern portions of the system with lower (ca. 10 ppb) mean concentrations. Within the marsh (Figure 6.3b), the model likewise generally stays within the range of observed data, in an area ranging from high (ca. 50 ppb) to low (<10 ppb) ambient concentrations.

Figure 6.3 (following 2 pages). Example plots of time series and Cumulative Frequency Distributions (CFD) of simulated and observed phosphorus concentrations in canal (Figure 6.3a) and marsh (Figure 6.3b) sites.

The constant dashed line indicates the TP field sampling Detection Limit (DL = 4 ug l^{-1} *for the model period of record), which was the minimum value used for observed data in plots and statistics. To enable equivalent comparisons, any simulated value which was below the DL was set equal to the DL. The model grid cell column and row locations (col_row) or canal reach identifier (single integer) are shown in parentheses of each plot's title.*

Time series plots: All data were aggregated into arithmetic mean values by wet and dry seasons within water years; the continuous lines pass through mean of all daily data points for each season; the mean of paired simulated and observed values are shown in red boxes and black diamonds, respectively; the 95% Confidence Interval (CI) of the paired means are shown by the "__" symbols in the red for the model and black for the observed data.

Cumulative Frequency Distributions: The CFDs of the simulated and observed (raw, un-aggregated) data; the 95% confidence interval for observed data is shown in the dashed black lines. Note that only paired simulated and observed data points are used.

Appendix B. The complete set of graphics for all monitoring sites in the greater Everglades is provided in Appendix B.

Figure 6.3a. Time series and CFDs of simulated and observed phosphorus concentrations for <u>canal</u> sites with high concentrations (L40-2, WCA-1) and low concentrations (S12-D, flowing into Everglades National Park). The time series plots have different scales.

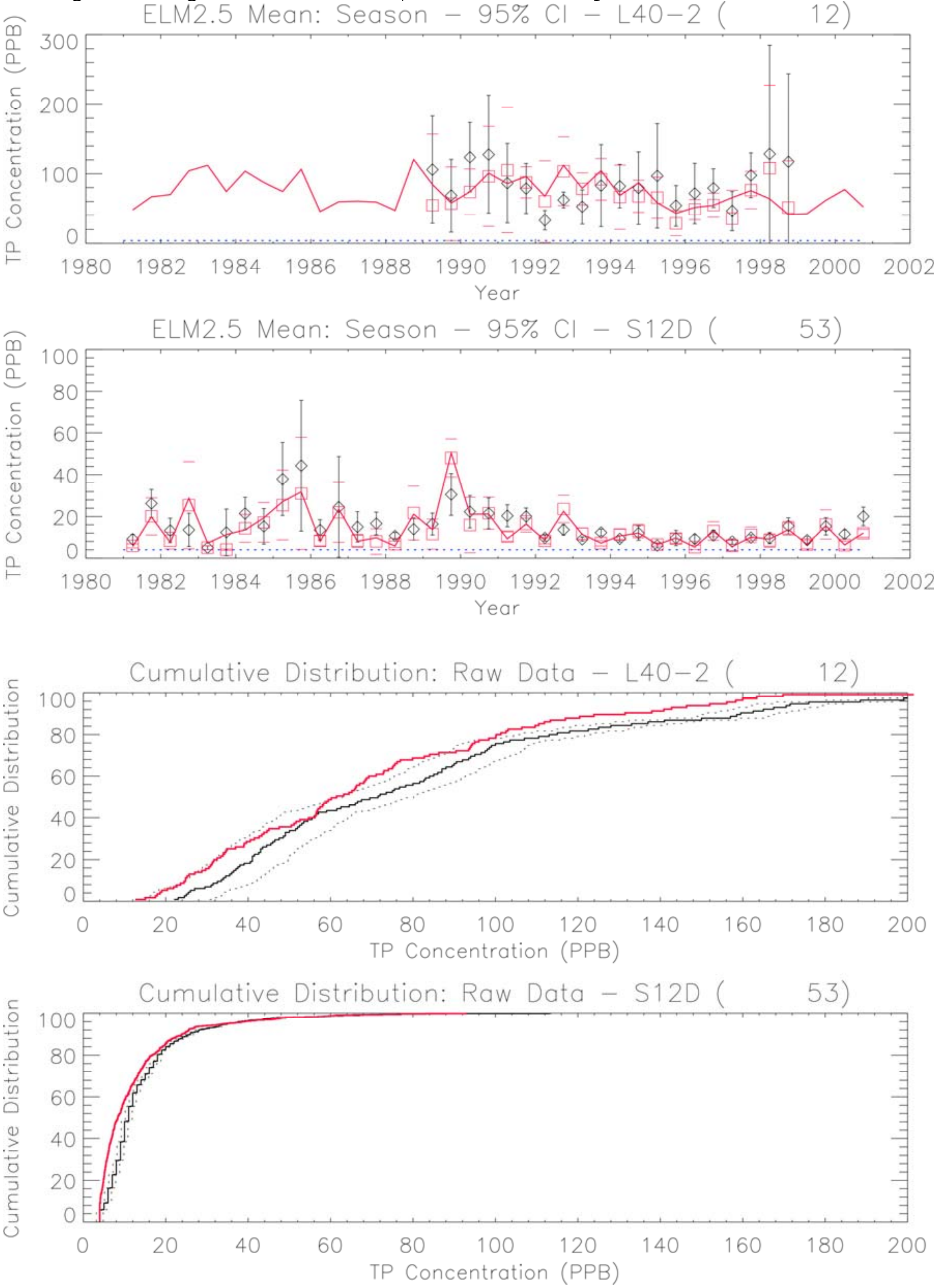
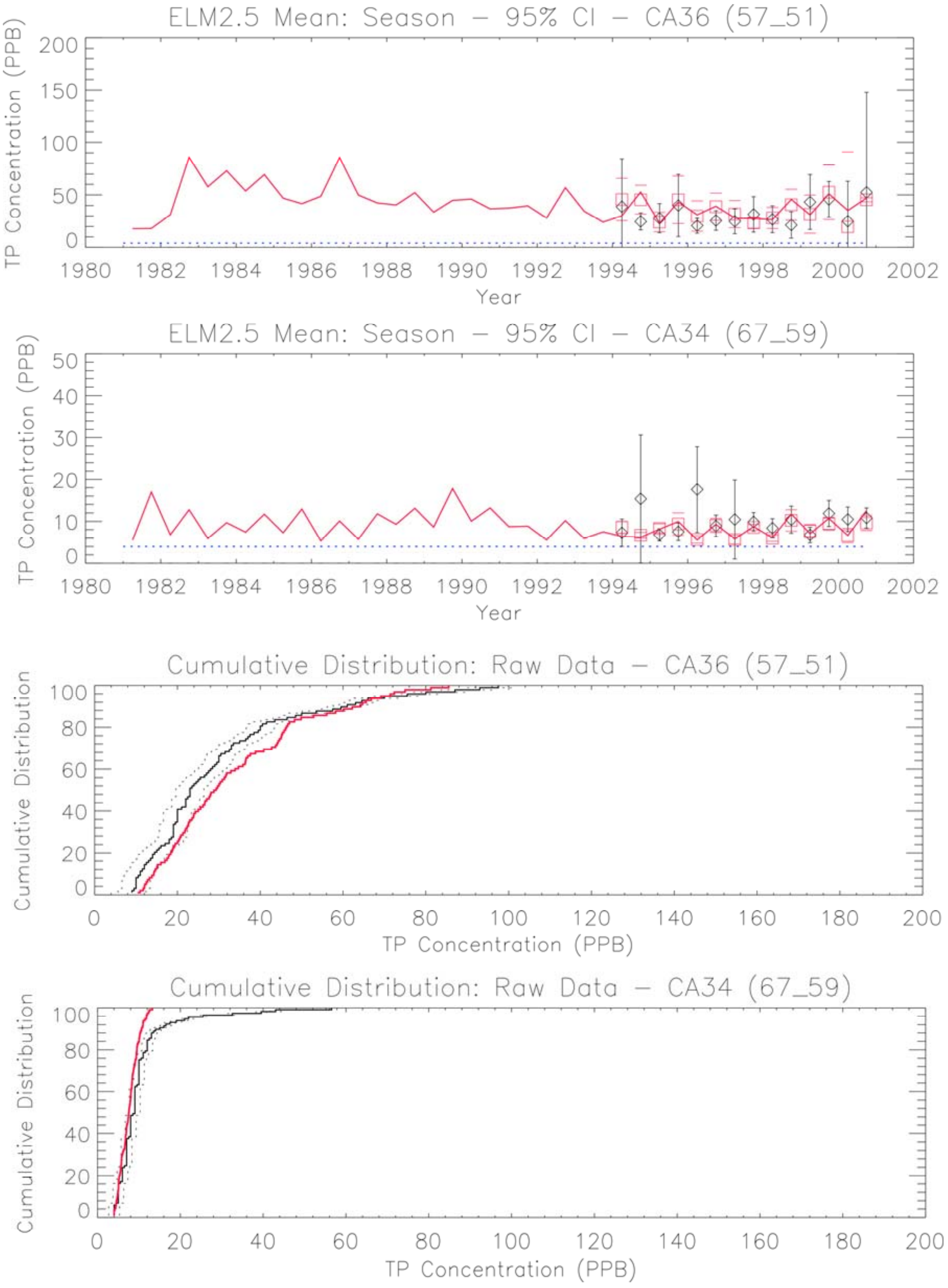


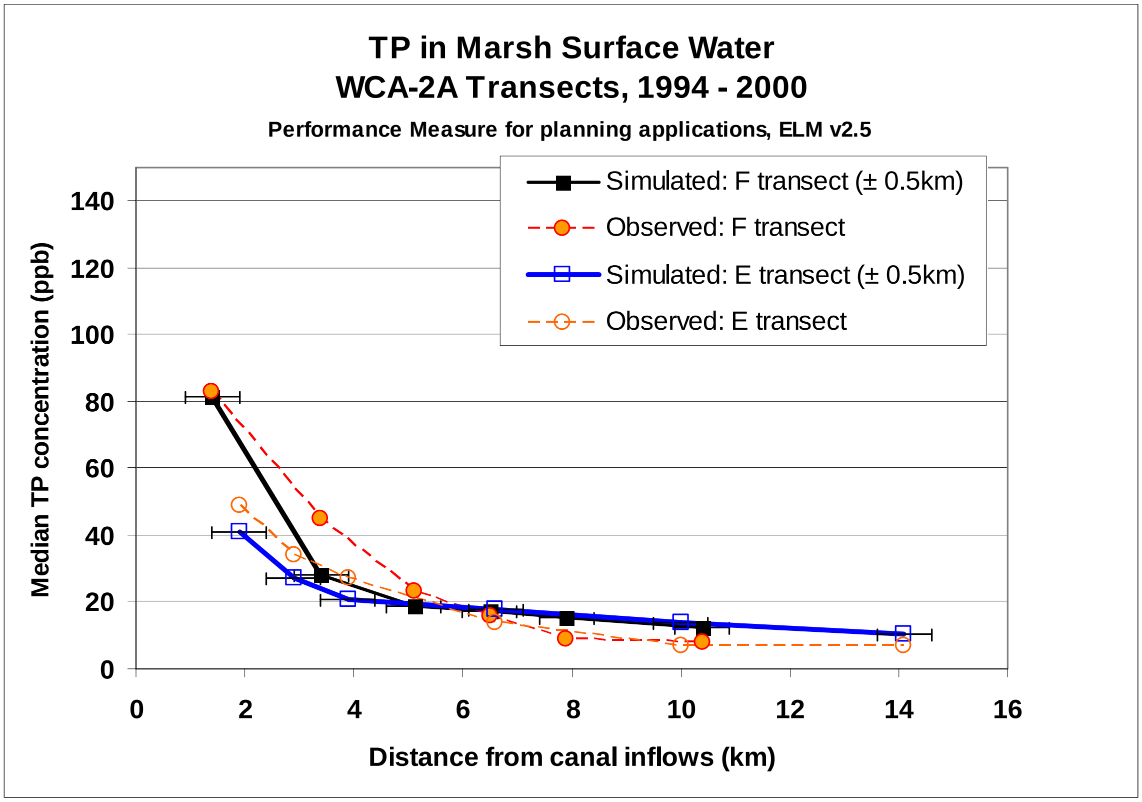
Figure 6.3a. Time series and CFDs of simulated and observed phosphorus concentrations for <u>marsh</u> monitoring sites with high mean concentrations (CA-36, WCA-3A) and low mean concentrations (CA-34, WCA-3A). The time series plots have different scales.

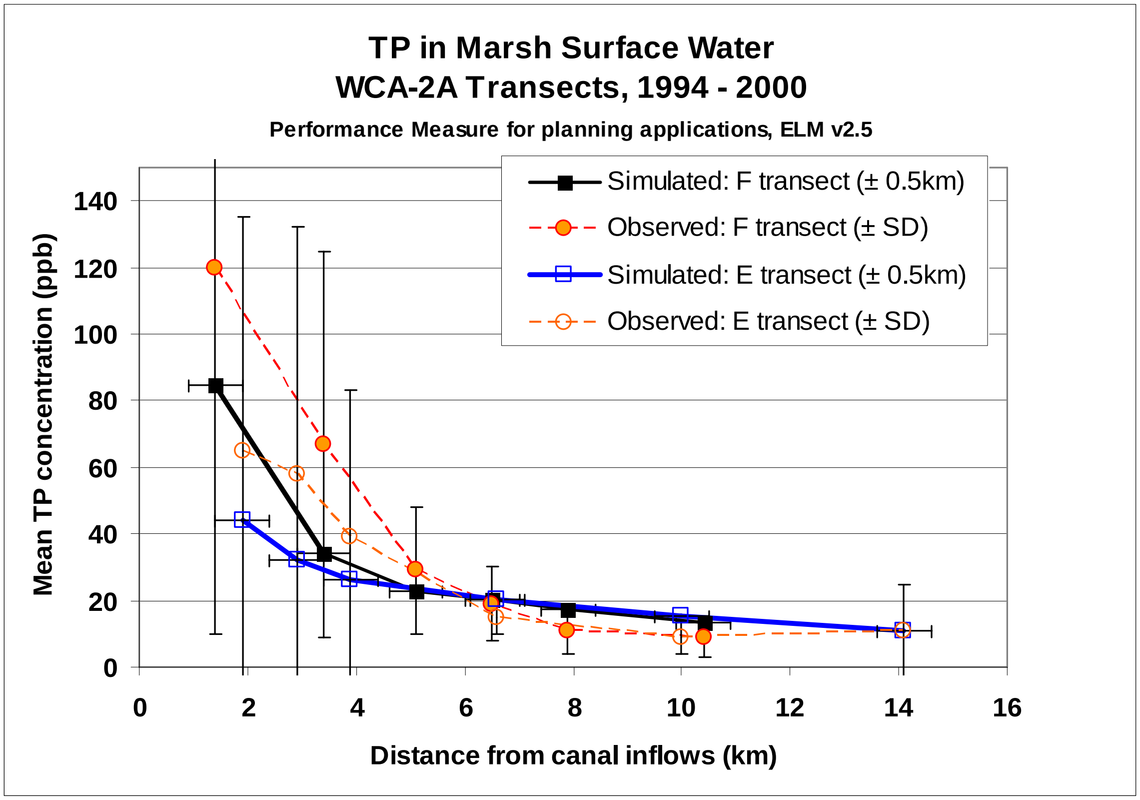


6.6.1.3 Surface water P concentration: transect evaluations

A subset of the monitoring locations analyzed above are actually sites that were established along specific eutrophication gradients in Water Conservation Area 2A. Each of these “E” and the “F” transects were monitored at six sites, from near the inflow “points” adjacent to canal inflows, into interior points 10-15 km downstream. At high ambient P concentrations near the inflows, there was high variability as evidenced in large standard deviations about the mean. The median values of modeled and observed concentrations were very closely matched along the gradient that ranged from approximately 80 to approximately 10 ppb concentrations.

Figure 6.4 Surface water phosphorus concentration along the “E” and the “F” transects in the WCA-2A eutrophication gradient. Sampling started in 1994.

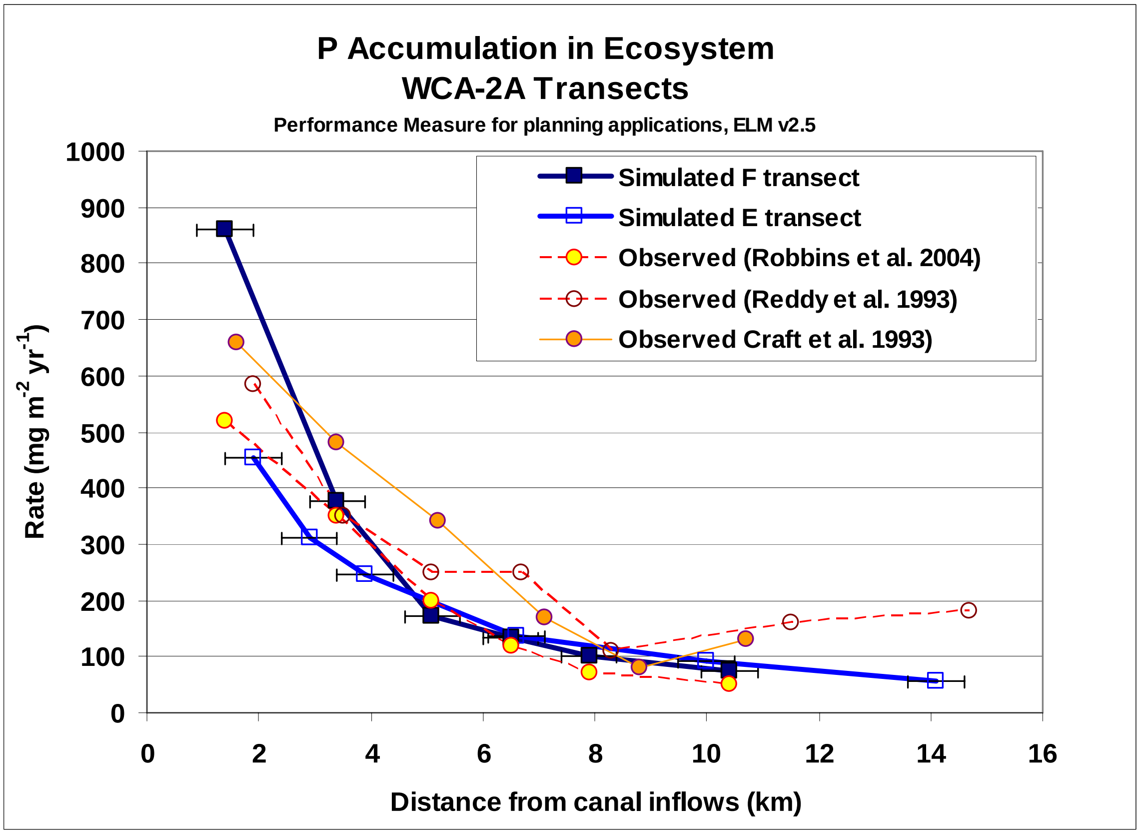




6.6.1.4 Phosphorus accumulation rate: transect evaluations

The accumulation rates of phosphorus are an integrated measure of the actual net nutrient load to which the ecosystem is responding. There was variability among studies and locations in estimated long term P accumulation from radionuclide tracers, but simulated data generally had strong concordance to the spatial trends in observed data.

Figure 6.5 Net phosphorus accumulation along the WCA-2A gradient. Observed data were summarized from Craft et al. (1993), Reddy et al. (1993) and Robbins et al. (2004).



6.6.2 Hydrologic performance

6.6.2.1 Stage: statistical metrics

The marsh stage monitoring locations used in evaluating the model performance are mapped in Figure 6.6. Table 6.2 shows the statistical performance metrics for the simulated vs. observed stage data at each location during the 1981-2000 historical simulation period. The median bias of predicted stages was -1 cm. The median Nash-Sutcliffe Efficiency statistic was 0.56 for the simulation. The spatial distribution of the annual hydroperiod (Figure 6.7) indicates relatively lengthy inundation periods in Water Conservation Areas and large slough features draining to the southwest and south in Everglades National Park. Biases do not appear in any spatial trend, but boundary conditions along the model periphery resulted in higher biases in and immediately adjacent to canals and estuarine regions.

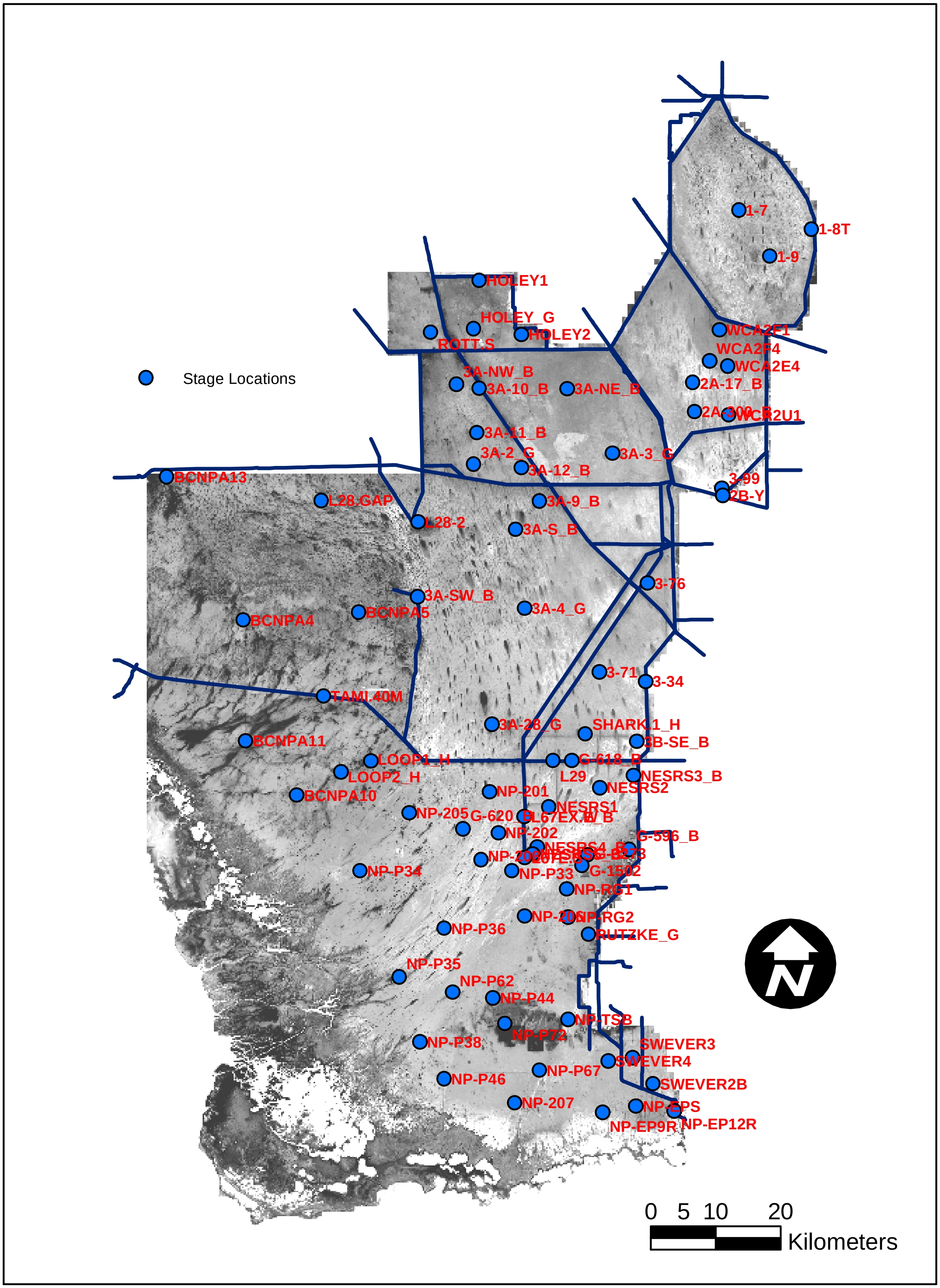


Figure 6.6. Map of stage monitoring site locations..

Figure 6.7 Map of statistical bias in model predictions of observed water stage elevations in marsh locations. Background map is the simulated mean annual hydroperiod during 1981-2000. Statistics are detailed in Table 6.2.

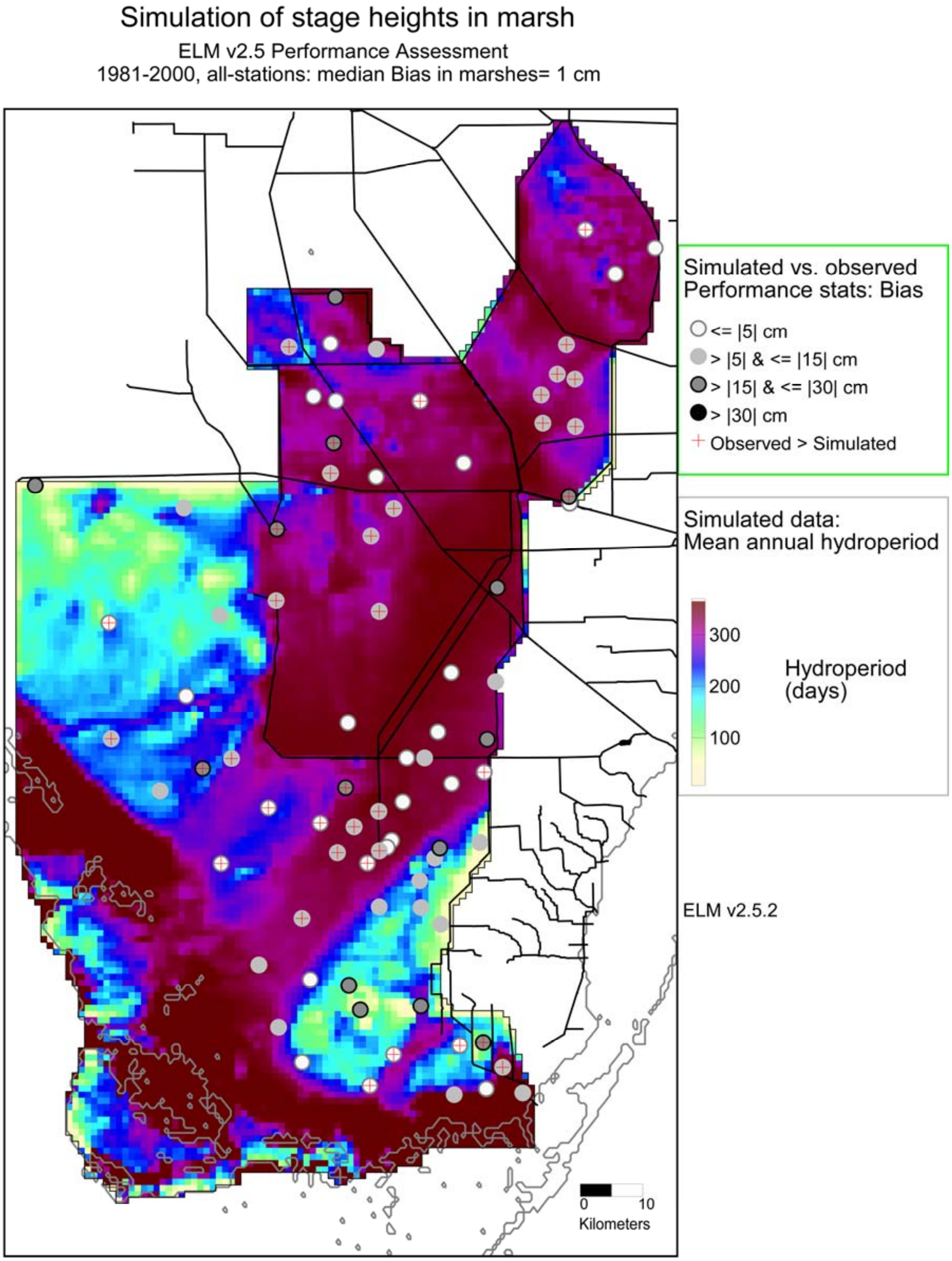


Table 6.2. Statistical evaluation of simulated vs. observed stage, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are meters.

				Stage 1981-2000		
Site	Basin	N	Bias (m)	RMSE (m)	R2	NS Eff.
_1-7	WCA1	7046	0.05	0.15	0.72	0.27
1-8T	WCA1	6869	-0.05	0.15	0.76	0.55
_1-9	WCA1	6879	-0.03	0.14	0.74	0.46
WCA2F1	WCA2A	2259	0.11	0.18	0.82	0.57
WCA2F4	WCA2A	1941	0.08	0.15	0.77	0.64
WCA2E4	WCA2A	2260	0.09	0.18	0.77	0.56
2A-17_B	WCA2A	7305	0.05	0.16	0.75	0.65
2A-300_B	WCA2A	7278	0.06	0.19	0.69	0.64
WCA2U1	WCA2A	2150	0.13	0.25	0.69	0.37
3A-NW_B	WCA3A	7035	-0.02	0.14	0.73	0.72
3A-10_B	WCA3A	6445	-0.03	0.13	0.75	0.58
3A-NE_B	WCA3A	6813	0.02	0.21	0.70	0.69
3A-11_B	WCA3A	6487	0.23	0.25	0.85	-0.56
3A-3_G	WCA3A	7305	-0.02	0.15	0.86	0.86
3A-2_G	WCA3A	7145	0.05	0.12	0.87	0.83
3A-12_B	WCA3A	6738	-0.02	0.16	0.65	0.56
3A-9_B	WCA3A	6969	0.15	0.18	0.86	0.59
L28-2	WCA3A	4007	0.18	0.21	0.84	0.06
3A-S_B	WCA3A	6871	0.12	0.16	0.86	0.61
3A-4_G	WCA3A	7305	0.12	0.18	0.85	0.68
3A-28_G	WCA3A	7295	-0.02	0.13	0.82	0.82
_3-99	WCA2B	3338	0.23	0.32	0.55	0.04
2B-Y	WCA2B	5515	-0.01	0.32	0.77	0.73
_3-76	WCA3B	3390	-0.16	0.22	0.61	-1.27
_3-71	WCA3B	3454	-0.02	0.12	0.63	0.50
_3-34	WCA3B	1633	-0.11	0.14	0.81	0.48
SHARK.1_H	WCA3B	6684	-0.04	0.12	0.84	0.78
3B-SE_B	WCA3B	6029	-0.15	0.23	0.83	0.50
HOLEY1	Holey L.	4041	-0.16	0.21	0.63	0.11
HOLEY_G	Holey L.	5599	-0.02	0.22	0.49	-0.49
HOLEY2	Holey L.	4046	-0.12	0.20	0.56	0.30
ROTT.S	Roten. T.	5208	0.12	0.17	0.60	0.24
BCNPA13	BCNP	1923	-0.18	0.26	0.37	-0.16
L28.GAP	BCNP	6393	-0.09	0.18	0.53	0.31
3A-SW_B	BCNP/3A	6641	0.08	0.13	0.86	0.68
BCNPA5	BCNP	3636	-0.13	0.21	0.42	0.02
BCNPA4	BCNP	3601	0.03	0.20	0.53	0.38
TAMI.40M	BCNP	7305	-0.01	0.18	0.72	0.66
BCNPA11	BCNP	3549	0.15	0.27	0.33	-0.01

Table continued on next page...

Table 6.2 continued. Statistical evaluation of simulated vs. observed stage, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are meters.

		Stage 1981-2000 (continued)				
Site	Basin	N	Bias (m)	RMSE (m)	R2	NS Eff.
G-618_B	ENP	7124	-0.05	0.14	0.72	0.66
L29	ENP	7305	0.00	0.13	0.69	0.67
LOOP1_H	ENP	5938	0.12	0.17	0.68	0.32
LOOP2_H	ENP	5972	0.17	0.23	0.70	0.24
NESRS3_B	ENP	5579	0.02	0.14	0.67	0.65
NESRS2	ENP	6228	-0.03	0.09	0.76	0.74
NP-201	ENP	5723	0.16	0.19	0.82	0.50
BCNPA10	ENP	3637	-0.10	0.17	0.53	0.24
NESRS1	ENP	6536	-0.02	0.09	0.74	0.72
NP-205	ENP	7149	0.04	0.14	0.80	0.78
L67EX.W	ENP	6319	0.05	0.18	0.74	0.59
L67EX.E_B	ENP	6187	-0.03	0.11	0.74	0.70
G-620_B	ENP	6264	0.01	0.11	0.79	0.79
NP-202	ENP	7069	0.08	0.15	0.74	0.61
NESRS4_B	ENP	4854	-0.03	0.10	0.71	0.63
G-596_B	ENP	7282	-0.13	0.23	0.60	0.16
NESRS5_B	ENP	4953	-0.01	0.08	0.76	0.70
G-3273	ENP	6137	-0.18	0.25	0.75	0.44
L67E.S	ENP	3631	0.10	0.19	0.55	0.34
NP-203	ENP	7049	0.05	0.13	0.74	0.68
G-1502	ENP	7305	-0.13	0.22	0.75	0.61
NP-P33	ENP	7147	0.02	0.13	0.60	0.57
NP-P34	ENP	6971	0.03	0.16	0.82	0.64
NP-RG1	ENP	1570	-0.09	0.14	0.85	0.67
NP-206	ENP	6641	-0.08	0.21	0.76	0.69
NP-RG2	ENP	1502	-0.11	0.16	0.85	0.63
NP-P36	ENP	6952	0.07	0.13	0.71	0.55
RUTZKE_G	ENP	2369	-0.05	0.20	0.79	0.21
NP-P35	ENP	6851	-0.14	0.20	0.79	-0.11
NP-P62	ENP	6851	-0.03	0.13	0.80	0.79
NP-P44	ENP	6440	-0.21	0.30	0.80	0.51
NP-TSB	ENP	7299	-0.16	0.22	0.79	0.56
NP-P72	ENP	7186	-0.20	0.29	0.75	0.47
NP-P38	ENP	6896	-0.09	0.14	0.87	0.44
SWEVER3	ENP	5330	0.20	0.25	0.68	-2.47
SWEVER4	ENP	5582	0.04	0.19	0.75	-0.58
NP-P67	ENP	7107	0.04	0.11	0.78	0.72
NP-P46	ENP	6680	-0.02	0.13	0.71	0.42
SWEVER2B	ENP	5488	0.14	0.17	0.58	-0.33
NP-207	ENP	6755	0.05	0.10	0.86	0.71
NP-EPS	ENP	5240	-0.02	0.06	0.70	0.67
NP-EP12R	ENP	2828	-0.07	0.09	0.76	0.22
NP-EP9R	ENP	2608	-0.12	0.13	0.75	-0.09
	Median:	6356	-0.01	0.17	0.75	0.56

6.6.2.2 Stage: graphical indicators

These visualizations of the temporal trends in simulated and observed data are an important component of understanding the model performance, particularly with respect to recognizing any unique aspects of the data dynamics at a particular site. Figure 6.8 shows an example of the time series of stage hydrographs in long and in short hydroperiod areas. The model effectively captured the spatial differences between southern Everglades marl prairie region that is periodically flooded, and a Water Conservation Area 3A location that is virtually always inundated with relatively deep surface water.

Figure 6.8 (following page). Example plots of time series and Cumulative Frequency Distributions (CFD) of simulated and observed stage in short hydroperiod (NP-206, Everglades National Park) and long hydroperiod (3A-28, WCA-3A) sites.

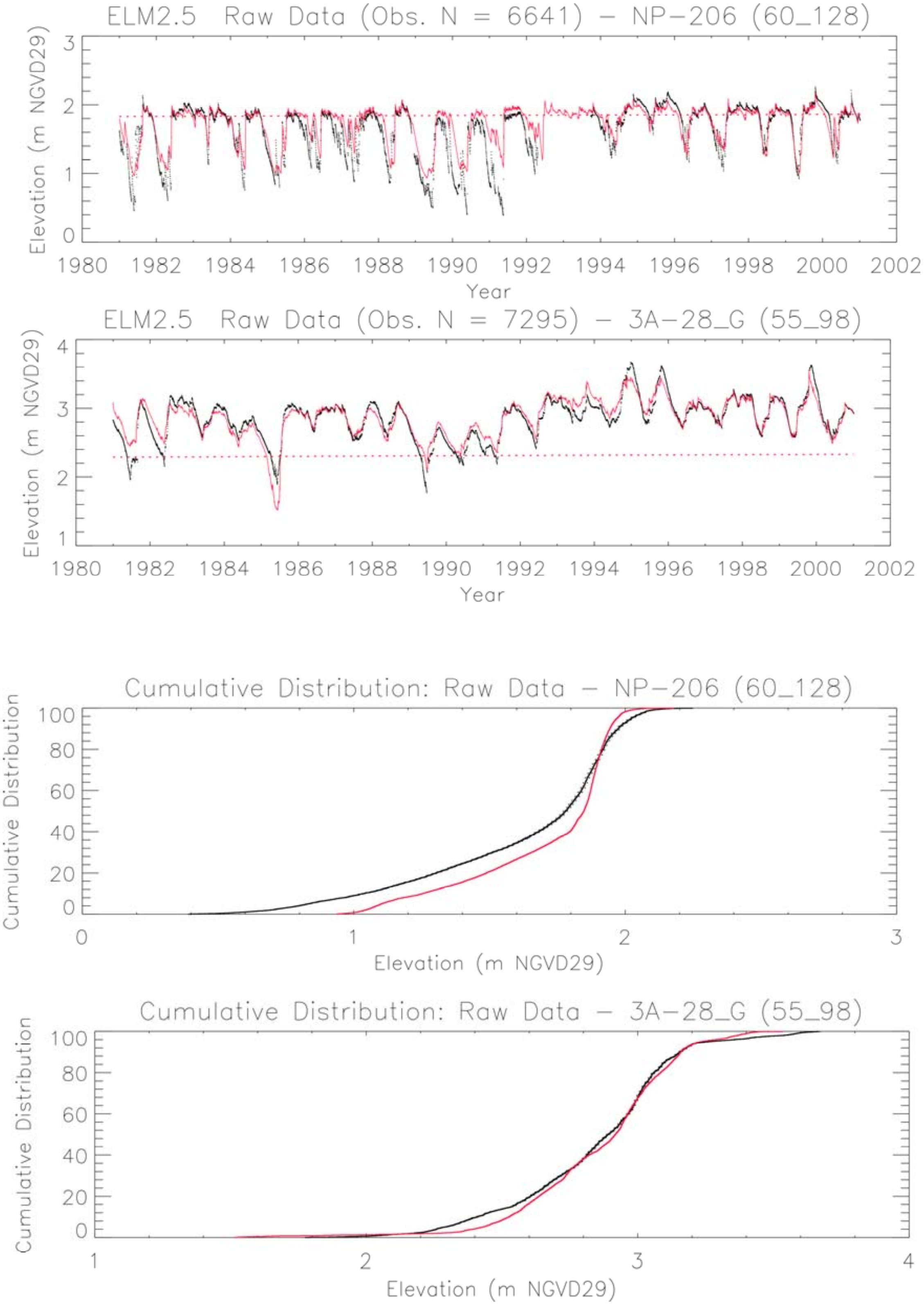
The red dashed line in the stage hydrographs is the model grid cell's land surface elevation, which is a time-varying output variable of the model. The model grid cell column and row locations are shown in parentheses (col_row) of each plot's title.

Time series plots: All data, with no temporal aggregation, of daily observations (black dots) and model results (red line).

Cumulative Frequency Distributions: The CFDs of the simulated and observed (raw, un-aggregated) data; the 95% confidence interval for observed data is shown in the dashed black lines. Note that only paired simulated and observed data points are used.

Appendix C. The complete set of graphics for all monitoring sites in the greater Everglades is provided in Appendix C.

Figure 6.8. Time series and Cumulative Frequency Distributions of simulated and observed stages for long and short hydroperiod sites. See full Figure legend above.



6.6.2.3 Consistency: inter-model water budget indicators

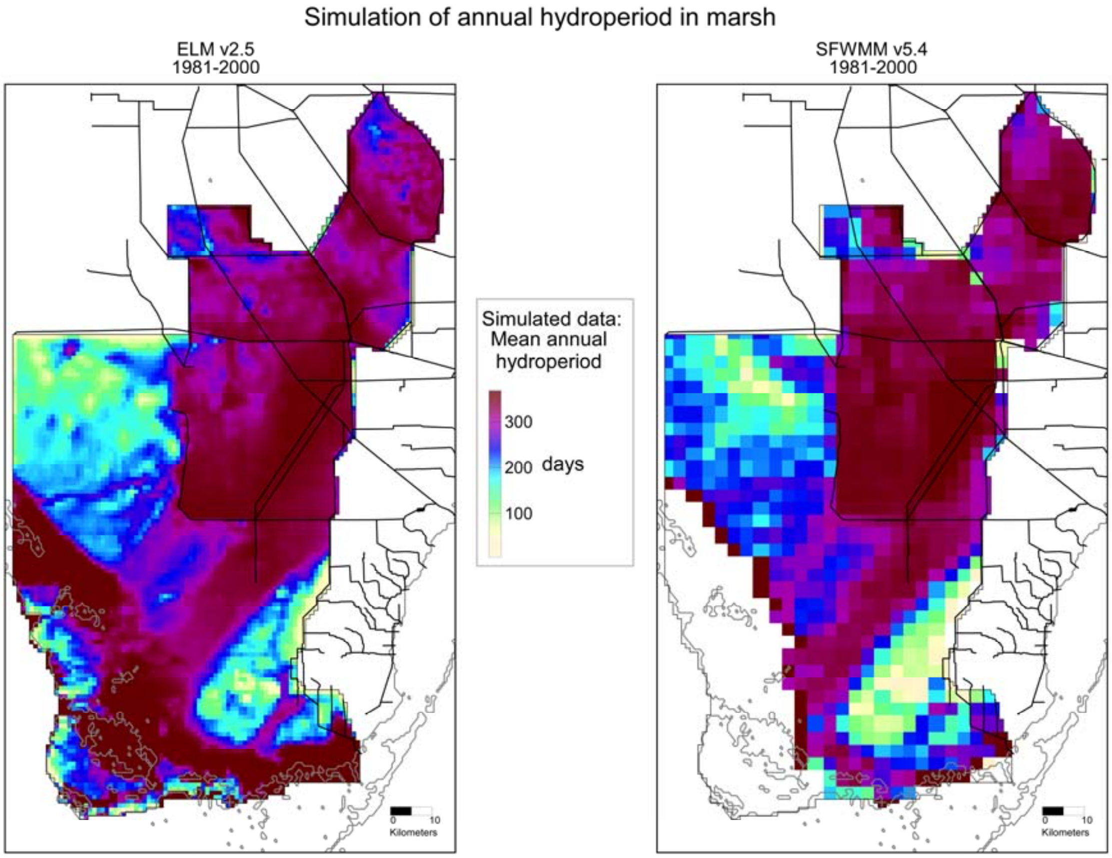
The water budgets of the ELM were generally similar to those of the SFWMM. For each of the major hydrologic basins, we compared the annual flows into and out of each Water Conservation Area. Figure 6.9 shows an example of such a comparison. Very minor differences in rainfall are due to the different spatial scales and discretization of grid cells. Other differences are observable in some years for other flows, but do not represent significant volumes (relative to the size of the basin). For each Water Conservation Area, Appendix D provides the actual hydrologic budgets for ELM, and the differences between the SFWMM and ELM.

Figure 6.9. Insert 3A budget comparison

6.6.2.4 Consistency: inter-model hydroperiod indicators

Another indicator of consistency between the ELM and the SFWMM is a comparison of the maps of the mean annual hydroperiod that is simulated by each model. Figure 6.10 indicates that the ELM generally mimics the distribution of hydroperiods, with some differences in the ELM capturing finer scaled features (largely due to finer scaled land surface elevation input data).

Figure 6.10. Mean annual hydroperiod simulated by the ELM and by the SFWMM, displaying only the portion of the SFWMM domain that overlaps with that of the ELM. The SFWMM grid cells are approximately 10.4 km^2, compared to the 1 km^2 grid resolution of the ELM. (As indicated, the SFWMM domain does not extend to the southwestern mangrove-dominated region along the Gulf of Mexico).



6.6.2.5 Consistency: Flow tracer (chloride) indicators

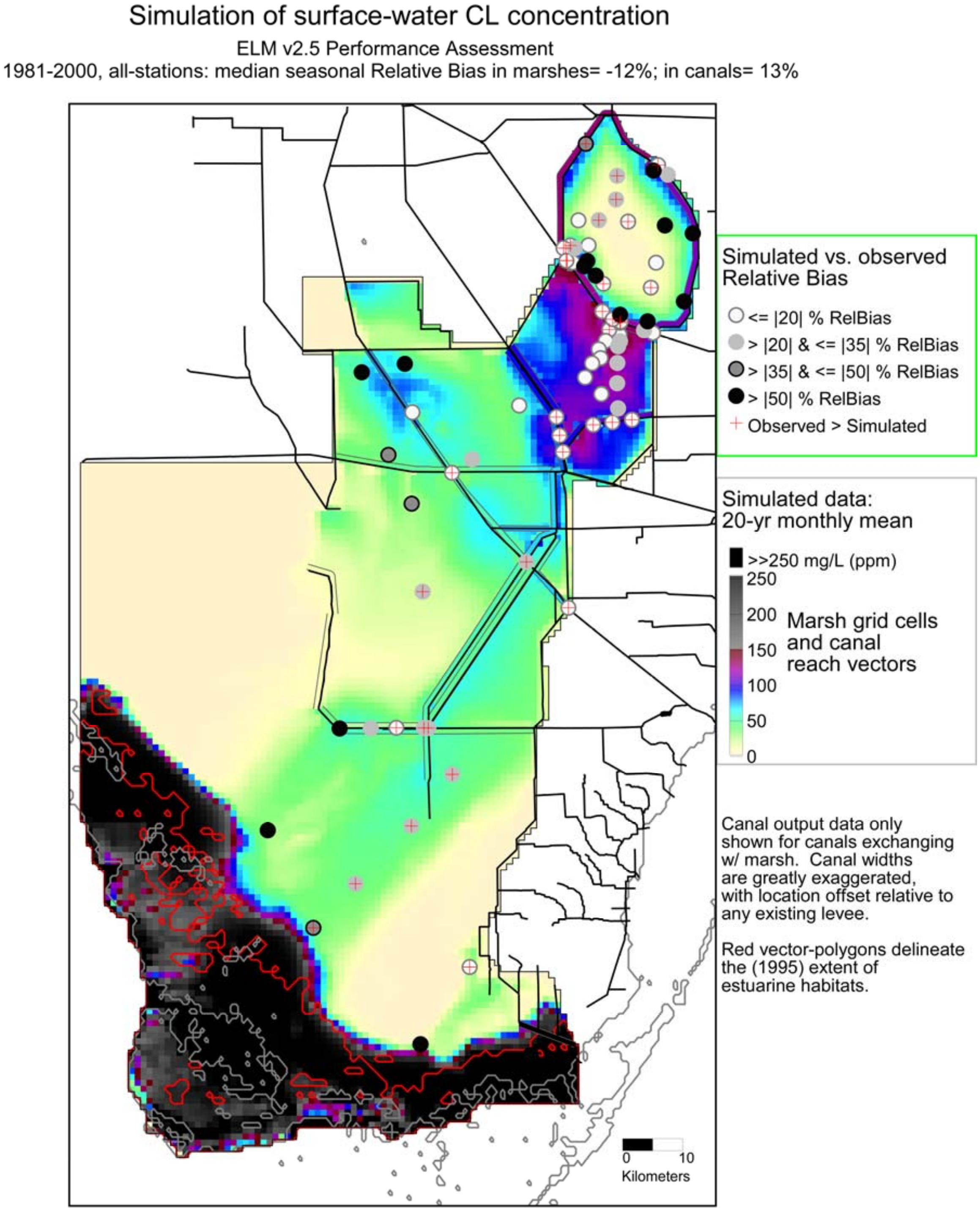
The distribution of chloride (CL) concentrations throughout the freshwater Everglades showed patterns of long-term flow regimes that were consistent with our understanding of major flow paths (Figure 6.11), most notably the “ring” of higher CL encircling WCA-1, and large inputs into WCA-2A. Other canal inputs within WCA-3A transported the tracer into Everglades National Park[9]. The relative bias metric indicated a distribution of relative errors that tended to be higher in close proximity to higher concentrations in canals, similar to the trends of phosphorus concentrations. The median relative error of all stations was -12% in the marshes, and 13% in canals (Table 3).

Appendix E: Figures E.1 – E.78 show the sets of 1981-2000 time series of chloride concentrations at varying temporal aggregations, including each site’s cumulative frequency distribution. These visualizations of the temporal trends in simulated and observed data can be an important component of understanding the model performance, particularly with respect to recognizing any unique aspects of the data dynamics at a particular site.

---

[9] The distribution of CL concentrations go “off-the-freshwater-scale” in the estuarine southern Everglades, with CL concentrations that were << 1 parts per thousand roughly corresponding to the extent of mangrove and other estuarine habitat types.

Figure 6.11 Map of statistical relative bias in model predictions of observed chloride (CL) concentrations in marsh and canal locations. Background map is the simulated mean monthly CL concentration during 1981-2000. Statistics are detailed in Table 6.3.



Canal output data only shown for canals exchanging w/ marsh. Canal widths are greatly exaggerated, with location offset relative to any existing levee.

Red vector-polygons delineate the (1995) extent of estuarine habitats.

Table 6.3. Statistical evaluation of simulated vs. observed surface water chloride concentration, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are mg l^{-1} (ppm).

					1981-2000		
Site	Basin	Site type	N	ObsMean	RelBias	Bias	RMSE
LOX4	WCA1	Marsh	25	68	-0.83	-57	77
LOX3	WCA1	Marsh	24	37	0.34	12	38
LOX5	WCA1	Marsh	26	18	0.34	6	12
LOX9	WCA1	Marsh	26	14	0.33	4	7
LOX10	WCA1	Marsh	24	28	-0.12	-3	29
LOX8	WCA1	Marsh	30	15	0.07	1	8
LOX7	WCA1	Marsh	30	29	-0.89	-26	35
LOX6	WCA1	Marsh	30	44	-1.20	-52	63
LOX11	WCA1	Marsh	29	13	-0.05	-1	7
LOX12	WCA1	Marsh	28	28	0.02	1	15
LOX13	WCA1	Marsh	29	12	0.01	0	6
LOX14	WCA1	Marsh	29	21	-2.97	-61	67
LOX15	WCA1	Marsh	29	48	-0.57	-28	42
LOX16	WCA1	Marsh	28	14	-3.60	-51	56
CA33	WCA3A	Marsh	38	53	-0.81	-43	56
CA35	WCA3A	Marsh	35	33	-0.79	-26	38
CA32	WCA3A	Marsh	46	50	-0.14	-7	43
CA36	WCA3A	Marsh	36	70	-0.10	-7	26
CA38	WCA3A	Marsh	51	31	-0.49	-16	28
CA34	WCA3A	Marsh	53	58	-0.29	-17	42
CA311	WCA3A	Marsh	45	29	-0.37	-11	26
CA315	WCA3A	Marsh	51	34	0.25	9	20
NE1	ENP	Marsh	107	78	0.25	20	32
P33	ENP	Marsh	113	71	0.21	15	29
P34	ENP	Marsh	69	22	-1.15	-26	39
P36	ENP	Marsh	108	72	0.26	19	34
P35	ENP	Marsh	103	131	0.48	63	223
TSB	ENP	Marsh	98	39	0.01	1	24
P37	ENP	Marsh	79	30	-1.59	-48	105
EP	ENP	Marsh	82	206	-64.21	-13229	17364
X1	WCA1	Mar. Trans.	55	122	0.12	15	29
X2	WCA1	Mar. Trans.	55	102	0.05	5	44
X3	WCA1	Mar. Trans.	55	86	-0.30	-26	55
X4	WCA1	Mar. Trans.	54	50	-0.19	-10	50
Y4	WCA1	Mar. Trans.	55	51	-0.86	-44	67
Z1	WCA1	Mar. Trans.	57	125	0.12	15	31
Z2	WCA1	Mar. Trans.	54	108	-0.09	-10	32
Z3	WCA1	Mar. Trans.	59	67	-0.55	-37	63
Z4	WCA1	Mar. Trans.	57	36	-0.92	-33	50
E1	WCA2A	Mar. Trans.	83	149	-0.01	-1	94
E2	WCA2A	Mar. Trans.	78	125	-0.24	-30	55
E3	WCA2A	Mar. Trans.	75	124	-0.23	-28	56
E4	WCA2A	Mar. Trans.	90	121	-0.26	-31	59
E5	WCA2A	Mar. Trans.	91	114	-0.32	-36	67
F1	WCA2A	Mar. Trans.	82	162	0.05	8	61
F2	WCA2A	Mar. Trans.	101	151	-0.11	-16	58
F3	WCA2A	Mar. Trans.	97	143	-0.12	-18	62
F4	WCA2A	Mar. Trans.	85	137	-0.12	-16	61
F5	WCA2A	Mar. Trans.	92	143	-0.08	-11	62
U1	WCA2A	Mar. Trans.	99	102	-0.28	-28	60
U2	WCA2A	Mar. Trans.	97	129	-0.05	-6	51
U3	WCA2A	Mar. Trans.	96	133	-0.10	-14	58

Table continued on next page...

Table 6.3 continued. Statistical evaluation of simulated vs. observed surface water chloride concentration, 1981 – 2000. Units of Bias (observed minus simulated) and RMSE are mg l^{-1} (ppm).

			1981-2000 (continued)				
Site	Basin	Site type	N	ObsMean	RelBias	Bias	RMSE
L7	WCA1	Canal	53	228	0.45	103	167
L40-1	WCA1	Canal	119	132	0.20	26	54
L40-2	WCA1	Canal	118	80	-0.33	-26	59
S10A	WCA1	Canal	94	95	-0.22	-21	56
S10C	WCA1	Canal	100	131	0.11	14	53
S10D	WCA1	Canal	198	145	0.17	24	56
S39	WCA1	Canal	251	106	-0.17	-18	56
S10E	WCA1	Canal	80	141	0.17	24	50
X0	WCA1	Can. Trans	60	131	0.18	24	38
Z0	WCA1	Can. Trans	59	133	0.19	25	40
E0	WCA2A	Can. Trans	108	128	0.01	1	37
F0	WCA2A	Can. Trans	110	132	0.04	5	41
S144	WCA2A	Canal	165	127	0.08	11	45
S145	WCA2A	Canal	206	121	0.07	8	44
S146	WCA2A	Canal	164	117	0.02	2	45
S11A	WCA2A	Canal	171	118	0.16	19	43
S11B	WCA2A	Canal	192	122	0.18	22	44
S11C	WCA2A	Canal	258	117	0.15	18	41
C123SR84	WCA3A	Canal	97	75	0.19	14	24
S151	WCA3A	Canal	229	98	0.25	24	39
S12A	WCA3A	Canal	320	29	-0.81	-24	33
S12B	WCA3A	Canal	345	39	-0.33	-13	28
S12C	WCA3A	Canal	350	54	0.04	2	33
S12D	WCA3A	Canal	367	69	0.24	16	37
S333	WCA3A	Canal	319	77	0.31	24	40
S31	WCA3B	Canal	109	89	0.01	1	60
		Median All:	80	80	-0.05	-3	44
		Median Canal:	165	118	0.13	14	43
		Median Marsh:	55	62	-0.12	-12	47

6.6.3 Ecological consistency

Beyond the above model application "water quality" Performance Measures, and the indicators of hydrologic consistency, below we provide some further indicators that the model adequately captures ecosystem dynamics in the regional landscape.

6.6.3.1 Consistency: Integrated ecosystem responses

The rate of peat accretion is a central integrator of the biological responses to water quality and hydrology. Using data from the "E" and "F" transects in WCA-2A, Figure 6.12 shows a strong correspondence of simulated and observed peat accretion, indicating a useful degree of balance between soil oxidation, plant mortality, and their hydrologic and nutrient drivers.

Macrophyte growth (and biomass) responds directly to porewater phosphorus availability, along with hydrologic variations. Simulated patterns of total macrophyte biomass were consistent with expected trends, particularly along nutrient gradients (Figure 6.13). Generally on longer time scales than those of macrophyte biomass changes, (and the even more transient porewater nutrients), phosphorus concentration in the soils[10] is a commonly used indicator of the eutrophication status of the Everglades wetlands. The simulated spatial pattern of the soil phosphorus concentrations (Figure 6.13) are consistent with our understanding of the trends in the Everglades, particularly downstream of known nutrient inflows such as those in WCA-2A. Also shown in that Figure, cattail succession as a result of (water levels and) eutrophication gradient in WCA-2A is generally consistent with the observed cattail distribution in 1995.

---

[10] While the upper 10 cm, and especially the surficial floc layer, of the soil is usually used in describing (recent) soil phosphorus status, the ELM does not stratify the soils beyond separating the floc and the 0-30 cm layers. There are often significant differences among soil layers (often with lower concentration in deeper 10-20 or 20-30 cm layers).

Figure 6.12 Simulated and observed rates of peat accretion along the WCA-2A eutrophication gradient. Data are summarized from Craft et al. (1993), Reddy et al. (1993).

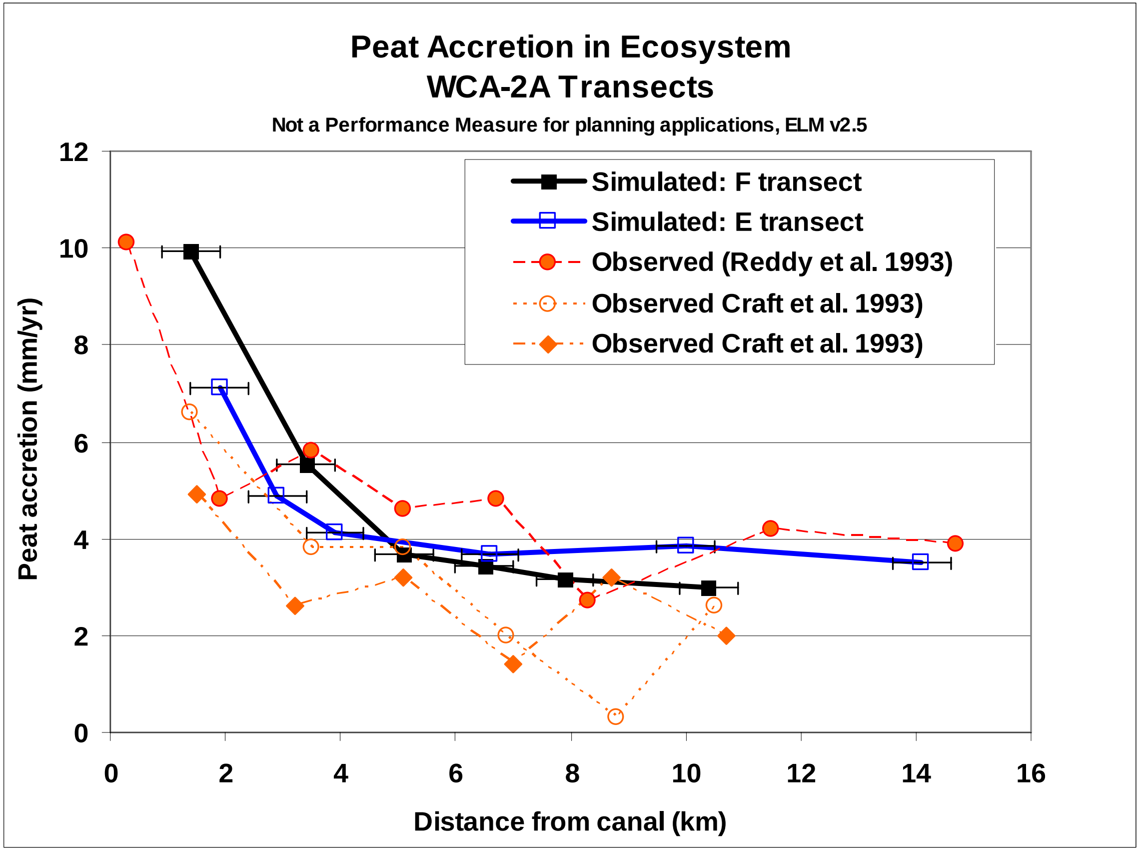
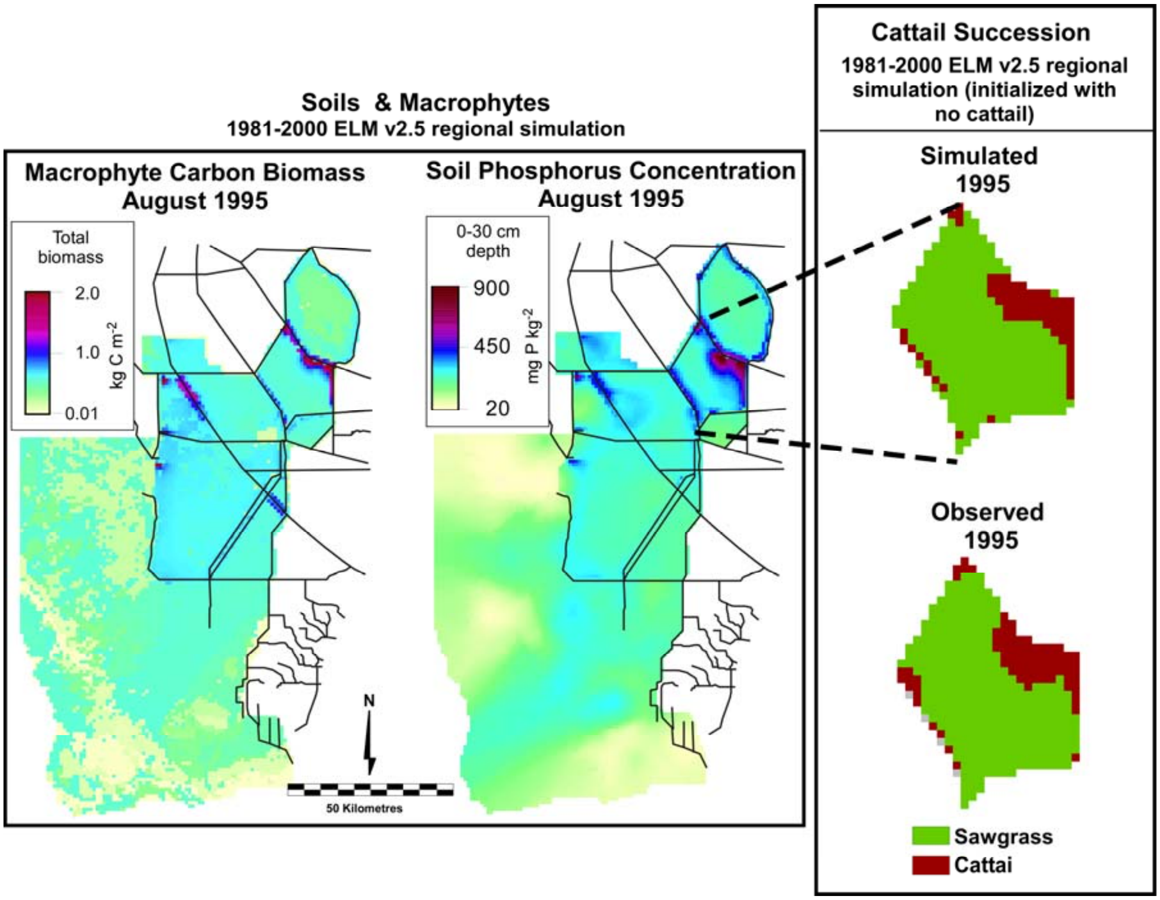


Figure 6.13 Simulated distribution of macrophyte biomass (left), in a snapshot of the mean during the month of August 1995. Soil phosphorus concentration during the same period, with the simulated cattail distribution at that time, compared to the observed distribution of that habitat (data summarized from (Rutchey and Vilchek 1999)).



6.6.4 Validation

With an extension to the period of simulation (to include 1996-2000), the interim ELM v2.2 results demonstrated a “classical validation” of the hydrologic algorithms and data used in ELM. Table 6.4 shows that the median of all (four) statistics comparing simulated to observed stages were similar during the (1981-1995) calibration and (1996-2000) validation periods. Moreover, the (theoretically) improved boundary condition data used to drive ELM v2.2 appeared to somewhat improve the model’s performance during the calibration period, as evidenced in the improved median statistics for the calibration of ELM v2.2 relative to v2.1 (Table 6.4).

As with the “classical” validation of stage predictions, the water column phosphorus predictions were “classically” validated in ELM v2.2. Table 6.5 shows that the median of both statistics comparing simulated to observed surface water phosphorus concentrations were similar during the (1981-1995) calibration and (1996-2000) validation periods. In updating the boundary condition data from ELM v2.1 to v2.2, there was generally little difference in the overall summary of the model’s performance, as evidenced in the similar median statistics for the calibration of ELM v2.2 relative to v2.1 (Table 6.5).

Table 6.4. Statistical evaluation of simulated vs. observed stages during the calibration period of ELM v2.1 and ELM v2.2, and during the validation period of ELM v2.2.

	ELM v2.1 stage calibration				ELM v2.2 stage calibration				ELM v2.2 stage validation			
Site	Bias	RMSE	R2	EFF	Bias	RMSE	R2	Eff	Bias	RMSE	R2	Eff
1-7	0.06	0.16	0.73	0.33	0.05	0.15	0.70	0.30	0.01	0.15	0.72	0.30
1-8T	0.04	0.23	0.67	0.06	0.07	0.19	0.72	0.28	0.02	0.18	0.75	0.39
1-9	0.00	0.15	0.72	0.50	0.02	0.14	0.68	0.42	-0.01	0.14	0.74	0.44
2A-17_B	-0.04	0.24	0.65	0.43	-0.14	0.24	0.69	0.22	-0.16	0.25	0.67	0.12
2A-300_B	-0.05	0.23	0.56	0.46	-0.14	0.25	0.69	0.42	-0.15	0.25	0.67	0.38
3-34	-0.09	0.16	0.84	-1.70	0.18	0.23	0.69	-0.18	0.18	0.22	0.72	-0.20
3-71	-0.09	0.14	0.68	0.35	0.16	0.20	0.60	-0.26	0.22	0.25	0.53	-1.31
3-76	-0.07	0.12	0.66	0.46	0.12	0.16	0.63	0.04	0.15	0.17	0.66	-0.36
3A-10_B	0.06	0.34	0.64	0.51	-0.05	0.14	0.76	0.53	-0.03	0.13	0.76	0.60
3A-11_B	-0.24	0.34	0.78	-1.25	0.21	0.24	0.85	-0.34	0.22	0.25	0.85	-0.46
3A-12_B	0.10	0.25	0.66	0.25	-0.03	0.19	0.59	0.46	-0.04	0.17	0.64	0.51
3A-2_G	0.06	0.26	0.59	0.53	0.02	0.12	0.87	0.85	0.04	0.12	0.87	0.84
3A-28_G	0.29	0.31	0.83	-0.19	-0.14	0.19	0.87	0.65	-0.13	0.17	0.88	0.69
3A-3_G	0.16	0.22	0.87	0.68	-0.07	0.16	0.88	0.85	-0.04	0.15	0.87	0.86
3A-4_G	0.10	0.17	0.84	0.75	0.07	0.14	0.86	0.80	0.08	0.14	0.87	0.80
3A-9_B	-0.02	0.16	0.83	0.82	0.09	0.16	0.86	0.74	0.10	0.15	0.86	0.72
3A-NE_B	0.07	0.25	0.68	0.59	0.00	0.23	0.68	0.67	0.01	0.21	0.71	0.70
3A-NW_B	-0.07	0.25	0.63	0.38	-0.04	0.15	0.75	0.70	-0.03	0.14	0.75	0.72
3A-S_B	0.01	0.20	0.85	0.44	0.07	0.15	0.86	0.71	0.09	0.15	0.86	0.69
3A-SW_B	0.11	0.17	0.82	0.49	0.03	0.11	0.86	0.75	0.04	0.11	0.87	0.79
3B-SE_B	0.03	0.31	0.56	0.46	0.07	0.26	0.71	0.46	0.07	0.23	0.70	0.52
G-1502	0.11	0.28	0.57	0.39	-0.16	0.25	0.74	0.54	-0.10	0.23	0.65	0.55
G-3273	0.15	0.26	0.67	0.39	-0.23	0.30	0.71	0.28	-0.16	0.26	0.64	0.38
G-618_B	0.10	0.18	0.60	0.02	-0.10	0.17	0.71	0.54	-0.07	0.15	0.69	0.61
G-620_B	0.11	0.16	0.80	0.57	-0.07	0.13	0.83	0.73	-0.05	0.11	0.84	0.79
HOLEY_G	0.24	0.29	0.55	-1.48	0.04	0.24	0.63	-0.48	-0.04	0.24	0.46	-0.74
HOLEY1	0.23	0.26	0.64	-0.53	-0.13	0.19	0.75	0.43	-0.20	0.24	0.59	-0.24
HOLEY2	0.19	0.23	0.67	-0.11	-0.12	0.19	0.69	0.48	-0.18	0.24	0.55	0.01
NESRS1	0.02	0.15	0.48	0.43	-0.06	0.12	0.67	0.56	-0.03	0.11	0.63	0.60
NESRS2	0.09	0.18	0.63	0.39	-0.07	0.13	0.70	0.53	-0.05	0.11	0.67	0.59
NESRS3_B	0.07	0.26	0.60	0.29	-0.03	0.21	0.62	0.39	0.01	0.18	0.59	0.45
NP-202	-0.06	0.12	0.81	0.71	-0.01	0.10	0.83	0.83	0.01	0.09	0.85	0.85
NP-203	0.00	0.10	0.79	0.77	-0.04	0.10	0.84	0.80	-0.02	0.09	0.85	0.84
NP-205	0.05	0.19	0.67	0.64	0.02	0.14	0.81	0.80	0.02	0.14	0.80	0.79
NP-206	0.14	0.29	0.57	0.45	-0.15	0.27	0.71	0.54	-0.11	0.23	0.71	0.60
NP-207	-0.05	0.14	0.79	-0.35	0.04	0.10	0.86	0.74	0.04	0.10	0.85	0.71
NP-P33	0.04	0.16	0.55	0.42	-0.06	0.14	0.69	0.57	-0.04	0.12	0.71	0.66
NP-P34	0.10	0.23	0.70	0.29	-0.05	0.17	0.85	0.60	-0.05	0.16	0.85	0.64
NP-P35	0.19	0.25	0.69	-0.95	-0.15	0.22	0.74	-0.36	-0.17	0.23	0.75	-0.59
NP-P36	0.04	0.18	0.47	0.38	0.01	0.11	0.76	0.72	0.03	0.10	0.78	0.74
NP-P38	0.08	0.19	0.70	-0.03	-0.10	0.16	0.84	0.35	-0.11	0.16	0.85	0.29
NP-P44	0.34	0.42	0.68	0.07	-0.37	0.43	0.77	0.02	-0.34	0.41	0.76	0.07
NP-P46	-0.06	0.17	0.63	0.59	-0.03	0.14	0.66	0.34	-0.05	0.14	0.66	0.31
NP-P62	0.12	0.20	0.72	0.32	-0.09	0.16	0.81	0.69	-0.08	0.15	0.80	0.72
NP-P67	0.03	0.13	0.71	0.63	0.02	0.10	0.79	0.77	0.02	0.10	0.80	0.79
NP-P72	0.37	0.44	0.66	-0.67	-0.42	0.47	0.79	-0.34	-0.40	0.44	0.78	-0.28
ROTT.S	0.03	0.15	0.62	0.25	0.18	0.21	0.71	-0.53	0.15	0.20	0.66	0.01
RUTZKE_G	-0.12	0.35	0.48	-1.36	-0.10	0.23	0.73	-0.42	0.00	0.27	0.73	-0.44
SHARK.1_H	-0.02	0.15	0.68	0.64	0.10	0.16	0.76	0.62	0.13	0.18	0.76	0.49
TAMI.40M	0.11	0.29	0.55	-14.55	-0.05	0.22	0.74	0.56	-0.03	0.20	0.75	0.57
Median:	0.06	0.21	0.67	0.39	-0.04	0.17	0.74	0.54	-0.03	0.17	0.75	0.56

Table 6.5. Statistical evaluation of simulated vs. observed phosphorus concentrations in surface waters during the calibration period of ELM v2.1 and ELM v2.2, and during the validation period of ELM v2.2.

		V2.1 Calibration		V2.2 Calibration		V2.2 Validation	
Site	Site type	Bias	RMSE	Bias	RMSE	Bias	RMSE
CA311	Marsh	-0.002	0.009	-0.001	0.001	-0.002	0.003
CA315	Marsh	0.002	0.007	0.002	0.002	0.001	0.001
CA32	Marsh	-0.001	0.008	0.002	0.002	0.003	0.003
CA33	Marsh	-0.017	0.011	-0.012	0.014	-0.005	0.008
CA34	Marsh	0.000	0.003	0.002	0.005	0.004	0.005
CA35	Marsh	-0.020	0.012	-0.031	0.032	-0.019	0.021
CA36	Marsh	-0.023	0.022	-0.021	0.024	-0.008	0.012
CA38	Marsh	-0.002	0.011	0.000	0.002	-0.001	0.005
EP	Marsh	-0.002	0.010	-0.004	0.006	-0.007	0.008
LOX10	Marsh	-0.001	0.008	0.006	0.007	0.005	0.006
LOX11	Marsh	0.004	0.004	0.000	0.001	0.001	0.002
LOX12	Marsh	-0.012	0.002	-0.021	0.021	-0.023	0.024
LOX13	Marsh	0.003	0.003	-0.003	0.004	-0.002	0.003
LOX14	Marsh	-0.014	0.005	0.001	0.002	0.002	0.002
LOX15	Marsh	-0.018	0.007	-0.015	0.015	-0.017	0.017
LOX16	Marsh	-0.016	0.007	-0.006	0.007	-0.008	0.008
LOX3	Marsh	0.000	0.011	0.001	0.005	-0.003	0.004
LOX4	Marsh	-0.022	0.004	-0.001	0.002	0.001	0.004
LOX5	Marsh	0.004	0.006	0.005	0.005	0.003	0.004
LOX6	Marsh	-0.004	0.006	-0.004	0.007	-0.007	0.008
LOX7	Marsh	-0.001	0.006	0.003	0.003	0.004	0.004
LOX8	Marsh	0.002	0.005	0.002	0.003	0.003	0.004
LOX9	Marsh	0.005	0.007	0.006	0.007	0.004	0.004
NE1	Marsh	0.006	0.009	0.007	0.009	0.004	0.004
P33	Marsh	0.002	0.009	0.002	0.005	0.000	0.002
P34	Marsh	-0.008	0.008	-0.004	0.005	-0.006	0.006
P35	Marsh	0.008	0.016	0.011	0.019	0.006	0.009
P36	Marsh	0.028	0.030	0.025	0.041	0.003	0.005
P37	Marsh	-0.012	0.009	-0.002	0.003	-0.004	0.005
TSB	Marsh	-0.002	0.017	-0.002	0.005	-0.004	0.005
C123SR84	Canal	0.004	0.038	0.024	0.028	0.026	0.028
COOPERTN	Canal	0.001	0.006	0.004	0.004	0.005	0.006
L40-1	Canal	-0.001	0.033	0.011	0.029	0.051	0.055
L40-2	Canal	0.017	0.049	0.039	0.050	0.063	0.066
L7	Canal	-0.023	0.047	0.054	0.072	0.000	0.000
S10A	Canal	0.002	0.033	-0.004	0.032	-0.003	0.015
S10C	Canal	0.037	0.064	0.026	0.044	0.021	0.026
S10D	Canal	0.060	0.072	0.061	0.071	0.041	0.045
S10E	Canal	0.050	0.101	0.068	0.078	0.042	0.046
S11A	Canal	-0.013	0.010	-0.013	0.027	0.005	0.012
S11B	Canal	0.008	0.014	0.011	0.028	0.010	0.016
S11C	Canal	0.018	0.034	0.030	0.039	0.028	0.029
S12A	Canal	0.006	0.009	0.009	0.024	0.006	0.009
S12B	Canal	0.005	0.008	0.005	0.017	0.005	0.009
S12C	Canal	0.004	0.005	0.004	0.008	0.003	0.004
S12D	Canal	0.004	0.005	0.005	0.008	0.003	0.004
S144	Canal	0.003	0.009	0.008	0.015	0.008	0.011
S145	Canal	0.000	0.006	0.007	0.012	0.006	0.008
S146	Canal	0.001	0.008	0.007	0.013	0.004	0.007
S151	Canal	0.009	0.016	0.020	0.026	0.013	0.015
S31	Canal	0.010	0.021	0.013	0.019	0.015	0.019
S333	Canal	0.006	0.006	0.006	0.010	0.005	0.006
Median:		0.001	0.009	0.003	0.010	0.003	0.006

Nevertheless, the strict, “classical validation” of a model is ephemeral. As soon as any improvement to the model is made based on scientific advances, the model is no longer truly validated in the classical sense. Classically, a new independent data set must be used to validate the model again. Perhaps more importantly, extending a model simulation period by another year (or 6 months, or 5 years) with an “independent” data set may or may not increase the confidence that users place in the model. As discussed elsewhere, any increased confidence in the model capabilities is largely dependent on how different the new boundary condition forcing data are from those previously input to the model. Instead of attempting to classically validate models, we argue that the most important criteria for user-confidence involves the demonstration of sufficient model performance under an extreme range of conditions – relative to the objectives of the model. Regardless of this debate (see Uncertainty Chapter for discussion of the utility of classical model validation), the ELM performance was enhanced under improved boundary conditions, and the overall performance of the ELM was comparable (if not improved) during the validation period that was driven by input data that were independent of the calibration period.

6.7 Discussion

6.7.1 Model performance summary

Multiple methods were used to evaluate the performance characteristics of this model of greater Everglades ecology. The following summarizes those performance evaluations:

6.7.1.1 Model Objectives – Phosphorus Performance Measures

- P concentration: median bias in predicting surface water TP concentrations was 2 ug l^{-1} for 78 marsh and canal locations in the greater Everglades, whose mean concentrations ranged from less than 10 to more than 100 ug l^{-1}
- P accumulation: along extreme eutrophication gradients, predicted rates of P accumulation in the ecosystems corresponded to field measurements

6.7.1.2 Model Consistency - Hydrology

- Water stage: median bias in predicting stage elevations was -1 cm for 82 marsh locations in the greater Everglades, whose hydroperiod ranged from continuously flooded to rarely flooded; other statistical metrics were comparable to the SFWMM
- Water flows: basin-wide flow budgets were in concordance with those of the SFWMM;
- Water flows: distribution of chloride (CL) concentrations throughout the freshwater Everglades showed patterns of long-term flow regimes that were consistent with our understanding of major flow paths, with a median relative error of -12% in marshes.

6.7.1.3 Model Consistency – Other Ecological Dynamics

- Peat accretion: along extreme eutrophication gradients, predicted rates of peat soil accretion in the ecosystems corresponded to field measurements
- Landscape patterns: regional patterns of macrophyte biomass, soil P concentrations, and (at least subregional) cattail succession corresponded to patterns of observed data

We note here that we have not evaluated the model performance within the mangrove-dominated region (that is delineated in the results map of the CL tracer regional). Thus, application of these ELM Performance Measures within that specific region have an undocumented level of accuracy.

6.7.2 Uncertainty & expectations

As discussed in more detail in the Uncertainty Chapter of this document, there are many factors that result in imperfect agreement of "point-to-point" comparisons between simulated and observed data. Particularly for "water quality" modeling, a critical consideration is the spatial and temporal quality of the inflow boundary nutrient loads, particularly in this managed system that is largely driven by such point sources. The frequency of observed data used to determine nutrient loading to the Everglades system is very sparse relative to the actual water flows; this imposes limits on the ability to simulate short term fluctuations in nutrient dynamics within the system.

At regional scales, it is possible for an improperly structured model to introduce spatial trends in predictive errors. However, such systematic spatial (or temporal) patterns of error were not observed during our extensive calibration process. Moreover, while a simulated value of phosphorus concentration is actually a mean concentration in one square kilometer (of the model grid), the measured phosphorus concentration is an instantaneous observation at a point location, and may not represent the average condition in a heterogeneous area that is subjected to a variety of random processes.

Because of these random errors in data observations, an exact match between simulated and observed "point" monitoring of phosphorus is difficult, and indeed is inappropriate when considering the data quality and expectations. When the number of observation is large, random samples do not increase bias, and thus random errors can be canceled out by aggregation. We thus used temporal aggregation to reduce the effects of random errors in observed data, in order to make the most effective use of the data in understanding long term dynamics: with available data, seasonal to annual (or coarser) temporal scales appear to be the most appropriate scale of aggregation for Everglades water quality dynamics. Decadal responses of the ecosystem are ultimately what we seek to understand and predict in planning for regional Everglades restoration.

6.7.3 Performance refinements

There are limits to model performance that are supported by input data that drive the model, as discussed in the Uncertainty Chapter. However, we also acknowledge that the current version can (and will) be improved within this boundary of expectations. In the Model Refinement Chapter, the near-term and long-term steps in model refinement are presented. We know of a number of relatively straightforward steps that can and will be taken to improve the model performance in the near term.

The overall statistical summaries presented were influenced by a small number of locations where stage or water quality performance is significantly lower than other, even adjacent, locations. In this version, we did not take the time to correct isolated performance "problems" at a handful of locations.

- Big Cypress region: Stage predictions in a number of sites in the Big Cypress National Preserve were generally not simulated as well as other regions in the model domain, likely due to our use of untested land topography data (that was different from that used in the SFWMM).
- WCA-1: While the topography in the marsh of this region is well sampled, we do not know of data that quantifies the magnitude of the topographic berms and associated dense brush vegetation along the edge of this canal; canal-marsh exchanges are significantly effected by these features, which we hope to better quantify. The unique hydraulics associated with this uninterrupted canal encircling the basin are sensitive to relative topographic differences along this feature.
- Mangrove region in south and southwest: Tidal boundary conditions are extremely aggregated in both space and time. Spatial distributions of tidal amplitude are not accounted for in our implementation, nor does the monthly-mean tide, repeating every year, accommodate the observed fluctuations at both fine temporal scales, nor among years.

Importantly, we have not completed our efforts to improve upon the parameter estimates used in the model (see the Uncertainty Chapter, which includes an evaluation of model sensitivity to parameter modifications). Nevertheless, the existing code and data support sufficient model performance to enable users to have reasonable confidence in applying model results to long term planning under new managed conditions.

6.7.4 Conclusions

The ELM performance was rigorously quantified in the greater Everglades system for a multi-decadal period of record (1981 through 2000). The primary Performance Measures intended for ELM v2.5 applications involve those of water quality: phosphorous concentrations and net accumulation throughout the greater Everglades region. Quantitative performance assessments provided strong, cumulative evidence that ELM could be effectively used to evaluate relative differences in those Performance Measures within the regional system. With other predicted ecological attributes and rates being consistent with available observations, there is cumulative, strong evidence of model skill in predicting phosphorus trends in the regional Everglades landscape at the relevant decadal time scales.

6.8 Literature cited

Craft, C. B., and C. J. Richardson. 1993. Peat accretion and N, P, and organic C accumulation in nutrient-enriched and unenriched Everglades peatlands. Ecological Applications **3**:446-458.

ELM_Team. 2002. Calibration performance of ELM v2.1a: 1979-1995 water quality and hydrology. SFWMD, West Palm Beach, FL. http://www.sfwmd.gov/org/wrp/elm/results/cal_ver/elm2.1/ELMcalibAnalysis_draft.pdf.

Fitz, H. C., E. B. DeBellevue, R. Costanza, R. Boumans, T. Maxwell, L. Wainger, and F. H. Sklar. 1996. Development of a general ecosystem model for a range of scales and ecosystems. Ecological Modelling **88**:263-295.

Fitz, H. C., and F. H. Sklar. 1999. Ecosystem analysis of phosphorus impacts and altered hydrology in the Everglades: a landscape modeling approach. Pages 585-620 *in* K. R. Reddy, G. A. O'Connor, and C. L. Schelske, editors. Phosphorus Biogeochemistry in Subtropical Ecosystems. Lewis Publishers, Boca Raton, FL.

Hargrove, W. W., and F. M. Hoffman. 2005. Potential of multivariate quantitative methods for delineation and visualization of ecoregions. Environmental Management **34**:S39 - S60.

Hoffman, F. M., W. W. Hargrove, D. J. Erickson, and R. J. Oglesby. 2005. Using clustered climate regimes to analyze and compare predictins from fully coupled general circulation models. Earth Interactions **9**:1-27.

Janssen, P. H. M., and P. S. C. Heuberger. 1995. Calibration of process-oriented models. Ecological Modelling **83**:55-66.

Limno-Tech, I. 1997. Everglades Water Quality Model calibration report. Final report prepared for the South Florida Water Management District, West Palm Beach, FL.

Mayer, D. G., and D. G. Butler. 1993. Statistical validation. Ecological Modelling **68**:21-32.

McCormick, P. V., S. Newman, S. L. Miao, D. E. Gawlik, D. Marley, K. R. Reddy, and T. D. Fontaine. 2002. Effects of anthropogenic phosphorus inputs on the Everglades. Pages 83-126 *in* J. W. Porter and K. G. Porter, editors. The Everglades, Florida Bay, and Coral Reefs of the Florida Keys: An ecosystem sourcebook. CRC Press, Boca Raton, FL.

Raghunathan, R., T. Slawecki, T. D. Fontaine, Z. Chen, D. W. Dilks, V. J. Bierman, and S. Wade. 2001. Exploring the dynamics and fate of total phosphorus in the Florida Everglades using a calibrated mass balance model. Ecological Modelling **142**:247-259.

Reddy, K. R., R. D. Delaune, W. F. Debusk, and M. S. Koch. 1993. Long-term nutrient accumulation rates in the Everglades. Soil Science Society of America Journal **57**:1147-1155.

Robbins, J. A., S. Newman, C. W. Holmes, and K. R. Reddy. 2004. Phosphorus accumulation in soils along a nutrient gradient in Water Conservation Area 2A, South Florida. *in* G. R. Best and D. J. Hayes, editors. First National Conference on Ecosystem Restoration. US Geological Survey, US Army Corps of Engineers, Lake Buena Vista, FL.

Rutchey, K., and L. Vilchek. 1999. Air photointerpretation and satellite imagery analysis techniques for mapping cattail coverage in a northern Everglades impoundment. Journal of Photogrammetric Engineering and Remote Sensing **65**:185-191.

USACE, and SFWMD. 1999. Central and Southern Florida Project, Comprehensive Review Study, Final Integrated Feasibility Report and Programmatic Environmental Impact Statement. US Army Corps of Engineers and South Florida Water Management District.

Villa, F., A. A. Voinov, H. C. Fitz, and R. Costanza. 2004. Calibration of large spatial models: a multi-stage, multi-objective optimization technique. Pages 77-116 *in* R. Costanza and A. A. Voinov, editors. Landscape Simulation Modeling: A Spatially Explicit, Dynamic Approach. Springer Verlag, New York, New York.

6.9 *Appendix A: Computational methods for statistics*

Although numerous methods exist for analyzing and summarizing model performance, there is no consensus in the modeling community on a standard analytical suite for hydrology and ecological (incl. water quality) models. It appears most useful to use a variety of methods to evaluate model performance, as no single statistic can fully capture all of the important characteristics of a comparison between the simulated and observed data. We employed the below methods to estimate Bias, RMSE, R^2, and NS Efficiency in assessing some aspects of the model performance relative to observed data.

Bias:

$$\text{Bias} = \frac{\sum (x - y)}{n}$$

Where x is the field-observation values, y is the model-prediction values, and n is the number of observations.

Bias is calculated as the mean differences between paired modeled and observed values. It is a measure of how biased the overall values simulated by the model from the observed values. The bias should be as close to zero as possible.

Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{\sum (y - x)^2}{n}}$$

Where x is the field-observation values and y is the model-prediction values.

RMSE is the square root of the average values of the prediction errors squared. RMSE measures the discrepancy between modeled and observed values on an individual level to indicate accuracy of model predictions. Because of the quadratic term, RMSE gives greater weight to larger discrepancies than smaller ones. The RMSE should be as close to zero as possible.

Pearson product-moment correlation coefficient (R^2):

$$R^2 = \left(\frac{\sum (y - y_m)(x - x_m)}{\sqrt{\sum (y - y_m)^2 \sum (x - x_m)^2}} \right)^2$$

Where x_m is the observed mean of x (calculated as Σx/n), and y_m is the model-predicted mean of observed y (calculated as Σy/n).

The R^2 measure the degree of linear association between x and y (i.e., field observation and model predictions). It represents the amount of variability of one variable that is explained by correlating it with another variable. Depending on the strength of the linear relationships, the R^2 varies from 0.0 to 1.0, with 1.0 indicating a perfect fit.

Nash-Sutcliffe Efficiency (Eff):

$$\text{Eff} = 1 - \frac{\sum (y - x)^2}{\sum (x - x_m)^2},$$

Where x_m is the mean of the observed x, and y is the model prediction.

Like correlation coefficient, model efficiency is another overall indication of goodness of fit (Mayer and Butler 1993, Janssen and Heuberger 1995). Efficiency is equal to one minus the sum of squared prediction errors divided by the sum of squared deviation of observed values from the mean. It represents the amount of variability of one variable that is explained by modeled values. A model efficiency of 1.0 indicates a perfect fit between modeled and observed values, and a efficiency of 0.0 indicates the fit to $y = x$ is no better than $x = x_m$.

6.10 Appendix B: Time series & CFDs: TP (separate pdf)

Figures B.1 – B.78. Time series plots of water column total phosphorus (TP) concentration and their associated Cumulative Frequency Distributions (CFD) for the period of record 1981-2000 at each monitoring location. The sequence of the figures is based on geographic location of marsh sites, starting in northwest, moving towards the southeast; following the set of plots of all marsh sites, the canal monitoring sites are similarly sequenced. A map of all sites is provided in the Model Performance Chapter.

The constant dashed line indicates the TP field sampling Detection Limit (DL = 4 ug l^{-1} for the model period of record), which was the minimum value used for observed data in plots and statistics. To enable equivalent comparisons, any simulated value which was below the DL was set equal to the DL. The model grid cell column and row locations (col_row) or canal reach identifier (single integer) are shown in parentheses of each plot's title.

a) All data were aggregated into arithmetic mean values by wet and dry seasons within water years; the continuous lines pass through mean of all daily data points for each season; the mean of paired simulated and observed values are shown in red boxes and black diamonds, respectively; the 95% Confidence Interval (CI) of the paired means are shown by the "__" symbols in the red for the model and black for the observed data.

b) All data aggregated into arithmetic mean values by water year, with the same treatment as in plot a).

c) The CFDs of the simulated and observed (raw, un-aggregated) data; the 95% confidence interval for observed data is shown in the dashed black lines. Note that only paired simulated and observed data points are used.

6.11 Appendix C: Time series & CFDs: stage (separate pdf)

Figures C.1 – C.82. Plots of stage hydrographs and their associated Cumulative Frequency Distributions (CFD) for the period of record 1981-2000 at each monitoring location. The sequence of the figures is based on geographic location, starting in the northwest, moving towards the southeast. A map of all sites is provided in the Model Performance Chapter.

The red dashed line in the stage hydrographs is the model grid cell's land surface elevation, which is a time-varying output variable of the model. The model grid cell column and row locations are shown in parentheses (col_row) of each plot's title.

a) All data, with no temporal aggregation, of daily observations (black dots) and model results (red line).

b) All data were aggregated into arithmetic mean values by wet and dry seasons within water years; the continuous lines pass through mean of all daily data points for each season; the mean of paired simulated & observed values are shown in red boxes and black diamonds, respectively; the 95% Confidence Interval (CI) of the paired means are shown by the "__" symbols in the red for the model and black for the observed data.

c) All data aggregated into arithmetic mean values by water year, with the same treatment as in plot b).

d) The cumulative frequency distributions of the simulated and observed (raw, un-aggregated) data; the 95% confidence interval for observed data is shown in the dashed black lines. Note that only paired simulated and observed data points are used.

6.12 Appendix D: Water budgets, ELM & SFWMM

Figures D.1 – D.5. Budget comparisons between ELM and SFWMM for the following basins: WCA-1, WCA-2A, WCA-2B, WCA-3A, and WCA-3B. Each numbered figure contains four graphs:

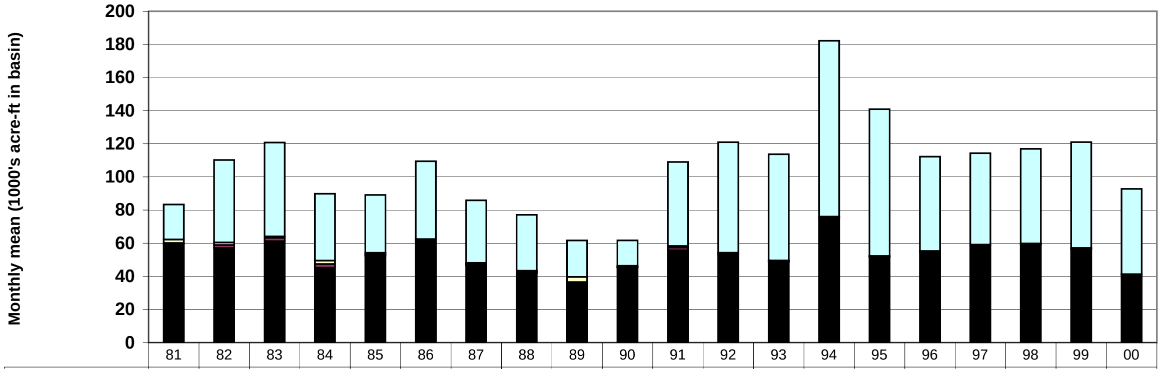
a) ELM inflows

b) ELM outflows.

c) Differences, inflows to SFWMM & ELM

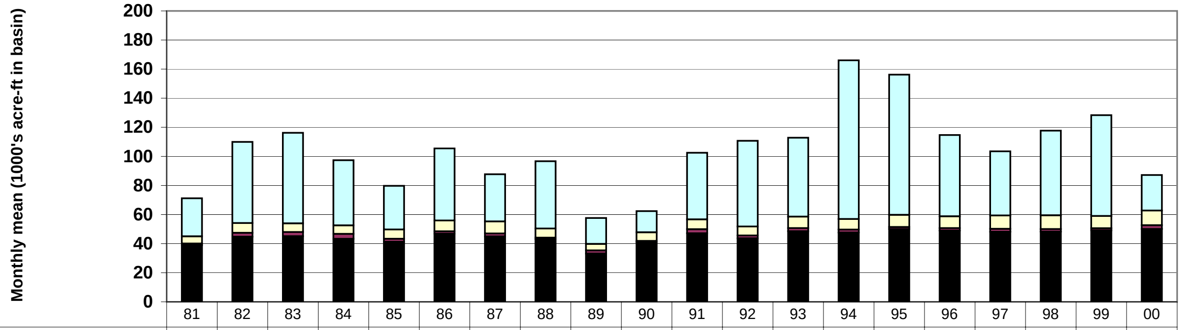
d) Differences, outflows from SFWMM & ELM

WCA1 ELM Inflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟦 Structure In 1	21	50	57	40	35	47	38	34	22	15	51	67	64	106	89	57	55	57	64	51
🟨 Groundwater In 1	2	2	1	2	1	0	1	1	3	1	1	1	0	0	0	0	0	0	0	0
🟪 Seepage In 1	1	2	2	2	1	1	1	1	1	0	2	1	1	1	1	1	0	1	1	0
■ Rain 1	60	57	61	45	52	61	47	42	36	45	56	53	49	75	52	55	59	59	56	41

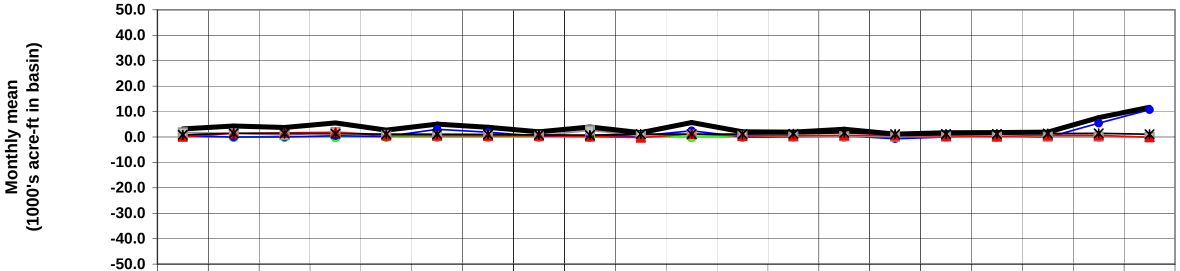
WCA1 ELMOutflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟦 Structure Out 1	26	56	62	45	30	50	32	46	18	15	46	59	54	109	96	56	44	58	69	24
🟨 Groundwater Out 1	5	7	6	6	6	7	8	6	4	6	7	6	8	7	8	8	9	9	8	10
🟪 Seepage Out 1	1	3	3	3	2	2	2	1	2	1	3	2	2	2	1	2	2	2	2	3
■ Evapotranspiration 1	39	45	45	43	41	47	45	43	33	41	47	44	49	48	50	49	48	48	49	50

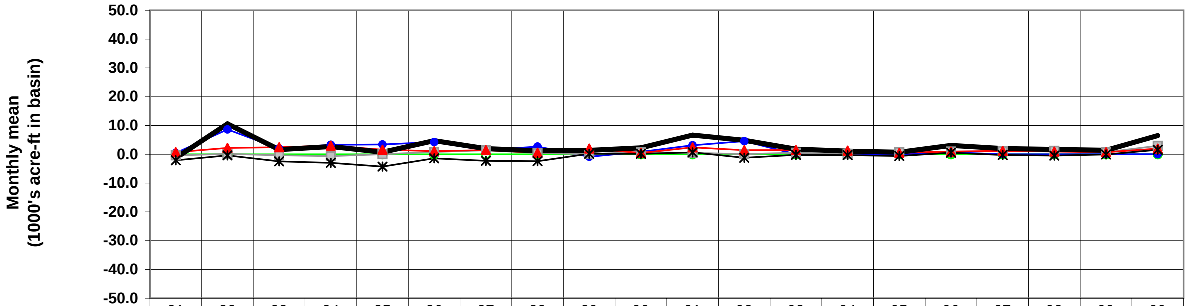
WCA1 ELM-SFWMM Inflow Differences

WCA1 ELM-SFWMM Inflow Differences



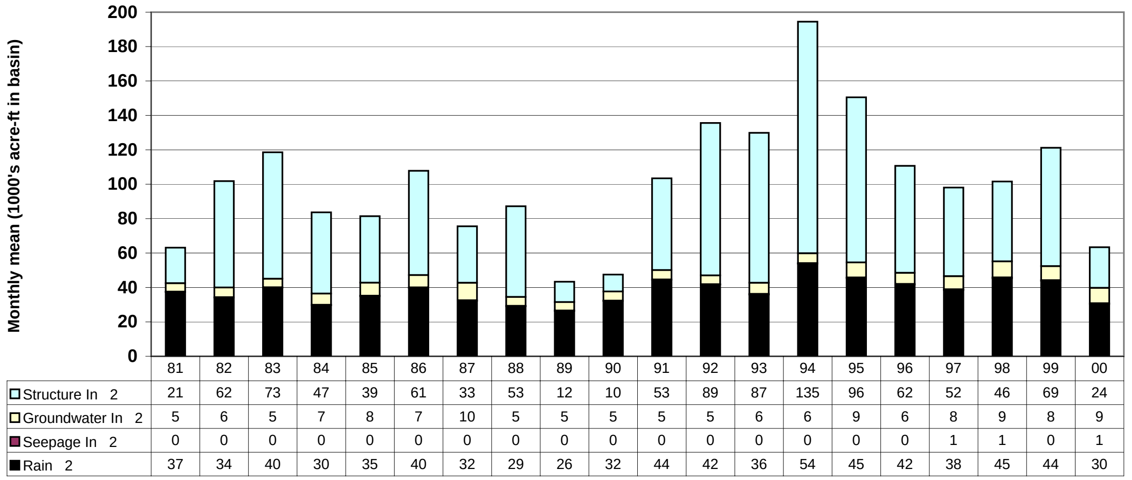
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
–○– Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
–●– Structure	0.4	0.0	0.0	0.5	0.2	3.0	1.9	0.3	0.4	0.4	2.4	0.0	0.0	0.5	-0.7	0.0	0.2	-0.1	5.4	10.7
–■– Groundwater	1.8	1.4	0.7	2.0	0.8	0.4	0.5	0.4	2.6	0.4	0.9	0.5	0.2	0.3	0.1	0.1	0.1	0.1	0.1	0.0
–▲– Seepage	0.1	1.4	1.6	1.8	0.5	0.6	0.5	0.5	0.2	-0.2	1.2	0.4	0.5	0.6	0.6	0.4	0.2	0.6	0.6	-0.1
–✳– Rain	0.8	1.5	1.4	1.3	1.1	1.1	1.0	0.8	0.7	1.1	1.1	1.1	1.2	1.6	1.1	1.1	1.2	1.3	1.4	1.1
━ Sum	3.2	4.3	3.7	5.5	2.7	5.1	3.8	2.0	3.9	1.7	5.7	2.0	1.9	3.0	1.1	1.7	1.7	1.9	7.6	11.7

WCA1 ELM-SFWMM Outflow Differences



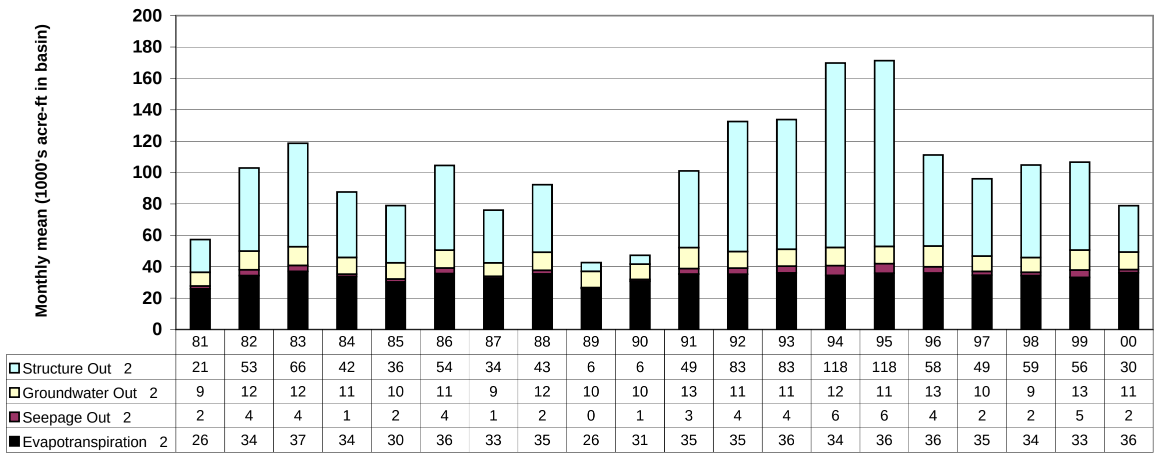
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
–○– Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
–●– Structure	0.4	8.6	2.0	3.3	3.4	4.2	1.4	2.7	-0.9	0.9	3.1	4.6	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.0
–■– Groundwater	-0.4	0.1	-0.3	-0.6	0.0	0.8	1.4	0.4	0.1	0.8	0.5	0.1	0.5	0.1	0.8	1.0	1.0	1.2	0.7	2.9
–▲– Seepage	0.7	2.2	2.5	3.0	1.6	1.1	1.4	0.5	2.0	0.5	2.4	1.4	1.4	1.3	0.5	0.9	1.2	0.9	0.7	2.0
–✳– ET	-2.1	-0.4	-2.5	-3.0	-4.3	-1.4	-2.3	-2.4	0.2	0.2	0.7	-1.2	-0.2	-0.3	-0.6	0.5	-0.3	-0.5	-0.1	1.6
━ Sum	-1.3	10.6	1.6	2.7	0.7	4.7	2.0	1.2	1.4	2.2	6.7	4.9	1.8	1.1	0.7	3.1	2.0	1.7	1.4	6.5

WCA2A ELM Inflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
□ Structure In 2	21	62	73	47	39	61	33	53	12	10	53	89	87	135	96	62	52	46	69	24
□ Groundwater In 2	5	6	5	7	8	7	10	5	5	5	5	5	6	6	9	6	8	9	8	9
■ Seepage In 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1
■ Rain 2	37	34	40	30	35	40	32	29	26	32	44	42	36	54	45	42	38	45	44	30

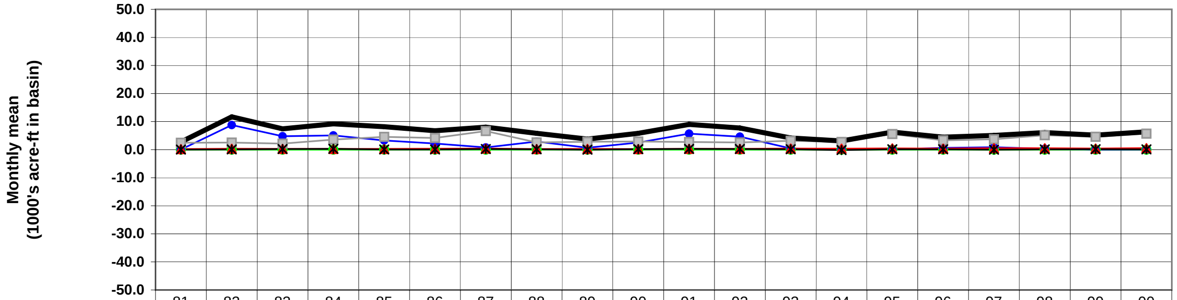
WCA2A ELMOutflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
□ Structure Out 2	21	53	66	42	36	54	34	43	6	6	49	83	83	118	118	58	49	59	56	30
□ Groundwater Out 2	9	12	12	11	10	11	9	12	10	10	13	11	11	12	11	13	10	9	13	11
■ Seepage Out 2	2	4	4	1	2	4	1	2	0	1	3	4	4	6	6	4	2	2	5	2
■ Evapotranspiration 2	26	34	37	34	30	36	33	35	26	31	35	35	36	34	36	36	35	34	33	36

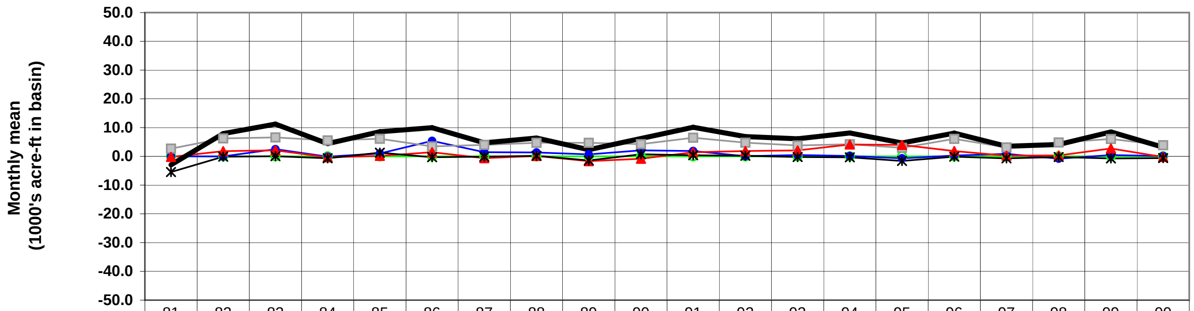
WCA2A ELM-SFWMM Inflow Differences

WCA2A ELM-SFWMM Inflow Differences



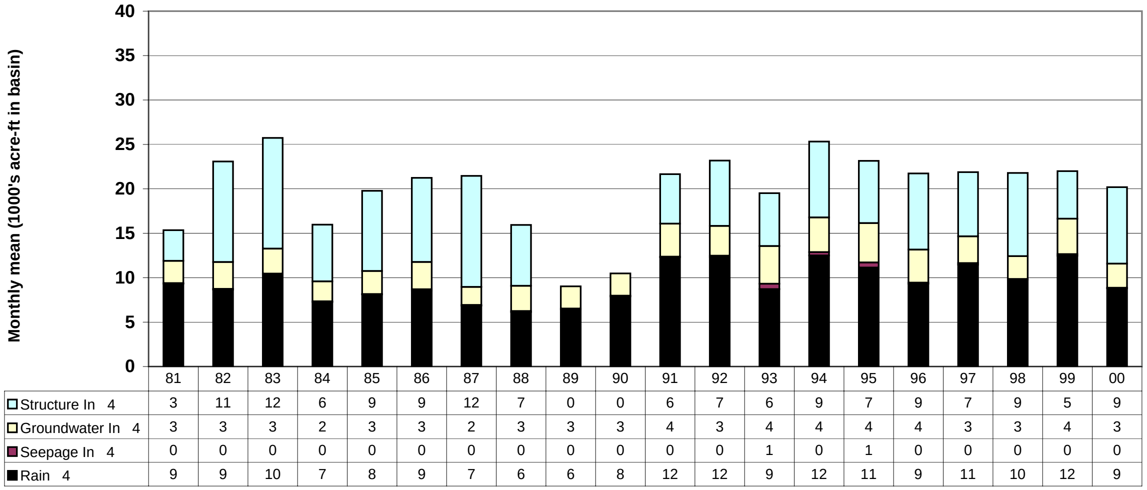
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	8.8	4.8	5.1	3.3	2.2	0.8	2.9	0.7	2.5	5.7	4.6	0.4	0.1	0.1	0.6	0.9	0.3	0.0	0.0
Groundwater	2.5	2.5	2.2	3.6	4.6	4.2	6.6	2.6	2.8	2.9	2.7	2.6	3.2	2.7	5.6	3.3	3.7	5.1	4.6	5.7
Seepage	0.2	0.3	0.3	0.3	0.3	0.3	0.4	0.2	0.3	0.2	0.3	0.3	0.4	0.4	0.5	0.4	0.5	0.5	0.4	0.5
Rain	0.0	0.1	0.2	0.3	0.1	0.1	0.2	0.1	0.0	0.2	0.3	0.2	0.1	-0.1	0.1	0.1	0.0	0.1	0.2	0.1
Sum	2.7	11.7	7.4	9.2	8.2	6.7	8.0	5.8	3.8	5.8	9.0	7.7	4.2	3.1	6.3	4.5	5.1	6.1	5.2	6.4

WCA2A ELM-SFWMM Outflow Differences



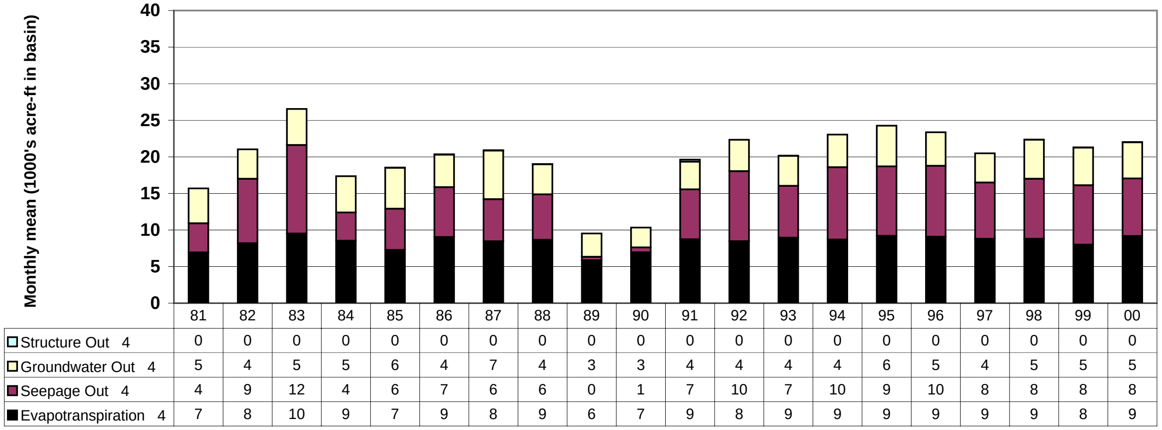
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	0.0	2.5	-0.1	1.0	5.4	1.5	1.4	0.7	2.1	1.9	0.1	0.5	0.1	-0.7	0.3	0.9	-0.8	0.5	0.1
Groundwater	2.7	6.2	6.6	5.5	6.1	3.5	4.0	4.7	4.7	4.2	6.5	4.8	3.8	4.2	3.0	6.1	3.1	4.8	6.0	3.9
Seepage	-0.2	1.8	2.1	-0.3	0.2	1.4	-0.6	0.1	-1.7	-0.8	1.4	1.9	2.1	4.1	4.0	1.8	0.2	0.3	2.7	-0.2
ET	-5.5	-0.1	0.0	-0.7	1.3	-0.3	-0.1	0.2	-1.4	0.7	0.4	0.2	-0.2	-0.3	-1.6	-0.1	-0.7	-0.2	-0.7	-0.6
Sum	-3.0	7.9	11.2	4.4	8.6	10.0	4.7	6.4	2.3	6.2	10.1	6.9	6.1	8.1	4.7	8.1	3.5	4.1	8.5	3.2

WCA2B ELM Inflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
☐ Structure In 4	3	11	12	6	9	9	12	7	0	0	6	7	6	9	7	9	7	9	5	9
☐ Groundwater In 4	3	3	3	2	3	3	2	3	3	3	4	3	4	4	4	4	3	3	4	3
■ Seepage In 4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
■ Rain 4	9	9	10	7	8	9	7	6	6	8	12	12	9	12	11	9	11	10	12	9

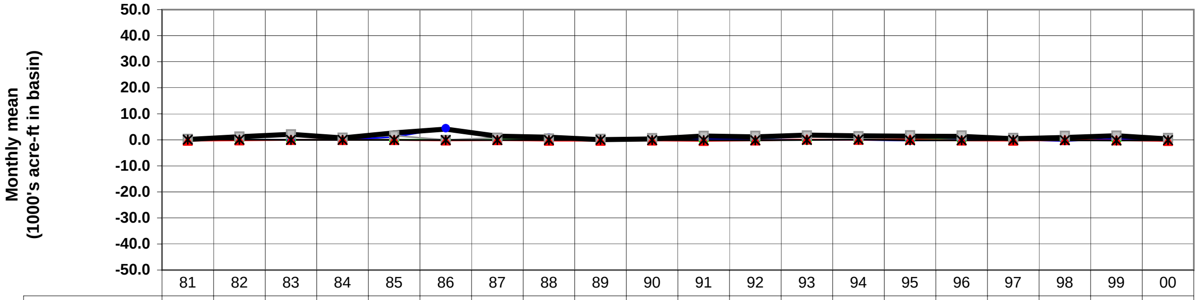
WCA2B ELMOutflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
☐ Structure Out 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
☐ Groundwater Out 4	5	4	5	5	6	4	7	4	3	3	4	4	4	4	6	5	4	5	5	5
■ Seepage Out 4	4	9	12	4	6	7	6	6	0	1	7	10	7	10	9	10	8	8	8	8
■ Evapotranspiration 4	7	8	10	9	7	9	8	9	6	7	9	8	9	9	9	9	9	9	8	9

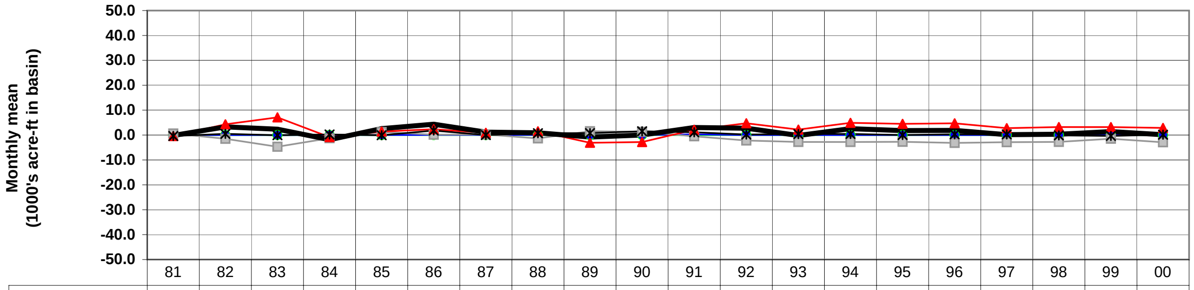
WCA2B ELM-SFWMM Inflow Differences

WCA2B ELM-SFWMM Inflow Differences



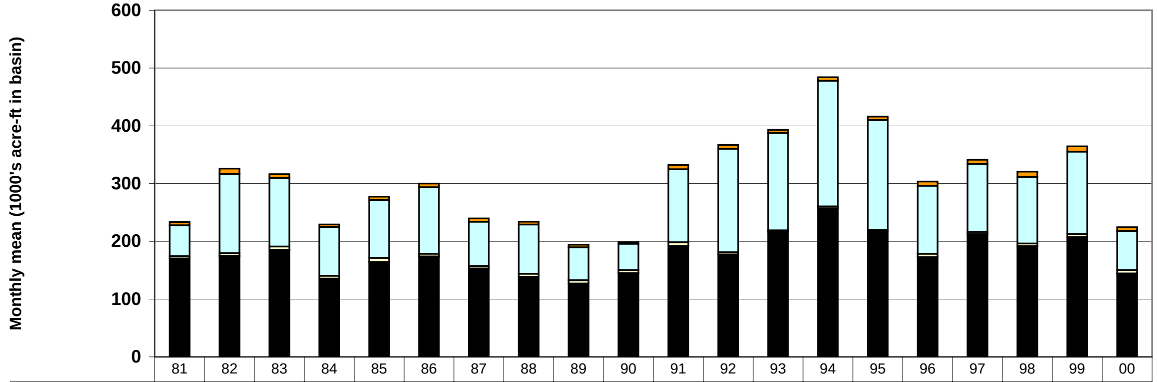
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	0.0	0.0	-0.1	1.1	4.5	0.7	0.9	0.0	0.0	0.4	0.0	0.0	0.0	-0.3	0.0	0.0	-0.4	0.4	0.1
Groundwater	0.4	1.3	2.1	0.8	1.6	0.0	0.8	0.5	0.4	0.7	1.6	1.6	1.7	1.4	1.8	1.7	0.8	1.6	1.6	0.7
Seepage	-0.3	-0.2	0.0	0.0	0.0	-0.2	-0.1	-0.3	-0.3	-0.2	-0.4	-0.2	0.1	0.0	0.1	-0.2	-0.3	-0.1	-0.2	-0.4
Rain	0.1	0.1	0.0	0.0	0.0	-0.1	-0.1	-0.1	0.0	-0.1	-0.1	-0.2	0.0	0.1	-0.2	-0.2	0.0	-0.2	-0.3	-0.1
Sum	0.1	1.2	2.1	0.7	2.7	4.1	1.4	1.0	0.1	0.3	1.5	1.1	1.8	1.5	1.4	1.3	0.4	0.9	1.6	0.4

WCA2B ELM-SFWMM Outflow Differences



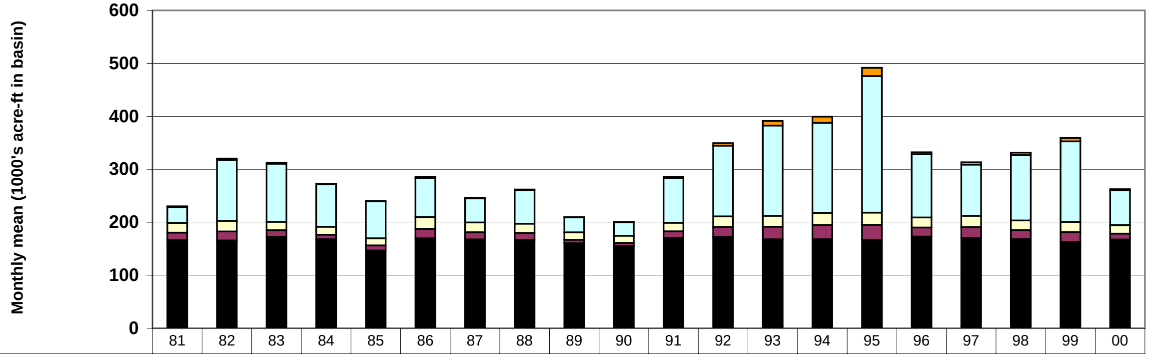
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Groundwater	0.6	-1.5	-4.7	-1.3	1.2	0.2	0.2	-1.3	1.5	1.3	-0.5	-2.2	-2.8	-2.8	-2.7	-3.2	-2.9	-2.7	-1.5	-2.9
Seepage	-0.4	4.3	7.1	-0.8	1.4	2.3	0.9	1.5	-3.1	-2.8	2.1	4.7	2.2	4.9	4.5	4.7	2.8	3.2	3.2	2.9
ET	-0.4	0.5	-0.1	0.2	-0.1	1.8	0.0	0.7	0.8	1.5	1.1	0.2	0.4	0.4	0.0	0.3	0.3	-0.1	-0.3	0.1
Sum	-0.2	3.3	2.3	-1.9	2.6	4.4	1.2	0.9	-0.8	0.0	3.0	2.8	-0.2	2.6	1.8	1.9	0.2	0.3	1.4	0.1

WCA3A ELM Inflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟧 Overland In 3	6	9	7	4	6	6	6	5	4	3	7	7	6	6	6	7	7	9	9	6
□ Structure In 3	53	137	119	85	100	115	77	85	57	45	126	179	168	217	190	118	118	115	143	67
□ Groundwater In 3	4	4	6	5	7	5	4	5	6	5	6	4	3	3	3	6	4	5	6	6
🟪 Seepage In 3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
■ Rain 3	170	175	185	135	164	174	153	139	127	145	192	177	216	257	217	173	212	192	207	144

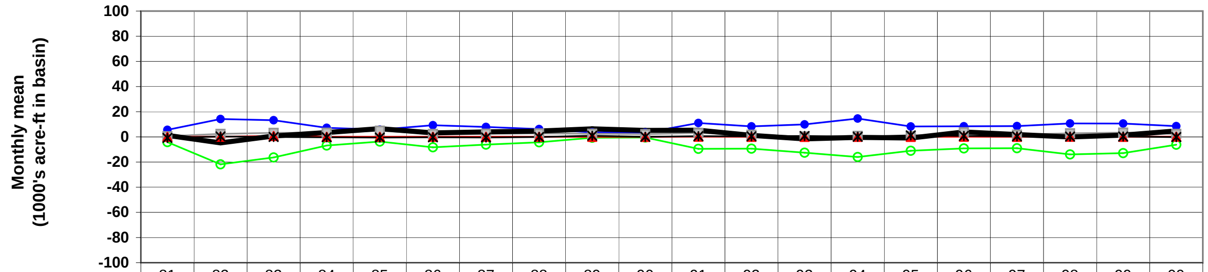
WCA3A ELMOutflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟧 Overland Out 3	1	3	2	1	1	2	1	1	0	0	2	5	9	12	16	4	4	5	6	2
□ Structure Out 3	30	115	110	80	70	74	45	63	28	26	84	134	170	170	258	120	97	123	152	66
□ Groundwater Out 3	18	20	16	15	13	22	18	17	14	13	16	20	21	23	23	19	21	18	19	16
🟪 Seepage Out 3	14	17	13	9	10	18	14	13	7	6	12	19	24	27	28	17	20	17	19	11
■ Evapotranspiration 3	166	166	172	168	147	170	168	167	160	155	171	172	168	168	167	173	171	168	163	167

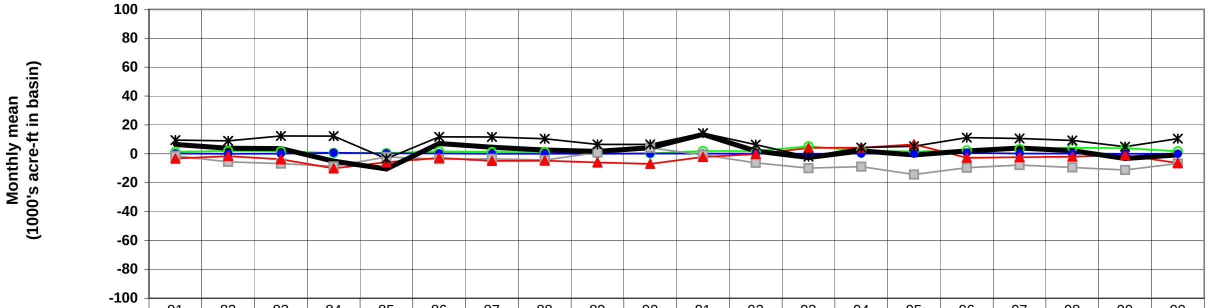
WCA3A ELM-SFWMM Inflow Differences

WCA3A ELM-SFWMM Inflow Differences



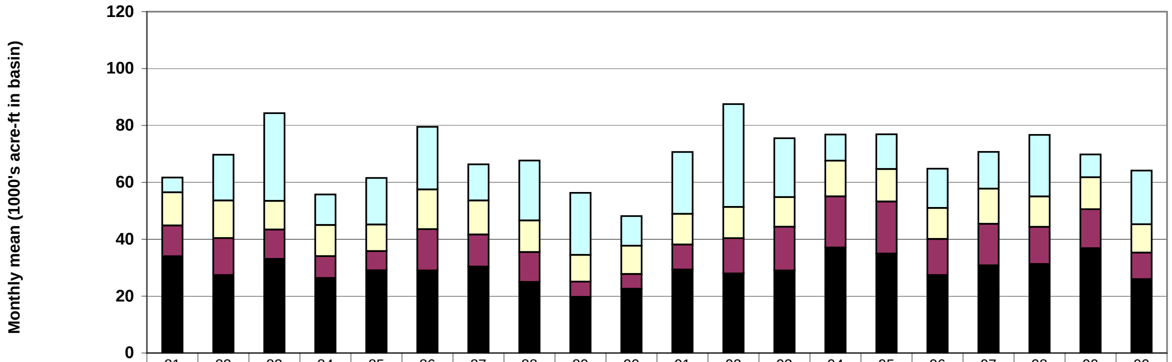
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	-4	-22	-16	-7	-4	-8	-6	-4	-1	0	-9	-9	-13	-16	-11	-9	-9	-14	-13	-6
Structure	6	14	13	7	6	9	8	6	4	3	11	8	10	15	8	8	9	11	11	9
Groundwater	1	2	3	3	5	2	3	3	3	2	3	2	0	1	1	4	2	3	3	2
Seepage	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rain	-1	0	0	0	-1	0	0	0	1	0	1	0	1	0	1	1	0	0	0	0
Sum	1	-5	1	3	7	3	4	5	6	5	5	1	-2	0	-1	4	2	0	1	5

WCA3A ELM-SFWMM Outflow Differences



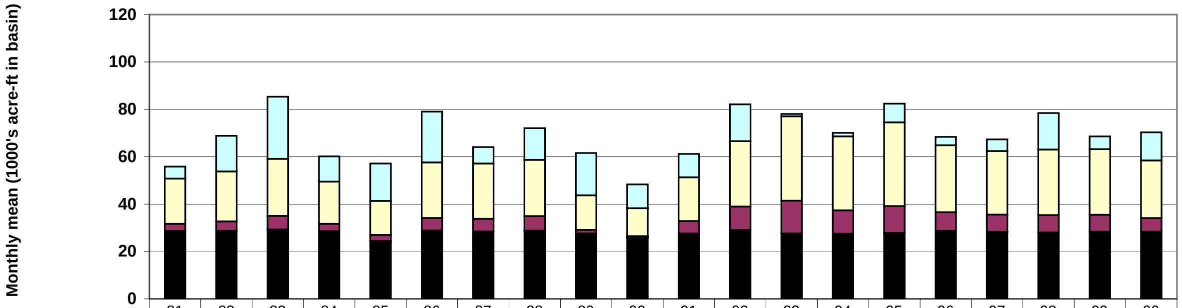
	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	1	2	2	1	1	2	1	1	0	0	2	2	5	2	2	2	3	4	4	2
Structure	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Groundwater	-1	-6	-7	-9	-2	-4	-3	-4	1	4	-1	-6	-10	-9	-14	-10	-8	-9	-11	-7
Seepage	-3	-2	-4	-10	-6	-3	-5	-5	-6	-7	-2	0	4	4	7	-3	-2	-2	-1	-6
ET	9	9	12	12	-3	12	12	10	7	7	14	6	-2	4	5	11	11	9	5	10
Sum	6	4	4	-5	-10	7	5	3	2	4	13	2	-3	2	-1	2	4	2	-3	-1

WCA3B ELM Inflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟧 Overland In 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
🟦 Structure In 5	5	16	31	11	16	22	13	21	22	10	22	36	21	9	12	14	13	22	8	19
🟨 Groundwater In 5	12	13	10	11	9	14	12	11	9	10	11	11	10	13	11	11	12	11	11	10
🟪 Seepage In 5	11	13	10	8	7	15	11	11	5	5	9	12	15	18	18	13	15	13	14	9
⬛ Rain 5	34	27	33	26	29	29	30	25	20	23	29	28	29	37	35	27	31	31	37	26

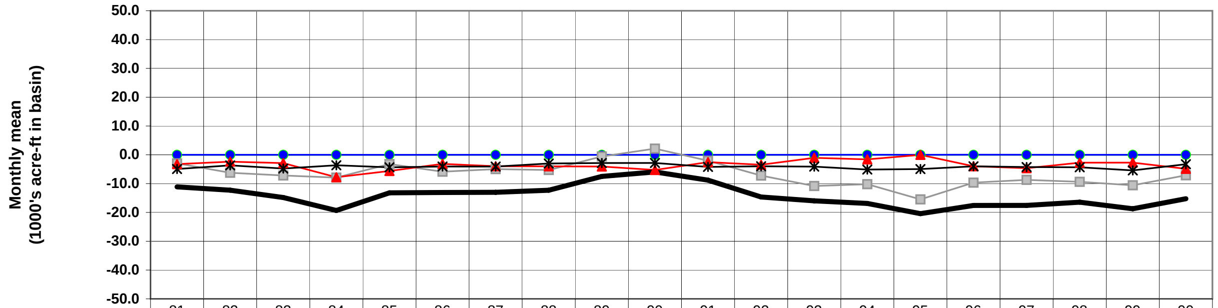
WCA3B ELMOutflows



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
🟧 Overland Out 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
🟦 Structure Out 5	5	15	26	11	16	21	7	13	18	10	10	16	1	1	8	4	5	15	5	12
🟨 Groundwater Out 5	19	21	24	18	14	23	23	24	15	12	18	28	36	31	35	28	27	28	28	24
🟪 Seepage Out 5	3	4	6	3	3	5	5	6	1	1	5	10	14	10	11	8	7	7	7	6
⬛ Evapotranspiration 5	29	29	29	29	24	29	28	29	28	26	28	29	28	28	28	29	28	28	28	28

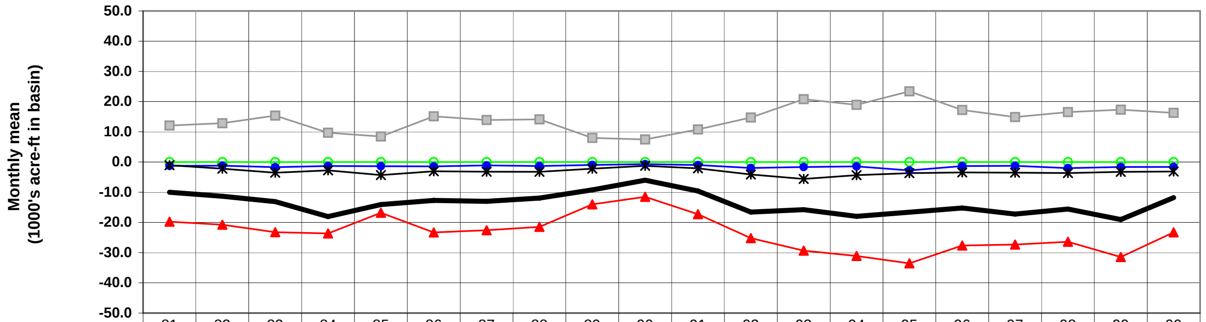
WCA3B ELM-SFWMM Inflow Differences

WCA3B ELM-SFWMM Inflow Differences



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Groundwater	-3.0	-6.2	-7.2	-7.9	-3.2	-5.9	-5.0	-5.3	-0.6	2.1	-2.1	-7.2	-10.8	-10.2	-15.5	-9.7	-8.7	-9.4	-10.6	-7.1
Seepage	-3.2	-2.4	-2.9	-7.8	-5.6	-3.1	-4.0	-4.0	-4.1	-5.3	-2.5	-3.5	-1.1	-1.6	0.0	-4.0	-4.6	-2.7	-2.7	-4.9
Rain	-4.9	-3.7	-4.8	-3.6	-4.4	-4.1	-4.1	-3.0	-2.8	-2.8	-4.2	-4.0	-4.1	-5.1	-4.9	-4.0	-4.2	-4.3	-5.4	-3.3
Sum	-11.1	-12.3	-14.8	-19.3	-13.2	-13.1	-13.0	-12.3	-7.5	-5.9	-8.8	-14.6	-16.0	-16.9	-20.4	-17.6	-17.6	-16.4	-18.7	-15.2

WCA3B ELM-SFWMM Outflow Differences



	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	00
Overland	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Structure	-1.3	-1.2	-1.7	-1.3	-1.4	-1.4	-1.1	-1.3	-1.0	-0.7	-1.0	-2.0	-1.7	-1.4	-2.8	-1.3	-1.3	-2.0	-1.6	-1.6
Groundwater	12.1	12.8	15.4	9.7	8.4	15.1	13.9	14.1	8.0	7.5	10.8	14.7	20.8	18.9	23.4	17.2	14.9	16.5	17.3	16.3
Seepage	-19.8	-20.8	-23.2	-23.6	-16.8	-23.3	-22.6	-21.5	-14.0	-11.5	-17.3	-25.2	-29.3	-31.1	-33.5	-27.6	-27.3	-26.4	-31.4	-23.3
ET	-1.1	-2.2	-3.5	-2.8	-4.3	-3.1	-3.2	-3.2	-2.2	-1.3	-2.1	-4.1	-5.6	-4.4	-3.7	-3.5	-3.5	-3.7	-3.3	-3.2
Sum	-10.0	-11.3	-13.1	-18.1	-14.1	-12.7	-13.0	-11.9	-9.2	-6.0	-9.6	-16.6	-15.8	-18.0	-16.6	-15.2	-17.2	-15.6	-19.0	-11.8

6.13 Appendix E: Time series & CFDs: CL (separate pdf)

Figures E.1 – E.78. Time series plots of water column chloride (CL) concentration and their associated Cumulative Frequency Distributions (CFD) for the period of record 1981-2000 at each monitoring location. The sequence of the figures is based on geographic location of marsh sites, starting in northwest, moving towards the southeast; following the set of plots of all marsh sites, the canal monitoring sites are similarly sequenced. A map of all sites is provided in the Model Performance Chapter.

The model grid cell column and row locations (col_row) or canal reach identifier (single integer) are shown in parentheses of each plot's title.

a) All data were aggregated into arithmetic mean values by wet and dry seasons within water years; the continuous lines pass through mean of all daily data points for each season; the mean of paired simulated & observed values are shown in red boxes and black diamonds, respectively; the 95% Confidence Interval (CI) of the paired means are shown by the "__" symbols in the red for the model and black for the observed data.

b) All data aggregated into arithmetic mean values by water year, with the same treatment as in plot a).

c) The cumulative frequency distributions of the simulated and observed (raw, un-aggregated) data; the 95% confidence interval for observed data is shown in the dashed black lines. Note that only paired simulated and observed data points are used.