Incorporation of spectral data into multivariate geostatistical models to map soil phosphorus variability in a Florida wetland

R.G. Rivero a,b,1, S. Grunwald a,b,⁎, G.L. Bruland c

a GIS Research Laboratory, Soil and Water Science Department, University of Florida, Institute of Food and Agricultural Sciences, 2169 McCarty Hall, P.O. Box 110290, Gainesville, FL 32611, United States
b School of Natural Resources and Environment, University of Florida, 103 Black Hall, P.O. Box 116455, Gainesville, FL 3261, United States
c Natural Resource and Environmental Management Department, College of Tropical Agriculture and Human Resources, University of Hawai‘i at Mānoa, 1910 East-West Road, Honolulu, HI 96882, United States

Available online 4 June 2007

Abstract

Hybrid geostatistical prediction methods incorporate (i) spatially-explicit soil observations and exhaustive grids of ancillary environmental variables (e.g. derived from remote sensing), (ii) spatial autocorrelation, (iii) spatial covariance, and/or (iv) combinations of the above. In numerous studies of terrestrial soils it has been shown that hybrid geostatistical methods outperform univariate spatial and regression (aspatial) methods. However, hybrid methods have rarely been employed to predict soil properties in wetlands. In this study we used spectral data and derived indices from two remote sensors (Landsat ETM+ and ASTER), with different spatial resolutions, from different seasons, but with similar spectral range, ancillary environmental data, as well as floc and soil total phosphorus (TP) observations from 111 sites. The specific objective of our study was to evaluate the performance of aspatial methods (multivariate regressions — REG), univariate spatial (Ordinary Kriging — OK) and hybrid/multivariate geostatistical methods (Regression Kriging — RK and Co-kriging — CK) in predicting the spatial variability and distribution of floc and soil TP in a subtropical wetland, WCA-2A, in the Florida Everglades. Measured floc TP ranged from 194 to 1865 mg kg−1 with a median of 751 mg kg−1 and standard deviation (SD) of 381 mg kg−1. According to cross-validation, predictions for floc TP based on the root mean square prediction error (RMSE) were best in the following order: RKquadratic (134.9)> RKmultivariate (201.1)> OK (206.1)> CK (212.1)> REGquadratic (220.3)> REGmultivariate (218.3)> REGlinear (264.4); and based on the mean prediction error (ME) followed the order RKmultivariate (0.9)> RKquadratic (1.1)> CK (−6.7)> REGmultivariate (18.2)> REGlinear (25.1)> OK (−27.3)> REGquadratic (27.3). The Normalized Difference Vegetation Index (NDVI)-green derived from Landsat ETM+ showed the largest predictive power for floc TP. Measured soil TP ranged from 155 to 1702 mg kg−1 with a median of 433 mg kg−1 and standard deviation of 316 mg kg−1. Predictions for soil TP based on RMSE were best in the following order: RKASTER (200.1)> CKASTER (238.2)> CKETM (239.0)> OK (258.0)> REGETM (279.2)> REGASTER (281.8)> REGETM (356.1); and based on ME followed the order: CKASTER (0.1)> CKETM (0.2)> RKASTER (−5.2)> RKETM (−31.5)> OK (−41.8)> REGASTER (94.4)> REGETM (133.7). The NDVI showed the largest predictive power for soil TP. This comparative study in a subtropical wetland demonstrated the benefits of incorporating remote sensing data into floc and soil TP prediction models. Overall, hybrid geostatistical methods (CK and RK) performed better than regressions and spatial univariate models (OK) in the prediction of floc and soil TP. Depending on the strength of the spatial covariance between primary and secondary variables (CK) and the ability of the regression model in RK to explain the variability of a target variable (e.g., floc or soil TP), either CK or RK performed best. Our findings in this wetland confirmed results from earlier studies on terrestrial soils indicating the superior performance of hybrid geostatistical methods in predicting soil properties.

Published by Elsevier B.V.

Keywords: Phosphorus; Wetland; Hybrid modeling; Ordinary kriging; Co-kriging; Geostatistics; Regression kriging; Remote sensing

1. Introduction

Recently, numerous digital soil prediction models have integrated soil measurements with digital elevation models (DEM), spectral reflectance and indices derived from satellite imagery, and other ancillary environmental landscape attributes

[⁎ Corresponding author. Current address: Soil and Water Science Department, University of Florida, 2169 McCarty Hall, P.O. Box 110290, Gainesville 32611-0290, United States. Tel.: +1 352 392 1951x204; fax: +1 352 392 3902. E-mail address: sbugru@ufl.edu (S. Grunwald).
1 Current address: Everglades Foundation, 18001 Old Cutler Road, Suite 625, Miami FL 33157, United States.]

Available online 4 June 2007

doi:10.1016/j.geoderma.2007.04.026
(Grunwald and Lamsal, 2006). A variety of deterministic and stochastic methods, based in both statistics and geostatistics, and mixed (hybrid) models have been applied to quantify relationships between soil properties and corresponding environmental variables (McBratney et al., 2000, 2003; Grunwald, 2006). The success of hybrid methods stem from the fact that they incorporate (i) spatially-explicit soil observations and exhaustive grids of ancillary environmental variables (e.g. derived from remote sensing), (ii) spatial autocorrelation, (iii) spatial covariation, and/or (iv) combinations of the above. Thus, hybrid geostatistical methods have the ability to overcome shortcomings of univariate methods such as Ordinary Kriging (OK); in particular, hybrid models have been suggested to predict soil properties for cases where naturally occurring soil-forming processes are complex and anthropogenic impacts have severely altered a landscape.

Numerous studies demonstrated that hybrid geostatistical methods or combinations of kriging and correlation with auxiliary environmental data outperformed OK and plain regressions to predict various target soil properties in terrestrial soil-landscapes (Knotterts et al., 1995; Odeh et al., 1995; Dobos et al., 2000; Odeh and McBratney, 2000; Hengl et al., 2004). Commonly used hybrid methods include Regression Kriging (RK) (Odeh et al., 1995; Chilès and Delfiner, 1999), Co-kriging (CK) (Webster and Oliver, 2001), and other multivariate geostatistical methods (Goovaerts, 1997). The integration of geostatistics and remote sensing in order to explicitly take into account spatial autocorrelation and covariation has been demonstrated by several studies (Odeh et al., 1995; Marceau and Hay, 1999; Odeh and McBratney, 2000).

However, the vast majority of hybrid geostatistical soil prediction models have focused on terrestrial soil-landscapes. Commonly, univariate geospatial methods are used to model the distribution and variation of soil properties in wetland ecosystems. For example, Newman et al. (1997) predicted soil NH$_4$-nitrogen, soluble reactive phosphorus, total phosphorus (TP), and total nitrogen (TN) in a subtropical wetland in Florida; DeBusk et al. (2001) predicted TP in Water Conservation Area 2A (WCA-2A), Everglades; Litaor et al. (2003) predicted soil phosphorus, iron, aluminum, calcium and degree of phosphorus saturation in a semiarid wetland; and Bruland and Richardson (2005) predicted soil organic matter and sand content in four paired created/restored and natural wetland sites in the Coastal Plain of North Carolina. Since anthropogenically-induced phosphorus inputs stimulate microbial, periphyton and macrophyte growth in naturally oligotrophic wetlands, it is critical to describe spatial patterns of soil phosphorus to better understand processes of phosphorus retention, availability and release. Evaluation of phosphorus retention capabilities must consider both biotic (assimilation by vegetation, plankton, periphyton, and microorganisms) and abiotic processes (sedimentation, adsorption by soils, precipitation, and exchange processes between soil and the overlying water column) (Reddy et al., 2005). Both the introduction of nutrient-enriched waters and hydrologic controls in the Florida Everglades, such as WCA-2A, has induced changes in vegetation communities, periphyton and soils that have been documented by many authors (DeBusk et al., 1994; Newman et al., 1997; Reddy et al., 1998; Grunwald et al., 2004; Bruland et al., 2006; Rivero et al., 2007).

Although spatially-explicit mapping has been conducted in numerous wetlands, very few studies have employed hybrid geospatial methods (Grunwald et al., 2007). Reasons for this might be due to (i) the complexity of wetland ecosystems where soil–water–periphyton–macrophytes (vegetation) form a vague mix, (ii) topographic patterns (ridges and sloughs) are less pronounced and challenging to map with digital techniques (e.g. Light Detection and Ranging — LIDAR), (iii) difficulties to access submerged soils and to collect a sufficient amount of soil samples in a given wetland, (iv) distinguishing floc, detritus and soil materials since boundaries are difficult to detect, (v) hydrologic patterns (dry–wet periods) creating complex spatio-temporal patterns, (vi) extreme, naturally occurring stresses such as fires or hurricanes that severely alter soil characteristics, and (vii) anthropogenic stresses (e.g. nutrient inputs) that are point/location specific. Rutchev and Vilchek (1999) used SPOT imagery and aerial photography describe the confounding factors and their impact on mapping of land cover in the Everglades. The spatial and temporal/seasonal variability in tropical and subtropical wetland ecosystems, such as the Everglades, is a result of complex interactions between different wetland components (periphyton, floc, soil, vegetation, and water) that have been discussed in detail by McCormick et al. (1998), Noe et al. (2001) and Gaiser et al. (2005).

In this study we used spectral data and derived indices from two remote sensors, with different spatial resolutions, from different seasons, but similar spectral range, spatial ancillary environmental data and floc and soil TP observations. The specific objective of our study was to evaluate the performance of aspatial methods (multivariate regressions — REG), univariate spatial methods (OK), and hybrid/multivariate geostatistical methods (RK and CK) in order to predict the spatial distribution of floc and soil TP from field observations, spectral data, indices and ancillary geospatial landscape properties in a subtropical wetland in the Florida Everglades.

2. Study area: Water Conservation Area 2-A

Water Conservation Area 2A covers 43,281 ha and is located in the northern portion of the Greater Everglades, Florida. It accounts for about 6.5% of the total area of the Everglades. Soil and vegetative patterns in WCA-2A are influenced by wet and dry periods, nutrient influx, and introduction of invasive species, fire and other stresses (Wu et al., 1997; Porter and Porter, 2002). The Everglades climate is characterized by an average yearly temperature of 20 °C and average yearly precipitation between 1175 and 1550 mm (DeAngelis, 1994). The Everglades, including WCA-2A, are one of the few areas in the U.S. for which no soil survey has been conducted. Taxonomic soil descriptions in this area are rare. According to Gleason (1974) and Davis (1994) soils in the study area are Histosols and encompass the Everglades and Loxahatchee peat formations that make up the ridge and slough system. Everglades peats are characteristic of the higher elevated areas (ridges) that are generally brown to black with minimal mineral
content, while Loxahatchee peats are characteristic of the sloughs or lower elevated areas that are composed of remains of the roots and rhizomes of *Nymphaea* spp. (Brown et al., 1991).

In WCA-2A, elevation ranges from 2.0 to 3.6 m above sea level (Wu et al., 1997) generating slow sheet flow running approximately north-east to south-west. Surface hydrology is controlled by a system of levees and water control structures along the perimeter of WCA-2A. The S-10A, C, and D structures have been the most significant inflows with respect to nutrient loading to WCA-2A (South Florida Water Management District, 2005). In proximity to water control structures *Typha domingensis* (cattail) is dominant, whereas *Cladium* sp. (sawgrass) communities are found elsewhere within the marsh (Jensen et al., 1995). McCormick et al. (1999) indicated the replacement of endemic periphyton communities by algal species typical of more eutrophic waters in WCA-2A.

### 3. Methods

#### 3.1. Data

The spatial soil phosphorus data used in this study was obtained from a previous study conducted by the University of Florida Wetland Biogeochemistry Laboratory. For details on the data collection and analytical methods contact: Dr. K.R. Reddy (krr@ufl.edu), Soil and Water Science Department, University of Florida. This prior investigation sampled the Greater Everglades system in its entirety, including all Water Conservation Areas, the Holeyland and Rotenberger tracts, The Everglades National Park, Big Cypress National Preserve, and a collection of public and private lands west of ENP collectively termed the Model Lands. A detailed description of the methods of collection and analysis can be found in Bruland et al. (2006), Corstanje et al. (2006), and Rivero et al. (2007).

Two different satellite images were used as ancillary environmental data in the analysis: (i) Landsat 7 Enhanced Thematic Matter (ETM+) and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). The Landsat 7 ETM+ (scene path: 15/42, date: 02/13/2003) includes seven multi-spectral bands with 30 m spatial resolution for all bands except band 6, which is a thermal infrared (TIR) band with 60 m spatial resolution, and an additional 15 m panchromatic band. The visible bands are ETM+ 1 (blue: 0.450–0.520 μm), ETM+ 2 (green: 0.525–0.605 μm), and ETM+ 3 (red: 0.630–0.690 μm). The near-infrared band (NIR) is ETM+ 4 (0.760–0.900 μm) and the mid-infrared bands are ETM+ 5 (1.550–1.750 μm) and ETM+ 7 (2.080–2.350 μm). The TIR band is ETM+ 6 (10.400–12.500 μm). The panchromatic band covers 0.520–0.900 μm. This image was obtained, pre-processed and rectified as ancillary environmental variables into the analysis. Spectral data (bands) from each sensor, NDVI, NDVig and NDWI, and principal component scores for each sensor were extracted for each of the 111 sampling locations. Additionally, spatial data including (i) distance of each sampling site to water control structures (WCS), (ii) distance to tree islands (TI), (iii)*x* and *y* geographic coordinates, and (iv) land cover/vegetation data (Florida Fish and Wildlife Conservation Commission — Stys et al., 2004) were assembled as ancillary environmental data using ArcGIS Vers. 9.1 (Environmental Systems Research Institute — ESRI, Redlands, CA).

#### 3.2. Analysis

Four methods were used to predict floc TP and soil TP across WCA-2A, respectively, including: (i) Simple linear and
multivariate stepwise regression — REG (Ott and Longnecker, 2001), (ii) Log-normal ordinary kriging — OK (Webster and Oliver, 2001), (iii) Collocated co-kriging — CK (Wackernagel, 2003), and (iv) Regression kriging — RK (Odeh et al., 1995; Grunwald, 2007). The statistical analysis was performed using SPSS statistical package (SPSS v. 14.0 Lead Technologies), geospatial analysis in ArcGIS Vers. 9.1 (ESRI, Redlands, CA), and geostatistical analysis in ISATIS (Geovariances Inc., Houston, TX). All predictions were made on a raster with 30 m resolution.

Linear and non-linear stepwise regression models were derived to predict floc TP and soil TP, respectively, using ancillary environmental datasets (i.e., spectral data and indices, ancillary environmental spatial data) as predictor variables. Separate REG models to predict TP in floc and surface soil were developed using either Landsat ETM+ or ASTER satellite data as well as ancillary environmental variables.

To conduct OK, floc TP and soil TP data were log-transformed to stabilize variances and better comply with stationarity assumptions according to Webster and Oliver (2001). Experimental semivariograms were derived using log-floc TP and log-soil TP data. Back-transformations of log-transformed floc TP and soil TP predictions were made according to a method suggested by Webster and Oliver (2001; page 180).

Co-kriging is the multivariate extension of kriging that combines a sparsely measured primary variable (or target variable) with a denser set of ancillary data as secondary variable (e.g. remote sensing data) that is spatially cross-correlated (Grunwald, 2006). Collocated CK addresses a heterotopic situation when the variable of interest is known at a few points and the auxiliary variable is known everywhere in the domain (Wackernagel, 2003). We used the linear coregionalization model as implemented by the ISATIS software (Geovariances Inc., Houston, TX). Using this approach a model is derived which must match all the semivariograms simultaneously, i.e., the ones for each individual variable and the cross-semivariogram. The principle of this model is to define a set of basic structures so that any simple or cross-variogram must be expressed as a linear combination of these structures. The coefficients must ensure that any CK matrix is definitely positive (ISATIS, 2001). To conduct CK, numerous experimental cross-semivariograms were tested plotting spectral data, indices and environmental ancillary data vs. log-floc TP and log-soil TP, respectively. The strongest spatial cross-semivariogram structures were found for the following pairs: (i) Landsat ETM+ NDVIg—log-floc TP, and (ii) Landsat ETM NDVI—log-soil TP. Back-transformations of log-floc TP and log-soil TP predictions derived from CK were made according to Webster and Oliver (2001).

Regression kriging combines the advantages of the traditional regression methods to model the trend or drift in conjunction with geostatistical methods that model the residuals from the regression (Grunwald, 2006). In RK, also called kriging after detrending (Goovaerts, 1997), the trend or drift and the residuals from the regression are fitted separately and then summed afterwards, as has been described by Odeh et al. (1995). The deterministic component is modeled using multiple linear regression or other regression variants. The stochastic component, that represents the spatially varying but dependant component, is modeled using OK or simple kriging of the regression residuals (Odeh et al., 1995). To conduct RK, the best performing linear and non-linear regression models developed under (i) were used to predict floc TP and soil TP, respectively. Residuals were derived by subtracting the floc TP predictions (derived from REG) from observed floc TP at each sampling point. The same procedure was adopted to derive soil TP residuals. Ordinary kriging was performed on log-floc TP and log-soil TP residuals, which were back-transformed according to Webster and Oliver (2001). The floc TP residuals were added to the floc TP predictions derived from REG to produce the final floc TP prediction map. Likewise soil TP residuals were added to the soil TP predictions derived from REG to produce the final soil TP prediction map.

The performance of each TP prediction model was evaluated using: (i) a qualitative visual assessment at fine scale comparing the predicted values in impacted or phosphorus-enriched zones (i.e., near canals, WCS, and tree islands) and non-impacted areas (i.e., marsh interior); (ii) quantitative measures of precision (measures of the residual variability in prediction) and accuracy, that measures the closeness of the prediction to the true conditions (Mueller et al., 2001). We used the coefficient of determination ($R^2$), mean error (ME), and RMSE to evaluate model performances in cross-validation mode. The smallest RMSE and a ME close to zero indicate the most accurate predictions. The ME and RMSE were derived according to Eqs. (4) and (5).

$$\text{ME} = \frac{1}{N} \sum_{i=1}^{N} \{z(x_i) - \hat{z}(x_i)\}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \{z(x_i) - \hat{z}(x_i)\}^2}$$

with

$N$ Number of samples
$z(x_i)$ Observed value at point $x_i$
$\hat{z}(x_i)$ Predicted value at point $x_i$

4. Results and discussion

The regression models used to predict log-floc TP and log-soil TP are summarized in Tables 1 and 2. The highest $R^2$ of 0.75 (model REGm) was achieved using NDVIg derived from Landsat ETM+ (ETM-NDVig), distance of each sampling site to water control structures (WCS) and $Y$-coordinates. The two latter variables contributed only a very small proportion to the overall prediction of log-floc TP. At the WCS nutrient influx is highest with a nutrient gradient that extends for about 7 km into the marsh interior (Noe et al., 2001). The $Y$-coordinate indicates...
prominent flow patterns across WCA-2A. Both the linear (model REGl) and quadratic (model REGq) regression models used ETM-NDVIg as predictor variable with $R^2$ of 0.68 and 0.71 and RMSE of 264 and 220 mg kg$^{-1}$, respectively. The regression models confirmed our assumption that ETM-NDVIg responds to a mix of carotenoids, chlorophyll $a$ and $b$, of the land cover of this wetland that is correlated to the phosphorus content in floc. Sites that showed high floc TP (max. of 1865 mg kg$^{-1}$) showed high ETM-NDVIg values (max. of 0.10), whereas low floc TP sites (min. of 194 mg kg$^{-1}$) corresponded to very low ETM-NDVIg values (min. of −0.18). A similar relationship was found between floc TP and NDVIg derived from ASTER, however, not as strong as between floc TP and NDVIg derived from the Landsat ETM+ image. Considering that measured floc TP showed a range of 1671 mg kg$^{-1}$ the RMSEs were moderately high for the linear, quadratic and multivariate regression models (Table 1). All three regression models underpredicted floc-TP according to MEs.

The regression models predicting soil TP showed weaker relationships as indicated by $R^2$, ME and RMSE (Table 2) when compared to floc TP regression models. The $R^2$ for the linear regression to predict log-soil TP was 0.46 using NDVI derived from the Landsat ETM+ image (ETM-NDVI) and 0.39 using NDVI derived from the ASTER (ASTER-NDVI) image. Interestingly, the NDVI had more predictive power to infer on vegetation that showed a functional relationship to soil TP. The ETM-NDVI (model REG_ETM) showed a stronger relationship to soil TP when compared to ASTER-NDVI (model REG_ASTER) (Table 2), which may be explained by the different seasons of the images. The Landsat ETM+ image was from the spring season (February) when periphyton was absorption band of healthy green vegetation. Jensen et al. (1995) found in a remote sensing study conducted in WCA-2A that Typha reflected more NIR radiant flux, while absorbing similar amounts of red radiant flux than the other vegetation and land cover types (Cladium and mixtures of Typha/Cladium).

In contrast, the green band incorporated in the NDVIg not only responds to a mix of carotenoids, chlorophyll $a$ and $b$, but it also has some ability to penetrate into water and therefore provided better inference on floc TP; but not on soil TP. Wetland soils are overlain by floc, detritus, floating periphyton mats with different depth and density and water, and therefore showed no affinity to the green band. In contrast, the NDVI responds to the amount of green biomass and chlorophyll content in vegetation (Jensen, 2000). In WCA-2A dense Typha stands were mapped south-southwest of the S-10 water control structures extending into the marsh interior migrating into transitional zones of sawgrass/freshwater marsh vegetation (Stys et al., 2004).

Rivero et al. (2007) found that Typha zones in WCA-2A showed elevated soil TP at 0–10 depths with means of 667 mg kg$^{-1}$ (SD: 150 mg kg$^{-1}$) in Typha contrasted by means of 394 mg kg$^{-1}$ (SD: 100 mg kg$^{-1}$) in Cladium stands. Similarly, the spatial association between higher levels of soil TP and Typha have been documented in varied hydrologic units within the Everglades by Wu et al. (1997), Newman et al. (1998), and Noe et al. (2001). These findings suggest that NDVI provided a proxy to infer on vegetation that showed a functional relationship to soil TP.

The ETM-NDVI (model REG_ETM) showed a stronger relationship to soil TP when compared to ASTER-NDVI (model REG_ASTER) (Table 2), which may be explained by the different seasons of the images. The Landsat ETM+ image was from the spring season (February) when periphyton was

Table 1: Regression models to predict log-floc TP (mg kg$^{-1}$) in WCA-2A, Everglades

<table>
<thead>
<tr>
<th>Models</th>
<th>$r^a$</th>
<th>$R^b$</th>
<th>Adjusted $R^2$</th>
<th>Standard error of estimate</th>
<th>Regression equations</th>
<th>Sig.</th>
<th>ME$^c$</th>
<th>RMSE$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear (REGl)</td>
<td>0.83</td>
<td>0.68</td>
<td>0.67</td>
<td>0.119</td>
<td>$1.842+1.952*ETM-NDVIg$</td>
<td>0.00</td>
<td>25.1</td>
<td>264.4</td>
</tr>
<tr>
<td>Quadratic (REGq)</td>
<td>0.84</td>
<td>0.71</td>
<td>0.70</td>
<td>0.114</td>
<td>$0.713+6.236<em>ETM-NDVIg−3.954</em>ETM-NDVIg$</td>
<td>0.00</td>
<td>27.3</td>
<td>220.3</td>
</tr>
<tr>
<td>Multi-variate (REGmv)</td>
<td>0.75</td>
<td>0.75</td>
<td>0.105</td>
<td>0.105</td>
<td>$4.237+1.277<em>ETM-NDVIg−0.0000267</em>$</td>
<td>0.01</td>
<td>18.2</td>
<td>218.3</td>
</tr>
</tbody>
</table>

Variables: ETM-NDVIg=Normalized Difference Vegetation Index green derived from Landsat ETM+; WCS=Distance of each sampling site to water control structures; Y=coord=Y geographic coordinate in m (Albers Equal Area Conic map projection).

$^a$ $r$: Pearson correlation coefficient.

$^b$ $R^2$: Coefficient of determination.

$^c$ ME: Mean prediction error on back-transformed TP values (mg kg$^{-1}$).

$^d$ RMSE: Root mean square error on back-transformed TP values (mg kg$^{-1}$).

Table 2: Regression models to predict log-soil TP (mg kg$^{-1}$) in WCA-2A, Everglades

<table>
<thead>
<tr>
<th>Models</th>
<th>$r^a$</th>
<th>$R^b$</th>
<th>Adjusted $R^2$</th>
<th>Standard error of estimate</th>
<th>Regression equations</th>
<th>Sig.</th>
<th>ME$^c$</th>
<th>RMSE$^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear (REG_ETM)</td>
<td>0.68</td>
<td>0.46</td>
<td>0.46</td>
<td>0.167</td>
<td>$1.905+1.457*ETM-NDVI$</td>
<td>0.00</td>
<td>133.7</td>
<td>356.1</td>
</tr>
<tr>
<td>Linear (REG_ASTER)</td>
<td>0.63</td>
<td>0.39</td>
<td>0.39</td>
<td>0.177</td>
<td>$2.572+1.440*ASTER-NDVI$</td>
<td>0.00</td>
<td>94.4</td>
<td>281.8</td>
</tr>
</tbody>
</table>

Variables: ETM-NDVI=Normalized Difference Vegetation Index derived from Landsat ETM+; ASTER-NDVI=Normalized Difference Vegetation Index derived from ASTER satellite image.

$^a$ $r$: Pearson correlation coefficient.

$^b$ $R^2$: Coefficient of determination.

$^c$ ME: Mean prediction error on back-transformed TP values (mg kg$^{-1}$).

$^d$ RMSE: Root mean square error on back-transformed TP values (mg kg$^{-1}$).
submerged and did not affect the spectral values. In contrast, the ASTER image corresponded to the end of the wet season (September) when periphyton covered a good portion of the areas with floating mats, which might have contributed to presence of more pixels with “mixed” signatures due to a mix of water, periphyton, vegetation, and floc. These findings suggest that spectral data from the spring season are better suited to infer on floc and soil TP. Interestingly, the higher spatial resolution of the ASTER image (15 m) did not improve the predictive capabilities to infer on floc and soil TP when compared to the coarse (30 m) Landsat ETM+ satellite image.

The NDWI, derived from NIR and SWIR, was developed for aquatic systems and enhances soil/vegetation and land/water contrasts (Jensen, 2000). This index did not show any significant correlations to floc TP and soil TP. This might be due to the fact that only a small portion (<5% in 2003) of WCA-2A was covered by open water (Stys et al., 2004).

The interpolation parameters for various floc TP prediction models are presented in Table 3 and coregionalization matrix parameters (CK) in Table 4. A spherical model was fitted with nugget of 0.00268, sill of 0.04971, and range of 9669 m for log-floc TP that was employed to perform OK. The spatial autocorrelation range was high indicating regional spatial distribution patterns. The nugget was very low relative to the sill indicating that most of the fine-scale variability was explained by the semivariogram model. Cross-validation results for floc TP from three hybrid geostatistical methods are summarized in Table 5 including: (i) RK model based on the multivariate regression model (REGm) – RKm; (ii) RK model based on the quadratic regression model (REGq) – RKq; (iii) CK with ETM-NDVig as secondary variable. A ME departing from zero and high RMSE suggest that the prediction model was limited to capture the underlying spatial variability of the observed parameters highlighting uncertainties in the prediction process due to sampling design, measurement errors, spacing of observations and other factors. Predictions for floc TP in mg kg\(^{-1}\) based on RMSE were best in the following order: RKq (134.9) > RKm (201.1) > OK (206.1) > CK (212.1). The ME was highest for OK (−27.3) followed by CK (−6.7), the latter one was one order magnitude smaller when compared to the regression models predicting floc TP. The MEs for both RK models predicting floc TP were very small. The RMSE for OK was significantly smaller when compared to the three regression models to predict floc TP. However, the ME for OK was larger than for the three regression models to predict floc TP. This indicates that the univariate interpolation model (OK) over-predicted floc TP at the 111 sites used for cross-validation. But overall the prediction accuracy for floc TP was better for the simple interpolation method of OK than all considered regression models. This suggests that the modeling of the spatial autocorrelation structure was more important for the prediction of floc TP than the incorporation of secondary data (e.g. from satellite images) into regression models. Thus, for

Table 3
Summary of interpolation parameters for log-TP by layer (floc and soil surface; (mg kg\(^{-1}\)) and each prediction model in WCA-2A, Everglades

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>Variogram model</th>
<th>Nugget</th>
<th>Sill</th>
<th>Range (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floc TP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinary kriging (OK)</td>
<td>Spherical</td>
<td>0.00268</td>
<td>0.04971</td>
<td>9669</td>
</tr>
<tr>
<td>Regression kriging (RK) – Kriging of residuals</td>
<td>Spherical</td>
<td>0.007272</td>
<td>0.01279</td>
<td>9838</td>
</tr>
<tr>
<td>RKm (multiple regression)</td>
<td>Spherical</td>
<td>0.006616</td>
<td>0.01313</td>
<td>4453</td>
</tr>
<tr>
<td>RKq (quadratic regression)</td>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-kriging (CK)</td>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil TP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinary kriging (OK)</td>
<td>Exponential</td>
<td>0.0185</td>
<td>0.0512</td>
<td>7468</td>
</tr>
<tr>
<td>Regression kriging (RK) – Kriging of residuals</td>
<td>Spherical</td>
<td>0.0059</td>
<td>0.0323</td>
<td>2038</td>
</tr>
<tr>
<td>RKETM</td>
<td>Spherical</td>
<td>0.0020</td>
<td>0.0071</td>
<td>1903</td>
</tr>
<tr>
<td>RKASTER</td>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-kriging (CK)</td>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Co-kriging (CK) coregionalization matrix for log-TP and soil surface (mg kg\(^{-1}\) in WCA-2A, Everglades

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>Coregionalization matrix parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floc TP</td>
<td></td>
</tr>
<tr>
<td>Sill</td>
<td>0.0475 0.0173</td>
</tr>
<tr>
<td>Nugget</td>
<td>−0.0009 0.0008</td>
</tr>
<tr>
<td>Soil TP</td>
<td></td>
</tr>
<tr>
<td>Sill</td>
<td>0.0008 0.0006</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil TP</td>
<td></td>
</tr>
<tr>
<td>Sill</td>
<td>0.0009 0.0218</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5
Summary of cross-validation statistics (ME: mean prediction error and RMSE: root mean square prediction error) for each layer (floc and soil surface; mg kg\(^{-1}\)) and each prediction model in WCA-2A, Everglades

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>ME(^{a}) (mg kg(^{-1}))</th>
<th>RMSE(^{b}) (mg kg(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Floc TP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinary kriging (OK)</td>
<td>−27.3</td>
<td>206.1</td>
</tr>
<tr>
<td>Regression kriging (RK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RKm (multiple regression)</td>
<td>0.9</td>
<td>201.2</td>
</tr>
<tr>
<td>RKq (quadratic regression)</td>
<td>1.1</td>
<td>134.9</td>
</tr>
<tr>
<td>Co-kriging (CK)</td>
<td>−6.7</td>
<td>212.1</td>
</tr>
<tr>
<td>Soil TP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ordinary kriging (OK)</td>
<td>−41.8</td>
<td>257.5</td>
</tr>
<tr>
<td>Regression kriging (RK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RKETM</td>
<td>−31.5</td>
<td>279.2</td>
</tr>
<tr>
<td>RKASTER</td>
<td>−5.2</td>
<td>200.1</td>
</tr>
<tr>
<td>Co-kriging (CK)</td>
<td>0.2</td>
<td>238.8</td>
</tr>
<tr>
<td>CKETM</td>
<td>0.1</td>
<td>238.2</td>
</tr>
<tr>
<td>CKASTER</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}\) ME: Mean prediction error on back-transformed TP values (mg kg\(^{-1}\)).

\(^{b}\) RMSE: Root mean square error on back-transformed TP values (mg kg\(^{-1}\)).
modeling of floc TP the stochastic spatially-dependent component provided better predictive power when compared to modeling of the deterministic trend component alone. Yet the most accurate predictions for floc TP were made using RKq that incorporated both, the spatial autocorrelation structure and covariation with secondary variables (i.e., ETM-NDVig, WCS and Y-coordinate). The prediction quality for RKq was significantly better when compared to all other prediction models suggesting this model to be superior. In summary, multivariate hybrid/geostatistical models (RKmp, RKm, and CK) provided more accurate predictions for floc TP than the univariate model (OK), except for CK that showed a slightly higher RMSE than OK.

The final floc TP predictions maps derived from RK are shown in Figs. 1 and 2. The sill to nugget ratio was comparable in both RK floc TP residual models. However, the semivariograms of TP residuals showed different spatial autocorrelation ranges with 9838 m (RKm) and 4453 m (RKq) translating into long-range patterns of TP shown on the residual RKm map and medium-range patterns of TP shown on the residual RKq map. Thus, modeling of the stochastic, spatially dependent residual component contributed differently to mapping of the variation of floc TP in RKm and RKq mode across WCA-2A. The final prediction maps based on RKm (Fig. 1) and RKq (Fig. 2) both show higher floc TP in the eastern part and western edge coinciding with dense Typha stands. Elongated landscape features such as tree islands/hammocks and tails can be clearly distinguished from the matrix of the sawgrass/freshwater marsh vegetation (lighter colors). Orem et al. (2002) found TP concentrations of 1500 (dry head Gumbo Limbo) to 3000 (wet head Gumbo Limbo) mg kg\(^{-1}\), which is a magnitude of order higher when compared to pristine marsh areas of the Everglades ranging from 200 to 500 mg kg\(^{-1}\) (Koch and Reddy, 1992). Besides anthropogenic induced phosphorus inputs via WCS, tree islands are an additional source of phosphorus into the system. Orem et al. (2002) suggested guano from nesting wading birds, ground water upwelling, active transport of phosphorus from underlying limestone by tree roots, and concentration of phosphorus in the soils on tree island heads by fire all cause phosphorus to accumulate on tree islands. The RK floc TP prediction maps differ with lower TP floc values predicted by RKq than RKm in the marsh interior and along a NW–SE axis across WCA-2A. Like the RK-based floc TP prediction maps the CK-based map (Fig. 3) shows similar regional distribution patterns that are slightly more muted. The cross-semivariogram model was fitted close to the hull of perfect correlation with a long spatial autocorrelation range of 10,328 m.

The percent coverage of predicted area with floc TP<300 mg kg\(^{-1}\) was 0% (OK), 0.77% (CK), 0.92% (RKm).
and 5.58% (RK_q). Predictions in floc TP ranging from 300–600 mg kg^{-1} covered 45.1% (RK_q), 52.9% (OK), 55.2% (RK_m) and 55.3% (CK) of WCA-2A. Differences in floc TP predictions >600 mg kg^{-1} were found with percent coverage of 43.9% (CK and RK_m), 47.0% (OK) and 49.3% (RK_q). Only CK, RK_q and RK_m were able to predict high TP floc values exceeding 1500 mg kg^{-1} resembling some of the high observed TP measured in floc in impacted areas in proximity to WCSs.

The interpolation parameters for log-soil TP are presented in Tables 3 and 5. An exponential model was fitted with nugget of 0.0185, sill 0.0512, and range of 7469 m to describe the spatial variability of log-soil TP. Important to note is that the spatial autocorrelation range was large indicating that regional patterns for soil TP prevailed. The spatial dependence structure as indicated by the nugget to sill ratio was less strong when compared to the one for floc TP. Cross-validation results for soil TP derived from OK and four hybrid/geostatistical methods are summarized in Table 5 including: (i) RK model based on the linear regression model using ETM-NDVI as independent predictor variable (REG_{ETM})—RK_{ETM}; (ii) RK model based on the linear regression model using ASTER-NDVI as independent predictor variable (REG_{ASTER})—RK_{ASTER}; (iii) CK with ETM-NDVI as secondary variable — CK_{ETM}; and (iv) CK with ASTER-NDVI as secondary variable — CK_{ASTER}.

Predictions for soil TP in mg kg^{-1} based on RMSE were best in the following order: RK_{ASTER} (200.1) > CK_{ASTER} (238.2) > CK_{ETM} (239.0) > OK (258.0) > RK_{ETM} (279.2). Measured soil TP ranged from 155.3 to 1701.5 mg kg^{-1} with a median of 432.8 mg kg^{-1} (SD: 316.2 mg kg^{-1}). Interestingly, CK-based predictions of soil TP with both sensors performed almost the same suggesting that seasonal differences between the two sensors (ASTER and Landsat ETM+) did not play as important a role as they did for the predictions of floc TP. The reason for this may be that processes of soil TP accumulation are driven by macrophytes and inoculation of particulate phosphorus into the soil matrix, while floc TP processes are driven by nutrient-laden water and periphyton (Reddy et al., 1997; Gaiser et al., 2005). However, this was contradicted by the different performance in RK models, where the RK_{ASTER} model performed superior when compared to the RK_{ETM} model in terms of ME and RMSE.

Both ME and RSME for soil TP regression models were higher than for CK_{ETM}, RK_{ETM}, CK_{ASTER} and RK_{ASTER} models. These results confirm the contribution of the spatial modeling of the residuals to the final predictions with lower ME and RSME in the hybrid/geostatistical models. The accuracy of predictions of soil TP derived from OK were superior when compared to the linear regression models.
suggesting that the modeling of spatial autocorrelation was more important to predict soil TP than incorporation of ancillary environmental variables (e.g. NDVI). However, it was also shown that models (CK and RK) that incorporated both, spatial covariation and autocorrelation structures into the modeling process, performed best to predict soil TP. Co-kriging performed well to predict soil TP based on small MEs with 0.14 for CKASTER and 0.18 for CKETM, respectively. Although the accuracy of the best soil prediction model was less than the best model for floc TP, it could be shown that soil models were much improved by incorporating exogenous spectral data. Since remote sensing data are cost-effective and readily-available it is critical to incorporate them into prediction models to accurately characterize the nutrient status across a wetland. Although topographic variables and seasonal-variable water depth might also have power to predict floc and soil TP they are more difficult to derive for wetlands. In the field it is challenging to distinguish floc from soil and it is even more challenging to accurately measure bathymetry in wetlands because the soil surface boundary is not well defined.

The percent coverage of predicted area with soil TP <300 mg kg⁻¹ was smallest for RKETM with 6.3% followed by OK (7.9%), CKETM (8.9%), CKASTER (9.3%) and RKASTER (17.0%). These values can be compared to soil TP observations where 13.6% of 111 samples showed <300 mg kg⁻¹, 53.6% 300–600 mg kg⁻¹, and 32.7% exceeded 600 mg kg⁻¹. Differences among prediction methods were found in the soil TP prediction range of 300–600 mg kg⁻¹ that covered 61.7% (RKASTER), 67.4% (CKASTER), 76.1% (OK), 77.5% (CKETM) and 82.8% (RKETM) of WCA-2A. Variations in soil TP predictions that exceeded 600 mg kg⁻¹ were also found with 10.9% (RKETM), 13.5% (CKETM), 16.0% (OK), 21.3% (RKASTER), and 23.2% (CKASTER). Ordinary kriging failed to predict TP soil values >1200 mg kg⁻¹ that were measured within the impacted zones of WCA-2A. Other methods predicted percent coverage of soil TP larger than 1200 mg kg⁻¹ on 0.13% (CKASTER), 0.54% (RKASTER), 0.64% (RKETM) and 0.68% (CKETM) of the wetland.

To solely base model evaluations on error assessment can be misleading because geographic distributions of over- and underpredictions are masked in these error metrics. Thus, we conducted an in-depth visual comparison of observed and predicted floc TP values in an impacted (i.e., phosphorus-enriched area) and non-impacted area (i.e., marsh interior) (Figs. 4 and 5). In particular for ecological assessment it is important to predict high TP values accurately to assess phosphorus storage across the wetland and subsequent implications for biogeochemical cycling (P, C, and N-cycles). Interestingly, measured floc TP was extremely variable across the impacted zone with values ranging from about 880 to 1865 mg kg⁻¹ that were smoothed out by OK predictions (Fig. 4). Regression kriging (RKₘ) and CK seemed to
overpredict in areas near the WCS structures, while RKq predicted more accurately in this nutrient-enriched area. In general, all three hybrid/geostatistical models tended to under-predict floc TP in the impacted area. In the non-impacted area (Fig. 5) both RKm and CK were very similar and smooth in comparison with RKq that showed speckled patterns in floc TP.

Ordinary kriging smoothed floc TP since interpolations solely relied on observations. In particular RKm predicted poorly at sites with very low observed values of floc TP (e.g., site #183 with 194.1 mg kg\(^{-1}\)). High measured floc TP (>500 mg kg\(^{-1}\)) in this non-impacted area occurred near tree island tails, while areas in sloughs showed lowest TP floc (<400 mg kg\(^{-1}\)).

Fig. 4. Comparison of floc TP (mg kg\(^{-1}\)) prediction models in an impacted area of WCA-2A (highlighted red rectangle): (a) Regression Kriging with multivariate regression (RKm); (b) Regression Kriging with quadratic regression (RKq); (c) Co-kriging with Landsat Enhanced Thematic Mapper - Normalized Difference Vegetation Index green (ETM-NDVIg) as secondary variable; (d) Ordinary Kriging (OK) with superimposed measured floc TP values (mg kg\(^{-1}\)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
differences are not as noticeable in the $R_{Km}$, $R_{Kq}$ and CK maps. Residuals in floc TP in this area are relatively small ($<100$ mg kg$^{-1}$), therefore contributing less to prediction errors.

Predictions of soil TP based on OK, RK and CK in impacted and non-impacted areas are shown in Figs. 6 and 7. The $CK_{ETM}$ model showed much higher soil TP predictions in the impacted area when compared to $CK_{ASTER}$ in particular along the Hillsboro canal (north-east of study area). The $RK_{ASTER}$ model followed by $CK_{ASTER}$ and $RK_{ETM}$ performed best to predict soil TP accurately in this phosphorus-enriched area. Because the variability in soil TP was large in this impacted area it was challenging to model the spatial distribution accurately. Measured soil TP values of 1260 mg kg$^{-1}$ were found in
close proximity to 380 mg kg$^{-1}$ showing abrupt changes possibly caused by flow patterns, elevation differences, or disproportionate nutrient influx from WCS. Few sites that showed elevated floc TP ($>1000$ mg kg$^{-1}$) showed relatively low soil TP (<400 mg kg$^{-1}$) indicating that phosphorus was either not transported from floc into the surface soils, was resuspended from the soil, or removed by macrophytes. Most of the sites that were elevated in floc TP showed

![Diagram](image-url)
also high soil TP. In the non-impacted areas (Fig 7) CK\textsubscript{ETM} showed good predictions of soil TP. There was a gradient in the observed soil TP values from 734 mg kg\(^{-1}\) in the NE corner to center site with 433 mg kg\(^{-1}\) and 367 mg kg\(^{-1}\) (SW). Although the general phosphorus gradient was captured by all models (even by the OK model), CK\textsubscript{ETM} and RK\textsubscript{ASTER} were the ones that provided the best representation of small landscape features (e.g. tree islands, and tree island tails) along with accurate

---

**Fig. 7.** Comparison of soil TP (mg kg\(^{-1}\)) prediction models in an unimpacted area in the interior of WCA-2A (highlighted blue rectangle): (a) Regression Kriging based on linear regression model using Landsat ETM\(^{+}\) Normalized Difference Vegetation Index (NDVI) as independent predictor variable — RK\textsubscript{ETM}; (b) Regression Kriging based on linear regression model using ASTER NDVI as independent predictor variable — RK\textsubscript{ASTER}; (c) Co-kriging with Landsat Enhanced Thematic Mapper-Normalized Difference Vegetation Index (ETM-NDVI) as secondary variable — CK\textsubscript{ETM}; (d) Ordinary Kriging (OK) with superimposed measured soil TP values (mg kg\(^{-1}\)) and (e) Co-kriging with ASTER-NDVI as secondary variable — CK\textsubscript{ASTER}. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
predictions of soil TP. This confirms that the incorporation of remote sensing data that capture land cover/vegetation improves soil prediction models.

Our findings that hybrid/geostatistical methods performed better than univariate interpolations (i.e., OK) and regression models, are in line with other studies performed on terrestrial soils. Knotters et al. (1995) found that RK performed better than CK and OK to predict soft layer depths (peat and unripe clay) using the bulk soil electrical conductivity as auxiliary variable. Hudak et al. (2002) performed a similar comparison of methods to integrate LIDAR and Landsat ETM+ data to predict forest canopy height. They called aspatial models those based on statistical methods (stepwise multiple regression), the spatial models those based on OK and integrated methods, similar to what we have called hybrid methods, integrating the residuals from the regression models that were used for kriging. They found that most of the biases in the regression estimates were eliminated in the integrated models, where the regression residuals were subsequently kriged and added back to the regression surface. Odeh and McBratney (2000) have demonstrated the superiority of RK to other prediction methods such as OK, universal kriging, multiple-linear regression and CK across various spatial scales. They also demonstrated the ability to improve predictions by including remotely sensed data that is available at denser locations than the soil target variable. Dungan (1998) compared three geostatistical methods (OK, CK, and stochastic simulation) with a traditional linear regression model, using remote sensing data, in order to predict vegetation parameters such as biomass, canopy cover, and stems per area unit for a forestry study. Dungan (1998) indicated that when correlation coefficients between sample and ancillary data did not exceed 0.89, the best model (lower RMSE) was obtained from CK. However, Knotters et al. (1995) obtained different results, indicating that in the presence of correlations lower than 0.70, CK and kriging combined with linear regression produced comparable results. In our study, correlation coefficients between remote sensing indices and floc TP reached a maximum of 0.87 with the multivariate regression model (REGm), followed by 0.84 with the quadratic regression (REGq), and 0.83 with the linear regression model (REG). These strong relationships translated into excellent performance of the RKq model that performed better than CK to predict floc TP in WCA-2A. In contrast, correlations to soil surface TP showed lower correlation coefficients with 0.68 (REGETM) and 0.63 (REGASTER) for linear regressions resulting in good performance of CKASTER and CKETM models. The stronger relationships between the remote sensing data and floc TP produced stronger regression models with $R^2$ up to 0.75 when compared to a maximum $R^2$ of 0.46 to predict soil TP. In our study, CKASTER and CKETM models performed best considering the ME and RKASTER performed best considering the RMSE to predict soil TP in WCA-2A. If the regression model in RK is strong, indicated by a high $R^2$, much of the variability of a target variable is explained; thus, the spatial modeling of the residuals contributes only a small portion to improve the final predictions (and vice versa). Depending on the strength of the spatial covariance between primary and secondary variables (CK) and the ability of the regression model in RK to explain the variability of a target variable (e.g., floc or soil TP), either CK or RK perform best. Odeh et al. (1995) pointed out that the better performance of CK occurs in those cases where the target soil variable is relatively more correlated with the predictor variable than those variables that vary systematically across the landscape. Odeh et al. (1995) argued that models perform different depending on particular cases, and it could be assumed that there is no single best method for all variables and landscape settings.

Based on numerous studies (Knotters et al., 1995; Odeh et al., 1995; Dungan, 1998; McBratney et al., 2000) complex hybrid/geostatistical models have been shown to perform better than univariate spatial methods and aspatial methods (regressions). While much focus has been to develop predictive soil models for terrestrial ecosystems our study focused on wetland soils. Since our study is one of the first to compare various statistical and hybrid/geostatistical methods to predict floc TP and soil TP in subtropical wetlands it provides critical knowledge for future studies. Our study area located within the Everglades is unique because of its hydrologic conditions (e.g., water pump stations, management to control hydropatterns within hydrologic units of the Everglades), slow sheet flow that has produced pronounced slough/ridge patterns, anthropogenic impact (nutrient impact) and mix of natural and invasive vegetation species (e.g. Typha sp., Melaleuca quinquenervia, Schinus terebinthifolius — Brazilian Pepper). Given these unique landscape conditions where biogeochemical variables in floc and soils show distinct spatial variability (DeBusk et al., 2001; Bruland et al., 2006; Corstanje et al., 2006; Bruland et al., in press; Rivero et al., 2007) more research is needed to identify soil prediction models for wetlands customized towards specific properties. This study showed to what extent remote sensing improves floc/soil TP prediction models.

5. Conclusions

This comparative study in a subtropical wetland proved the benefits of incorporating remote sensing data into floc and soil TP prediction models. Hybrid/geostatistical methods (CK and RK) performed best to predict floc and soil TP when compared to aspatial (regression) and spatial univariate models (OK). Univariate methods rely solely on the observed soil property values, ignoring the spatial variability of other components of this ecosystem such as periphyton, the ridge and slough dynamic and tree islands that can be characterized by remote sensors. These landscape features can be captured at high spatial resolution by remote sensors such as Landsat ETM+ or ASTER providing continuous, exhaustive data across the whole wetland. Results from this study confirmed that remote sensing can play an important role in improving predictions of phosphorus levels in WCA-2A, and that several characteristics of a wetland area, despite limitations, can be captured by these sensors. Depending on the strength of the relationships between target soil variable and ancillary environmental variables (REG, CK, and RK), the spatial variability of soil observations (OK, CK) and residuals (RK) models differ in performance. While similar studies were performed on terrestrial soils, our study focused on wetland soils that are more complex due to soil/water interactions, transport processes, wet/dry cycles that impact pedogenesis and
biogeochemical cycling, microbial induced transformations, diverse periphyton assemblages and soil/vegetation interaction. In addition, WCA-2A is impacted by nutrient influx and frequent natural events such as fire and tropical storms/hurricanes. Despite the complexity of ecosystem conditions in WCA-2A, our study showed that hybrid/geostatistical modeling of floc-soil TP has value for the mapping of the distribution of ecologically-important soil properties at fine resolution with high accuracy. Since there are, and will be limitations about the numbers of soil samples that can be collected in any given study, the combination of remote sensing and geostatistics can improve or extend the prediction capabilities beyond sampled points. These prediction models can be used to improve monitoring efforts in similar wetland areas, taking advantage of existing data from remote sensors that are collected with a higher frequency than soil samples. There are numerous future opportunities to incorporate such cost-effective remote sensing into digital soil mapping providing numerous possibilities to conduct change detection analysis.

Acknowledgments

This research was supported with Funds from the SFWMD and various other sources. The authors thank Drs. Sue Newman, South Florida Water Management District, Todd Osborne, and K. R. Reddy, Wetland Biogeochemistry Laboratory for allowing us to use their spatial soil phosphorus data in this multivariate study. We also wish to thank Michael Binford for assistance with interpretation of remote sensing data.

References


Gleason, P.J., 1974. The Chemical Quality of Water in Conservation Area 2A and Associated Canals. Central and South Florida Food Control District, West Palm Beach, FL.


