Evidence and implications of the background phosphorus concentration of submerged aquatic vegetation wetlands in Stormwater Treatment Areas for Everglades restoration

John M. Juston[1,2] and Thomas A. DeBusk[3]

Received 16 March 2010; revised 19 September 2010; accepted 22 October 2010; published 27 January 2011

[1] The limits of phosphorus (P) removal from the 18,120 ha Stormwater Treatment Areas (STAs) for Everglades restoration depend largely on the performance of submerged aquatic vegetation (SAV) wetlands, as SAV treatment cells now provide final stage treatment for 85% of the STA project. A long-term internal P profile in STA-2 cell 3 (STA2C3), one of the longest-running and best performing SAV cells, demonstrated no further net removal in the back quarter of the cell once total P (TP) levels approached 15 μg L^{-1}. Inflow-outflow performance data from STA2C3 were analyzed at monthly and annual scales and were pooled with data from an additional eight STA SAV treatment cells. The pooled data allowed inference of background TP concentrations in SAV treatment cells using existing Bayesian methods. Results showed a central tendency of 16 μg L^{-1} (13–17, 90% bounds), insensitivity to P loads less than $\sim$1.7 g m^{-2} yr^{-1}, and interannual variability outside these bounds. Internal data from the STA2C3 profile provided validation. Background P concentrations of 7 and 6 μg L^{-1} were identified for dissolved organic and particulate P fractions in the data pool, respectively, again similar to values in the STA2C3 gradient. Existing simulation modeling approaches for STA evaluations were identified as ineffective at or near background TP concentrations. Instead, we use an empirical and probabilistic approach based on full-scale data from STAs that produces annual risk of exceedance statistics and is easy to update. The current analysis suggests tangible risks for exceeding proposed annual discharge criteria from the STAs in the range of 16–20 μg L^{-1}.

Citation: Juston, J. M., and T. A. DeBusk (2011), Evidence and implications of the background phosphorus concentration of submerged aquatic vegetation wetlands in Stormwater Treatment Areas for Everglades restoration, *Water Resour. Res., 47*, W01511, doi:10.1029/2010WR009294.

1. Introduction

[2] The Everglades Construction Project was initiated in the 1980s to reduce phosphorus (P) concentrations in waters flowing into the Everglades Protection Area in southern Florida [*Goforth*, 2001; *Chimney and Goforth*, 2001]. The project was one important aspect of larger efforts focused on restoration of the greater Everglades ecosystem as mandated by the state of Florida in the Everglades Forever Act [*Perry*, 2004]. The South Florida Water Management District (SFWMD) has managed the design, construction, and maintenance of six constructed wetlands, known as Stormwater Treatment Areas (STAs), as a cornerstone of this project. The scale and scope of this constructed wetlands project is huge and unprecedented in terms of size, cost, and scientific challenges. To date, the six STAs have cost over $1 billion in capital expenditures, occupy over 18,000 ha of effective treatment area, and have retained over 1000 metric tons of P since 1994 that would have otherwise entered the Everglades Protection Area [*Pietro et al.*, 2008].

[3] Each STA is typically divided into three or four parallel treatment or process trains, and each process train is typically divided into two or three serially connected treatment cells. Research has shown that the type of vegetative communities in STA treatment cells has crucial influence on the P removal performance of the STAs. In subscale, prototype, and full-scale systems, submerged aquatic vegetation (SAV) wetlands have demonstrated higher P removal rates and lower outflow P concentrations than conventional emergent vegetated wetlands [*Dierberg et al.,* 2002; *Nungesser and Chimney,* 2001; *Juston and DeBusk,* 2006]. Research as to why SAV cells outperform emergent vegetation cells in the STAs is ongoing, but at least some differences seem to be attributable to the formation of calcium-rich marl sediments by SAV that may be more effective for P sequestration than muck soils that underlie conventional emergent vegetation wetlands [*Dierberg et al.*, 2002]. Inflow waters to the STAs are generally calcium rich, and in-column photosynthesis in SAV communities tends to elevate pH, which in turn mediates calcium carbonate precipitation.

[4] As of this writing, SAV-dominated treatment cells occupy approximately 9900 ha, or approximately 55%, of

[1]Juston Konsult, Uppsala, Sweden
[2]Department of Land and Water Resources Engineering, KTH Royal Institute of Technology, Stockholm, Sweden.
[3]DB Environmental, Rockledge, Florida, USA.

0043-1397/11/2010WR0092942011

the net STA area [*Pietro et al.*, 2008]. The remaining STA areas are emergent-dominated treatment cells plus a few small prototype cells for testing a periphyton-based treatment technology. Eleven of the existing 19 STA flow paths incorporate the conceptual design of sequential treatment through an emergent-dominated front end followed by an SAV-dominated back end. One motivation for this layout has been to provide reduced inflow P concentrations to the downstream SAV communities so that they operate in a P regime where they are more ecologically competitive. There are three additional single-cell SAV flow paths in the STAs that have historically received lower inflow P concentrations. In total, SAV cells are now in the effluent-producing position in 14 of the 19 treatment paths. These 14 treatment paths occupy approximately 85% of the existing STA area, so another perspective is that SAV cells provide final stage treatment for 85% of the STA project.

[5] The regulatory criterion for discharge P concentrations from the STAs remains contentious among the numerous vested stakeholders, as noted by *Rizzardi* [2001]. A legal limit has not yet been specified, but there are indications the STAs will be regulated with a not-to-exceed outflow limit of 16–17 μg L^{-1} as an annual flow-weighted mean (FWM) [*Payne et al.*, 2005]. Modeling and analysis tools based on annual and longer scales are required to support these objectives.

[6] A discharge criterion in this range requires the STAs to operate in a domain of the lowest reported P outflows from constructed wetlands [*Kadlec*, 1999a]. At these levels, the background P concentrations in wetlands are a serious concern and possible constraint. Background P (C^*) is achieved when there is no further net removal or conversion of the chemical constituent [*Kadlec and Wallace*, 2008]. Processes that contribute to apparent background P concentrations include atmospheric P deposition, internal hydraulic patterns, and an effective balance in the biogeochemical cycle between P removal and recycle [*Kadlec*, 1999a]. The net effect is an apparent lower floor in outflow P concentrations, independent of increasing wetland sizes or decreasing loading rates. Given the current state of knowledge in wetlands, C^* is considered an empirical and stochastic parameter [*Kadlec*, 1999a; *Kadlec and Wallace*, 2008].

[7] *Juston and DeBusk* [2006] analyzed 15 years of outflow P data pooled from four full-scale STA cells to estimate a central tendency in C^* in SAV systems of 16 μg L^{-1}. This assessment was based on a limited data set and did not investigate stochasticity. As the C^* in SAV cells largely determines the C^* from the STAs, the C^* in SAV cells remains a critical issue for Everglades restoration, with considerable regulatory and practical implications [*Rizzardi*, 2001]. A more thorough treatment with alternate perspectives is warranted. We present new field-collected data of the internal P gradient in STA-2 cell 3 (STA2C3), including a breakdown of P speciation. This treatment cell is arguably the most important in operation for this investigation as it is among the longest continually operating STA cells and also has regularly produced some of the lowest observed outflow P concentrations from the STAs. We revisit an expanded annual-scale data pool that is about double the size of that of *Juston and DeBusk* [2006], this time also including new perspectives on P speciation. We model these data using an approach similar to the probabilistic risk assessment of *Qian and Richardson* [1997]. We cross-reference data from the STA2C3 internal profile to these stochastic model calibrations from inflow-outflow data for validation and more robust assessments. Furthermore, we illuminate problematic issues with dynamic simulation approaches at the lower limits of P concentrations from SAV cells, including evaluation of the Dynamic Model for Stormwater Treatment Areas (DMSTA), the principal design tool used for simulating STA performance for design and regulatory purposes. We demonstrate that the stochastic model calibrated herein can provide useful alternative perspectives on outflow P expectations from the STAs and to protected waters.

2. Methods

2.1. Study Sites

[8] The STAs are situated between Lake Okeechobee and the Everglades Agricultural Area to the north and the Everglades Protection Area to the south (Figure 1). To varying degrees, inflow waters to the STAs have been a mixture of managed lake releases and agricultural runoff. STA inflows are controlled by pumps and/or flow control structures in regional canals but still generally follow a pulsed pattern dictated by the May–October wet season. Occasionally, the wet season includes tropical storms and/or hurricanes. On an annual scale, pumped inflow volumes generally exceed rainfall by an order of magnitude or two. There currently exist 17 SAV treatment cells in six STAs (Figure 1).

[9] We have conducted extensive field research in STA2C3, which occupies approximately 920 ha, with flow passing from north to south (Figure 2). The southeast corner of the cell, occupying approximately 25% of the surface area, is remnant mixed marsh from a historic wetland in the region. The remainder of the cell was previously farmed and has been managed by SFWMD for open water to encourage a viable SAV community.

2.2. STA2C3 Internal P Gradient Data

[10] Internal phosphorus concentrations were field sampled in STA2C3 28 times from May 2004 through July 2008 (Figure 3). The internal grid for P sampling consisted of nine longitudinal transects (Figure 2). Grab samples were collected at five stations along each of nine internal transects (labeled A–I in Figure 2) and at inflow and outflow locations. Samples were analyzed separately for individual stations along transects A, C, E, G, and I for 20 sample dates and were field composited for transects B, D, F, and H. Samples were field composited along all transects on the remaining eight sample dates. Total P (TP) was measured using rigorous quality methods [*U.S. Environmental Protection Agency* (*EPA*), 1983] with a minimum detection limit of 3 μg L^{-1}. Constituent P fractions of soluble reactive P (SRP), dissolved organic P (DOP), and particulate P (PP) were also assessed on 25 of the 28 sample dates. Soluble reactive P analyses were performed using the ascorbic acid–molybdenum blue method. Total P and total soluble P (TSP) analyses included persulfate digestion and neutralization prior to the colorimetric procedure (EPA 365.2) [*EPA*, 1983]. Using the approach reported by

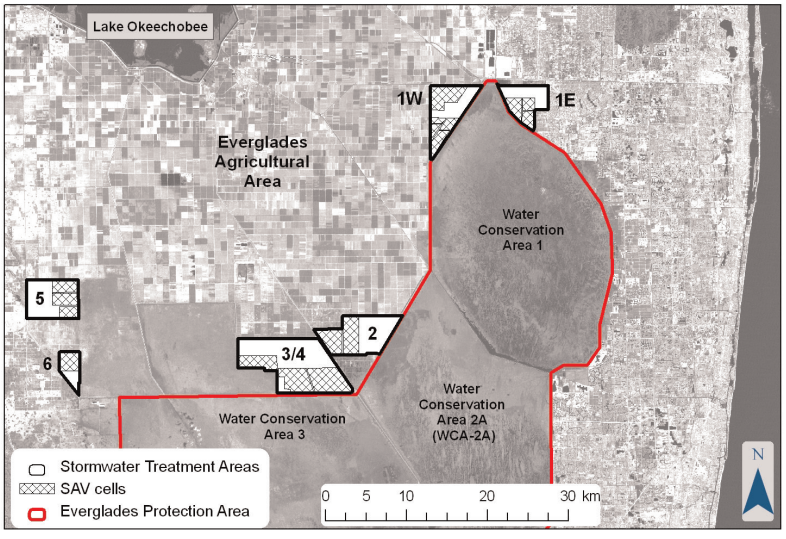

Figure 1. Locations of six Stormwater Treatment Areas (STAs) for Everglades restoration. This map represents an ~6200 km^2 region of southern Florida in southeast USA.

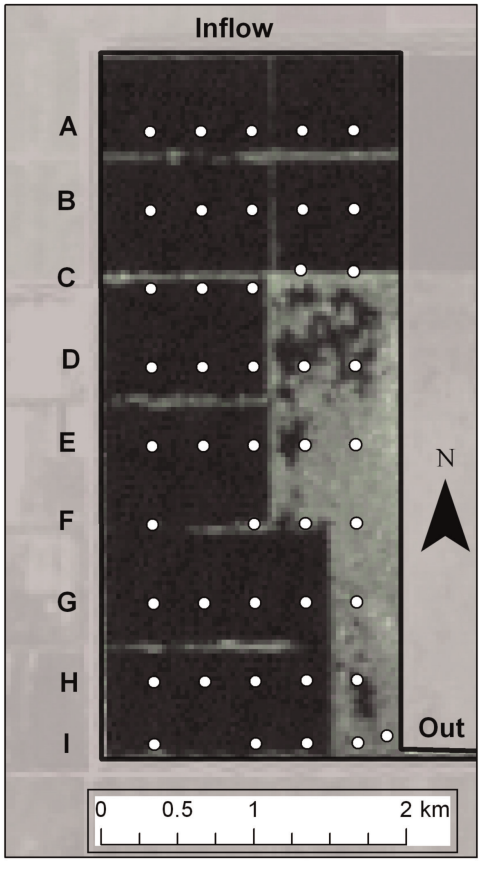

Figure 2. STA-2 cell 3 and the internal sampling grid that was used for sampling P gradients and vegetation. Satellite image is from January 2007. Darker areas indicate open water areas suitable for submerged aquatic vegetation (SAV). Lighter areas indicate remnant emergent aquatic vegetation from a historic wetland in the region.

Dierberg et al. [2002], PP, DOP, and SRP fractions were calculated as PP = TP − TSP and DOP = TSP − SRP.

[11] Transect means were calculated for the non–field composited transects (A, C, E, G, and I). There was no evidence of stationary bias in the time series of residuals relative to transect mean values. Two sampling events were screened from the data set, both of which occurred within 10 days of extremely disruptive tropical storms (Hurricanes Jeanne and Wilma in 2004 and 2005, respectively). A single representative profile for long-term average conditions in STA2C3 was calculated by flow weighting the 26 samples according to averaged daily inflow rates from a 2 week period preceding each longitudinal profile sampling date. Similar methods were used for an impacted zone in the Everglades Protected Area by *Walker* [1995]. Inflow and outflow P concentrations from this approach and effective flow and loading rates were within 10%–15% of values derived from the continuous long-term data record maintained by SFWMD, thus suggesting good representation. Uncertainties in the weighted mean values were estimated with a similar flow-weighting procedure based on standard deviations in grab sample data along transects A, C, E, G, and I.

2.3. SAV Inflow-Outflow Data Pool

[12] Daily flow and weekly or biweekly water quality data (e.g., TP, SRP, TSP, and calcium) for all STA treatment cells are collected by SFWMD personnel and are publicly available through the SFWMD online data portal, DBHYDRO. Following methods detailed by *Juston and DeBusk* [2006], monthly and annual-scale FWM inflow and outflow P concentrations and P mass loading rates (MLR) were estimated from these raw data for the treatment cells and data years of interest to this study. Additionally, similar series were calculated for SRP, TSP, and calcium (Ca) constituents. P fractions were calculated in the same way described for transect data.

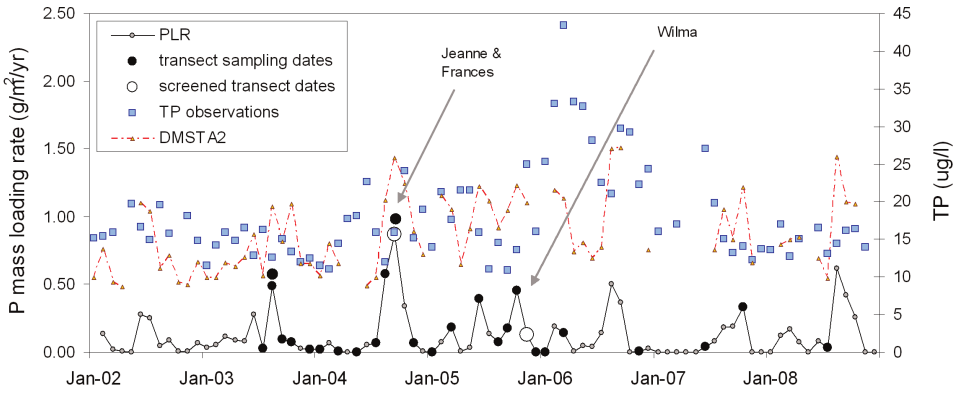

Figure 3. Seven year history of monthly total phosphorus loading rates and flow-weighted mean outflow P concentrations from STA-2 cell 3 (STA2C3). Sampling dates of longitudinal P gradients in STA2C3 are shown relative to monthly PLR pulses. Two months had two sampling dates, and these are shown with a second dot above the first. Total P (TP) observations and simulated DMSTA (case 1) values are monthly flow-weighted means. Gaps in DMSTA time series are for months when the model simulated no outflow.

[13] The treatment cells and data years of interest were defined as all STA cells managed for SAV that were past start-up and that did not include periods of extensive rehabilitation, vegetation management, and/or construction activities. Rehabilitation and other management activities are well documented in annual reports produced by SFWMD [e.g. *Pietro et al.*, 2008]. We operationally defined start-up as the initial 2.5 year period after flooding on the basis of vegetation cover surveys (personal observations). As a result of data screening, data from nine SAV treatment cells provided 28 net data years of information for analysis (Table 1). There were an additional eight SAV cells that were in various phases of start-up and were not included in this study. Two treatment cells provided half of the annual-scale data used here, while the remaining six cells provided the other half (Table 1).

[14] For additional quality control on data usage, we monitored vegetation patterns in these nine SAV cells with Landsat satellite imagery and airboat surveys of SAV communities along internal grids. Spatial analysis of one clear-sky Landsat image per year was used for annual monitoring of open water fractions in SAV-managed cells (Table 1). Open water fractions vary between cells and over time because of SFWMD management strategies regarding emergent vegetation. Regular airboat surveys were also conducted in most SAV cells in this study along extensive internal grids (see Figure 2) to quantify that SAV was, indeed, occupying open water areas in at least 80% of internal grid stations visited (Table 1).

2.4. Statistical Analyses

[15] Relationships between time series of water quality constituents were assessed using least squares regression. The slopes of regressions were checked against zero using single-tailed *t* tests and a 95% probability level for rejecting the null hypothesis (slope = 0).

2.5. Modeling Approaches

[16] Two modeling approaches were evaluated at different time scales: (1) simulation modeling of the STA2C3 inflow-outflow time series at monthly scale using conceptual mass balance approaches and (2) empirical modeling of outflow P relations to inflow mass loads at annual scale based on the SAV data pool. The utility of these approaches for eliciting and reproducing *C** behavior in SAV systems was compared with predictions of the

Table 1. Characteristics of SAV Treatment Cell Data Pool[a]

STA	Cell	Area (km^2)	Position	Data Interval	Data Years	Percent Open	Airboat Surveys
STA-1E	4N	2.6	middle	2008	1	88	2/1
	4S	3.0	back end	2008	1	93	2/1
	6	4.2	back end	2008	1	88	2/1
STA-1W	4	1.4	back end	1997–2003	7	84	3/1
	5b	9.3	back end	2002–2008	4	83	6/3
STA-2	3	9.2	single	2002–2008	7	81	18/6
STA-3/4	2b	11.7	back end	2007–2008	2	86	1/1
	3b	9.8	back end	2007–2008	2	79	0
STA-5	1b	4.9	back end	2003–2008	3	96	0

[a]Position refers to the location along a Stormwater Treatment Area (STA) flow path. Data interval column indicates span of post start-up operation. Data years indicate number of years used in data pool after screening. Percent open indicates mean open water fraction from spatial analysis of annual Landsat imagery. Airboat surveys indicates total airboat surveys over number of years (not all years surveyed) that quantitatively confirmed submerged aquatic vegetation (SAV) coverage in open water areas.

long-term P profile in STA2C3. The observed data for the internal P profile were effectively an independent data set, excepting data from the outflow station, and were used for validation and to move beyond black box inflow-outflow perspectives.

2.6. Dynamic Mass Balance Models

[17] A two-compartment model structure was investigated for dynamic simulations based on conceptual P removal, recycle, and burial processes (Figure 4). Models of this type have been successfully applied to P dynamics in eutrophied lakes [e.g., *Chapra and Canale*, 1991; *Seo and Canale*, 1996; *Ruley and Rusch*, 2004]. Several variants of model flux representations were investigated (Table 5). One variant warrants detailed description, as it was developed explicitly for analyses of Stormwater Treatment Areas for Everglades restoration by consultants of the U.S. Department of the Interior and U.S. Army Corps of Engineers [*Walker and Kadlec*, 2008]. In recent years the Dynamic Model for Stormwater Treatment Areas has been a cornerstone of most major considerations of STA performance evaluations [*Goforth*, 2007; *Goforth and CH2M Hill*, 2008], modifications [*Burns & McDonnell*, 2003], and interim regulations (e.g., G. Goforth et al., Technical support document for the STA-2 TBEL, 2007; available at http://www.garygoforth.net/technical%20reports.htm). One purpose of our investigation was to assess constraints to using the model at or near the background P levels in SAV systems. DMSTA and variants were tested with calibration to a 7 year time series of outflow P from STA2C3. The STA2C3 time series and internal P profile are the best existing data sets for this assessment.

[18] DMSTA maintains dynamic water balance, hydraulic routing, and P cycle routines (Figure 4). The model was conceptualized as a dynamic extension of a steady state settling rate model with background concentration [*Kadlec and Wallace*, 2008]. P removal from the water column is posed as autobiotic [*Kadlec*, 1997], in that removal is proportional to the product of water column P and labile P storage (Figure 4). The labile pool represents P storage in wetland vegetation and sediments. The P recycle flux in DMSTA is posed as quadratic to the labile pool. Phosphorus removal to permanent burial is first order from labile storage. It can be shown that long-term predictive trends from DMSTA (on the order of decades) are almost exactly reproducible ($R^2 = 0.99$; J. M. Juston, unpublished data, 2009) with the steady state NKC* settling rate model of *Kadlec and Wallace* [2008].

[19] The model has three key rate constants (K_1, K_2, and K_3) that scale P removal, recycle, and burial fluxes (Figure 4). These are the principal fitting parameters for calibration to historic data. Following the convention of the developers [*Walker and Kadlec*, 2008], these parameters were linearly transformed to a different parameter space (K, $C0$, and $C1$), where K was defined as the net settling rate at steady state, $C0$ was the water column concentration at zero labile storage ($S = 0$ in Figure 4), and $C1$ was the water column concentration when labile storage equaled 1 g m^{-2}. The interested reader is referred to model documentation source materials [*Walker and Kadlec*, 2008] for more details on these transformations and other aspects of DMSTA. We used six tanks in series in the model's hydraulic routine (Figure 4), which was consistent with tracer study results in STA2C3 (DB Environmental, unpublished data, 2004) as well as the physical compartmentalization in the cell (Figure 2). DMSTA algorithms were coded exactly and validated against "official" code with numerous case studies.

[20] The model was run at daily scale, and results were aggregated to monthly FWM concentrations. Model parameters were calibrated to STA2C3 time series at monthly scale to maximize a model efficiency index while also constraining <10% mass balance error. The model efficiency index (R_{eff}) was given by [*Nash and Sutcliffe*, 1970]

$$R_{\text{eff}} = 1 - \frac{\sigma^2_{\text{residuals}}}{\sigma^2_{\text{observations}}}. \quad (1)$$

[21] R_{eff} has a value of 1.0 for perfect model fits and has negative values for model fits with greater variance (i.e., less efficient) than the observed mean. In this way, the utility of dynamic simulation was evaluated within the constraint that the model should reasonably (within 10%) reproduce the long-term flow-weighted mean concentration (17.4 μg L^{-1}). The parameter space was searched with random Monte Carlo sampling using large sample sizes ($n \geq$ 20,000). Parameter and predictive uncertainties were generated but are not presented. A single "best" simulation was chosen as representative. Three additional performance indices were calculated for the best solution, including root-mean-square error, mean absolute error, and the

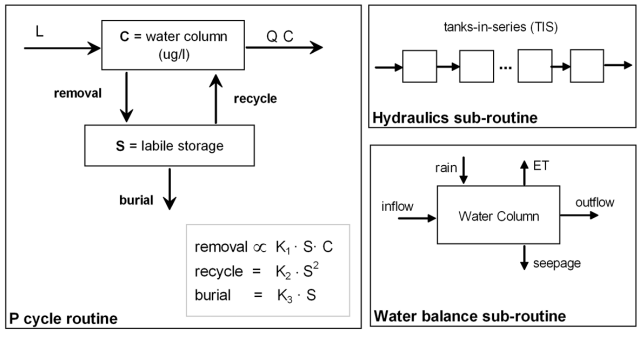

Figure 4. Schematic diagrams of key routines in dynamic P storage models evaluated in this study. The key governing equations for DMSTA [*Walker and Kadlec*, 2008] are shown in the bottom right of the P cycling routine.

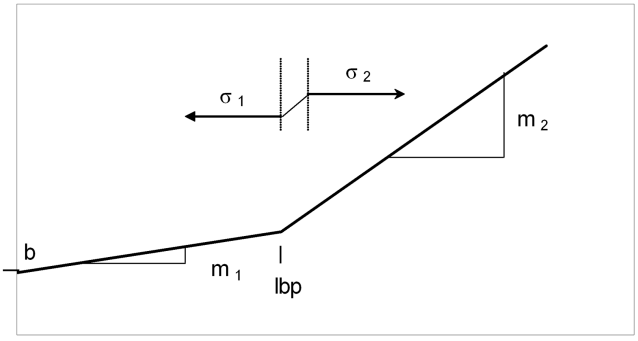

Figure 5. Piecewise linear model (PLR model) used to model annual outflow P relationship to P mass load. See Table 2 for parameter definitions.

coefficient of determination (R^2), all evaluated at monthly scale. A long-term longitudinal P profile was calculated from model results as the flow-weighted mean concentration in each tank in series (Figure 4). This was used for comparison to the observed longitudinal P profile in the cell.

2.7. Empirical Model of SAV Data Pool at Annual Scale

[22] Following *Qian and Richardson* [1997], a piecewise linear model was used to investigate relations between inflow MLR and annual FWM outflow P in the SAV data pool (Figure 5). From here forward, we refer to this as the PLR model. An alternate consideration for this task would be the steady state NKC* model [*Kadlec*, 1999a]. Preliminary investigations with the NKC* model (not reported here) to the same SAV data pool suggested C^* estimates consistent with the PLR model results reported in detail. The PLR model has six parameters (Table 2): four that define the piecewise linearity and two stochastic parameters. Separate calibrations were conducted for outflow TP, SRP, DOP, and PP data in the SAV data pool.

[23] Model parameters were estimated with uncertainty analysis. Markov chain Monte Carlo (MCMC) analysis was used to simulate the posterior of Bayesian inference using formal statistical likelihoods. This method was first applied to environmental models in the 1990s [*Qian and Richardson*, 1997; *Kuczera and Parent*, 1998]. The application here is by now fairly routine and straightforward. The Metropolis-Hastings algorithm was used for generating Markov chains. This algorithm and its use with Bayesian inference has been well described elsewhere [*Chib and Greenberg*, 1995; *Gelman et al.*, 2003].

Table 2. Parameters for the Piecewise Linear PLR Model

Parameter	Units	Description	Prior Range
b	$\mu g\ L^{-1}$	y intercept of first segment	0–25
lbp	$g\ m^{-2}\ yr^{-1}$	P load breakpoint	0–5
m_1	$\mu g\ L^{-1}\ (g\ m^{-2}\ yr^{-1})^{-1}$	slope of first line segment	−0.5–5
m_2	$\mu g\ L^{-1}\ (g\ m^{-2}\ yr^{-1})^{-1}$	slope of second segment	0–30
σ_1	dimensionless	log-transformed error variance in lower range	0–5
σ_2	dimensionless	log-transformed error variance in upper range	0–5

[24] Data and PLR model predictions were log transformed. The variance in the error model was specified as stepwise in correspondence to observed scatter in data above and below the apparent mass load breakpoint (Figure 5). *Qian and Richardson* [1997] used a similar approach. The combined effect of these two measures was to stabilize the normality and homoscedasticity of standardized residuals, as verified with visual diagnostics of the posterior. Error variances were treated as calibration parameters (Table 2) and inferred directly within the MCMC analysis. Uniform prior distributions were used for all parameters (Table 2). Thus, inferences were based on information in the data only. Following *Gelman et al.* [2003], warm-up chains were conducted to separate burn-in effects and to provide information for tuning the jump distribution in the final chains. Final chains were initiated in the region of highest probability identified in the warm-up chain. Covariance structure in parameter jump distributions were tuned for posterior acceptance rates between 25% and 45%. Final chains were run for 200,000 steps.

[25] Posterior distributions for model parameters were calculated from every tenth sample of the second half of the final chain ($n = 10{,}000$). Two categories of predictive uncertainty were estimated from PLR calibration. Model uncertainty is uncertainty in predictions due to uncertain model parameters. Total uncertainty adds to this the stochastic element from calibrated error variance parameters. Thus, model uncertainty represents an uncertainty in the central tendency of response while total uncertainty reflects probabilities for what actually might be observed at a particular instance in time, here at annual scale.

3. Results

3.1. STA2C3 Performance Data

[26] Inflow TP loads to STA2C3 pulsed seasonally over the 7 year history studied (Figure 3). Inflow P loads varied between 0.0–0.9 $g\ m^{-2}\ month^{-1}$ at a monthly scale and 0.8–2.1 $g\ m^{-2}\ yr^{-1}$ at an annual scale and had a 7 year average of 1.46 $g\ m^{-2}\ yr^{-1}$. Inflow P loads were significantly correlated to inflow flow rates and TP, PP, and SRP constituents (Table 4). Inflow TP (not shown) varied between 15–217 $\mu g\ L^{-1}$ at a monthly scale and 56–144 $\mu g\ L^{-1}$ at an annual scale and had a 7 year FWM of 99 $\mu g\ L^{-1}$. On average, inflow TP consisted of approximately 70% SRP, 8% DOP, and 21% PP (Table 3).

Table 3. Summary of Inflow Water Quality to SAV Treatment Cells, Calculated as Long-Term Flow-Weighted Means Over Data Years Used in This Study[a]

STA	Cell	Data Years	HLR (cm d^{-1})	TP ($\mu g\ L^{-1}$)	SRP ($\mu g\ L^{-1}$)	DOP ($\mu g\ L^{-1}$)	PP ($\mu g\ L^{-1}$)	Ca ($mg\ L^{-1}$)	P Load ($g\ m^{-2}\ yr^{-1}$)
STA-1E	4N	1	14	108	68	13	19	69	5.4
	4S	1	15	27	7	8	7	59	1.5
	6	1	7	145	75	19	35	85	4.1
STA-1W	4	7	15	98	35	10	37	82	4.5
	5b	4	7	153	101	12	40	81	3.9
STA-2	3	7	4	99	70	8	21	101	1.5
STA-3/4	2b	2	4	23	4	8	11	92	0.3
	3b	2	5	19	2	8	8	98	0.3
STA-5	1b	3	5	143	98	13	27	56	2.7

[a]HLR, hydraulic loading rate; TP, total P; SRP, soluble reactive P; DOP, dissolved organic P; PP, particulate P.

[27] Outflow TP from STA2C3 varied between 10–43 $\mu g\ L^{-1}$ at a monthly scale (Figure 3) and 14–27 $\mu g\ L^{-1}$ at an annual scale and had a 7 year FWM of 17 $\mu g\ L^{-1}$. Dynamics in outflow TP at a monthly scale had no significant relationship to inflow P loading rate ($R^2 = 0.02$), TP ($R^2 = 0.01$), SRP ($R^2 = 0.00$), DOP ($R^2 = 0.00$), or Ca ($R^2 = 0.03$) (see Table 4). Outflow TP exhibited a significant negative correlation to water flow rates (i.e., stagnation conditions tended to produce elevated outflow TP; $R^2 = 0.09$) and a positive correlation to inflow PP ($R^2 = 0.06$), but neither relation had substantial explanatory power.

[28] The internal TP profile in STA2C3 appeared to flatten about three fourths of the way through the cell (Figure 6). TP concentrations decreased steadily until reaching an apparent plateau just above 15 $\mu g\ L^{-1}$, according to the aggregated internal profile data. The coefficient of variation from weighted transect means and standard deviations was 0.25 at transect A (Figure 2) and ~0.14 at transects G and I, suggesting that TP values in the flattened tail were observed with more certainty than in the gradient. SRP, PP, and DOP fractions also appeared to individually flatten in internal profiles (Figure 6). SRP had the highest inflow fraction and lowest outflow fractions and was effectively removed to 3 $\mu g\ L^{-1}$, just above the analytical detection limit. Together, DOP and PP represented about 80% of outflow TP in the profile data.

Table 4. Coefficient of Correlation (R) for Monthly Average Water Quality Parameters in STA2C3 Waters From 7 Year Time Series Data[a]

	Inflow Waters						
	PLR	Flow	Ca	SRP	DOP	PP	TP
PLR	***x***	***0.90***	***0.53***	***0.86***	*0.12*	***0.22***	***0.80***
Flow	**0.84**	**0.96**	0.48	0.68	*−0.03*	*0.10*	***0.60***
Ca	**0.50**	**0.49**	**0.63**	***0.61***	*−0.03*	*0.30*	***0.59***
SRP	0.30	**0.42**	**0.53**	0.35	***0.22***	*0.36*	***0.97***
DOP	−0.44	**−0.50**	**−0.76**	−0.45	−0.30	*0.01*	***0.31***
PP	−0.26	**−0.18**	**−0.30**	−0.11	−0.05	0.14	***0.56***
TP	−0.08	**−0.23**	−0.17	0.04	0.05	**0.24**	0.11

[a]Cells with italic values show correlations between inflow constituents (constituent names in the first column represent the same inflow waters). Cells that do not contain italic values indicate correlations between inflow and posttreatment outflow time series (constituent names in the first column represent outflow waters). Bold values are statistically significant. Only 2 years of data were used for relations to outflow SRP and DOP due to a change in SFWMD analytical detection limits in 2007 (4 $\mu g\ L^{-1}$ before, 2 $\mu g\ L^{-1}$ after).

3.2. Simulation Modeling to STA2C3 at Monthly Scale

[29] Results with DMSTA are presented first, followed by alternative model configurations (Table 5). Following *Walker and Kadlec* [2008], we initially reduced the dimensionality of DMSTA calibration by fixing the $C0$ (3 $\mu g\ L^{-1}$) and $C1$ (22 $\mu g\ L^{-1}$) parameters. We refer readers to source documentation for the rationale for this approach [*Walker and Kadlec*, 2008]. The net effect was that the calibration problem was reduced to seeking a single K value to minimize the sum of log errors for monthly simulations.

[30] Modeling the dynamics of the STA2C3 monthly P time series was not successful with this approach (Figure 4). The best fit simulation with DMSTA had $K = 42$ m yr^{-1}. The coefficient of determination for the simulated TP time series was zero ($R^2 = 0.0$) and the Nash-Sutcliffe model efficiency was negative ($R_{eff} = -0.35$), suggesting the model fit was a worse predictor than the long-term mean of the data. We experimented with releasing constraints on $C0$ and $C1$ and calibrating all three DMSTA parameters to STA2C3 time series. Here results were somewhat improved (Table 5, trial 2), but performance indices were still unacceptably low ($R^2 = 0.28$, $R_{eff} = 0.27$). Variations on flux equations in the two-compartment model provided no further improvement in performance (Table 5). A final calibration trial using input-output calcium time series from STA2C3, rather than TP, was satisfactory and much more successful than any of the TP simulations (Table 5, trial 6). In this simulation, monthly FWM calcium concentrations varied between 38 and 108 $mg\ L^{-1}$ and were generally well above background concentrations for this constituent.

[31] All trials exhibited problematic validations to the STA2C3 internal P profile (Figure 6). If the calibration targets for DMSTA were changed from the input-output time series (Figure 4) to the internal profile data (Figure 6), then reasonably good simulations to the internal data could be achieved. This is sensible given DMSTA collapses to the NKC* model at steady state, and it is established that NKC* is suited to modeling longitudinal profiles in treatment wetlands [*Kadlec and Wallace,* 2008]. However, when the dynamic model (DMSTA) was tuned to the long-term steady state profile, its dynamic behavior worsened substantially (e.g., $R_{eff} \approx -1.0$). This suggested a fundamental incompatibility with the model structure and these multiple objectives. Although results are not presented in detail here, site-specific calibration of the simpler steady

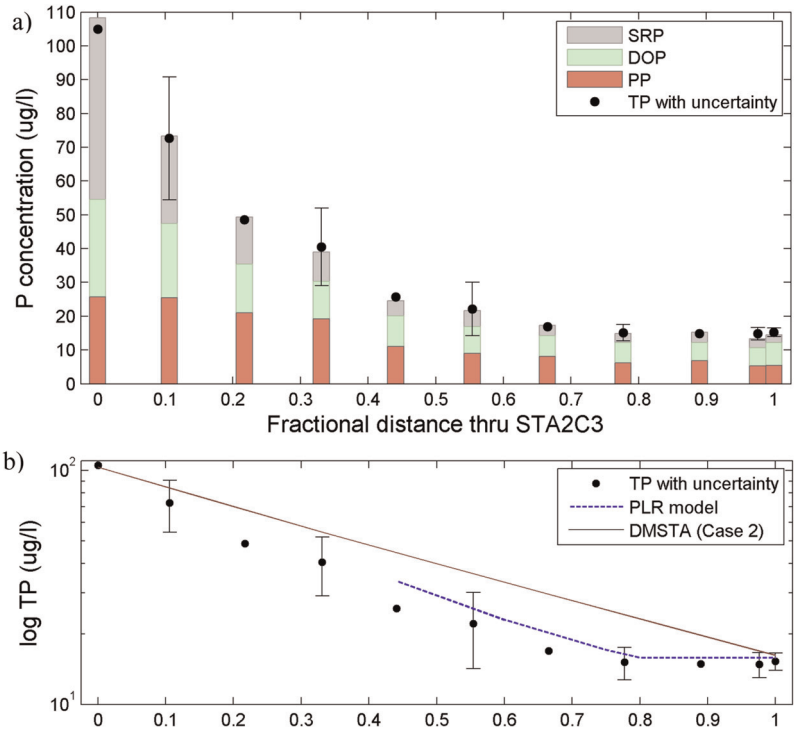

Figure 6. (a) Estimated long-term average longitudinal TP and P species gradients in STA2C3 based on 26 sampling events over a 6 year period and (b) model-produced P profiles. Error bars indicate estimated measurement uncertainties for TP. The DMSTA-produced profile resulted from calibration to STA2C3 inflow-outflow data at monthly scale (trial 2 in Table 5). Other calibration trials to monthly scale data (Table 5) exhibited similar nonflattened response in the back end as the case shown. The PLR-produced profile (shown without model uncertainty bands) resulted from calibration to the annual-scale cross-platform data pool.

state NKC* model to these data (Figure 6) yielded a probability density for C^* with 90% limits of 9–16 μg L^{-1} when fitted to all data points in the profile and 11–17 μg L^{-1} when fitted to focus on reproducing the tail only (last seven points in Figure 6).

[32] It is interesting to consider additional aspects of this model's predictive capability. Simulated monthly P outflow concentrations in DMSTA trial 1 (Table 5) were significantly correlated to monthly inflow P loading rates ($R^2 = 0.64$) in comparison to no correlation in the observed data ($R^2 = 0.01$; Figure 7). Similar trends have been observed with DMSTA predictions at an annual scale (Goforth et al., technical support document, 2007). As a point of comparison, correlations between inflow Ca mass loads and outflow Ca concentrations were evident in the STA2C3 data ($R^2 = 0.47$), and this time series was rather successfully simulated with similar model structure (Table 5, trial 6). At the P levels in the STA2C3 time series, a predictive relationship between P load and outflow TP breaks down (Figure 7). It appears the conceptual framework of DMSTA is not suited to a "near-background concentration" domain (i.e., there is no near-linear information in these data to effectively

Table 5. Summary of Calibration Trials for Dynamic Simulation of STA2C3 Outflow Time Series at Monthly Time Scale[a]

			Model Flux Representation			Model Performance[b]			
Trial	Model	Calibration Parameters	Removal	Recycle	Burial	R_{eff}	RMSE	MAE	R^2
1	DMSTA	K	autobiotic	second order	first order	−0.35	8.4	6.6	0.02
2	DMSTA	K, $C0$, $C1$	autobiotic	second order	first order	0.27	6.4	4.8	0.28
3	K-C*	K, C^*	first order	zeroth order	-	0.01	6.9	5.2	0.03
4	CC1	$k1$, $k2$, $k3$	first order	first order	first order	0.28	6.5	5.1	0.29
5	CC2	$k1$, $k2$, $k3$, zero-order term	first order	first plus zeroth orders	first order	0.28	6.4	5.0	.29
6	calcium	$k1$, $k2$, $k3$	first order	first order	first order	0.71	11.6	9.8	0.71

[a]The first five trials modeled TP, and the sixth modeled calcium. Model flux representations refer to Figure 4. Trial 4 (CC1) used a model structure similar to *Chapra and Canale* [1991]. Trial 5 (CC2) added to that a zero-order recycle flux. Performance indices are the best identified, with the goal of maximizing R_{eff} while constraining <10% mass balance error. For reference, long-term FWM outflow concentrations were 17 μg L^{-1} for P and 73 mg L^{-1} for calcium.

[b]RMSE, root-mean-square error; MAE, mean absolute error.

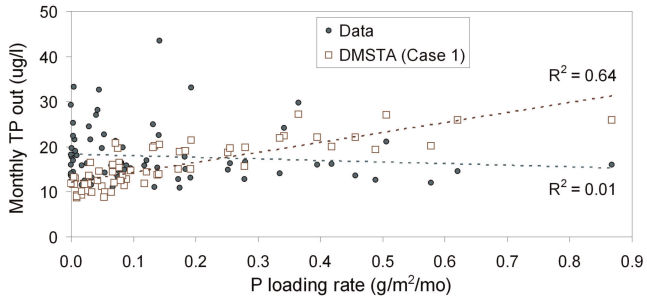

Figure 7. Correlations in observed and simulated P concentrations to dynamic P loads at monthly scale in STA2C3. DMSTA values were from the trial 1 calibration (Table 5). A similar trend in the STA2C3 data is evident at annual scale in the cross-platform data pool (Figure 8b).

inform the model). As a consequence, the long-term mass balances in these approaches were calibrated at a point but were not constrained or accurate at other points. This is clearly evident in internal P gradients simulated with DMSTA (Figure 6b). If the simulated P profile from DMSTA were extended to the right (e.g., in Figure 6b), effectively decreasing the P loading rate to the system, it would eventually reach an asymptote below 5 μg L^{-1}, or a concentration approximately 10 μg L^{-1} lower than has been observed in the field. Similarly, at higher mass loads (effectively moving to the left along the x axis in Figure 6b), DMSTA simulations were not accurate compared to long-term average data at internal stations. This would remain the case even if the model were calibrated with probabilistic methods. Taken together, these characteristics demonstrate that using DMSTA or similar model structures for predictions at the lowest limits of P concentration in SAV cells is problematic.

3.3. SAV Data Pool

[33] Inflow water quality varied within the broader SAV data pool (Table 3). In general, the fraction of inflow SRP decreased with decreasing inflow TP concentrations. The lowest observed annual FWM outflow TP from any SAV cell to date has been 14 μg L^{-1}, of which there have been a total of five occurrences among four different SAV cells.

[34] Annual outflow TP in the SAV data pool demonstrated an apparent breakpoint relationship to inflow P loading rates (Figures 8a and Figures 8b). This was confirmed with calibrations of the piecewise PLR model to these data. The MCMC calibration algorithm was allowed to explore possible breakpoints between 0 and 5 g m^{-2} yr^{-1} (Table 2). A breakpoint of $\sim$1.7 g m^{-2} yr^{-1} was identified in two variations of treatment for the independent variable (i.e., annual mean and biennial mean P load, Figure 9b). Both of these treatments excluded one data year from STA2C3 because of documented extensive internal damage to SAV communities following Hurricane Wilma in 2005 (Figure 8b). Below the $\sim$1.7 g m^{-2} yr^{-1} breakpoint, the slope of the first line segment tended toward zero (Figure 9c), which implied a C^* in SAV treatment cells in the range of 13–17 μg L^{-1} (90% limits, Figure 9a) and a most likely value of $\sim$16 μg L^{-1}. We consider this a more robust finding than the site-specific estimate from section 3.2, which was complementary, as it represents additional information at very low P loads and from a multiplatform data pool (Figure 8b).

[35] The PLR model response was validated against the internal P profile in ST2C3. Effective P loads were prescribed to TP values from the nine internal transects in STA2C3 and plotted in the context of the inflow-outflow data pool and PLR model uncertainty bands (Figure 8b). Similarly, PLR model predictions were plotted as a function of longitudinal position within the uncertainty estimates of the long-term average P profile in STA2C3 (Figure 6b). Breakpoint and leveling behavior identified with Bayesian MCMC calibration of the PLR model to inflow-outflow data in the SAV data pool were well matched to trends in the longitudinal P profile of STA2C3.

[36] There were 11 net years of data with biennial mean P loads less than $\sim$1.7 g m^{-2} yr^{-1}, including five annual values from STA2C3. Among these data ($n = 11$), there were no significant correlations between annual outflow TP to inflow P loads (annual or biennial mean), hydraulic loading rate, or annual inflow SRP, DOP, PP, TP, or Ca constituents.

[37] Outflow P fractions in the SAV data pool (Figures 8c–Figures 8e) generally confirmed observations from internal gradients in STA2C3 (Figure 6a). The lowest observed annual outflow DOP and PP constituents from any SAV treatment cell have been 5 and 4 μg L^{-1}, respectively. The PLR model structure was calibrated to these data for deeper insights. Breakpoints were not clearly identified in all of these calibrations (Figure 9b); however, ranges for lowest achievable concentrations (the b parameter in Figure 5) were. Calibrations of the PLR model suggested most likely values of 2, 7, and 6 μg L^{-1} for SRP, DOP, and PP, respectively, for the lowest observed values of these constituents in SAV systems as annual FWM values (Figure 9a).

[38] It is interesting to compare and contrast the performance of an STA treatment cell that is not in the SAV data pool. STA-2 cell 1 is an emergent-vegetated treatment cell that was a remnant fragment of a historic wetland before it was incorporated in the STA-2 footprint (Figure 1). It lies approximately 2 km east of STA2C3. Results from *Juston and DeBusk* [2006] demonstrated significantly lower outflow P concentrations from emergent STA treatment cells that were historic wetlands in comparison to those constructed on recently farmed soils. Indeed, STA-2 cell 1 had a 5 year interval wherein it was consistently flooded and produced annual outflow P concentrations in the range of 8–11 μg L^{-1} with annual P loads ranging from 0.9 to 1.7 g m^{-2} yr^{-1}. The most discernable difference between the performance of cell 1 during those years and the SAV data pool was in PP removal (Figures 8c–8e). Mean outflow PP was typically 2–3 μg L^{-1} from STA-2 cell 1 during those years compared to the lowest calculated value of 4 μg L^{-1} in the SAV data pool and a more typical value of 6 μg L^{-1} (Figure 9a).

3.4. Probabilities for Exceeding Outflow Criteria

[39] Probabilities for exceeding several hypothetical not-to-exceed annual criteria were explored using the calibrated stochastic parameters in the PLR model. *Qian and Richardson* [1997] produced risk assessments from a

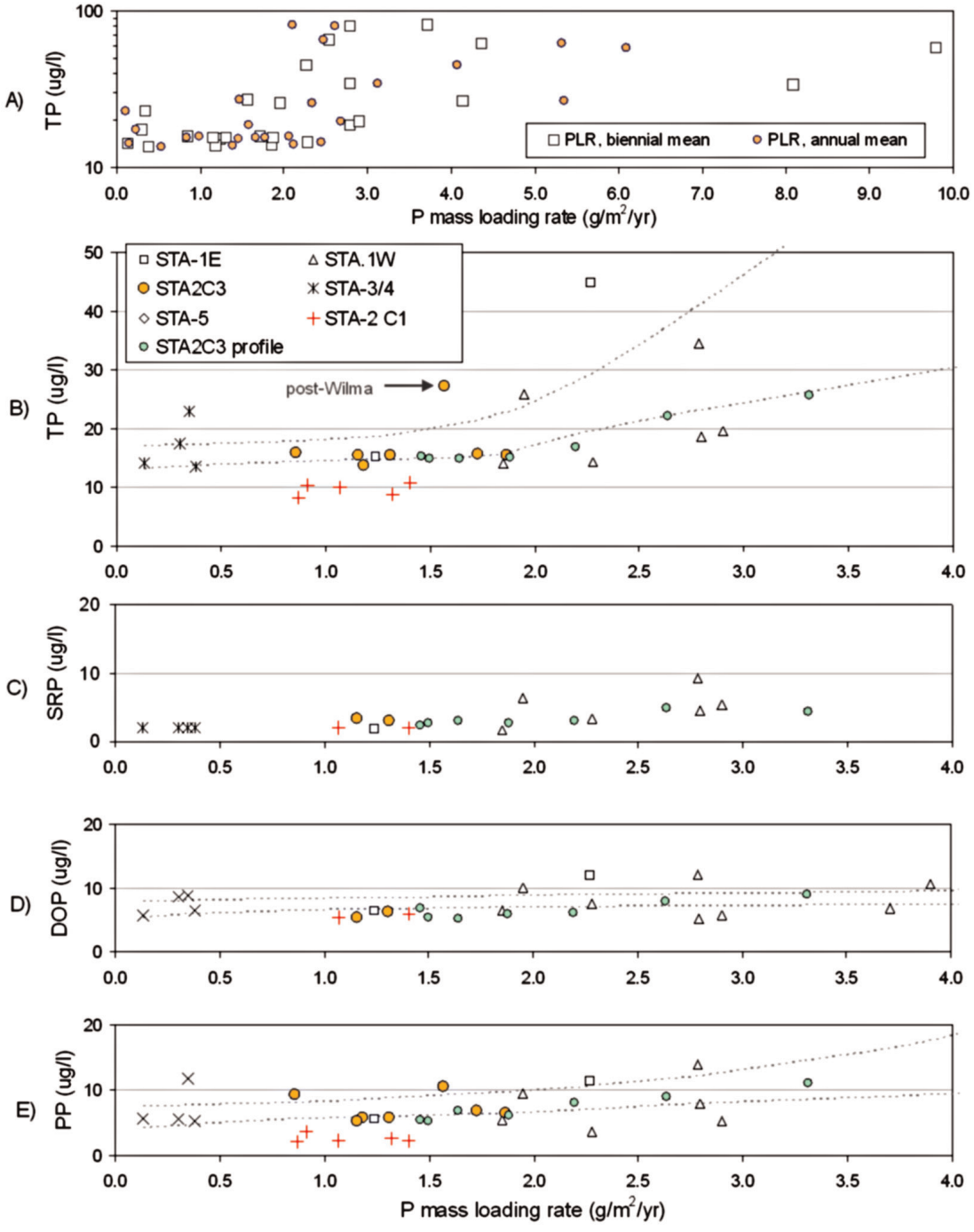

Figure 8. Annual P data as a function of total phosphorus loading rate in the STA SAV data pool. (a) TP as a function of both annual and biennial average loading rates plotted on log linear scale (x axis cut off at 10 g m^{-2} yr^{-1}, one annual mean data point not shown). Annual outflow (b) TP, (c) soluble reactive P (SRP), (d) dissolved organic P (DOP), and (e) particulate P (PP) as a function of biennial loading, with the legend shown in Figure 8b. Dotted lines indicate 90% confidence intervals of model uncertainty from PLR model calibrations with Bayesian inference. Data from STA-2 C1 and STA2C3 P profiles (from Figure 6) are shown for reference but were not used for model fitting. STA-2 C1 is a historic emergent vegetation wetland and is not an SAV treatment cell.

probabilistic piecewise linear model in the same fashion. Following that approach, probabilities were generated for exceeding various annual outflow P thresholds for several scenarios of the recent MLR history of SAV cells (Figure 10). Calculations represent total uncertainties in PLR model prediction on the basis of adding interannual stochasticity to model parameter uncertainties. For instance, there is a 44% risk of exceeding a 17 μg L^{-1} limit if all final stage SAV

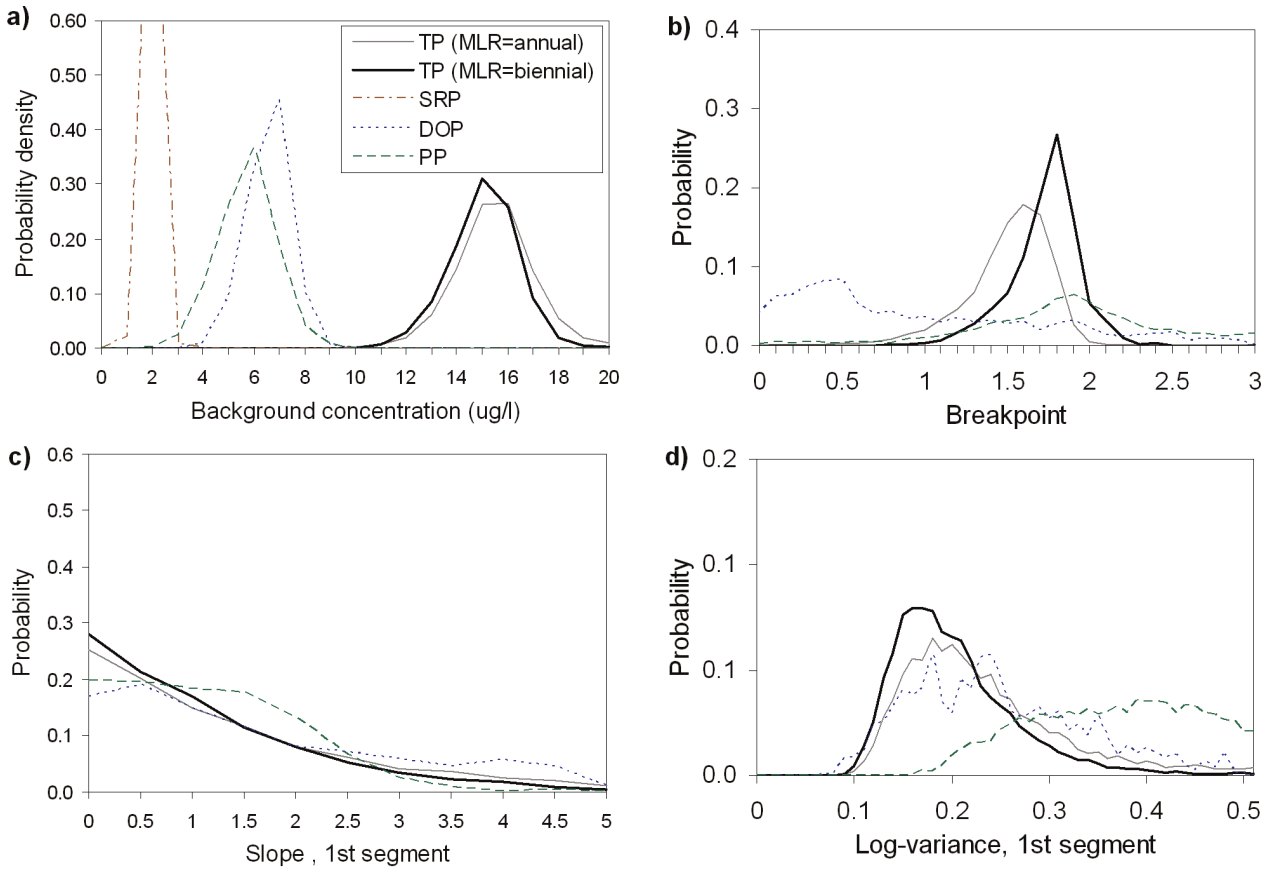

Figure 9. Posterior marginal distributions for PLR model parameters. (a) Background concentration (b) breakpoint, (c) slope of the first segment, and (d) log variance of the first segment. SRP not shown in Figures 9b–9d as much of the SRP data set was at the 2 μg L^{-1} detection limit.

treatment cells were loaded at 1.0 g m^{-2} yr^{-1}. Risk of exceedance increased with the inclusion of the hurricane-influenced data year in the analysis. There appeared to be a benefit in using a longer-term (2 year) perspective on mean P loads to SAV cells (Figure 10), as risks were somewhat reduced.

4. Discussion and Conclusions

[40] A probability density of background phosphorus concentration in SAV treatment cells for Everglades restoration has been identified and has 90% bounds of 13–17

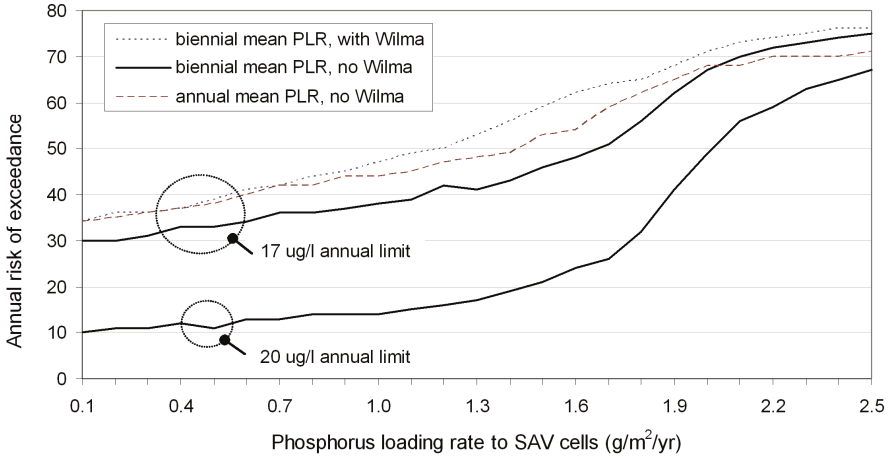

Figure 10. Risk of exceedance of STA effluent criteria based on loading conditions to back-end SAV treatment cells. Risks of exceeding two criteria are shown (17 and 20 μg L^{-1}), each representing a hypothetical not-to-exceed annual flow-weighted mean. The effects of including data from the year following Hurricane Wilma are also depicted for the 17 μg L^{-1} criterion.

μg L^{-1} and a most likely value of 16 μg L^{-1}. This was assessed with modeling analysis of a pooled database of annual inflow-outflow data from all existing full-scale SAV treatment cells in Everglades Stormwater Treatment Areas. Strong corroborative evidence was collected by sampling water quality transects in one of the longest-running and best performing SAV treatment cells (STA-2 cell 3). These internal data were independent from the data pool used for model calibration and demonstrated no further removal in the back quarter of the cell as P levels approached 15 μg L^{-1}. Data on P fractionations along this transect were also corroborative of trends observed in the data pool between inflow mass loads and outflow P fractions and predicted by modeling. In combination, these analyses pointed toward what appears to be a robust assessment of C^* and associated threshold behavior in the SAV systems.

[41] Previous analysis of an SAV data pool of "natural" Florida aquatic systems seemed to suggest a floor in annual outflow TP of around 25 μg L^{-1}, with some scatter below [*Knight et al.*, 2003], but those natural systems had generally higher hydraulic loads and lower SAV densities than the engineered systems in the STAs. Performance of experimental SAV mesocosms, comprising over 50 data years, demonstrated minimum annual outflow P concentrations of 13–20 μg L^{-1} over a broad range of P loading rates (DB Environmental, unpublished data, 2009). Earlier analysis of STA data suggested a C^* of 16 μg L^{-1} in full-scale SAV cells [*Juston and DeBusk*, 2006], and those results were largely confirmed by the analysis here. The critical difference is that the present approach is now at a level appropriate for decision support for Everglades restoration. SAV treatment cells are the effluent-producing treatment technology in most of the current STA flow ways, and the estimated range of C^* (13–17 μg L^{-1}) lies directly in the middle of annual not-to-exceed regulatory criteria that are being considered for STA discharges. The current analysis suggests tangible risks for exceeding possible annual discharge criteria in the range of 16–20 μg L^{-1} from the STAs. (Figure 10).

[42] The internal P speciation profile in STA2C3 (Figure 6a) demonstrated that SRP was reduced to the analytical detection limit (2 μg L^{-1}) at a location approximately three quarters along the fractional cell length, while PP and DOP attained background concentrations of approximately 6 μg L^{-1} at 0.8 of the length of the cell. Soluble reactive P is known to be sequestered rapidly and completely in SAV wetlands [*Dierberg et al.*, 2002], and this was borne out by the internal STA2C3 transect data. The higher background levels and more modest removal rates of PP and DOP may be due to the incomplete removal of recalcitrant constituents in the inflow waters. The wet prairies and sloughs of the downstream Everglades commonly attain 10 μg L^{-1} and lower TP water concentrations, as their biota have the ability to hydrolyze labile and moderately recalcitrant organic P compounds [*Wright and Reddy*, 2001]. However, the taxonomic composition of SAV-dominated back-end STA cells differs markedly from that of native Everglades communities, and these SAV wetlands may not support the organisms capable of hydrolyzing the more recalcitrant organic P forms [*Pant et al.*, 2002]. An alternative explanation may lie in the internal generation of both PP [*Dierberg and DeBusk*, 2008] and DOP [*Pant et al.*, 2002] within the SAV community. It is possible, for example, that SAV cells have inherently higher levels of water column particulate matter than natural, emergent marshes. *Dierberg and DeBusk* [2008] noted a difference in both concentration and composition of particles in the outflows of C3 (SAV) and C1 (emergent) flow paths of STA2. Research is ongoing in this area.

[43] This study also raised issues about the utility of dynamic mass balance models, including DMSTA, a prominent STA design tool, at near-background levels for the constituent of interest. DMSTA and variants were generally not successful in reproducing outflow time series of STA2C3 (Table 5), nor did they capture the observed internal concentration gradients in this cell (Figure 6b). We have doubts as to whether more complicated wetland P models [e.g., *Wang and Mitsch*, 2000] would do better given no explicit representation of C^* phenomena and a paucity of data on internal fluxes and states in SAV wetlands that might aid with more detailed process descriptions. We found the use of complementary data on internal constituent profiles extremely valuable in moving past black box inflow-outflow perspectives in this assessment. We also found it very useful to compare modeling results for P and Ca constituents (one near background, one not; Table 5), and we recommend that this approach be extended to modeling efforts of other data sets in the STAs and elsewhere for validation of these principles.

[44] Modeling at annual scale provided an alternate perspective and was important for regulatory considerations. It also aggregated removal, recycle, and burial dynamics that were difficult to capture at monthly scale. We induced an empirical piecewise linear model at annual scale from the SAV data pool, in a fashion similar to *Richardson et al.* [1997], and calibrated model parameters within a stochastic framework as suggested by *Qian and Richardson* [1997]. The efforts of *Qian and Richardson* [1997] occurred before benefits of SAV communities for P removal were fully documented, so their focus was on the performance of emergent vegetated wetlands for future STAs. They combined information from a cross-sectional database of wetlands across North America and site-specific data from the P gradient in Water Conservation Area 2A (Figure 1) to assess a C^* of 18–24 μg L^{-1} and a P load threshold of $\sim$1.0 g m^{-2} yr^{-1} for the assimilative capacity in future STAs. *Kadlec* [1999b] criticized *Qian and Richardson* [1997] partly for their use of the cross-sectional data, which consisted of vastly different wetlands. Here the SAV data pool we have used (Table 1) consisted of largely similar wetlands that are located in close proximity, situated on previously farmed muck soils, contained very similar vegetation communities, and have been closely monitored and frequently surveyed to assure representation in our analysis (Table 2). It is our belief that the results from this pooled data analysis are the best current basis for predicting the behavior in individual SAV cells at or near background concentrations. As more data become available, Bayesian hierarchical mixed models may be appropriate to refine stochastic predictions for individual cells [e.g., *Stow et al.*, 2009].

[45] It is important to note that the results presented are based on currently available data and can and should be regularly updated. Furthermore, SAV-dominated STA cells

occupy massive footprints and can be adversely impacted by periodic large-scale weather events such as hurricanes and severe droughts. Such phenomena will most likely increase the long-term mean outflow P levels over the 13–17 μg L^{-1} noted above. Finally, the identified breakpoint in performance at $\sim$1.7 g m^{-2} yr^{-1} applies to SAV cells only and not to the STAs as a whole. Most SAV treatment cells have emergent-vegetated front-end cells, and it is the combination of the two that delivers net STA performance. Clearly, our results provide insights on what performance is required from front-ends cells to enable optimal SAV back-end performance, but more work is also required to generate integrated probabilistic perspectives for the STAs.

[46] **Acknowledgments.** We acknowledge the South Florida Water Management District for their efforts in the design, construction, maintenance, data collection, and optimization of the STAs. We also acknowledge efforts of five anonymous reviewers who provided constructive insights at various stages.

References

Burns & McDonnell (2003), Final report: Everglades Protection Area tributary basins, long term plan for achieving water quality goals, South Fla. Water Manage. Dist., West Palm Beach. (Available at http://www.sfwmd.gov/portal/page/portal/xrepository/sfwmd_repository_pdf/waterquality_0.pdf.)

Chapra S. C., and R. P. Canale (1991), Long-term phenomenological model of phosphorus and oxygen for stratified lakes, *Water Res.*, *25*(6), 707–715.

Chib S., and E. Greenberg (1995), Understanding the Metropolis-Hastings algorithm, *Am. Stat.*, *49*(4), 327–335.

Chimney M. J., and G. Goforth (2001), Environmental impacts to the Everglades ecosystem: A historical perspective and restoration strategies, *Water Sci. Technol.*, *44*(11-12), 93–100.

Dierberg F. E., and T. A. DeBusk (2008), Particulate phosphorus transformations in south Florida stormwater treatment areas used for Everglades protection, *Ecol. Eng.*, *34*(2), 100–115.

Dierberg F. E., T. A. DeBusk, S. D. Jackson, M. J. Chimney, and K. Pietro (2002), Submerged aquatic vegetation-based treatment wetlands for removing phosphorus from agricultural runoff: Response to hydraulic and nutrient loading, *Water Res.*, *36*(6), 1409–1422.

Gelman A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2003), *Bayesian Data Analysis*, 2nd ed., Chapman and Hall, Boca Raton, Fla.

Goforth G. (2001), Surmounting the engineering challenges of Everglades restoration, *Water Sci. Technol.*, *44*(11-12), 295–302.

Goforth G. (2007), Updated STA phosphorus modeling for the 2010 planning period, report, work order CN040902-WO03R2, South Fla. Water Manage. Dist., West Palm Beach.

Goforth G., and CH2M Hill (2008), Technical memorandum—Water quality performance evaluation of the pre-2006 projects, report, work order 4600000851-WO01, South Fla. Water Manage. Dist., West Palm Beach.

Juston J. M., and T. A. DeBusk (2006), Phosphorus mass load and outflow concentration relationships in stormwater treatment areas for Everglades restoration, *Ecol. Eng.*, *26*(3), 206–223.

Kadlec R. H. (1997), An autobiotic wetland phosphorus model, *Ecol. Eng.*, *8*(2), 145–172.

Kadlec R. H. (1999a), The limits of phosphorus removal in wetlands, *Wetlands Ecol. Manage.*, *7*(3), 165–175.

Kadlec R. H. (1999b), Response to the Richardson and Qian comments on "Limits of phosphorus removal in wetlands," *Wetlands Ecol. Manage.*, *7*(4), 239–245.

Kadlec R. H., and S. D. Wallace (2008), *Treatment Wetlands*, 2nd ed., CRC Press, Boca Raton, Fla.

Knight R. L., B. Gu, R. A. Clarke, and J. A. Newman. (2003), Long-term phosphorus removal in Florida aquatic systems dominated by submerged aquatic vegetation, *Ecol. Eng.*, *20*(1), 45–63.

Kuczera G., and E. Parent (1998), Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm, *J. Hydrol.*, *211*(1-4), 69–85.

Nash J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, *10*(3), 282–290.

Nungesser M., and M. Chimney (2001), Evaluation of phosphorus retention in a south Florida treatment wetland, *Water Sci. Technol.*, *44*(11-12), 109–115.

Pant H. R., K. R. Reddy, and F. E. Dierberg (2002), Bioavailability of organic phosphorus in a submerged aquatic vegetation-dominated treatment wetland, *J. Environ. Qual.*, *31*(5), 1748–1756.

Payne G., K. Weaver, and F. Nearhoof (2005), Derivation of the water quality based effluent limit (WQBEL) for phosphorus discharges to the Everglades Protection Area, draft report, Fla. Dep. of Environ. Prot., Tallahassee.

Perry W. (2004), Elements of south Florida's comprehensive Everglades restoration plan, *Ecotoxicology*, *13*(3), 185–193.

Pietro K., R. Berzotti, G. Germain, and N. Iricanin (2008), STA performance, compliance, and optimization, in *2009 South Florida Environmental Report*, chap. 5, pp. 5-1 to 5-132, South Fla. Water Manage. Dist., West Palm Beach. (Available at https://my.sfwmd.gov/portal/page/portal/pg_grp_sfwmd_sfer/portlet_sfer/tab2236041/volume1/chapters/v1_ch_5.pdf.)

Qian S. S., and C. J. Richardson (1997), Estimating the long-term phosphorus accretion rate in the Everglades: A Bayesian approach with risk assessment, *Water Resour. Res.*, *33*(7), 1681–1688.

Richardson C. J., S. S. Qian, C. B. Craft, and R. G. Qualls (1997), Predictive models for phosphorus retention in wetlands, *Wetlands Ecol. Manage.*, *4*, 159–175.

Rizzardi K. (2001), Translating science into law: Phosphorus standards in the Everglades, *J. Land Use Environ. Law*, *17*(1), 149–167.

Ruley J. E., and K. A. Rusch (2004), Development of a simplified phosphorus management model for a shallow, subtropical, urban hypereutrophic lake, *Ecol. Eng.*, *22*(2), 77–98.

Seo D. I., and R. P. Canale (1996), Performance, reliability, and uncertainty of total phosphorus models for lakes—1. Deterministic analyses, *Water Res.*, *30*(1), 83–94.

Stow C. A., C. Lamon, S. S. Qian, P. A. Soranno, and K. H. Reckhow (2009), Bayesian hierarchical/multilevel models for inference and prediction using cross-system lake data, in *Real World Ecology: Large-Scale and Long-Term Case Studies and Methods*, edited by S. L. Miao, S. Carstenn, and M. Nungesser, pp. 111–136, Springer-Verlag, New York.

U.S. Environmental Protection Agency (EPA) (1983), Methods for chemical analysis of water and wastes, Rep. EPA-600/4-79/020, Washington, D. C.

Walker W. W. (1995), Design basis for Everglades stormwater treatment areas, *J. Am. Water Resour. Assoc.*, *31*(4), 671–685.

Walker W. W., and R. H. Kadlec (2008), Dynamic Model for Stormwater Treatment Areas, model version 2, 9/30/2005, documentation update 01/17/2008, U.S. Dep. of Inter., Washington, D. C. (Available at http://www.wwwalker.net/dmsta/)

Wang N., and W. J. Mitsch (2000), A detailed ecosystem model of phosphorus dynamics in created riparian wetlands, *Ecol. Modell.*, *126*(2-3), 101–130.

Wright A. L., and K. R. Reddy (2001), Phosphorus loading effects on extracellular enzyme activity in Everglades wetland soils, *Soil Sci. Soc. Am. J.*, 65(2), 588–595.

T. A. DeBusk, DB Environmental, Rockledge, FL 32955, USA.

J. M. Juston, Juston Konsult, Malma Stensvag 6, Uppsala, SE-75645, Sweden. (juston@telia.com)