MODELING OF SOIL PROPERTIES IN THE SANTA FE RIVER WATERSHED USING EXHAUSTIVE SPATIAL ENVIRONMENTAL DATA

By

JINSEOK HONG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

To my family, friends and GOD

ACKNOWLEDGMENTS

I would like to express my gratitude to my chair, Dr. Scot E. Smith, and co-chair, Dr. Sabine Grunwald, for patience, support, and guidance during my pursuit of the Ph.D. degree. I also acknowledge my supervisory members, Dr. Bon A. Dewitt, Dr. David W. Gibson, Dr. Grenville Barnes and Dr. Pierre Goovaerts for their expertise in different aspects of this research. I also thank Dr. Amr Abd-Elrahman for his time and suggestions for my defense.

I thank my colleagues, Sanjay Lamsal, Rosanna Rivero, Brent Myers, Deoyani Sarkhot, Brandon Hoover, Aja Stoppe, William Deich IV, Wade Ross, Miyoun Ahn, Ho-young Kwon, Jongsung Kim, Siripon Kamontum, Zoltan Szantoi, Benjamin Wilkinson, and Sowmya Selvarajan from the Soil and Water Science Department and Geomatics program at the University of Florida for their help and sharing their knowledge. I am very grateful to offer special thanks to Dr. Gustavo Vasques for his great help and friendship.

All my family and friends in Korea and the United States have encouraged me to pursue this amazing journey with great love and support. They always trusted and supported my decisions. I give very special thanks to my wife, Dr. Eunkyung Lee, and my daughter, Ellen L. Hong for their love and encouragement every day whether it was a pleasant day or not. This dissertation would not have been completed without them.

Last but not least, I would like to express my deepest gratitude to my God's never-ending grace and love. He has given me eyes and ears to watch and listen to this world with different points of view so that I can share my blessing with other people in this world. He gave me this amazing special opportunity to understand other people better and to become a spiritually stronger person.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS 4

LIST OF TABLES 7

LIST OF FIGURES 9

ABSTRACT 11

CHAPTER

1 INTRODUCTION 14

2 DESCRIPTION OF STUDY AREA AND ENVIRONMENTAL DATA 22

Study Area 22
Santa Fe River Watershed 22
Santa Fe River Ranch Beef Unit 23
Soil Samples and Laboratory Analysis 24
Santa Fe River Watershed 24
Santa Fe River Ranch Beef Unit 25
Exhaustive Spatial Environmental Data 26
Santa Fe River Watershed 26
Santa Fe River Ranch Beef Unit 26

3 INTRODUCTION TO SOIL-LANDSCAPE MODELING 40

Soil Mapping History and Geostatistical Modeling 40
Applications of Soil-Landscapes Modeling 45
Remote Sensing in Soil Mapping 47
Limitations and Research Gaps 50

4 DIGITAL SOIL MAPPING OF SOIL PHOSPHORUS IN THE SANTA FE RIVER WATERSHED, FLORIDA 54

Materials and Methods 55
Soil Sampling, Laboratory Analysis, and Exhaustive Spatial Environmental Data 55
Exploratory Analysis of Soil Data: Comparison of Soil MP at Four Depths 55
Quantify Relationships Between Soil Phosphorus and Environmental Variables 56
Geostatistical Modeling of Soil Phosphorus 56
Model Validation 59
Results and Discussion 60
Exploratory Data Analysis 60
Relationships between Soil MP and Environmental Variables 61
Geostatistical Modeling of Soil Phosphorus and Model Validation 68

5 MODELING OF SOIL PHOSPHORUS USING SATELLITE IMAGERY AND ANCILLARY SPATIAL ENVIRONMENTAL DATA IN THE SANTA FE RIVER RANCH BEEF UNIT 95

Materials and Methods 97
Soil Sampling, Laboratory Analysis, and Exhaustive Spatial Environmental Data 97
Comparison of Soil Phosphorus at Four Depths 97
Supervised Land Use Classification from Landsat ETM+ and IKONOS Image 98
Evaluation of the Benefits of Different Spatial Resolutions on Prediction Map of Soil Properties 98
Geostatistical Modeling of Soil Properties 100
Results and Discussion 102
Exploratory Data Analysis 102
Land Use Classification from Landsat and IKONOS Image 103
Evaluation of the Benefits to Use Different Spatial Resolutions For Soil Maps 107
Relationships Between Soil Properties and Environmental Variables 111
Geostatistical Modeling of Soil Properties and Model Validation 112

6 SUMMARY AND SYNTHESIS 145

LIST OF REFERENCES 148

BIOGRAPHICAL SKETCH 162

LIST OF TABLES

Table page

1-1 Total phosphorus (TP) loads from the Santa Fe River (tons/year) into the Gulf of Mexico from 2001 to 2005 (Hornsby, 2007, personal communication) 20

2-1 The number of soil samples used at different depths in the Santa Fe River Watershed (SFRW) and the Santa Fe River Ranch Beef Unit (SFRRBU). 28

2-2 Detailed information on environmental data used to model the global trend of $\log_e$-transformed Mehlich phosphorus (MP) and total phosphorus (TP) in the Santa Fe River Watershed 29

2-3 Detailed information on environmental data used to model the global trend of $\log_e$-transformed Mehlich phosphorus (MP) and total phosphorus (TP) in the Santa Fe River Ranch Beef Unit. 31

3-1 Overview of remote sensors' characteristics (adopted from National Aeronautics and Space Administration) 53

4-1 Descriptive statistics of Mehlich phosphorus (MP) at each depth and total phosphorus (TP) at the top layer in the Santa Fe River Watershed 79

4-2 Pair-wise comparison of $\log_e$-transformed Mehlich phosphorus (MP) in µg/g among the different four depths and total phosphorus (TP) in µg/g at top layer in the Santa Fe River Watershed. 81

4-3 Analysis of variance (ANOVA) tests and homogeneous groups of $\log_e$-transformed Mehlich phosphorus in µg/g at L1 (0-30 cm) and at L4 (120-180 cm) according to either Tukey's test or Dunnett's T3 test at the 0.05 significance level in the Santa Fe River Watershed 82

4-4 Global trend model of $\log_e$-transformed Mehlich phosphorus (LNMP) in µg/g at four different layers and total phosphorus (LNTP) in µg/g at L1 represented by the stepwise multiple linear regression models and regression coefficients in the Santa Fe River Watershed 84

4-5 Spearman correlation of selected variables for cokriging in the Santa Fe River Watershed 85

4-6 Summary of semivariogram parameters for the best spatial prediction method for $\log_e$-transformed Mehlich phosphorus (MP) in µg/g at different four layers and total phosphorus (TP) in µg/g at L1 in different layers across the Santa Fe River Watershed 86

4-7 Error statistics of the best prediction methods for Mehlich phosphorus (MP) in µg/g and total phosphorus (TP) in µg/g across the Santa Fe River Watershed. 87

5-1 Descriptive statistics of Mehlich phosphorus and total phosphorus data at each layer in the Santa Fe River Ranch Beef Unit 121

5-2 Pair-wise comparison of log_e-transformed Mehlich phosphorus (MP) in µg/g among the different four depths and total phosphorus (TP) in µg/g at top layer in the Santa Fe River Ranch Beef Unit 123

5-3 Error matrix and accuracy assessment of the 2003 Landsat ETM+ supervised classification in the Santa Fe River Ranch Beef Unit 124

5-4 Error matrix and accuracy assessment of the 2006 IKONOS supervised classification in the Santa Fe River Ranch Beef Unit 125

5-5 Comparisons of areal extent based on the 2003 Landsat ETM+ and the 2006 IKONOS supervised classification in the Santa Fe River Ranch Beef Unit 126

5-6 Goodness-of-fit statistics of the stepwise multiple linear regression models (SMLR) and error statistics of regression kriging with SMLR using coarse (30 m) and fine resolution (4 m) dataset in the Santa Fe River Ranch Beef Unit 127

5-7 Summary of semivariogram parameters derived from residuals of global trend model with stepwise multiple linear regression for Mehlich phosphorus (MP) in µg/g and total phosphorus (TP) in µg/g in the Santa Fe River Ranch Beef Unit 128

5-8 Summary of semivariogram parameters by the lognormal kriging interpolation method for Mehlich phosphorus at different four depths in the Santa Fe River Ranch Beef Unit 129

5-9 Global trend model of Mehlich phosphorus (MP) in µg/g and total phosphorus (TP) in µg/g represented by the stepwise multiple linear regression models (SMLR) with land use, topographic, and spectral information from coarse and fine datasets. 130

5-10 Analysis of variance (ANOVA) and homogeneous groups of loge-transformed Mehlich phosphorus at 0-30cm in the Santa Fe River Ranch Beef Unit according to the Dunnett's T3 test at the 0.05 significance level 131

5-11 Global trend model of Mehlich phosphorus (MP) in µg/g at four different layers and total phosphorus (TP) in µg/g at L1 represented by the stepwise multiple linear regression models with all coarse data in the Santa Fe River Ranch Beef Unit. 132

5-12 Spearman correlation of selected variables for cokriging in the Santa Fe River Ranch Beef Unit 133

5-13 Error statistics of the best prediction methods for Mehlich phosphorus and total phosphorus across the Santa Fe River Ranch Beef Unit data. 134

LIST OF FIGURES

Figure page

1-1 Phosphorus cycle (adopted from Shober, 2008) 20

1-2 Total water withdrawals (69,273 million liters per day (ML/D)) in Florida by source in 2005 (Marella, 2009) 21

1-3 Fresh ground and surface water usage in Florida by category in 2005 (Marella, 2009) ... 21

2-1 Location of the Santa Fe River Watershed (SFRW) within the State of Florida and 137 sampling sites within the watershed 33

2-2 Land use classes in the Santa Fe River Watershed derived from 2003 Landsat ETM+ image (Data sources: Watershed boundary: Florida Department of Environmental Protection; land use: Florida Fish and Wildlife Conservation Commission) 33

2-3 Soil orders in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; soil order: Soil Data Mart, Natural Resource Conservation Service) 34

2-4 Soil texture classes in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; texture: Soil Data Mart, Natural Resource Conservation Service) 34

2-5 Drainage classes in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; soil drainage: Soil Data Mart, Natural Resource Conservation Service) 35

2-6 Elevations in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; elevation: National Elevation Dataset (30 m), US Geological Survey) 35

2-7 Slope in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; slope derived from the National Elevation Dataset (30m), US Geological Survey) 36

2-8 Surficial geology in the Santa Fe River Watershed (Data sources: Geology; Watershed boundary: Florida Department of Environmental Protection) 36

2-9 Locations of study area and 144 soil sampling sites in the Santa Fe River Ranch Beef Unit (SFRRBU) 37

2-10 Land use classes in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from 2003 Landsat ETM+ image 37

2-11 Land use classification in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from 2006 IKONOS image 38

2-12 Soil orders in the Santa Fe River Ranch Beef Unit (SFRRBU) 38

2-13 Elevations in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from (a) National Elevation Dataset (NED) and (b) Light Detection and Ranging (LIDAR) 39

2-14 Slope in the Santa Fe River Ranch Beef Unit (SFRRBU) from (a) National Elevation Dataset (NED) and (b) Light Detection and Ranging (LIDAR) 39

3-1 Typical semivariogram with nugget, partial sill, sill, and range. 53

4-1 Histograms of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g datasets in the Santa Fe River Watershed 88

4-2 Semivariograms for the best spatial prediction method for $\log_e$-transformed Mehlich phosphorus (MP) in µg/g and a cross-semivariogram for total phosphorus (TP) in µg/g in different layers across the Santa Fe River Watershed.. 91

4-3 Scatter plots of measurements and predictions of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g derived by regression kriging with stepwise multiple linear regression in the Santa Fe River Watershed. 93

4-4 Final predictions in the Santa Fe River Watershed. 94

5-1 Histograms of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g datasets in the Santa Fe River Ranch Beef Unit 135

5-2 Semivariograms for kriging on residuals with coarse and fine data across the Santa Fe River Ranch Beef Unit. 138

5-3 Prediction maps of Mehlich phosphorus (MP) and total phosphorous (TP) by regression kriging with stepwise multiple linear regression (RK/SMLR) with coarse (COARSE) and fine (FINE) resolution dataset in the Santa Fe River Ranch Beef Unit. 140

5-4 Final maps of prediction in the Santa Fe River Ranch Beef Unit 143

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING OF SOIL PROPERTIES IN THE SANTA FE RIVER WATERSHED USING EXHAUSTIVE SPATIAL ENVIRONMENTAL DATA

By

Jinseok Hong

August 2011

Chair: Scot E. Smith
Co-chair: Sabine Grunwald
Major: Civil Engineering

Phosphorus (P) enrichment in surface waters has been documented in the Santa Fe River Watershed (SFRW), north-central Florida over the past decades. Yet the environmental factors that control P enrichment in soil, contributing to water quality impairment, are not well understood.

The main goal of this research was focused to quantify spatial patterns of Mehlich phosphorus (MP) and total phosphorus (TP) in soils across two spatially nested regions (SFRW-3,584 km^2) and the Santa Fe River Ranch Beef Unit (SFRRBU-5.58 km^2) in north-central Florida. The main objectives in both regions were to (a) identify those environmental properties which impart most control on the prediction of soil MP and TP, respectively, and (b) model the spatial patterns of soil MP and TP through statistical and geostatistical methods and assess model quality. Another set of objectives focused on assessment of scaling effects in the SFRRBU, specifically to (c) investigate the effect of environmental co-variates at fine (1-4 m) and coarse (15-30 m) spatial resolutions on predictive modeling of soil MP and TP, and (d) elucidate on the effect of aggregation (resampling) of environmental (spectral, topographic, and other) variables on prediction of soil MP and TP.

A stratified random sampling design was used to collect soil samples from four depths (0-30, 30-60, 60-120, and 120-180 cm) in both regions (137 sites in the SFRW and 150 sites in the SFRRBU). A suite of environmental co-variates was assembled to characterize topography/hydrology, land use/land cover, soils, geology, and climate. Readily available spatial environmental data were complemented with remote sensing images from Landsat ETM+ and IKONOS to derive vegetation properties and spectral indices, and a LIght Detection and Ranging (LIDAR) derived digital elevation model. All data were processed using a Geographic Information System (GIS). Site-specific soil phosphorus (P) data were fused with spatially-explicit environmental co-variates to develop models using univariate (lognormal kriging, LNK) and multi-variate methods (regression kriging, RK, and cokriging, CK).

The results from the SFRW showed that the incorporation of exhaustive environmental data into multivariate, hybrid geospatial models (RK and CK) improved the prediction of soil MP and TP when compared to the univariate method (LNK) that relies solely on soil measurements, but not in the SFRRBU. Among all tested environmental covariates (SFRW), land use, soil drainage, soil order, topographic information, and geologic unit showed the largest predictive power to build inferential models for soil MP. However, contrasting results were found in the SFRRBU indicating that spatial dependence (autocorrelation) of MP and TP in the model of LNK could better represent the distributions of MP and TP than models derived from ancillary environmental variables that link to soil properties in RK and CK. In the SFRRBU results showed that coarse resolution environmental data (30 m spatial resolution) had a smoothing effect within each pixel that better captured soil P global trends compared to fine resolution environmental data (4 m) that did not produce better MP and TP predictions.

Findings from this research contribute to better understand spatially-explicit interactions between soil properties and other environmental variables (e.g., land use, soil drainage, soil order, etc.) facilitating to improve land resource management minimizing adverse risks to the environment, while preserving ecosystem structure and functioning.

CHAPTER 1
INTRODUCTION

Soils, as part of an ecosystem, are crucial to life on Earth as the natural medium for the growth of plants, regulator of water supplies, recycler of raw materials, habitat for soil organisms and engineering medium (Soil Taxonomy, 1999; Brady and Weil, 2002). It has been subjected to environmental factors of climate, and macro- and microorganisms, conditioned by relief, acting on parent material over a period of time (Soil Science Society of America, 1997). These external factors condition soil and contribute to genetic development of soils. Along with air and water, soil is one of the most precious and diminishing resources on the earth because of its dynamic, physical, chemical and biological functions.

Soil contamination typically can be caused by rupture of underground hazardous storage/septic tanks, excessive accumulation of nutrients from fertilizers and farming operations, percolation of contaminated surface water to subsurface strata, untreated storm water runoff, pesticides derived from agriculture management, and wastes leaching from landfills among many others (Ng Kee Kwong et al., 2002). Soil contamination may also lead to underground/surface water contamination as pollutants from soils and surface water runoff leach into an aquifer and/or well field (Parry, 1998). Dissolved phosphorus (P) from runoff and particulate P from erosion are the main factors affecting the P transport to underground/surface waters (Randall et al., 2002). The excessive P in water accelerates eutrophication that causes abnormal production of algae and aquatic vegetation and depletion of dissolved oxygen and may lead to poor water quality. A depiction of the phosphorus cycle can be found in Figure 1-1.

Soil contamination by excessive P is not an issue limited to the State of Florida. There have been numerous research studies in the U.S. which aimed to identify source of excessive P and P load reduction in North Carolina (Cahoon and Ensign, 2004), Oklahoma (Peters and Basta,

1996) and Texas (Somenahally et al., 2009) as well as in Florida (DeBusk et al., 2001; Bruland et al., 2006b; Rivero et al., 2007). At international level P impacts on environmental systems were investigated in India (Solim and Wanganeo, 2008), Canada (Bochove et al., 2006 and 2007), Thailand (Phupaibul et al., 2004), and Argentina (Garcia and Iorio, 2003).

Water is one of the most valued resources in Florida. The State of Florida has more than 1,700 streams and rivers, 7,800 fresh water lakes, 700 springs, and 44,515 km^2 of wetland (Fernald and Purdum, 1998). Surface waters (lakes, springs, and rivers) are attracting residents and tourists for recreational activities and groundwater is the principal source of water for public water supply, agriculture, irrigation purposes, industry, and domestic use in Florida (Raulston et al., 1998). Marella (2009) states that about 69,273 million liters per day (ML/D) (37% as freshwater and 63% as saline surface water) was withdrawn in Florida in 2005 (Figure 1-2). Among 37% of fresh water withdrawn in 2005, 15,899 ML/D of ground water was provided for 90% of Florida's population (16.1 million people) with 9,842 ML/D of surface water provided to about 10% of the state's population (1.8 million people) for its drinking water (Figure 1-2). About 60% of the ground water (9,464 ML/D) withdrawn in 2005 was obtained from the Floridian aquifer system, while Lake Okeechobee and associated canals provided an estimated 4,164 ML/D (43%) of fresh surface water withdrawn in 2005 (Marella, 2009). Out of 15,899 ML/D of fresh ground water withdrawal in 2005, the largest users are for public supply (52%) and agricultural irrigation (31%), followed by commercial-, industrial-mining, and self-supplied (8.5%), domestic self-supplied (private wells) (4%), recreational irrigation (4%), and power generation (0.5%). Agricultural irrigation (56%) used the largest volume of fresh surface water in 2005, followed by power generation (20.5%), public supply (13%), recreational irrigation (6%), and commercial- industrial-mining self-supplied (4.5%) (Marella, 2009).

Ground water within the Suwannee River Water Management District (SRWMD) flows toward the Suwannee and Santa Fe River (SFR) corridors or to the coast, which is the major discharge area for the Floridian aquifer system (Raulston et al., 1998). The District is the least populated in the State and two large industries that have provided an important economic base are forest products and phosphate mining (Raulston et al., 1998).

To assess the quality and quantity of water in the Suwannee River District, the Water Assessment Regional Network (WARN) has been established in 1989 as a priority project of the Surface Water Improvement and Management (SWIM) program in 1987. The WANR was established to monitor river, lake, and ground water levels, river discharge, rainfall, and surface and ground water quality conditions to identify short and long-term trends, that can be used to assess future water quantity and quality conditions (SRWMD, 2002b and 2002c). As of 2004, the network includes a total of 67 monitoring stations (18 stations along the SFR) in the District (SRWMD, 2004). Increasing trends in total phosphorus (TP) concentrations in ground water in the middle Suwannee River Basin (SRB) received attention due to concerns for public and environmental health. Federal, state, regional and local governments, universities and industry associations formed the Suwannee River Basin Nutrient Management Working Group in 1999 that developed into the Suwannee River Partnership (SRWMD, 2002c) to address degradation of water quality in the Suwannee Basin. According to Raulston et al. (1998) and Cox et al. (1994) the most likely sources of elevated nutrients are ground water polluted from a combination of croplands, poultry operations, and septic tanks.

According to Table 1-1, in 2001 and 2002, 42 tons and 68 tons of TP, respectively, were transported into the Gulf of Mexico from the SFR (Hornsby, 2007, personal communication). However, in 2003 and 2005, 241 tons and 426 tons of TP were transported into the Gulf of

Mexico indicating an increasing trend in TP loads, except for 188 tons due to reduced discharge in 2004 (SRWMD, 2002a, 2003a, 2004; Hornsby, 2007, personal communication). Total phosphorus load is a function of rainfall and water flow, with increasing P loads due to increasing discharge. Subsurface water in the SFR area dominantly influences water quality and contains low dissolved minerals and is generally acidic due to low buffering. This allows for dissolution of P from the Hawthorne formation leaching into the underground (Florida Department of Environmental Protection, 2001; Hornsby, 2007).

Once an aquifer is contaminated it may be very difficult, if not impossible, to restore, and costs associated with such efforts can be extremely high so that the most efficient way of protecting ground water is by controlling land uses over sensitive aquifers (Wade, 1991). Therefore, monitoring and modeling of soil and water quality in the District is important for decision making in support of conservation and improvement of soil and water quality.

This research was focused on improving our understanding of spatial patterns of Mehlich phosphorus (MP) and TP across two spatially nested landscapes (Santa Fe River Watershed (SFRW) and the Santa Fe River Ranch Beef Unit (SFRRBU)). The aim was to investigate which of the environmental factors impart most control on MP and TP (i.e., to identify cause-effect relationships). Such MP prediction models have much value to delineate nutrient enriched areas that pose risk to degrade the quality of water resources. The research findings from this study can be used by decision makers, public and private stakeholders, and researchers to guide optimized land resource management in the selected geographic domain.

Hypotheses and Objectives. The main goal in this research is to build soil-landscape models that describe the spatial distribution and variability of MP and TP across a large watershed in north-central Florida (SFRW) and a nested smaller landscape unit (SFRRBU).

Unviariate (ordinary kriging-OK) and multivariate (regression kriging-RK and cokriging-CK) interpolation methods were compared and their accuracy to predict MP in four soil layers and TP in the topsoil evaluated. Special attention was given on incorporation of multi-scale spectral and topographic datasets into soil-landscape models to study spatial scaling behavior.

To achieve these objectives the following hypotheses were tested and are presented in Chapters 4 and 5.

Hypothesis one. The incorporation of exhaustive environmental datasets into multivariate, hybrid geospatial models (RK and CK) improves the prediction of MP and TP when compared to a univariate method (OK) that relies only on MP and TP measurements from soil samples.

Hypothesis two. Among environmental factors, land use ranks highest to predict MP in the topsoil, and geology ranks highest to predict MP in the subsoil.

Hypothesis three. Finer spatial resolution of remote sensing data (IKONOS and LIDAR) improves the accuracy of predicting MP across the landscape compared to those by coarser spatial resolution (Landsat and National Elevation Data (NED)) data.

The specific objectives were:

Objective one. To identify relationships between MP and TP and environmental landscape properties and identify those properties that impart most control on MP across different soil layers within the SFRW and the SFRRBU.

Objective two. To investigate the correlation of MP among different soil layers (0-30, 30-60, 60-120, and 120-180 cm) in the SFRW and the SFRRBU.

Objective three. To characterize the spatial distribution patterns of MP and TP using a variety of geostatistical methods (OK, RK, and CK) across the SFRW and the SFRRBU.

Objective four. To assess the accuracy and bias of each geospatial interpolation methods and identify the method that performs best to predict MP and TP in the SFRW and the SFRRBU.

Objective five. To evaluate the effectiveness of different spatial resolutions of remote sensing data for predictive modeling of MP and TP within the SFRRBU using:

1) Spectral bands, indices, and land use/land cover classification data derived from:

2003 Landsat Enhanced Thematic Mapper Plus (ETM+) (15 m and 30 m; panchromatic and multispectral, respectively) vs. 2006 IKONOS (1 m and 4 m; panchromatic and multispectral, respectively)

2) Topographic properties derived from:

National Elevation Dataset (30 m) vs. LIght Detection and Ranging, LIDAR (4 m)

Table 1-1. Total phosphorus (TP) loads from the Santa Fe River (tons/year) into the Gulf of Mexico from 2001 to 2005 (Hornsby, 2007, personal communication)

Years	2001	2002	2003	2004	2005
TP loads (tons/year)	42	67.6	240.6	187.6	425.8

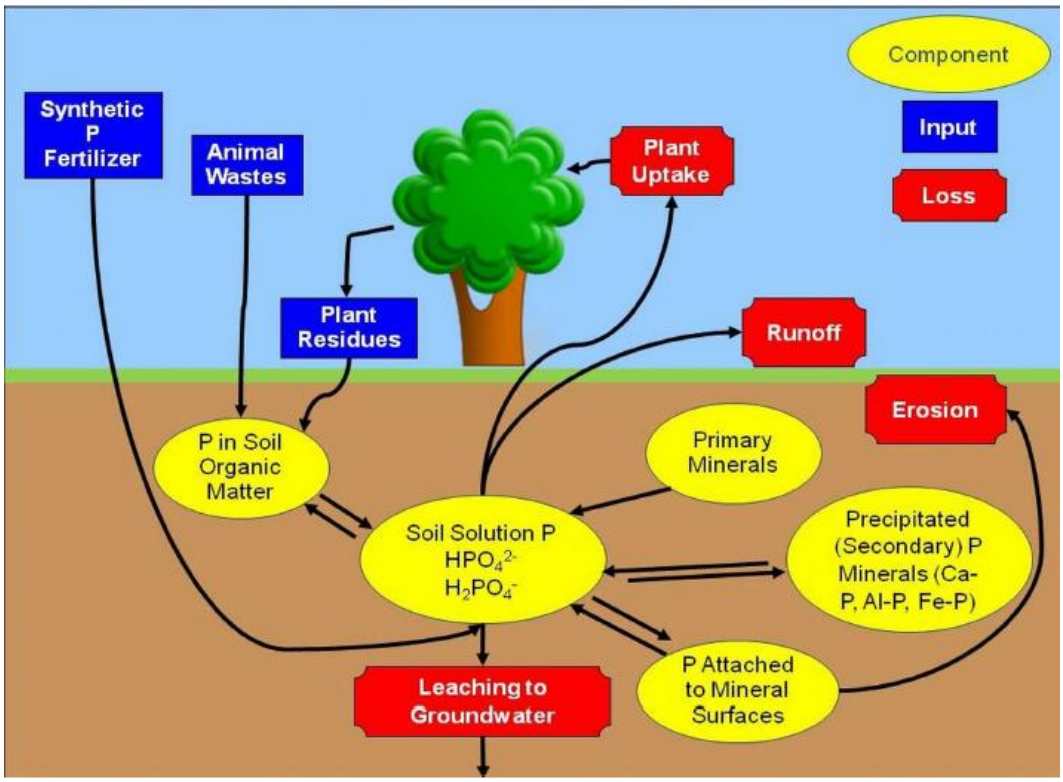


Figure 1-1. Phosphorus cycle (adopted from Shober, 2008)

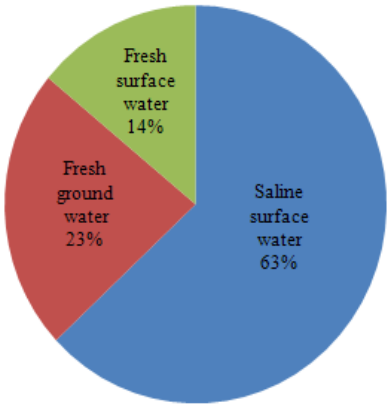


Figure 1-2. Total water withdrawals (69,273 million liters per day (ML/D)) in Florida by source in 2005 (Marella, 2009)

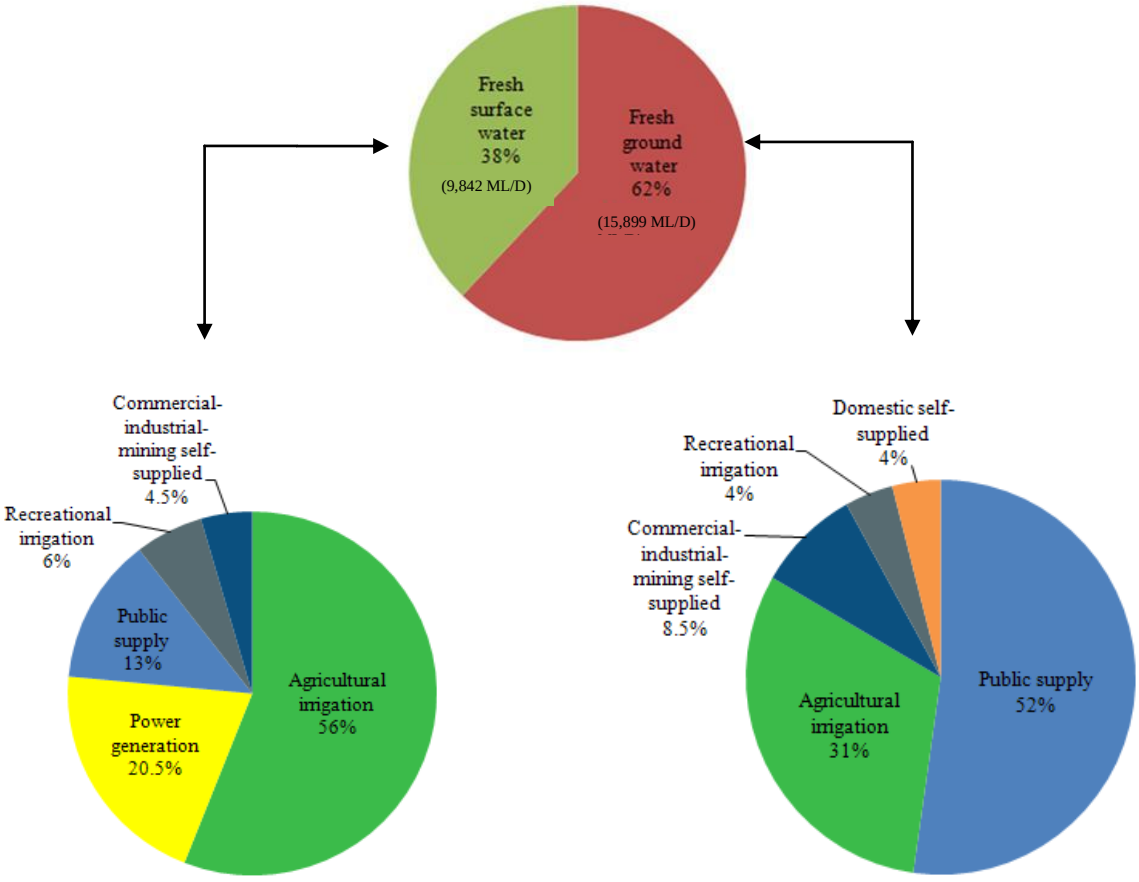


Figure 1-3. Fresh ground and surface water usage in Florida by category in 2005 (Marella, 2009)

CHAPTER 2
DESCRIPTION OF STUDY AREA AND ENVIRONMENTAL DATA

This research was conducted in two different study areas. One is in the SFRW in north-central Florida and the other one is the SFRRBU nested within the larger watershed. Descriptions of each study area and environmental data including soil sampling design, lab analysis methods, and environmental data are described below.

Study Area

Santa Fe River Watershed

The SFRW (Figure 2-1), a tributary of the Suwannee Basin that drains into the Gulf of Mexico, spans over an area of approximately 3,584 km^2 across nine counties including Alachua, Baker, Bradford, Clay, Columbia, Gilchrist, Putnam, Suwannee, and Union in north-central Florida (SRWMD, 2002).

Main land use/land covers within the watershed (Figure 2-2) are pine plantations (30.1%), wetlands (13.8%), improved pasture (12.8%), rangeland (12.0%), upland forest (11.3%), urban and barren (6.8%), agriculture (4.6%), and others (8.6%; water and extractive, etc.) (Florida Fish and Wildlife Conservation Commission (FFWCC), 2004). These Florida vegetation and land cover data were derived from 2003 Landsat ETM+ imagery by the FFWCC in 2004 (Stys et al., 2004). Soil information in the SFRW was derived from the Soil Survey Geographic (SSURGO) database (renamed to Soil Data Mart) (Natural Resource Conservation Service, 2006) and the dominant soil orders are Ultisols (46.5%), Spodosols (26.6%), and Entisols (16.5%). Less prominent soil orders within the watershed are Histosols (4.3%), Alfisols (2.8%) and Inceptisols (1.1%) with 2.2% unmapped area (Figure 2-3).

The soil textures are predominantly sand (88.2%), followed by muck (8.9%), clay (0.8%) and other (2%: loam and unmapped). Figure 2-4 depicts the spatial distribution of soil texture

within the SFRW. A soil drainage class map (Figure 2-5) describes that the western part of the SFRW is mainly composed of excessively or well drained soils, while poorly drained soils are prominent in the eastern part. About 64.6% of soils are classified as poorly drained, followed by well drained (22.9%), excessively drained (10.5%), and other (2%). Elevations derived from the NED (30m) fall in the range of 1.5 m and 91 m above mean sea level that is displayed in Figure 2-6 (US Geological Survey, 1999). The Lowlands and Uplands are separated by the Cody Scarp in the SFRW. The eastern portion of the watershed (Uplands) of the Cody Scarp show relatively high elevations (> 40 m), while lower elevations (< 20 m) are found in Lowlands west of the Cody Scarp (Figure 2-6). Soils in Lowlands are formed on P-rich parent material (Coosawhatchie formation) that is exposed in portions along the Cody Scarp (Grunwald et al., 2006). The slope ranges from level to gently sloping and undulating (0~5%) in most of the watershed area, except for moderate to deep slopes (5~29%) along the Cody Scarp (Figure 2-7). The climate is mostly humid subtropical in the SFRW. The mean annual precipitation is 1,224 mm and the mean annual temperature is 20.5^{o}C based on seven monitoring stations within the watershed (National Climatic Data Center, 2008).

Santa Fe River Ranch Beef Unit

The SFRRBU is a teaching and research facility of the University of Florida, approximately 5.58 km^2 in size and supports 200 heads of Angus-Brangus cows (Dept. of Animal Science, UF). The unit is located in the northern part of Alachua County and adjoins the SFR (Figure 2-9) and is nested within the SFRW.

Improved pasture (39%) is the main land use/land cover within the unit (Figure 2-10) followed by wetland forest (25%), rangeland (13%), upland forest (10%), agriculture (8%), pineland forest (5%) (FFWCC, 2004). The dominant soil orders in the SFRRBU (Figure 2-11) are Ultisols (78%) followed by Entisols (13%), Inceptisols (5%), Spodosols (2%), Alfisols (1%),

and unmapped (1%) (NRCS, 2006). National Elevation Data (NED) (30 m; US Geological Survey, 1999) and LIDAR data (4.5 m; 3001, Inc. New Orleans, LA) were used to derive elevations in the SFRRBU. The lower elevations were found in the northern area adjacent to the SFR, whereas the central area shows undulating elevation (Figure 2-12). Elevations and slopes derived from LIDAR show more detailed (finer-scale) variation than those derived from NED in the SFRRBU (Figure 2-13 and Figure 2-14).

Soil Samples and Laboratory Analysis

To characterize the spatial variability of MP and TP stratified random sampling designs were used to select observation sites in both the SFRW and the SFRRBU. This design and soil samples were collected as outlined in Grunwald et al. (2006). Briefly, soil samples were collected randomly within the identified land use/land cover-soil strata which were homogeneous areas. Random selection within the strata ensures unbiased sampling. Combinations of land use/land cover and soil orders were used to assign the number of samples randomly proportional to their areas.

Santa Fe River Watershed

Based on the stratified random sampling design, with combinations of land use/land cover and soil orders as strata, a total of 137 sample sites were identified within the watershed. Four samples at each site at four different depths (L1: 0-30 cm, L2: 30-60 cm, L3: 60-120 cm, and L4: 120-180 cm) were collected in September 2004. However, not all sample sites had all four depths samples due to field conditions (e.g., high water table during rainy season). Samples were analyzed for available MP in all four depths and TP at L1. The actual number of samples used in this research were 137 (L1), 137 (L2), 135 (L3), and 131(L4). Coordinates of each site within the SFRW were recorded with a Trimble® differential Global Positioning System (GPS) Pathfinder® Pro XR with sub-meter accuracy (Trimble Navigation Limited, Sunnyvale, CA).

For MP analysis (Mehlich-1), 5 g of soil and 20 mL of Mehlich extracting solution were added to a 50 mL centrifuge tube, shaken for five minutes and filtered through a Whatman # 42 filter paper to extract clear solution. The measurement of MP concentrations was done by the method described by Murphy and Riley (1962). The samples (extracted solutions with Reagent B1 and double deionized (DDI) water) were prepared and left to stand for at least 30 minutes for stable color development. Absorbance was read at 880 nm on a UV1 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Absorbance of standard curves at 0, 50, 100, 200, 400, 600, 800, and 1200 ppb (with R^2 were 0.995 or greater) were used to predict the P concentration (µg/g of soil) of samples by relating sample absorbance to the standard curve.

For TP analysis, 0.2 g of ball-milled soils from each site and 2 mL of sulfuric acid were put into glass tubes and digested in the block at 340^{o}C for 30 minutes. The process of cooling-down for 15 minutes and adding 0.5 mL of 30% diluted hydrogen peroxide were repeated until the solutions were clear or white. Filtered solutions were mixed with DDI water and Reagent B1 and set for at least 30 minutes for bluish color development. Once samples were prepared for absorbance reading, the remaining procedures were the same as the ones in the MP analysis.

Santa Fe River Ranch Beef Unit

A stratified random sampling design was also used in the SFRRBU. Eight land uses, five soil orders and two elevation ranges were the stratifications. Samples from 150 sites were collected at four depths (L1: 0-30 cm, L2: 30-60 cm, L3: 60-120 cm, and L4: 120-180 cm) in March, 2005. Like in the SFRW, not all sites had four samples at four depths due to field conditions and 144 (L1), 144 (L2), 131 (L3), and 114 (L4) were used to be analyzed for MP and TP. Extraction procedures of MP and TP were the same as the methods in the SFRW.

Exhaustive Spatial Environmental Data

Various data were used to model spatial trends and identify relationships between soil properties and environmental data in the SFRW and the SFRRBU. The data derived as Geographic Information Systems (GIS) layers were attained from various agencies (e.g., U.S. Geological Survey, Natural Resources Conservation Service, and Florida Fish and Wildlife Conservation Commission).

Santa Fe River Watershed

The numerous data used in SFRW included soil related information (e.g., soil taxonomic order, soil drainage class, and soil hydrological group), climate related information (e.g., precipitation and temperature), land use/land cover information, satellite related information (e.g., reflectance bands, vegetation indices, and principal components of the reflectance bands from Landsat ETM+), topography related information (e.g., elevation, slope, aspect, and compound topographic index from NED), and geological related information (e.g., ecological regions, environmental geology, and surficial geological formation unit). A complete list of the data used in the SFRW can be found in Table 2-2.

Santa Fe River Ranch Beef Unit

From the data used in the SFRW some (e.g., geological formation unit, precipitation and temperature) were not used in the SFRRBU study because the entire SFRRBU lies in Coosawhatchie formation and is small enough to have same climate related information. To prepare data for the SFRRBU study most data used in the SFRW were clipped by using the SFRRBU boundary with SFRW GIS data layers in ArcGIS® 9.3. Land use/land cover and satellite related information (e.g., reflectance bands) derived from IKONOS and topography related information (e.g., elevation) derived from LIDAR were added to assess the effectiveness

of different resolutions. A comprehensive list of the data used in the SFRRBU is presented in Table 2-3.

Table 2-1. The number of soil samples used at different depths in the Santa Fe River Watershed (SFRW) and the Santa Fe River Ranch Beef Unit (SFRRBU).

	Layers for MP				Layer for TP
	L1 (0-30 cm)	L2 (30-60 cm)	L3 (60-120 cm)	L4 (120-180 cm)	L1 (0-30 cm)
SFRW	137	137	135	131	137
SFRRBU	144	144	131	114	144

Abbreviations: MP: Mehlich phosphorus; TP: Total phosphorus

Table 2-2. Detailed information on environmental data used to model the global trend of log_e-transformed Mehlich phosphorus (MP) and total phosphorus (TP) in the Santa Fe River Watershed.

Environmental Property	Format	Data type	Source	Date	Original scale/spatial resolution (m)
Soil taxonomic order	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil drainage class	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil hydric rating	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil hydrologic group	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil AWC (cm/cm)	Vector	Continuous	USDA/NRCS	1995	1:24,000
DTW (cm)	Vector	Continuous	USDA/NRCS	1995	1:24,000
KSAT (µm/s)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Clay content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Silt content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Sand content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Mean annual Temperature (^oF)	Raster	Continuous	NCDC/NOAA	1993-2005	N/A
Mean annual precipitation (in)	Raster	Continuous	NCDC/NOAA	1993-2005	N/A
Ecological regions	Vector	Categorical	FDEP	1995	1:250,000
Physiographic divisions	Vector	Categorical	FDEP	2000	1:2,000,000
Land use	Raster	Categorical	FFWCC	2003	N/A / 30
Total population per census tract	Vector	Continuous	USCB	2000	1:100,000
Environmental geology	Vector	Categorical	FDEP	2001	1:250,000
Surficial geology	Vector	Categorical	FDEP	1998	1:100,000
DRASTIC index	Vector	Continuous	FDEP	1998	1:100,000
Hydrogeology	Vector	Categorical	FDEP	1998	1:100,000
ETM+ reflectance (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ NDVI (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ TNDVI (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ IR/R (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+IR-R (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ TC indices (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ PC (DN)	Raster	Continuous	USGS	2004	N/A / 30
Elevation (m)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
Slope (%)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
Aspect	Raster	Categorical	USGS/NED	1999	1:24,000 / 30
Catchment area (ha)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
CTI	Raster	Continuous	USGS/NED	1999	1:24,000 / 30

Abbreviations: AWC = available water capacity; CTI = compound topographic index; DN = digital number; DRASTIC = aquifer vulnerability index (depth to water, net recharge, aquifer media, soil media, topography, impact of the vadose zone, hydraulic conductivity of the aquifer); DTW = depth to water table; ETM+ = Landsat Enhanced Thematic Mapper Plus; FDEP = Florida Department of Environmental Protection; FFWCC = Florida Fish and Wildlife Conservation Commission; IR/R = infrared – red ration; IR-R = infrared-red difference; KSAT = saturated hydraulic conductivity; N/A = not applicable; NCDC = National Climatic Data Center; NDVI = normalized

difference vegetation index; NED = National Elevation Dataset; NOAA = National Oceanographic and Aeronautic Administration; NRCS = National Resources Conservation Service; PC = 6 principal components derived from ETM+; TC = Tasseled Cap; TNDVI = transformed NDVI; USCB = United State Census Bureau; USDA = United States Department of Agriculture; USGS = United States Geological Survey.

Table 2-3. Detailed information on environmental data used to model the global trend of log_e-transformed Mehlich phosphorus (MP) and total phosphorus (TP) in the Santa Fe River Ranch Beef Unit.

Environmental Property	Format	Data type	Source	Date	Original scale/spatial resolution (m)
Soil taxonomic order	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil drainage class	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil hydric rating	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil hydrologic group	Vector	Categorical	USDA/NRCS	1995	1:24,000
Soil AWC (cm/cm)	Vector	Continuous	USDA/NRCS	1995	1:24,000
DTW (cm)	Vector	Continuous	USDA/NRCS	1995	1:24,000
KSAT (µm/s)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Clay content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Silt content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Sand content (%)	Vector	Continuous	USDA/NRCS	1995	1:24,000
Ecological regions	Vector	Categorical	FDEP	1995	1:250,000
Environmental geology	Vector	Categorical	FDEP	2001	1:250,000
DRASTIC index	Vector	Continuous	FDEP	1998	1:100,000
Hydrogeology	Vector	Categorical	FDEP	1998	1:100,000
ETM+ land use	Raster	Categorical	FFWCC	2003	N/A / 30
ETM+ reflectance (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ NDVI (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ TNDVI (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ IR/R (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ IR-R (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ TC indices (DN)	Raster	Continuous	USGS	2004	N/A / 30
ETM+ PC (DN)	Raster	Continuous	USGS	2004	N/A / 30
IKO land use	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO reflectance (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO NDVI (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO TNDVI (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO IR/R (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO IR-R (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
IKO TC indices (DN)	Raster	Continuous	Geoeye, Inc.	2007	N/A / 4
NED elevation (m)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
NED slope (%)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
NED aspect	Raster	Categorical	USGS/NED	1999	1:24,000 / 30
NED catchment area (ha)	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
NED CTI	Raster	Continuous	USGS/NED	1999	1:24,000 / 30
LIDAR slope (%)	Raster	Continuous	3001 Inc.	2002	N/A / 4
LIDAR aspect	Raster	Continuous	3001 Inc.	2002	N/A / 4
LIDAR catchment area (ha)	Raster	Continuous	3001 Inc.	2002	N/A / 4
LIDAR CTI	Raster	Continuous	3001 Inc.	2002	N/A / 4

Abbreviations: AWC = available water capacity; CTI = compound topographic index; DN = digital number; DTW = depth to water table; DRASTIC = Aquifer vulnerability index (depth to water, net recharge, aquifer media, soil

media, topography, impact of the vadose zone, hydraulic conductivity of the aquifer); ETM+ = Landsat Enhanced Thematic Mapper Plus; FDEP = Florida Department of Environmental Protection; FFWCC = Florida Fish and Wildlife Conservation Commission; GIS Lab/SWS/UF = GIS laboratory at the Department of soil and water science, University of Florida; IKO = IKONOS; IR/R = infrared – red ratio; IR-R = infrared-red difference; KSAT = saturated hydraulic conductivity; LIDAR = Light detection and ranging; N/A = not applicable; NDVI = normalized difference vegetation index; NED = National Elevation Dataset; NRCS = National Resources Conservation Service; SDM = Soil Data Mart; TNDVI = transformed NDVI; USDA = United States Department of Agriculture; USGS = United States Geological Survey.

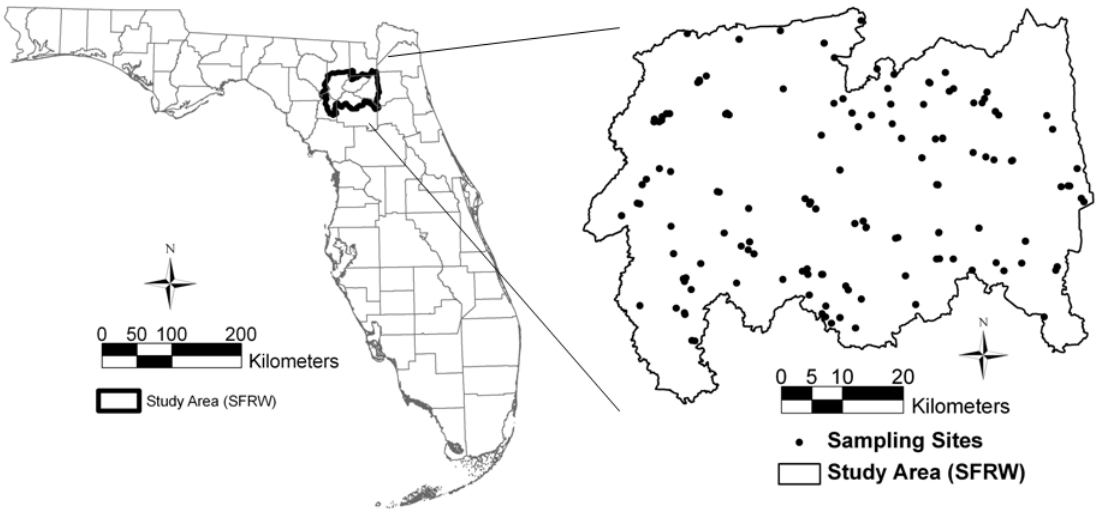


Figure 2-1. Location of the Santa Fe River Watershed (SFRW) within the State of Florida and 137 sampling sites within the watershed (Data sources: Florida State boundary: US Census Bureau; watershed boundary: Florida Department of Environmental Protection)

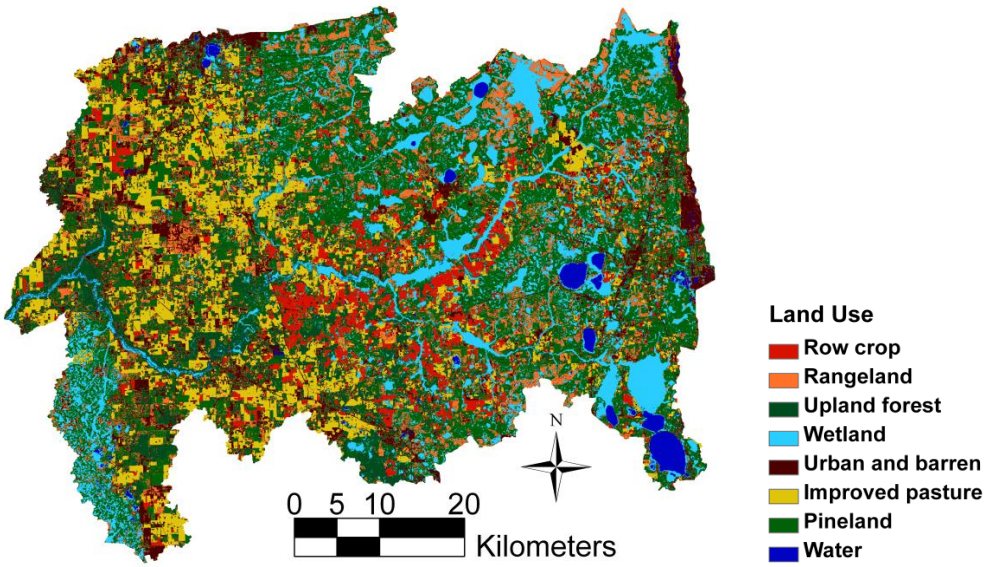


Figure 2-2. Land use classes in the Santa Fe River Watershed derived from 2003 Landsat ETM+ image (Data sources: Watershed boundary: Florida Department of Environmental Protection; land use: Florida Fish and Wildlife Conservation Commission)

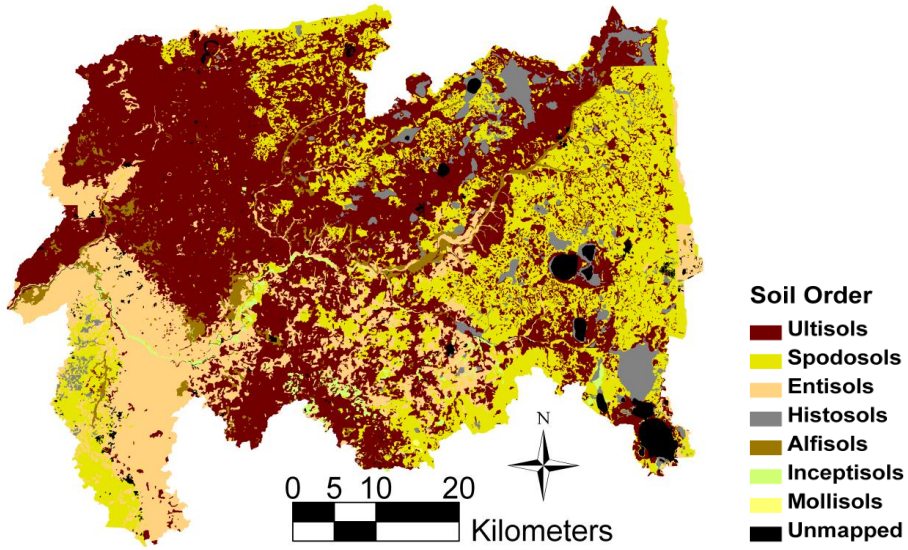


Figure 2-3. Soil orders in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; soil order: Soil Data Mart, Natural Resource Conservation Service)

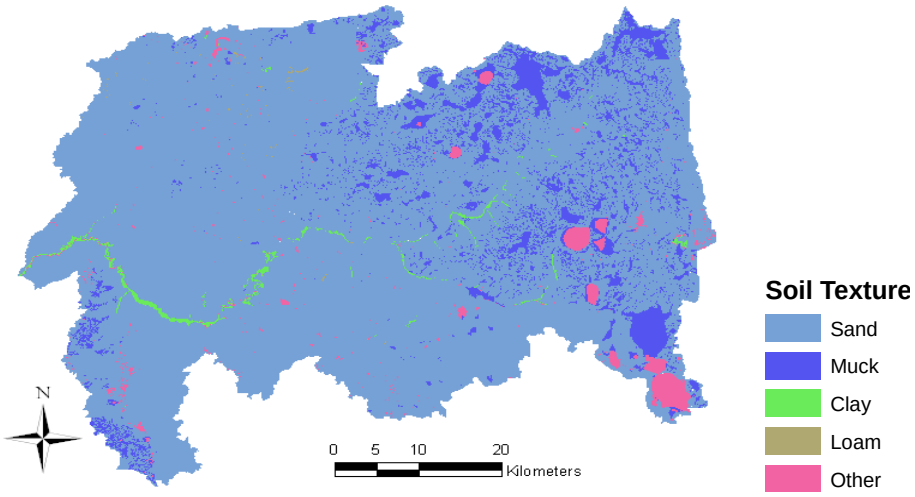


Figure 2-4. Soil texture classes in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; texture: Soil Data Mart, Natural Resource Conservation Service)

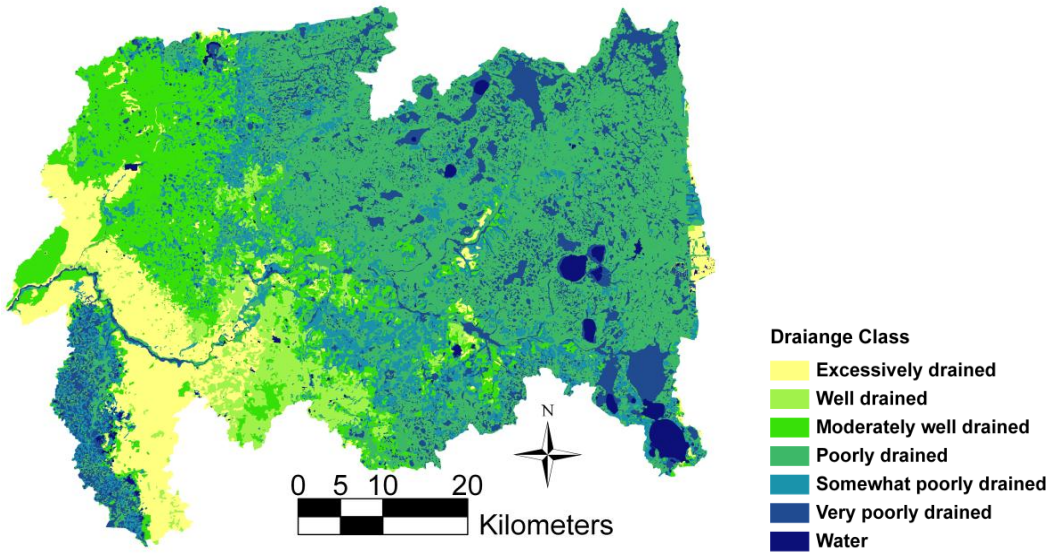


Figure 2-5. Drainage classes in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; soil drainage: Soil Data Mart, Natural Resource Conservation Service)

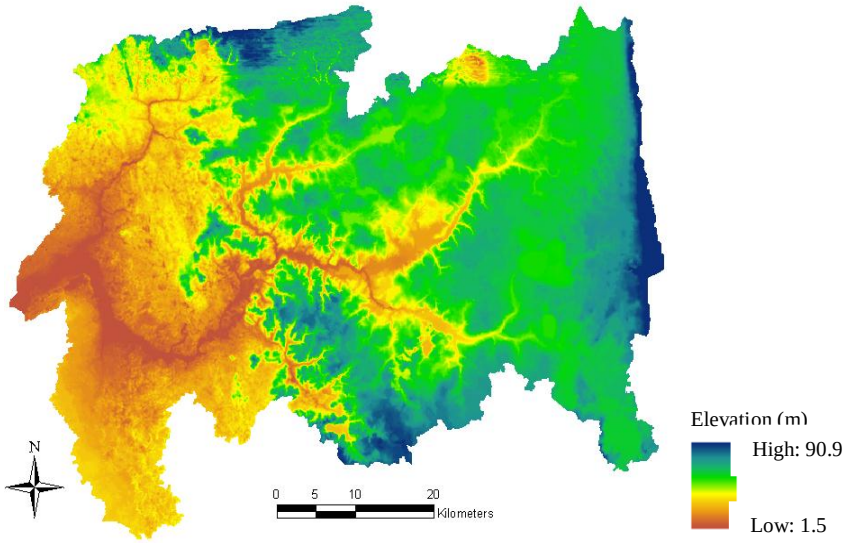


Figure 2-6. Elevations in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; elevation: National Elevation Dataset (30 m), US Geological Survey)

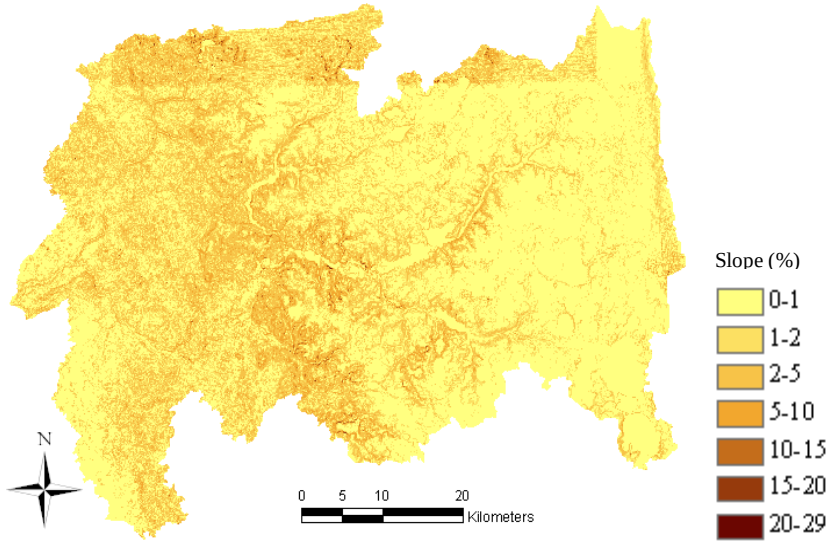


Figure 2-7. Slope in the Santa Fe River Watershed (Data sources: Watershed boundary: Florida Department of Environmental Protection; slope derived from the National Elevation Dataset (30m), US Geological Survey)

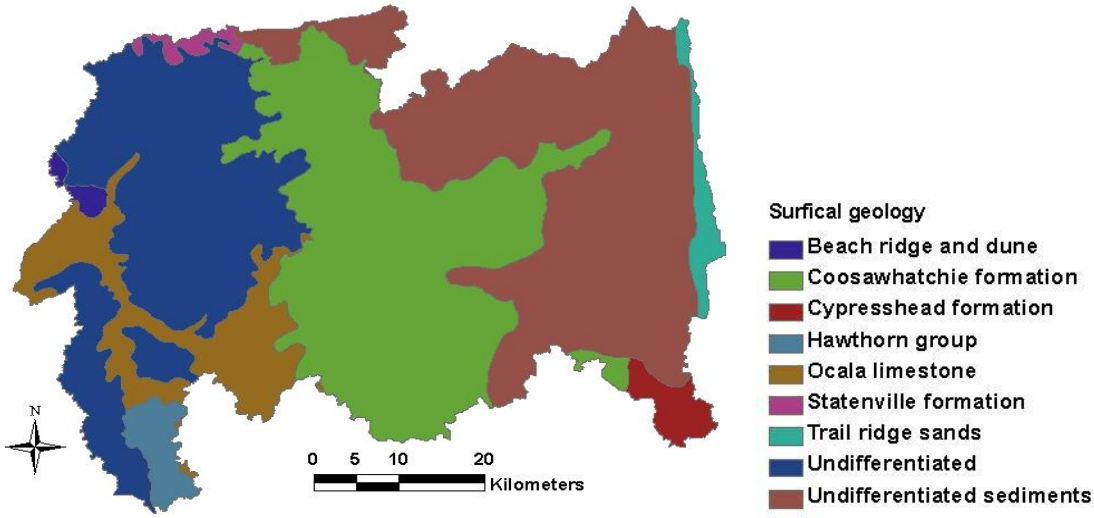


Figure 2-8. Surficial geology in the Santa Fe River Watershed (Data sources: Geology; Watershed boundary: Florida Department of Environmental Protection)

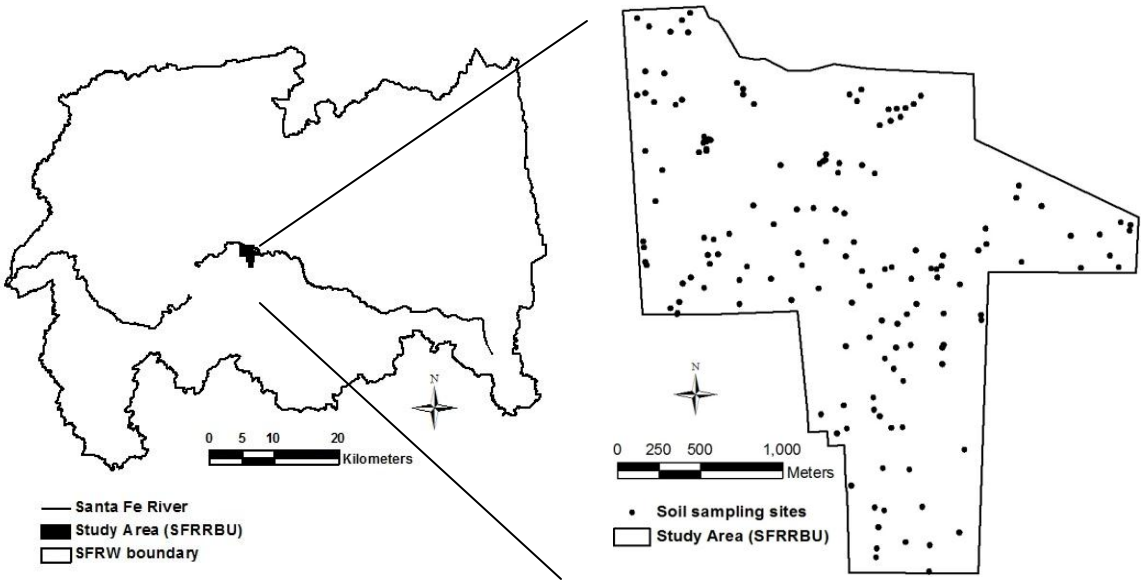


Figure 2-9. Locations of study area and 144 soil sampling sites in the Santa Fe River Ranch Beef Unit (SFRRBU) (Data sources: Florida state boundary: US Census Bureau, watershed boundary: Florida Department of Environmental Protection, SFRRBU boundary: Geographic Information System Laboratory, Soil and Water Science Department, University of Florida)

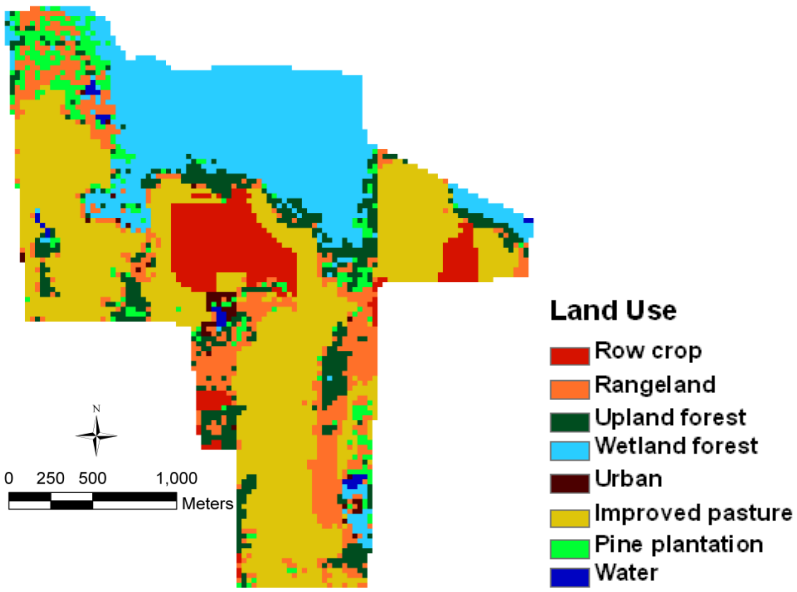


Figure 2-10. Land use classes in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from 2003 Landsat ETM+ image (Data sources: SFRRBU boundary: Soil and Water Science Department; University of Florida; land use: Florida Fish and Wildlife Conservation Commission)

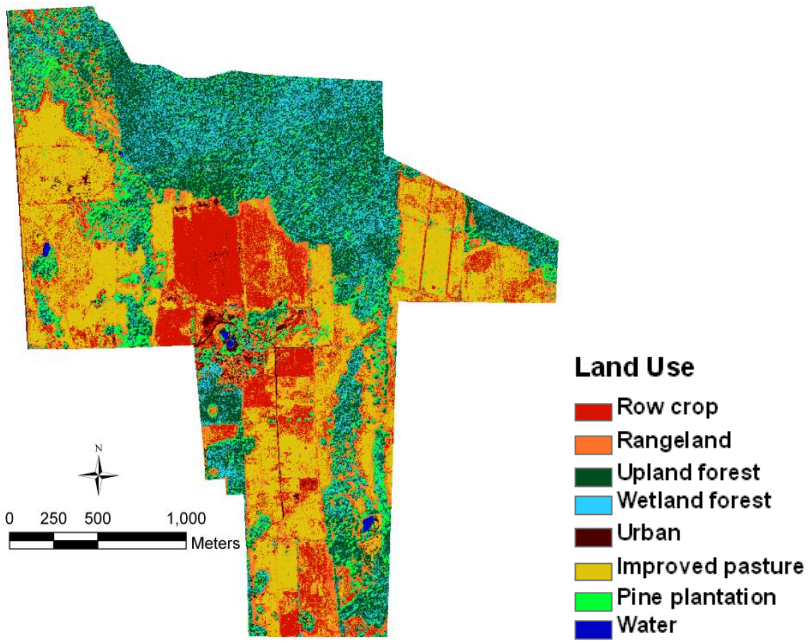


Figure 2-11. Land use classification in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from 2006 IKONOS image (Data sources: SFRRBU boundary; land use: Jinseok Hong, Geographic Information System Laboratory, Soil and Water Science Department, University of Florida)

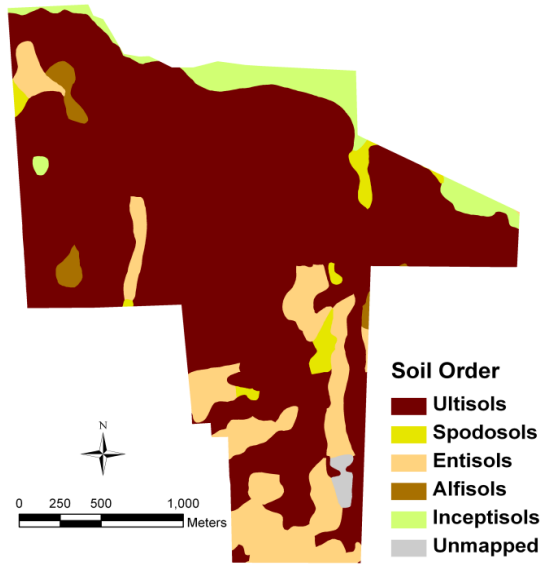
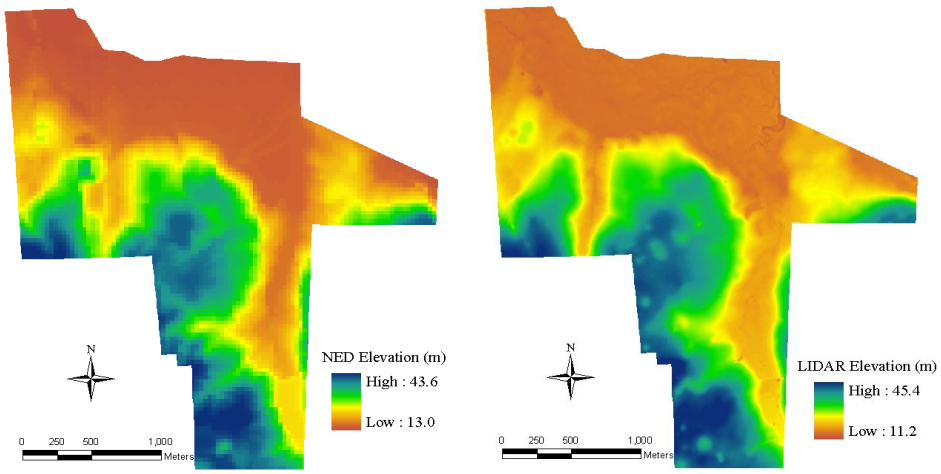
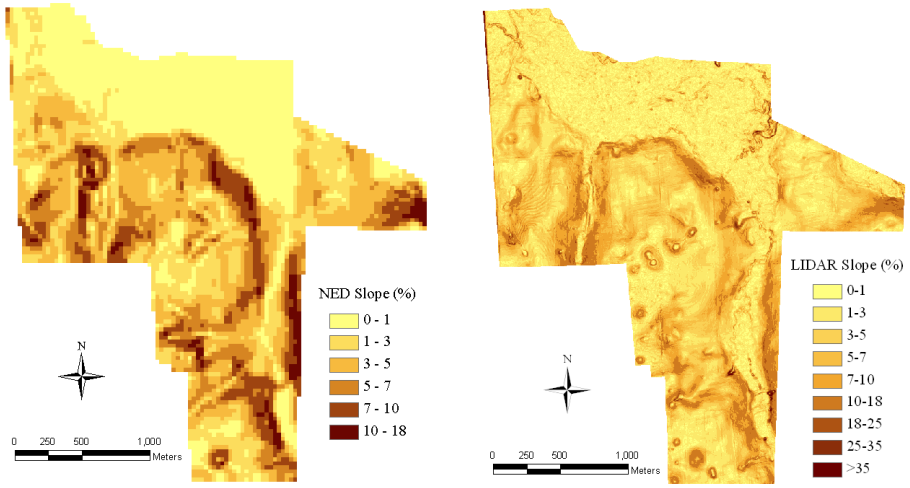


Figure 2-12. Soil orders in the Santa Fe River Ranch Beef Unit (SFRRBU) (Data sources: SFRRBU boundary: Geographic Information System Laboratory, Soil and Water Science Department, University of Florida; soil orders: Soil Data Mart, Natural Resource Conservation Service)



(a) (b)

Figure 2-13. Elevations in the Santa Fe River Ranch Beef Unit (SFRRBU) derived from (a) National Elevation Dataset (NED) and (b) Light Detection and Ranging (LIDAR) (Data sources: SFRRBU boundary: Geographic Information System Laboratory, Soil and Water Science Department, University of Florida; NED: US Geological Survey; LIDAR: 3001 Inc, New Orleans, LA)



(a) (b)

Figure 2-14. Slope in the Santa Fe River Ranch Beef Unit (SFRRBU) from (a) National Elevation Dataset (NED) and (b) Light Detection and Ranging (LIDAR) (Data sources: SFRRBU boundary: Geographic Information System Laboratory, Soil and Water Science Department, University of Florida; NED: US Geological Survey; LIDAR: 3001 Inc, New Orleans, LA)

CHAPTER 3
INTRODUCTION TO SOIL-LANDSCAPE MODELING

Phosphorus (P) is essential to all forms of life on earth and erosion, storm water runoff, and intensive agricultural operations are the major sources of P. High concentrations of P in soils and groundwater in Florida are mainly due to agricultural practices and Miocene P deposits of Florida (Randazzo and Jones, 1997)

Despite the high variability in soil properties in space and time farmers have treated land uniformly applying fertilizers without considering the spatial variation of soils (López-Granados et al., 2005). This has caused problems that under-treated zones do not reach optimum levels of production, while over-treated zones pose potential environmental pollution and increase in cost (Bouma, 1997). According to Brady and Weil (2002), environmental quality can be affected severely and widely by P, e.g. land degradation caused by too little available P and eutrophication caused by too much P. For environmental protection, cost reduction and optimization of crop yield, farmers need to know the spatial distribution of field deficiencies and abundances (López-Granados et al., 2005). However, only about 5% of the global agricultural research budget is allocated to soil research (Yaalon, 2000). Soil sampling is labor-, time-intensive and costly, and thus, it is challenging to collect dense soil datasets across large regions (e.g. watershed). Recently geospatial soil-landscape modeling has been suggested using exhaustive spatial environmental datasets along with field soil data (McBratney et al., 2003). In the following sections an overview of various soil-landscape modeling techniques is presented.

Soil Mapping History and Geostatistical Modeling

Soil-landscape modeling attempts to integrate soils, parent material, topography, land use and land cover, and human activities to gain an understanding of the spatial distribution of soil attributes, characteristics of soils, and their behavior through time (Grunwald, 2005). The

models can be grouped into three categories: Crisp, continuous, and conceptual approaches (McBratney et al., 2003).

Traditional soil-landscape models are qualitative/descriptive models that use crisp map units which are discrete, sharply delineated by polygons based on soil surveyor's prior knowledge, intuition and understanding of soils (Grunwald, 2006). The crisp mapping approach assumes homogeneity of soil properties within a map unit and efficiently delivers general information about soil distributions. Soil classification systems, which adopt the crisp soil maps concept, still prosper around the world (e.g., State Soil Geographic (STATSGO) Database and Soil Survey Geographic (SSURGO) Database) (Grunwald, 2006). However, the crisp map does not account for ambiguities and spatial autocorrelation within and across map units, and they are too rigid to take into account genuine uncertainty (McBratney and De Gruijter, 1992; Heuvelink and Webster, 2001). However, Burrough and McDonnell (1998) suggested that the continuous field model instead of the crisp mapping model is adequate for modeling natural phenomena that do not show obvious boundaries (e.g., soil). This spatial model has the ability to describe the gradual change of soil attributes formed by a variety of pedological processes within a domain. The demand for quantitative soil-landscape models focusing on soil property prediction rather than soil classes is increasing (Grunwald, 2009; Grunwald et al., 2011) and the shift from qualitative to quantitative soil-landscape modeling provides digital and more accurate, precise and nonbiased information about soils (Grunwald, 2006).

Dokuchaev, a Russian soil scientist, presented a conceptual soil-landscape-modeling approach based on soil forming factors in 1892. Jenny (1941) further developed Dokuchaev's model and popularized the soil factor model, which describes soil as a function of five soil forming factors and can be mathematically expressed as following:

$$S = f\,(cl, o, r, p, t) \quad (3\text{-}1)$$

where

S = soil properties (e.g. MP or TP)

cl = climate (e.g. temperature and precipitation)

o = organisms (e.g. vegetation, land use and human/biological activities)

r = relief (e.g. elevation, slope and aspect)

p = parent material (or geology)

t = time

McBratney et al. (2003) expanded Jenny's model by including spatial position and soil itself and named it SCORPAN model because soil can be predicted from its properties, or soil properties from its class or other properties and can be written as:

$$S_a \text{ or } S_c[x,y,\sim t] = f\,(S[x,y,\sim t], C[x,y,\sim t], O[x,y,\sim t], R[x,y,\sim t], P[x,y,\sim t], A[x,y,\sim t], N) \quad (3\text{-}2)$$

where

S_a = soil properties (e.g. MP or TP)

S_c = soil classes (e.g. soil order)

S = soil, other properties of the soil at a point (or site) (e.g. physical, chemical or biological soil properties derived from soil databases)

C = climate, climatic properties of the environment at a point (or site)

O = organisms, vegetation or human activity

R = topography

P = parent material

A = the time factor (age)

N = space (e.g. spatial position)

x, y = x and y coordinates

~t = approximate or vague time coordinate (or at about some time t)

Each factor is represented by a set of one or more continuous or categorical variables (e.g. C by average annual rainfall, temperature or climate class).

Jenny's model (CLORPT) and McBratney's model (SCORPAN) are examples of conceptual factorial models that provide the framework for many quantitative soil-landscape models.

In order to describe soil-landscape patterns geostatistical methods have been used widely (Corstanje et al., 2006; Bruland et al., 2006a; Rivero et al., 2007b; Grunwald and Reddy, 2008; Mora-Vallejo et al., 2008; Mishra et al., 2010; Schaefer et al., 2010).

Geostatistical methods encompass a variety of techniques, which are based on the underlying spatial correlations to guide interpolation (Grunwald, 2006). These spatial correlations in geostatistical models assume that observations obtained close to each other are more likely to be similar than observations collected further apart from each other (i.e., spatial autocorrelation) and can be expressed by semivariograms as following (Goovaerts, 1997; Webster and Oliver, 2007):

$$\gamma(h) = \frac{1}{2N}\sum_{i=1}^{N}[z(x_i) - z(x_i + h)]^2 \qquad (3\text{-}3)$$

where

γ: semivariance

h: distance (lag) in meter

N: total number of data pairs

z (x_i): an observation value at location i

The semivariogram is a measurement of spatial properties dependency (degree of dissimilarity of values) and provides basic information for the spatial analysis. As shown in Figure 3-1, the range defines the distance beyond which there is no spatial correlation between data pairs and the sill indicates the value at which the semivariograms model attains the range, and the nugget defines spatial sources of variation at distances smaller than the sampling interval (short-range variation) and/or measurement (random) errors that cannot be captured or explained by the model (Goovaerts, 1997; Webster and Oliver, 2007).

Geostatistical modeling, which is based on the theory of regionalized variables introduced by Matheron in 1965, incorporates spatial correlation between neighboring observations to predict attribute values at unsampled locations from available soil attribute observations that are spatially distributed throughout the soil mapping area (Goovaerts, 2000). Regionalized variable theory assumes that spatial variation of any variable Z can be described by the following equation.

$$Z(x_i) = m(x_i) + \varepsilon'(x_i) + \varepsilon'' \tag{3-4}$$

where

$Z(x_i)$: predicted soil property at certain location, x_i

$m(x_i)$: deterministic function describing the structural component of Z at x_i,

("deterministic trend")

$\varepsilon'(x_i)$: stochastic, locally varying but spatially dependent residual from $m(x_i)$,

that can be characterized by using the semivariogram i.e. the regionalized variable (autocorrelated component)

ε'': residual; a spatially independent noise term having zero mean and variance

Equation 3-4 also assumes that the mean is constant for all x_i, which is called stationarity assumption, meaning that the semivariance depends only on the separation distance *h* and not on the positions x_i and $x_i + h$ (Webster and Oliver, 2005).

Applications of soil-landscapes modeling

Many conceptual, qualitative and quantitative soil-landscape models have been applied to study the distribution, behavior, and genesis of soils, because understanding the spatial and temporal distribution of soils and soil characteristics is required for optimizing economic profits while minimizing adverse impacts on soil (Grunwald, 2006).

The increasing power of tools such as GIS, GPS, remote sensors, and data sources such as those provided by digital elevation models (DEM), facilitate development of more advanced soil-landscape modeling. Grunwald (2009) outlined the increasing use of soil data and remote sensors in digital soil mapping and modeling.

However, many countries worldwide such as Australia, Germany, Brazil, France, etc. lack soil spatial data infrastructures, mainly due to slow and expensive conventional soil survey methods, and complexity and dynamics of soil variation across geographic areas (Heuvelink, 2001; McBratney et al., 2003). Many researchers have shown various approaches (different sampling methods, interpolation methods and environmental datasets) to model soil properties such as nitrogen (Lamsal et al., 2006; Borůvka et al., 2007; Oorts et al., 2008), P (De Gryze et al, 2008; Rivero et al., 2009), carbon (Grimm et al., 2008; Zhang et al., 2008; Vasques et al., 2009), and soil organic matter (OM) (Regalado and Ritter, 2006; Liu et al., 2008; Ouimet, 2008; Wu et al., 2009) from site-specific spatial scale to large watershed or basin spatial scale (Hudson, 1992; Hewitt, 1993; Bui, 2003; Bruland et al., 2006a).

Recently, soil properties at landscape scales have been predicted by mixing deterministic and stochastic components, linking data collected at different spatial scales, and integrating sparse point datasets (e.g. soil samples) with exhaustively measured environmental datasets (e.g. DEM, satellite imagery, etc.) called 'hybrid' models. McBratney et al. (2000) compared 24 different univariate and multivariate statistical, geostatistical and hybrid geospatial soil prediction models at different spatial scales (field, sub-catchment, and regional scales) and found that hybrid geospatial techniques outperformed all other methods at all scales.

Statistical methods such as multiple regressions, kriging, artificial neural networks (ANNs) or classification and regression trees (CARTs) are also commonly used to predict soil attributes or soil classes (McKenzie and Ryan, 1999; Holmberg et al., 2006; Lamsal et al., 2006; Hengl et. al., 2007; Rivero et al., 2007b; Wu et al., 2009). Holmbert et al. (2006) practiced ANN successfully to model daily total organic carbon, total nitrogen and TP concentrations in stream water and predicted future fluxes with climate change based on simulated concentration. Lamsal et al. (2006) used a mixed approach (CART and RK) that combines exhaustive spatial environmental data and sparse, season-specific field observations to model soil nitrate-nitrogen (NO_3-N) at watershed scale and learned that their methodology provide a compromise between labor, costs, and accuracy for prediction of NO_3-N across the watershed. Rivero et al. (2007) applied geostatistical techniques (OK, RK, and CK) to predict spatial distribution and variability of floc and TP throughout Water Conservation Area (WCA) - 2A located in the northern portion of the Everglades. The study concluded that multivariate methods (RK and CK), which incorporated spatial covariation between TP and other environmental variables, are more accurate in predicting TP than a univariate method (OK).

Remote Sensing in Soil Mapping

The application of remote sensing for soil mapping dates back to the 1930s when panchromatic aerial photographs became available for the first time for civil purposes (Lillesand and Kiefer, 2008). However, the new era of soil mapping began with the launch of the Landsat series of satellite in 1970s (Abuzar and Ryan, 2001). Since then, remote sensing has been appreciated for cost-effective and accurate data acquisition to improve our understanding of the extent and distribution of Earth's resources. Various remote sensing techniques, such as multispectral, hyperspectral, Radio Detection and Ranging (RADAR) and LIDAR systems, on airborne and spaceborne platforms are growing commercially and offering important sources of timely, high-quality spatial information that can serve research and operational needs (Birk et al., 2003). Even though there are many different remote sensors available for different usages two main sensors (Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and IKONOS) used in this research are briefly discussed.

The first Landsat satellite was launched in July 1972 and Landsat ETM+ covers an area of 185 × 185 km, and has six multispectral bands with 30 m spatial resolution, one thermal infrared (TIR) band with 60 m spatial resolution and one panchromatic band with 15 m spatial resolution (Lillesand and Kiefer, 2008).

IKONOS, the Greek word for "image", is the world's first commercial earth observation satellite that collects high-resolution imagery at 1 m (panchromatic) and 4 m resolution (multispectral) launched in September 24, 1999. IKONOS has the ability to collect 20,000 km^2 in a single pass and orbits the Earth every 98 minutes at an altitude of about 680 km. It was launched into a sun-synchronous orbit and its revisit rate is 3 days (Geoeye, Inc., VA). Table 3-1 briefly summarizes Landsat ETM+ and IKONOS's characteristics such as band, wavelength and spatial resolution.

Many remote sensing applications focus on land use/land cover change detection, climate change, disease control mapping, natural disaster evaluation, map making and revision, water resources/urban growth planning and monitoring, coastal mapping, aviation safety and crop yield assessment (Goodchild, 1994; Barnes and Baker, 2000; Dobos et al. 2000; Song et al., 2001; Di et al., 2003; Masuoka, et al., 2003; López-Granados et al, 2005; Sullivan et al., 2005). Such applications provide base information that can be readily incorporated into soil surveys and soil-landscape models (e.g. CLORPT or SCORPAN models).

Dobos et al. (2000) demonstrated that combining a DEM and satellite radiometric data was useful for soil characterization in a large area using Advanced Very High Resolution Radiometer (AVHRR) with normalized difference vegetation index (NDVI) and terrain databases such as DEM, slope, aspect and curvature. The research showed that the purely AVHRR-based model resulted in a soil classification accuracy of 49.1%, while soil classification performance for integrated AVHRR-terrain database showed 87.3% accuracy (500 m resolution) and 70.1% accuracy (1-km resolution). López-Granados et al. (2005) found that combining primary data (soil samples) and secondary data (aerial color photograph) with a geostatistical technique (kriging with varying local means) accurately produced maps of soil properties (OM, pH, and potassium). Barnes and Baker (2000) used multispectral airborne (green, red, near infrared and thermal) and satellite data (Satellite Pour l'Observation de la Terre (SPOT) and Landsat Thematic Mapper) to derive surface soil texture maps in Maricopa, Arizona and achieved 81%, 88% and 92% of accuracy assessment results, respectively. Sullivan et al. (2005) evaluated the IKONOS multispectral data with kriging methods as a soil mapping tool to delineate soil property variability in Coastal Plain and Tennessee Valley in Alabama.

Numerous studies have been conducted to examine similarities and/or differences of datasets derived from different remote sensing techniques (Tian et al., 2002; Mumby et al., 2002; Sawaya et al., 2003; López-Granados et al., 2005; Netzban, 2005; Stefanov and Netzban, 2005; and Gallo et al., 2005). Bailloeul et al. (2003) used high resolution satellite images with existing GIS data as prior knowledge for updating of urban digital maps. The International Development Research Center (IDRC) funded a project to test and demonstrate remote sensing data and techniques that can be used to replace expensive traditional soil mapping methods for soil map inventory of arid and semi-arid environments in Morocco (IDRC, 1999). Leone et al. (1995) facilitated the preparation of a methodology integrating spectral signatures derived from Landsat TM and digital image processing to produce soil cartographic representations at a reconnaissance scale. Soil physical properties such as OM have been correlated to specific spectral responses (Shonk et al., 1991). Therefore, multispectral images have shown potential for the automated classification of soil mapping units (Leone et al., 1995).

A key to scaling remote sensing is to understand the magnitude of the effects resulting from processes acting at different scales (Tian et al., 2002). Rogan and Chen (2004) provided an overview of a variety of remote sensors to map land cover and land use that differed in their spatial resolutions ranging from 0.8 to 8,000 meters (QuickBird to GOES-8, 10). They suggested that resolutions (spatial, spectral, and radiometric), processing techniques and estimated cost per unit area should be considered to select appropriate remote sensing datasets for a given geographical environment. Marceau and Hay (1999) stated that the change in spatial resolution changes the patterns of reality and higher spatial resolution data lowered the overall classification and prediction accuracies.

However, since parameters and processes important at one scale are frequently not important or predictive at another scale it is essential to be familiar with advantages and/or disadvantages of remote sensors (e.g. optimal spatial, spectral, temporal and radiometric resolutions or costs) used for specific applications. For example, Wang et al. (2004) recognized that the IKONOS had higher variance and entropy values in all the compared bands, whereas the QuickBird image displayed a finer texture roughness in the forest canopy based on band by band examination. By examining only multispectral bands, the IKONOS image showed better spectral discrimination than QuickBird, while inclusion of panchromatic bands had no effect on the classification accuracy of either the IKONOS or QuickBird images. Rivero et al. (2007a and 2009) compared the ability of spectral signatures and indices from two remote sensors Advanced Space-borne Thermal Emission Reflectance Radiometer (ASTER) and Landsat ETM+ to capture the spatial variability of ecosystem components of WCA-2A and relate them to TP in floc (i.e., detritus layer overlaying the topsoil) and surface soil. The study concluded that integrating GIS and remote sensing technologies with multivariate geostatistical methods for describing spatial soil-vegetative interrelationships has the potential to improve prediction of soil properties at the landscape scale.

Limitations and Research Gaps

Even though numerous studies have been conducted on remote sensing applications there is still no single combination of sensor and interpretation procedure that is appropriate to all resource inventories and environmental monitoring applications. Furthermore, successful application of remote sensing is premised on the integration of multiple, interrelated data sources and analysis procedures (Lillesand and Kiefer, 2008).

Dobos et al. (2000) indicated that their results and conclusions were limited to their case study only due to specific climate, vegetation, soils and land uses and should be carefully

validated in other areas of the world. Kustas et al. (2004) concluded that remote sensing for soil mapping is limited because several other variables can potentially impact soil reflectance such as tillage practices and moisture content.

Remote sensing data must be at high enough pixel resolution to discriminate individual ground conditions. Weiss et al. (2001) noted that validation of coarse spatial resolution products is not straightforward due to ground-based measurements that cannot be easily compared to coarse resolution satellite sensor data. In addition, coarse-scale validation should rely on methods that avoid time-consuming procedures but preserve accuracy of mapped properties. This motivated data producers and users to utilize high resolution remote sensing data more and more and the commercial sector has developed a new generation of satellite-based sensors with finer spatial resolutions (Stefanov and Netzban, 2005). But it is not clear if and how much finer resolution remote sensing data can improve soil prediction models.

More research is required to gain knowledge on the impact of different spatial/spectral resolutions of environmental datasets on predicting soil properties. Jensen (2000) pointed out that data from commercial systems are costly compared to government-operated sensors and limited in both spatial and temporal coverage. More research is needed to assess the usefulness of improved spatial, spectral, radiometric and geometric resolution satellite data and advanced algorithms to process datasets (Giri et al., 2005). Sawaya et al. (2003) argued that high resolution satellite data is best suited for relatively small geographic areas (i.e., single images) or potentially north-south oriented areas that could fall along a single path due to the high cost of high-resolution imagery for large area assessments. Kustas et al. (2004) states that remotely sensed surface temperature is the key boundary condition by most remote sensing based energy balance models for estimating evapotranspiration and coarse spatial resolution (1 km or greater

pixel resolution) has limited capability to monitor the impact of land cover changes and evaluate evapotranspiration from different crop covers (soybean and corn crops) in central Iowa. However, higher resolutions (250 m) by thermal sharpening techniques of Moderate-Resolution Imaging Spectroradiometer (MODIS) for estimating surface temperature could provide enough spatial detail for discriminating evapotranspiration between soybean and corn fields. They also state that Landsat 7 and ASTER instruments on Terra provide remotely sensed surface temperature at pixel resolutions 60 m and 90 m, respectively but routine application is hindered by the low frequency of repeated coverage (~ 16 days) and cloud cover, which severely limits the utility of these sensors in providing routine monitoring of evapotranspiration.

Atmospheric correction and adjustments for differences between sensors in solar and satellite viewing geometry and spectral response functions of the sensors might be expected to improve accuracy of merging different satellite resolutions (Sawaya et al., 2003; Gallo et al., 2005).

Table 3-1. Overview of remote sensors' characteristics (adopted from National Aeronautics and Space Administration)

Landsat ETM+				IKONOS			
Band	Wavelength (μm)	Spectral location	Resolution (m)	Band	Wavelength (μm)	Spectral location	Resolution (m)
1	0.45-0.52	Blue	30	1	0.45-0.52	Blue	4
2	0.52-0.60	Green	30	2	0.52-0.60	Green	4
3	0.63-0.69	Red	30	3	0.63-0.69	Red	4
4	0.75-0.90	Near IR	30	4	0.76-0.90	Near IR	4
5	1.55-1.75	Mid IR	30	PAN	0.45-0.90	PAN	1
6	10.4-12.5	Thermal IR	60				
7	2.09-2.35	Mid IR	30				
8	0.52-0.90	PAN	15				

Abbreviations: ETM+ = Enhanced Thematic Mapper Plus, IR = Infrared, PAN = Panchromatic

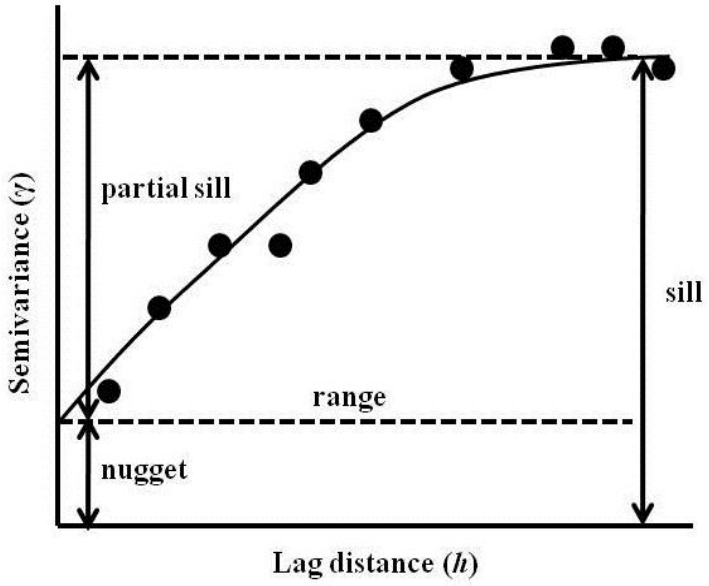


Figure 3-1. Typical semivariogram with nugget, partial sill, sill, and range.

CHAPTER 4
DIGITAL SOIL MAPPING OF SOIL PHOSPHORUS IN THE SANTA FE RIVER WATERSHED, FLORIDA

Excessive P by fertilizers, animal manure, runoff and erosion from surface soils are responsible for much of downstream eutrophication and underground water pollution (Bochove et al., 2006).

For example, one of the largest algal blooms which covered up to 311 square kilometers in the western quarter of the Lake Okeechobee, Florida for one month occurred in 1986 and P was the factor controlling algal growth in the lake (Rechcigl, 1997). Gray et al. (2005) stated that P level increased from 40 parts per billion (ppb) in 1970 to more than 130 ppb by the 1990s and continuing P loading from field crops, improved pasture, and dairy operations have resulted in 777 square kilometers of organic muck on the lake's bottom, which contains about 51,000 tons of P.

This chapter is focused on the regional assessment of P distributions at four different soil depths across a large subtropical watershed. The overall goal was to improve our understanding of regional distributions of MP and TP in the SFRW. The specific objectives were to (a) investigate the correlation of MP and TP among different soil layers (0-30, 30-60, 60-120, and 120-180 cm). (b) identify the strongest predictors out of a suite of environmental data representing SCORPAN factors. (c) compare univariate (OK) and multivariate (RK and CK) methods to characterize the spatial patterns of MP and TP in the topsoil and MP in the subsoil across the SFRW. (d) validate each geospatial interpolation method using an independent validation set and identify the best method that predicts MP and TP in the SFRW.

The following hypotheses were investigated in this research.

Hypothesis one. The incorporation of exhaustive environmental datasets into multivariate, hybrid geospatial models (RK and CK) improves the prediction of MP and TP when compared to a univariate method (OK) that relies only on soil P measurements.

Hypothesis two. Among environmental factors, land use and vegetation specific properties derived from spectral data rank highest to predict MP and TP in the topsoil and geology ranks highest to predict MP in the subsoil.

The soil MP and TP models which show the best prediction performance provide a comprehensive assessment of P enrichment within the watershed, which poses a risk to degrade water quality in the aquifer. Total P allows inference on the overall system state, whereas MP represents the more dynamic, available P which responds more readily to anthropogenic and natural forcing.

Materials and Methods

Soil Sampling, Laboratory Analysis, and Exhaustive Spatial Environmental Data

Detailed information about soil sampling and the extraction procedures of MP and TP can be found in Chapter 2. A complete list of the environmental data used in the study is summarized in Table 2-2, Chapter 2.

Exploratory Analysis of Soil Data: Comparison of Soil MP at Four Depths

Pearson's product moment correlations were calculated on natural logarithm-transformed (LN) MP values (L1-LNMP~L4-LNMP) and TP (L1-LNTP) to analyze correlations among different depths of MP and MP and TP at L1. Mann-Whitney test was conducted to investigate how the median concentrations of MP vary based on the different soil depth. The equality of variances in MP at different depths was examined by the Levene's test (Levene, 1960) to check the assumption of homogeneous variance.

Quantify Relationships between Soil Phosphorus and Environmental Variables

To test the hypothesis that land use and vegetation specific properties influence most of the MP concentration in the topsoil and geology influences most of the MP concentrations in the subsoil, stepwise linear multiple regression (SLMR) was conducted between MP and various environmental variables across the watershed. Some of environmental variables that were run in SLMR were information from satellite images (e.g. land use/land cover, NDVI, and reflectance), topographic indices, soil information (e.g. soil order and soil drainage class), geologic unit, depth to water table, and more. A complete list of the data used in the SFRW is presented in Table 2-2. Coefficients of each selected variable in the SLMR model may indicate the sign and magnitude of the effect of one variable upon another. To check the significance of targeted predictors (i.e., land use, soil drainage, soil order, and geologic unit), one-way analysis of variance (ANOVA) was used in case of equal variance groups based on Levene's test ($p > 0.05$) or either one-way Welch's ANOVA (if number of samples in each group was greater than 10) or Brown-Forsythe's ANOVA (if number of samples in each group was less than 10) (Myers and Well, 2003; Vasques et al., 2010). Further investigations to classify homogenous groups in land use, drainage and geology were conducted using either Tukey's test for equal variance groups or Dunnett's T3 test for unequal variance groups (Dunnett, 1980). These tests are based on the assumption of statistical independence of the observations.

Geostatistical Modeling of Soil Phosphorus

Univariate (OK) and multivariate methods (RK and CK) were employed to model spatial distributions of MP and TP across the SFRW. Log-Normal Kriging (LNK) is suitable for datasets with positively skewed distribution to approximate a normal distribution and stabilize variances (Webster and Oliver, 2007). To process LNK, the positively skewed distributed MP and TP data were transformed to logarithm base-e and OK was used to predict $\log_e$-transformed

concentrations of MP (LNMP) at L1, L2, L3, and L4, and TP (LNTP) at L1. The predicted LNMP and LNTP values ($\log_e$ µg/g) were then back-transformed to original units (µg/g) based on the formula suggested by Webster and Oliver (2007).

Ordinary Kriging is also known as Best Linear Unbiased Estimator (BLUE) and the most common practical type of Kriging because it aims to characterize the linear relationship between prediction and input data (linear), intends a mean error of zero (unbiased), and minimizes the variance of prediction residuals (best). Log-Normal Kriging used only MP and TP measurements with their spatial autocorrelations for the prediction of MP and TP, while RK and CK used not only MP and TP measurements but also various auxiliary environmental variables to map spatial distributions of MP and TP across the watershed. These environmental variables were assembled either at sampling sites or focal analysis based on the attributes in the neighborhood (e.g. 3 x 3 or 7 x 7 moving window) across the watershed by ArcGIS® 9.3.

Regression kriging constitutes three components; global trend, spatially dependent residuals (stochastic component), and spatially independent residuals (noise or error component). It assumes that the deterministic component can be explained by a regression model rooted in the SCORPAN conceptual model by relating predictor variables (i.e., SCORPAN factors) to a target variable (e.g., MP or TP). In this research, stepwise linear multiple regression (SLMR) was used as a regression model to identify the relationships between MP and TP and various environmental variables and to map the global trend of spatial MP and TP distributions across the watershed. In case of categorical variables (e.g., land use/land cover) among environmental variables they were converted into dummy variables (e.g., 1 as agriculture, 0 as non-agriculture) and used in SLMR. The coefficient of determination (R^2) was used to assess the quality of regression models. The residuals were computed by subtracting predicted values from observed

values and semivariograms and spatial autocorrelation structures were derived for each residual of soil properties and layer. Then OK was used to interpolate residuals. The predicted residuals were then added back to the predictions derived from global trend modeling to produce the final prediction map of soil properties (Odeh et al., 1995).

Cokriging can be used when a sparsely measured target (primary) variable is spatially cross-correlated with a denser ancillary variable (i.e., secondary variable) (Goovaerts, 1997, 2000; Wackernagel, 2003; Rivero et al., 2009). Collocated CK was used in this research since MP and TP data was sparsely measured at some locations and other environmental variables, including remote sensing data and DEM derived data, were available everywhere within the study area. One environmental variable which showed the highest correlation with MP at all depths and TP at L1 was selected and semivariograms and cross-semivariograms between soil properties (e.g., MP and TP) and selected variables were modeled and applied to predict concentrations of MP and TP across the watershed.

A number of software packages were used in this research. ArcGIS® 9.3 (Environmental System Research Institute, Redlands, CA) was used to convert site-specific (point) data to raster data and display the final continuous spatial distribution of MP and TP with a grid resolution of 30 meters. All spatial predictor (input) data were projected to Albers Equal Area Conic map projection.

SPSS® 17.0 (SPSS Inc., Chicago, IL) was used for statistical analysis including exploratory data analysis, correlation and significant tests, various group means and SMLR. All geostatistial analysis was conducted in ISATIS® 8.02 (Geovariances, Avon, France) and ERDAS IMAGINE® 9.1 (Leica Geosystems Geospatial Imaging, LLC, Norcross, GA) was used for focal analysis of raster data (e.g. reflectance of Landsat ETM+ bands, compound topographic index

(CTI), DEM, etc.) which is based on the attributes in the neighborhood (e.g. 3 x 3 or 7 x 7 moving window) across the watershed.

Model Validation

The total dataset at each depth was randomly divided into two groups. One was the calibration set (about 70%) for model development and the other was the validation set (about 30%) for independent validation of each predicted model (LNK, RK, and CK). Unlike cross-validation, this validation set is an independent dataset to evaluate the accuracy of predictions. Error statistics (differences between the observed and the predicted values) were calculated to compare each model to assess the accuracy of prediction by various kriging methods. The mean error (ME), root mean square error (RMSE), and residual prediction deviation (RPD) were used to evaluate the performance of each predicted model. The ME indicates whether the model is, on average, overestimating or underestimating the observed values, while RMSE indicates if predictions are close to the observed values (e.g., ME should be close to 0 if it is unbiased and RMSE should be as small as possible for accurate prediction (Elbir, 2003; Verfaillie, 2006; Mishra et al., 2010). The residual prediction deviation is the ratio between the standard deviation (SD) of the dataset and the standard error of prediction (William, 1987). The MP and TP prediction model at each depth was evaluated based on the ME, RMSE, and RPD and the equations are shown below in Equation 4-1, 4-2, and 4-3, respectively.

$$\text{ME} = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n} \tag{4-1}$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n}} \tag{4-2}$$

$$\text{RPD} = \frac{SD}{RMSE\sqrt{\frac{n}{(n-1)}}} \tag{4-3}$$

where y_i = measured values; $\hat{y}_i$ = predicted values; n = number of measured values with i = 1, 2,…., n.; SD = standard deviation.

Results and Discussion

Exploratory Data Analysis

Descriptive statistics of MP and TP datasets such as mean, median, maximum, minimum, standard deviation, skewness and kurtosis are summarized in Table 4-1. Overall, MP showed a positively skewed distribution at all depths due to some low and some high observed values and logarithm transformation (base-e) was used to approximate a normal distribution and stabilize variances (Figure 4-1). After the transformation was applied, the means and medians of MP at all depth and TP at L1 were much closer to each other and the skewness of the total MP raw dataset (2.43, 2.63, 3.07, 2.58 at L1, L2, L3, and L4, respectively) were reduced to -0.46, -0.67, -0.31, and -0.23, at L1, L2, L3, and L4, respectively. In the case of TP at L1, the skewness of the total dataset was reduced from 3.03 to 0.74.

Overall, measurements of MP were variable among four depths with a range from 0.028 µg/g to 189.12 µg/g, with a mean of 23.91 µg/g, and a median of 6.10 µg/g at L1, from 0.005 µg/g (below detection limit (BDL)) to 145.32 µg/g, with a mean of 15.44 µg/g, and a median of 3.28 µg/g at L2. The values of MP at L3 varied from 0.005 µg/g (BDL) to 271.78 µg/g, with a mean of 20.44 µg/g, and a median of 1.70 µg/g, while those of MP at L4 ranged 0.005 µg/g (BDL) to 258.97 µg/g, with a mean of 26.0 µg/g, and a median of 1.92 µg/g. Lastly, TP showed a range from 24.89 µg/g to 2,958.21 µg/g, with a mean of 333.86 µg/g and a median of 142.59 µg/g at L1.

Based on the means of LNMP at different depths, the highest mean of LNMP was found at L1 (L1-LNMP; 1.68 µg/g) followed by L2 (L2-LNMP; 0.88 µg/g), L4 (L43-LNMP; -0.16 µg/g) and L3 (L3-LNMP; -0.09 µg/g). Overall, it was shown that the means of LNMP decreased

at L2 and L3 from the top layer (L1) and slightly increased (from -0.09 µg/g to -0.16 µg/g) at the deepest layer (L4). At all four depths of MP and TP at L1, the means of the dataset were much greater than the medians of the dataset. This indicates that sample dataset was influenced by a few very high samples which caused the inflation of the mean.

Table 4-2 shows that LNMP at all four depths and LNMP and LNTP at L1 are significantly correlated at the 0.05 significance level. The highest correlation existed between L3 and L4 (0.81) and the lowest correlation could be found between L1 and L4 (0.37). The overall pattern of correlations at four depths showed that adjacent layers (i.e., L1 – L2, L2 – L3, and L3 – L4) had relatively higher correlations (0.80, 0.72, and 0.81, respectively), while the layers further from each other (i.e., L1 – L3, L1 – L4, and L2 – L4) had relatively lower correlations (0.44, 0.37, and 0.61, respectively). This indicates that the spatial distribution of MP at the upper layer might influence the adjacent layer beneath it through vertical transport processes. The median differences among MP layers from Mann-Whitney test also confirmed the trend that the median differences between the adjacent layers showed lower differences compared to median differences between further layers from each other's (i.e., between L1 and L2 (0.62), L1 and L3 (1.28), L1 and L4 (1.16), and L2 to L3 (0.66), and L2 and L4 (0.54)).

Vasques et al. (2010) also found similar results showing that total soil carbon (TC) was significantly correlated between depths at the 95% confidence level, with highest correlation of 0.76 (between L2 and L3), followed by 0.70 (between L1 and L2), 0.69 (between L3 and L4) in the SFRW.

Relationships between Soil MP and Environmental Variables

The hypothesis that the land use influences the spatial distribution of MP at the topsoil layer and geological unit influences MP distribution in the subsoil were evaluated based on significances of land use and geologic unit data. Significances of soil order and soil drainage on

LNMP were also evaluated because P leaching and runoff related to soil type and soil drainage can cause potential surface and underground water contamination.

According to Levene's test, LNMP at L1 in land use (p-value = 0.582), in soil order (p-value = 0.395) and LNTP at L1 in soil drainage (p-value = 0.347) had homogenous variances, thus, one-way ANOVA and Tukey's Honestly Significant Differences (HSD) test were used for these variables to separate homogeneous groups based on group means. On the other hand, LNMP at L1 in soil drainage (p-value = 0.008) and LNMP at L4 in geologic units (p-value = 0.017), LNTP at L1 in land use (p-value = 0.000) and LNTP at L1 in soil order (p-value = 0.000) had non-homogeneous variances, thus, either Welch's ANOVA or Brown-Forsythe's ANOVA was used depending on the number of samples in each variable's category and Dunnett's T3 test was used to classify homogeneous groups. Table 4-3 summarizes tests used with significance and group mean of homogeneous groups at the 0.05 significance level.

For the significance of land use on LNMP at L1 according to Tukey's HSD test, the highest mean LNMP was observed in agriculture (3.47 $\log_e$ µg/g) and decreased in the following order: Improved pasture (3.46 $\log_e$ µg/g), upland forest (2.39 $\log_e$ µg/g), urban (2.18 $\log_e$ µg/g), rangeland (1.99 $\log_e$ µg/g), pineland (0.87 $\log_e$ µg/g), and wetland (-0.65 $\log_e$ µg/g). For land use on LNTP improved pasture had the highest mean value (5.84 $\log_e$ µg/g) followed by agriculture (5.75 $\log_e$ µg/g), rangeland (5.49 $\log_e$ µg/g), wetland (5.33 $\log_e$ µg/g), upland forest (5.10 $\log_e$ µg/g), urban (5.08 $\log_e$ µg/g) and pineland (4.41 $\log_e$ µg/g).

Land use was likely to be a significant environmental variable influencing variability of soil's P fertility and LNMP in agriculture and improved pasture was statistically higher than LNMP in pineland and wetland, while LNTP in improved pasture, agriculture and rangeland was statistically higher than LNTP in upland forest, urban and pineland. These findings are

consistent with the literature. White and Hammond (2006) reported that from the 31.3 kiloton (kt)/year of P loads to the River Basin Districts (RBD) of the United Kingdom, 5.8 kt/year (18.6%) was apportioned to agriculture (improved grassland, field horticulture, livestock, etc.) and about 2.7 kt/year (9%) was apportioned to background (orchards, woodlands, and forests). They also found that about 11.8 kt/year (28.3%) and 2.7 kt/year (6.5%) of TP input (41.6 kt/year) contributions to the RBD in United Kingdom were from agriculture and background, respectively. Mander et al. (1998) found that crop-cultivated land and improved pasture had significantly higher MP than other lands. This enhanced P fertility can result from addition of nutrients and OM through application of mineral or organic fertilizer and deposition of plant debris. Organic matter has been reported to enhance water holding capacity and prevent nutrient leaching, improving structure and nutrient status of the soil (Berkelaar, 2001; Mafongoya et al., 2004; Fageria et al., 2005). Plants on land can physically protect the soil from weathering effect. There was negative correlation between the soil P fertility and the degree of soil weathering which lowered P in more weathered soil (Sharpley et al., 1987). In addition, plant roots can increase the available P pool by generating organic anions and hydrogen protons and result-in changes in soil pH (Hinsinger, 2001; Hopkins and Huner, 2004; Weintraub et al., 2007).

He et al. (2006) examined the concentrations and loads of various P forms in runoff water from 11 agriculture (four vegetable farms and seven citrus groves) sites and relations to soil P status and fertilizer P input in South Florida. They found that concentrations of P in runoff water were correlated with total and labile P in the soils as well as fertilizer P rate and the median of TP concentrations (1.0 mg/L) was much greater than the ecoregional critical levels (0.008 – 0.038 mg/L) set by the United States Environmental Protection Agency (USEPA, 2005).

LNMP values from upland forest, urban and rangeland were statistically higher than wetland but not significantly different among themselves and statistically lower LNMP was found in pineland and wetland and there were no significant differences between the two in the SFRW.

Similar outcomes were found by Kim et al. (2006) who conducted research to investigate P accumulation in surface soil based on land use conditions in Busan, Korea. They collected 128 soil samples from various land uses such as forest (70% - pine trees and deciduous trees), paddy fields (15%), orchards (8% - plum and peach), upland fields (6% - cultivating sesame, green perilla, pepper, carrot, and Chines cabbage) and residential area (1%) in their study area and found higher P contents in agricultural soils (530.5 of upland fields, 359.3 of orchards, and 210.2 mg/kg of paddy fields) compared to lower P content in forest soils (147.2 mg/kg). The elevated P values in upland fields, orchards and paddy fields were attributed to P inputs from chemical and organic fertilizers.

Davis (1994) compared amounts of total P and organic P in the topsoil (0 – 10 cm) collected from planted pine plantation (*Pinus radiata*; 10-year-old) with those of the topsoil from adjacent unplanted *Chionochloa rigida* grassland in New Zealand. Total P (688 µg/g) in grassland was significantly higher than TP (576 µg/g) in pine plantation at 0.05 significance level, while organic P (576 µg/g) in grassland showed significantly lower values than organic P (494 µg/g) in pine plantation at 0.05 significance level, largely due to more uptake by the young pine trees than by the grassland.

On the other hand, the opposite conversion from forest to pasture documented nutrient losses due to loss of plant cover and soil erosion and resulted in soil degradation (Sandoval-Pérez et al. (2009) and it may require more input of fertilizer to grow/maintain agronomic activities in the future. Vaithiyanathan and Correll (1992) found that the amount of P lost to the stream from

an agricultural watershed (2.41 kg/hectare/year) is eight times higher than that of P lost from forested watershed (0.30 kg/hectare/year) in the Rhode River watershed in Maryland. This issue can lead to a concern in the SFRW based on findings from Sabesan (2004). The author stated that the areal extent of pine plantations and wetland (low MP land use) from 1990 to 2005 in the SFRW showed a slight decrease (43.12 to 41.20 %), while that of agriculture (high MP land use) showed an increase (23.53 to 37.29 %). This may indicates that there would be a potential risk for soils in the SFRW to have an elevated MP level unless appropriate and effective land use management is implemented.

The soil drainage class also could significantly impact the mean difference of MP and TP values at L1. Based on Table 4-3, the mean MP concentrations at L1 relating to soil drainage classes decreased in the following order: Well drained > moderately well drained > somewhat poorly drained > excessively drained > poorly drained > very poorly drained. The mean of MP measurements in the well drained class (3.50 $\log_e$ µg/g) was significantly higher than the poorly drained (0.85 $\log_e$ µg/g) and very poorly drained class (-1.50 $\log_e$ µg/g), while other classes showed statistically no differences. These outcomes were similar to the research from Kuo and Baker (1982). They examined the impact of soil drainage class (well drained and poorly drained soil) on the amount of extractable P in long-term manure-amended soils in the State of Washington and observed more P extractable in the near surface layers (0 to 30 cm) from well drained soil than from poorly drained soil.

Mean TP at L1 based on soil drainage class also showed a similar trend suggesting higher TP in the well drained class (6.11 $\log_e$ µg/g) and lower TP in the very poorly drained (4.91 $\log_e$ µg/g) and the poorly drained class (4.72 $\log_e$ µg/g).

Two possible explanations for this could be offered. One is that the extent of drainage could influence soil P fertility through altering mineralization of OM. The release of inorganic phosphate from organic phosphates is controlled by microorganisms, whose activity can be suppressed by oxygen- deficient condition in poorly drained soil. The other one is the soil drainage also can affect soil P fertility through altering crystallinity of Fe oxide in soil. Reduced crystalliny was observed in poorly drained soil (Khalid et al., 1977; Kuo and Mikkelsen, 1979), which could make more inorganic P trapped in Fe-oxide and less extractable by Mehlich solution.

The impact of soil order on the variability of LNMP at L1 was evaluated and presented in Table 4-3. Based on Tukey's, the highest mean LNMP was found in Ultisols (2.19 $\log_e$ µg/g) followed by Entisols (1.96 $\log_e$ µg/g), others (1.78 $\log_e$ µg/g - Inceptisols, Histosols, and Alfisols), and Spodosols (0.73 $\log_e$ µg/g). The mean LNMP in Ultisols was significantly higher than the mean LNMP in Spodosols, while no statistical difference was found in Entisols and others at 0.05 significance level. Even though Ultisols are typically "highly weathered tropical soil with low total and available P content" (Friesen et al., 1997), however agriculture and improved pasture land uses prevailed on Ultisols showing higher means of LNMP in the SFRW. This may be caused by improper land management (e.g., excessive use of fertilizers) and may indicate that land use at the topsoil layer (L1) is recognized as a higher impact environmental variable on the variation of LNMP than soil order in this study area (Table 4-3).

This result could be the consequences of different soil properties such as drainage and textures in addition to excessive usage of fertilizer mentioned earlier. The capacity of clay to hold phosphate is greatly high due to negatively charged nature and high surface area per unit weight compared with sandy soil (Dong and Chesters, 1983; White, 2006). Low base saturation of Ultisols indicates low pH condition which is favorable condition for phosphate absorption

onto oxides of Fe and Al and some positively charged clay particles (Canfield Jr. and Hoyer, 1988).

On the other hand, the characteristics of Spodosols in Florida are poor to very poor drainage and sandy soils (Nair et al., 2004). Significant influence of the soil drainage class on the mean MP is shown in Table 4-3 indicating poorly drained soil had significantly lower MP than well drained soil. Low P fertility of sandy soil can be explained by low surface area per unit weight and a non-reactive chemical nature (White, 2006). Based on the results of the analysis the behaviors of MP and TP have showed somewhat similar patterns in responding to environmental variables (e.g., land use, soil drainage, and soil order) tested in the SFRW.

Various (physical, chemical and biological) properties of soils in Florida also depend on their different formations. Because P is likely the most important variable on geologic time scale (Tyrell, 1999) and geologic units (formations) are related to geologic time scale the impact of geologic unit on MP at L4 was investigated. According to Dunnett's T3 test at L4, the highest mean LNMP was identified in the Coosawhatchie formation (2.80 $\log_e$ µg/g) and decreased in the order of Ocala limestone (1.26 $\log_e$ µg/g), others (1.19 $\log_e$ µg/g which consists of Beach ridge and dunes, Cypresshead formation, Hawthorn group, Statenville formation, and Trail ridge sands), undifferentiated geology (-0.93 $\log_e$ µg/g), and undifferentiated sediments (-3.02 $\log_e$ µg/g). The formations of Coosawhatchie, Ocala limestone, undifferentiated geology, undifferentiated sediments originated during epochs of the Miocene, Ecocene, Pleistocene, and Pliocene, respectively.

Coosawhatchie formation had statistically higher mean LNMP compared to undifferentiated geology and undifferentiated sediments, but no significant differences with Ocala limestone and other geologic formations. This can be explained by the fact that the

Coosawhatchie formation is poorly to moderately consolidated, is affected by erosion and could potentially have phosphatic sands and variable clay, while undifferentiated geology and undifferentiated sediments are mainly located in Entisols (moderately high P) over karst terrain and Spodosols (low P), respectively (Ramnarine, 2003; Vasques, 2010). Latimer et al. (2002) analyzed 650 samples originating from different epochs (Miocene, Oligocene, and Eocene) in the southeastern Atlantic Ocean and found that much of the P were deposited in conjunction with terrigenous material and long-term P record showed elevated P concentrations in the late early Miocene.

Geostatistical Modeling of Soil Phosphorus and Model Validation

The RK model consists of modeling of structural component, $m(x_i)$, geostatistical modeling of the spatial autocorrelation of residuals, $\varepsilon'(x_i)$, and kriging of residuals, ε'' (Equation 3-4). The global trend models derived by SMLR with calibration data explained 66, 58, 52, and 68% of the variability of LNMP at L1, L2, L3, and L4, respectively (Table 4-4) and 68% in case of L1-LNTP. Each layer regression model describes the relationships between environmental variables and MP and TP, respectively. Some of the selected environmental variables in Table 4-3 could be confirmed as important predictor variables in the SMLR regression equations (Table 4-4). For example, land use (agriculture and wetland) and soil order (Ultisols) at L1-LNMP and geologic formation (Coosawhatchie formation) at L4-LNMP in the Table 4-4 were supporting the hypothesis that the most important environmental variables affecting MP were land use in the topsoil and geological information in the subsoil. The highest standardized coefficients of SMLR were mean reflectance of band blue (-1.566) from Landsat ETM+ followed by mean reflectance of band green (1.269) from Landsat ETM+, depth to water table (0.466), wetland (-0.266), and agriculture (0.250) at L1-LNMP and Coosawhatchie formation

(4.449) at L4-LNMP and they also confirmed the first hypothesis. In fact, the SMLR equations in Table 4-4 also showed that land use effects on soil MP occurred not only in the top layer but down to L2 and geologic formation influenced MP at all depths either positively or negatively. The mean reflectance of band 1 (blue) and 2 (green) from Landsat ETM+ at L1-LNMP were included as predictor variables in regression equations. Although these two spectral bands showed a statistical relationship to soil LNMP it should be noted that passive satellite cannot penetrate soil layers to relate directly to soil properties only the surface soil. Instead, interactions between these bands, which physically infer on chlorophyll content (Gitelson et al., 2008; Lim et al., 2009; Mahasandana et al., 2009), composition of vegetation (Solaimani et al., 2011) and LNMP express statistical relationships between the O factor and soil MP based on the SCORPAN.

The coefficients of the wetland (-2.491) and agriculture (1.243) at L2-LNMP showed that the influences of land use on the spatial distribution of MP were extended to deeper layers.

The variable 'east aspect' (direction of slope) at L3-LNMP could be a sign of more sunlight during day time stimulating vegetation growth leading to accumulation of OM and higher MP concentrations in soil. Compared with positive coefficient (0.799) of Ultisols at L1-LNMP Ultisols at L3-LNMP showed a negative coefficient (-1.323). This can be explained by the excessive usage of fertilizer (land management) which impacts mostly the topsoil (L1-LNMP) and natural characteristic of Ultisols at L3-LNMP which are highly weathered soils with low available P content.

The soil textures of clay and silt at L4-LNMP showed their association to soil drainage as relatively better drained soils (higher silt percentage) had a positive impact (coefficient of 0.637) and relatively poor drained soil (higher clay percentage) had a negative impact (coefficient of -

0.334) on MP in the study area. Moderately well or well drained soil (MWWDRAIN) at L3-LNMP, L4-LNMP, and L1-LNTP had positive coefficients of 2.254, 1.827, and 0.739, respectively, and these were consistent with the finding in Table 4-3. The positive coefficients of the Coosawhatchie formation of 1.543 at L2-LNMP and 1.706 at L1-LNTP extended the influences of geological unit on MP and TP up to L2 for MP and L1 for TP. The negative coefficient (-0.614) of sandy soil at L1-LNTP indicated the low capacity of sand to hold phosphate compared to clayey soils.

It was interesting to see that all regression equations in the Table 4-4 had variables that were somewhat related to soil drainage condition. Those variables included DTW and wetland at L1 and L2, SLOPEPCME33 and MWWDRAIN at L3, and CLAYPCT, PHYSHIGH, SPOPEPCME33, and MWWDRAIN at L4. In case of TP at L1, SPOPEPCME33, SAND, PHYSLOW, MWWDRAIN, and HYDGRPB were the variables selected related to soil drainage conditions. Table 4-4 shows coefficients of determination (R^2) of calibration and validation set and all other selected environmental variables in SMLR equations for LNMP and LNTP at each depth.

According to Table 4-5, the environmental variables most highly correlated with LNMP at the 0.01 significance level were DTW at L1 (Spearman correlation coefficient = 0.45) and L2 (0.47), and SLOPCME33 at L3 (0.40) and L4 (0.42). Thus, those were selected as secondary variables in the CK model and used to predict LNMP and LNTP in the study area. In the case of TP at L1, a correlation coefficient of 0.36 was found with SLOPCME33. Both DTW and SLOPCME33 can be related to soil drainage conditions which can explain the variability of MP and TP. A deeper depth to water table from ground surface and steeper mean slope represents relatively drier (well drained) soil conditions which somewhat favor plant growth, decomposition

of organic matter, and microorganism's mineralization mentioned earlier. Haygarth et al. (1999) states that permeable soils occupy steeper slopes and impermeable soils occupy gentler slopes.

To investigate the spatial structure of MP at each depth and that of TP at L1 in the SFRW semivariograms for each depth were generated. The parameters of semivariograms including the best fitted model, nugget, sill, and range for MP at each depth and TP at L1 are summarized in Table 4-6 and presented in Figure 4-2. The spherical model for MP at all depths and cubic models for TP at L1 were selected as best fitted semivariogram models and used to predict the variation of MP and TP in the study area.

The ranges of MP modeled by LNK varied from 10,897 m (L1) to 36,883 m (L4), while those of MP residuals modeled by RK/SMLR varied from 1,735 m (L1) to 18,785 m (L4). Note that the RK/SMLR semivariogram models were derived from residuals (MP observations - MP predictions based on SMLR). The range of TP at L1 modeled by CK was 38,954 m.

Compared to ranges of semivariograms by LNK at all depths, semivariograms from the residuals of log-transformed MP data tended to show much shorter ranges with 1,735 m (L1), 1,195 m (L2), 15,786 m (L3), and 18,785 m (L4) and stronger spatial continuity (i.e., nugget/sill ratio) of 14.3 (L1) and 8.1 (L2) and weaker spatial continuity of 49.2 (L3) and 82.2 of (L4). Rivero et al. (2007b) evaluated the performance of univariate (OK) and multivariate methods (RK and CK) to predict the spatial distribution of floc and soil TP in the Florida Everglades. They also showed that the spatial range of floc TP changed from 9,669 m in OK to 4,453 m in RK with quadratic regression and that of soil TP became shorter from 7,468 m in OK to 2,038 m in RK with Landsat ETM+ data and 1,903 m in RK with ASTER data, respectively.

It also showed that the sills at each depth became smaller (i.e., from 4.68 to 1.62 at L1, from 4.98 to 3.07 at L2, from 5.71 to 4.64 at L3, and from 6.70 to 4.66 at L4) in RK/SMLR

compared to the ones in LNK. This indicates that the global trends (SMLR models) explained some of the variability in MP leaving smaller residuals, which across all layers showed spatial autocorrelations. This is consistent with findings outlined by Grunwald (2008).

The nuggets of MP were smaller in RK/SMLR compared to the ones in LNK at L1 and L2, while they became larger at L3 and L4 (Table 4-6). A smaller nugget effect of MP at L1 and L2 (RK/SMLR models) may be due to the fact that they are residuals from raw measurements. This may explain that residuals of LNMP (RK/SMLR models at all four depths) show smaller variances compared to the raw measurements of MP. Furthermore, the secondary variable in CK facilitates to explain some of the short range variability in all soil P models, thus, reducing the nugget effect. Rivero et al. (2007b) also showed similar outcomes. Compared to the nugget of topsoil TP (mg/kg) in OK (0.0185) in their research the nugget became smaller with 0.0059 in RK with Landsat ETM + data and 0.0020 in RK with ASTER data.

The larger nugget effect from the MP residuals at L3 and L4 can be explained by the fact that the global trend model explained some of the long range variability but did little about capturing the short range variability. What was left was mostly noise in the short range adding to increase the nugget effect.

Table 4-7 shows error statistics (RMSE, ME, and RPD) calculated for both validation and calibration sets. Based on the RMSE calculated from the validation set RK/SMLR of MP at all depth and CK of TP at L1 were selected as the best methods to predict the spatial distribution of soil properties in the SFRW. Values of ME were quite different from zero for most of LNK and CK and they indicate that back transformation from logarithmic unit to original unit might have caused biased prediction results to produce larger RMSEs of LNK and CK at all depths. Results prove that multivariate methods (RK/SMLR and CK) outperformed the univariate method

(LNK) indicating that prediction of MP and TP maps in the SFRW can be improved by incorporating the spatial covariation between soil properties and auxiliary environmental properties compared to the univariate method (LNK) that uses only soil properties.

This is consistent with findings by other research studies. For example, Rivero et al. (2007a) used second variables including NDVI green derived from Landsat ETM+ (ETM – $NDVI_{green}$), distance of each sampling site to water control structures (WCS), Y geographic coordinates (Y-coord), NDVI derived from Landsat ETM+, and ASTER (ETM – NDVI and ASTER – NDVI, respectively) to predict the spatial distribution of floc TP and soil TP in WCA-2A in Florida. Based on RMSE values of floc TP prediction (i) RK with quadratic regression using ETM – $NDVI_{green}$ (134.9 mg/kg) was the best model followed by (ii) RK with multiple regression using ETM – $NDVI_{green}$, WCS, and Y-coord (201.2 mg/kg), and (iii) OK (206.1 mg/kg). For soil TP prediction, (i) RK with linear regression using ASTER – NDVI (200.1 mg/kg) was the best model followed by (ii) CK with ETM – NDVI (238.8 mg/kg), (iii) CK with ASTER – NDVI (238.2 mg/kg), and (iv) OK (257.5 mg/kg). They concluded that multivariate methods (RK and CK) outperformed the univariate (OK) for both floc TP and soil TP in their study. Heisel et al. (1999) experimented whether CK could perform better than OK to map weed density on a winter wheat field of 2.1 hectares and showed that CK (weed density as a primary variable and silt content as secondary variable) improved by up to 21% of the prediction variance (i.e., the average squared difference between actual and predicted weed density) when compared to OK. Triantafilis et al. (2001) compared various geostatistical methods including OK, RK, and CK to predict soil salinity (EC_e) within the Namoi Valley of northern New South Wales, Australia from a raw electromagnetic induction instrument (Type EM38) in a soil electrical conductivity (EC_a). Based on ME and RMSE to compare the least biased and most precise

method, RK (ME = -0.164 and RMSE = 1.240) was the best followed by CK (ME = -0.178 and RMSE = 1.279) and OK (ME = -0.185 and RMSE = 1.290). The superiority of RK was because of the high correlation (R^2 = 0.64) between estimated and predicted EC_e at calibration sites (R^2 = 0.61 for CK and R^2 = 0.60 for OK) and the environmental predictor to infer on soil salinity. That high correlation indicated low regression residuals between the measured and predicted EC_e.

Scatter plots between measurements and predictions of MP and TP at L1 (0-30 cm) are shown in Figure 4-3. The R^2 of 0.36 between measurements of MP and TP at L1 was increased to R^2 of 0.47 between MP and TP predictions, both derived using RK/SMLR. The points are more closely scattered along the line of equal values in graph B (Figure 4-3). This may indicate that secondary variables in SMLR explained some of the variability of MP and TP measurements at L1 so that predictions with RK/SMLR had stronger correlation compared with correlation in raw dataset.

The final prediction maps of MP and TP derived from the best performing methods at each depth are presented in Figure 4-4. The spatial distribution of MP and TP was reflected based on distributions of environmental predictor variables in the SMLR models at each depth. For example, higher MP at each depth could be found where the following mixture of environmental variables existed 1) agricultural land use, Ultisols, high DTW, and high green reflectance of Landsat ETM+ at L1, 2) high DTW, Coosawhatchie formation, high mean vegetation index, and agricultural land use at L2, 3) high SLOPCME33, east aspect, and moderately well or well drained soil at L3, and 4) Coosawhatchie formation, high silt percent in the soil, high SLOPCME33, and moderately well or well drained soil at L4. In case of TP at L1, higher TP could be found where a mixture of high SLOPCME33, Coosawhatchie formation,

undifferentiated sediment, lowlands, gaps and valley (PHYSLOW), and moderately well or well drained soil occurred.

Once significantly elevated MP and TP areas within the watershed are identified different management practices may be needed depending upon causes of elevated concentration of MP and TP. For example, buffer zones around areas where concentrations of MP and/or TP were high can be created and classified as vulnerable areas so that special attention may be given to monitor, detect, and prevent possible nutrient pollution of water resources in the SFRW. Appropriate education, awareness building, and regulations targeting the general public, decision makers and farmers can optimize land use activities and minimize environmental impact. There have been ordinances and programs already implemented to protect the environment at various levels of public organizations. For example, at County level in Florida, the Alachua County Fertilizer Code (Ordinance 09-06) and Landscape Irrigation Code (Ordinance 09-08) by the Alachua County Environmental Protection Department (ACEPD) have been adopted since 2009. Some of the sections of these codes are timing of fertilizer application (e.g., prohibited if rainfall greater than or equal to 2 inches in a 24-hour period), fertilizer free (e.g., 10 feet of any surface water body) and low maintenance zones (e.g., 6-foot from any ponds, lakes, streams, etc.), and training and licensing for commercial and institutional applicators (ACEPD, 2009). At State level, the new Florida Urban Turf Fertilizer Rule published in 2007 by the Florida Department of Agriculture and Consumer Services (FDACS) limits the amount of nitrogen and P used for homeowner's lawn, sport turf, and urban turf fertilization (e.g., 0.25 lb/1000ft^2 per application or 0.5 lb/1000ft^2 per year for low phosphate applications) (Dubberly, 2007). At national level, the National Pollutant Discharge Elimination System (NPDES) permit program and total maximum daily loads (TMDLs) by U.S. Environmental Protection Agency (EPA) regulate point sources

and nonpoint sources, respectively, which discharge pollutions into waters in the United States. Total maximum daily loads are the maximum amounts of particular pollutants that can be discharged into a water body and still meet its designated uses and applicable water quality standards (Hallas and Magley, 2008). The State of Florida is required by Section 303(d) of the Federal Clean Water Act to develop TMDLs for impaired waters that do not meet their designated uses (Florida Department of Environmental Protection, 2006). These efforts are focused on public education and regulatory efforts that guide and instruct individuals for best management practices (BMPs) including proper usage of fertilizer to protect our environment.

Additionally, continuous monitoring of soil nutrient levels from both soil and water bodies such as lake, river, and even ocean should be conducted because excessive soil nutrients will be transported by runoff and leaching process from soil to water bodies and finally into the ocean impacting coastal marine environments. Suzumura et al., (2000) compared phosphate concentrations in groundwater and seawater mixing at two sandy beaches in Tokyo Bay, Japan and learned higher concentrations (~32 μ mole (M)) were discovered in fresh groundwater, while saline water showed lower concentrations (0.58 to 2.8 μM). This study supports the idea that high phosphate concentrations in groundwater should be reduced to preserve ocean ecosystem.

These efforts will preserve our environment at some degree, but there are still certain key factors that cannot be easily amended. As an example, geological unit was one of the key factors impacting variation of soil properties in the SFRW which is a part of the State of Florida where P-rich geological parent materials naturally exist. Whereas land use activities, land cover, and surface characteristics of soils (e.g. through amendments, conservation management, etc.) can be adapted, there is little that can be done managing geologic formations of which some are naturally P-rich, except for mining operations that have ceased in many areas in Florida to

conserve the environment. Even with this limitation, understanding the spatial distribution of soil properties and applying different management practices will be helpful to optimize crop productivity and to minimize P loss from soil protecting the environment.

As conclusions in this chapter, 1) pair-wise comparisons of MP among the four depths and MP and TP at L1, 2) relationships between MP and TP and selected environmental variables (land use, soil drainage, soil order, and geologic unit), and 3) geostatistical modeling of MP at four depths and TP at L1 in the SFRW were investigated.

Pair-wise comparisons of MP among the four depths indicated stronger correlations (i.e., L1-L2 (0.80), L1-L3 (0.44), L1-L4 (0.37), L2-L3 (0.72), L2-L4 (0.61) and L3-L4 (0.81)) between close (adjacent) layers. These findings suggest that spatial distribution of MP from the upper layer might influence the adjacent layer beneath it through vertical transport processes.

Land use, soil drainage, soil order, and geologic unit were statistically significant factors at the 0.05 significance level to affect the variations of MP and TP. The global trend models derived by SMLR showed that land use may be the key factor to influence the variation of MP at the top layer (L1), while the geologic unit could be the major factor at the deepest layer (L4). This confirms the hypothesis that among environmental factors land use ranks highest to predict MP at the topsoil, and geology ranks highest to predict MP at the subsoil.

Based on RMSEs calculated on validation sets the best prediction methods of MP were RK/SMLR at all depths (i.e., 41.3 µg/g at L1, 24.9 µg/g at L2, 32.3 µg/g at L3, and 48.3 µg/g at L4) and CK was the most accurate for TP (404.9 µg/g) at L1. This suggests that multivariate methods (RK/SMLR and CK) outperformed the univariate method (LNK) indicating that incorporating the spatial covariation between soil and auxiliary environmental properties are more accurate than the univariate method (LNK) in the SFRW. For example, classification of

land use by the satellite imagery, soil drainage related variables, and various topographic variables played a key role to improve predictions of soil properties in the study area.

Once soil and surface/ground water are contaminated by excessive P they would take much time to recover; thus, conservation of their conditions consequently should be focused on prevention not treatment or recovery. Various policies and regulations already have been employed and they provide detailed instructions to safeguard our environment (e.g., timing and limits on amounts of fertilizer and irrigation).

Even though the key factors influencing the spatial distribution of MP and TP may be different at other regions the general approach including soil sampling design, lab analysis, geostatistical modeling, etc. adopted in this research is transferable to other regions with similar environmental characteristics. The spatially-explicit assessment of soil properties through incorporation of ancillary environmental variables provides a better understanding of the underlying heterogeneity and variation of soil properties to perform risk assessment, optimize fertilization for better crop production, and conserve ecosystems.

.

Table 4-1. Descriptive statistics of Mehlich phosphorus (MP) at each depth and total phosphorus (TP) at the top layer in the Santa Fe River Watershed.

Statistics	Total set	Calibration	Validation	Total set	Calibration	Validation
	L1-MP (µg/g) 0 – 30 cm			L1-LNMP (log_e µg/g) 0 – 30 cm		
Observations	137	98	39	137	98	39
Mean	23.91	25.03	21.10	1.68	1.59	1.88
Median	6.10	5.76	6.76	1.81	1.75	1.91
Std. deviation	39.71	40.85	37.02	2.09	2.21	1.78
Kurtosis	5.97	4.98	10.97	-0.37	-0.74	1.57
Skewness	2.43	2.25	3.13	-0.46	-0.36	-0.75
Minimum	0.028	0.045	0.028	-3.56	-3.10	-3.56
Maximum	189.12	189.12	187.10	5.24	5.24	5.23
Range	189.09	189.07	187.07	8.80	8.34	8.79
	L2-MP (µg/g) 30 – 60 cm			L2-LNMP (log_e µg/g) 30 – 60 cm		
Observations	137	98	39	137	98	39
Mean	15.44	17.86	9.36	0.88	0.92	0.78
Median	3.28	3.42	2.60	1.19	1.23	0.95
Std. deviation	28.20	30.88	18.97	2.50	2.62	2.19
Kurtosis	6.77	5.01	18.88	-0.03	-0.29	1.19
Skewness	2.63	2.32	4.04	-0.67	-0.61	-1.01
Minimum	BDL	BDL	BDL	-5.30	-5.30	-5.30
Maximum	145.32	145.32	106.73	4.98	4.98	4.67
Range	145.32	145.32	106.73	10.28	10.28	9.97
	L3-MP (µg/g) 60 – 120 cm			L3-LNMP (log_e µg/g) 60 – 120 cm		
Observations	135	97	38	135	97	38
Mean	20.44	24.52	10.01	-0.09	0.20	-0.82
Median	1.70	2.09	1.20	0.53	0.74	0.18
Std. deviation	45.50	49.78	30.21	3.40	3.41	3.30
Kurtosis	10.15	8.32	17.00	-1.06	-1.05	-1.18
Skewness	3.07	2.79	4.12	-0.31	-0.34	-0.32
Minimum	BDL	BDL	BDL	-5.30	-5.30	-5.30
Maximum	271.78	271.78	153.92	5.61	5.61	5.04
Range	271.78	271.78	153.91	10.90	10.90	10.33
	L4-MP (µg/g) 120 – 180 cm			L4-LNMP (log_e µg/g) 120 -180 cm		
Observations	131	94	37	131	94	37
Mean	26.00	30.07	15.63	-0.16	0.02	-0.61
Median	1.92	2.04	1.65	0.65	0.71	0.50
Std. deviation	52.41	55.61	42.14	3.71	3.80	3.49
Kurtosis	6.77	5.31	17.10	-1.31	-1.35	-1.22

Table 4-1. Continued.

Statistics	Total set	Calibration	Validation	Total set	Calibration	Validation
	L4-MP (µg/g) 120 – 180 cm			L4-LNMP (log_e µg/g) 120 -180 cm		
Skewness	2.58	2.30	3.94	-0.23	-0.26	-0.22
Minimum	BDL	BDL	BDL	-5.30	-5.30	-5.30
Maximum	258.97	258.97	223.02	5.56	5.56	5.41
Range	258.97	258.97	223.02	10.86	10.86	10.71
	L1-TP (µg/g) 0 – 30 cm			L1-LNTP (log_e µg/g) 0 – 30 cm		
Observations	137	98	39	137	98	39
Mean	333.86	349.76	293.91	5.14	5.17	5.05
Median	143.59	145.35	141.25	4.97	4.98	4.95
Std. deviation	519.71	522.20	517.97	1.06	1.10	0.98
Kurtosis	9.49	9.00	12.76	0.04	-0.29	1.60
Skewness	3.03	2.88	3.60	0.74	0.65	1.06
Minimum	24.89	24.89	33.86	3.21	3.21	3.52
Maximum	2958.21	2958.21	2447.00	7.99	7.99	7.80
Range	2933.32	2933.32	2413.14	4.78	4.78	4.28

Abbreviations: L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus; TP = total phosphorus; BDL (below detection limit) = 0.005 µg/g.

Table 4-2. Pair-wise comparison of log_e-transformed Mehlich phosphorus (MP) in µg/g among the different four depths and total phosphorus (TP) in µg/g at top layer in the Santa Fe River Watershed.

Pair	Number of samples	Pearson Correlation	Mann-Whitney Test Median difference (log_e µg/g)
L1-LNMP – L2-LNMP	137	0.80*	0.62*
L1-LNMP – L3-LNMP	135	0.44*	1.28*
L1-LNMP – L4-LNMP	131	0.37*	1.16*
L2-LNMP – L3-LNMP	135	0.72*	0.66*
L2-LNMP – L4-LNMP	131	0.61*	0.54*
L3-LNMP – L4-LNMP	131	0.81*	-0.12
L1-LNMP – L1-LNTP	137	0.60*	-3.16*

Abbreviations: L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus; TP = total phosphorus.

* Statistically significant at the 0.05 significance level.

Table 4-3. Analysis of variance (ANOVA) tests and homogeneous groups of log_e-transformed Mehlich phosphorus in µg/g at L1 (0-30 cm) and at L4 (120-180 cm) according to either Tukey's test or Dunnett's T3 test at the 0.05 significance level in the Santa Fe River Watershed.

Variables and factor	Levene	ANOVA	Welch	B-F	Homogeneous groups[1]	Group mean (log_e µg/g)
L1-LNMP by land use	0.8	*13.7*[*]	13.4[*]	13.6[*]	Agriculture [a] Improved pasture [a] Upland forest [a b] Urban [a b] Rangeland [a b] Pineland [b c] Wetland [c]	3.47 3.46 2.39 2.18 1.99 0.87 -0.65
L1-LNTP by land use	3.0[*]	6.4[*]	*8.7*[*]	5.7[*]	Improved pasture [a] Agriculture [a] Rangeland [a] Wetland [a b] Upland forest [b] Urban [b] Pineland [b]	5.84 5.75 5.49 5.33 5.10 5.08 4.41
L1-LNMP by soil drainage	3.3[*]	11.4[*]	27.8[*]	*17.8*[*]	Well drained [a] Moderately well drained [a b] Somewhat poorly drained [a b] Excessively drained [a b] Poorly drained [b] Very poorly drained [c]	3.50 2.81 2.58 1.80 0.85 -1.50
L1-LNTP by soil drainage	1.1	*6.0*[*]	4.7[*]	5.7[*]	Well drained [a] Moderately well drained [a b] Somewhat poorly drained [a b] Excessively drained [b] Very poorly drained [b] Poorly drained [b]	6.11 5.64 5.44 4.94 4.91 4.72
L1-LNMP by soil order	1.0	*4.09*[*]	4.12[*]	4.11[*]	Ultisols [a] Entisols [a b] [2] Other [a b] Spodosols [b]	2.19 1.96 1.78 0.73
L1-LNTP by soil order	8.5[*]	4.7[*]	*4.0*[*]	3.6[*]	[2] Other [a] Entisols [a b] Ultisols [a b] Spodosols [b]	5.69 5.33 5.19 4.68

Table 4-3. Cotinued.

Variables and factor	Levene	ANOVA	Welch	B-F	Homogeneous groups[1]	Group mean (log_e µg/g)
L4-LNMP by surficial geologic unit	3.1*	23.1*	*22.5**	28.3*	Coosawhatchie formation [a]	2.80
					Ocala limestone [a b]	1.26
					[3] Other [a b]	1.19
					Undifferentiated geology [b]	-0.93
					Undifferentiated sediments [c]	-3.02

Abbreviations: B-F = Brown-Forsythe's ANOVA test statistic; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; Levene = Levene's test statistic; MP = Mehlich phosphorus; TP = total phosphorus; Welch's = Welch's ANOVA test statistic.

[1] Columns with different letters have significantly different means at the 0.05 significance level, based on Tukey's test for L1-LNMP by land use, L1-LNTP by soil drainage, and L1-LNMP by soil order and Dunnett's T3 test for L1-LNTP by land use, L1-LNMP by soil drainage, L1-LNTP by soil order, and L4-LNMP by surficial geologic unit.
[2] Other = Inceptisols, Histols, and Alfisols.
[3] Other = beach ridge and dunes, Cypresshead formation, Statenville formation, and Trail ridge stands.

The preferred test based on the number of samples and equality of variance is shown in italics.

Table 4-4. Global trend model of log_e-transformed Mehlich phosphorus (LNMP) in µg/g at four different layers and total phosphorus (LNTP) in µg/g at L1 represented by the stepwise multiple linear regression models and regression coefficients in the Santa Fe River Watershed.

Property	R^2			Regression equation*
	Calib.	Valid.	Adj.	
L1-LNMP	0.66	0.13	0.63	14.486 + 0.020 * [DTW] + 1.245 * [AGRI] - 0.392 * [B1ME33] - 1.669 * [WET] + 0.258 * [B2ME33] + 0.799 * [ULT] - 1.449 * [OCALA]
L2-LNMP	0.58	0.26	0.55	-1.97 + 0.015 * [DTW] + 1.543 * [COOSA] - 2.491 * [WET] + 0.038 * [VEGIDXME33] + 1.243 * [ARGI]
L3-LNMP	0.52	0.18	0.49	0.326 - 2.888 * [UNDIFFSED] + 0.604 * [SLOPCME33] - 2.538 * [UNDIFF] + 1.322 * [ASPECTE] + 2.254 * [MWWDRAIN] - 1.323 * [ULT]
L4-LNMP	0.68	0.35	0.66	0.363 + 4.449 * [COOSA] - 0.334 * [CLAYPCT] + 0.637 * [SILTPCT] - 3.660 * [PHYSHIGH] + 0.480 * [SLOPCME33] + 1.827 * [MWWDRAIN]
L1-LNTP	0.68	0.24	0.65	4.180 + 0.151 * [SLOPCME33] - 0.809 * [UPFOREST] + 1.706 * [COOSA] - 0.614 * [SAND] + 1.339 * [UNDIFFSED] + 0.886 * [PHYSLOW] + 0.739 * [MWWDRAIN] - 0.363 * [HYDGRPB]

Abbreviations: Adj. = adjusted on calibration data; AGRI = agriculture land use classified from Landsat Enhanced Thematic Mapper Plus (ETM+); ASPECTE = east-facing slope; B1ME33 = mean reflectance of band 1 derived from Landsat ETM+ within a 3×3 pixel moving window; B2ME33 = mean reflectance of band 2 derived from Landsat ETM+ within a 3×3 pixel moving window; Calb. = calibration dataset; CLAYPCT = clay content percent in soil; COOSA = Coosawhatchie formation originated during the Miocene; D.F. = degree of freedom; HYDGRPB = moderate infiltrations rate in soil hydrologic group; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); MWWDRAIN = moderately well or well drained in soil drainage class; DTW = depth to water table (cm); OCALA = Ocala limestone formation originated during the Eocene; PHYSHIGH = highlands, uplands and ridges in general physiographic type class; PHYSLOW = lowlands, gaps and valleys in general physiographic type class; R^2 = coefficient of determination; SAND = sandy soil in particle size class; SILTPCT = silt content percent in soil; SLOPCME33 = mean slope (percent) within a 3×3 pixel moving window; TP = total phosphorus (µg/g); ULT = Ultisols in soil order; UNDIFF = Undifferentiated formation originated during the Pleistocene; UNDIFFSED = Undifferentiated sediment originated during the Pliocene; UPFOREST = upland forest land use classified from Landsat ETM+; VEGIDXME33 = mean vegetation index (infrared – red) derived from Landsat ETM+ within a 3×3 pixel moving window; Valid. = validation dataset; WET = wetland land use classified from Landsat ETM+.

* Statistically significant at the 0.10 significance level.

Table 4-5. Spearman correlation of selected variables for cokriging in the Santa Fe River Watershed.

Property	Selected variable	Spearman correlation*
L1-LNMP	DTW	0.45
L2-LNMP	DTW	0.47
L3-LNMP	SLOPCME33	0.40
L4-LNMP	SLOPCME33	0.42
L1-LNTP	SLOPCME33	0.36

Abbreviations: DTW = depth to water table (cm); L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus in µg/g; SLOPCME33 = mean slope (percent) within a 3×3 pixel moving window;. TP = total phosphorus in µg/g;

* Statistically significant at the 0.01 significance level.

Table 4-6. Summary of semivariogram parameters for the best spatial prediction method for log_e-transformed Mehlich phosphorus (MP) in µg/g at different four layers and total phosphorus (TP) in µg/g at L1 in different layers across the Santa Fe River Watershed.

Property	Interpolation method	Fitted model	Nugget	Sill	Range (m)	Nugget/sill (%)
L1-MP	LNK	Spherical	2.0540	4.6760	10,897	43.9
L1-LNMP [a]	RK/SMLR	Spherical	0.2317	1.6197	1,735	14.3
L2-MP	LNK	Spherical	1.3310	4.9820	11,326	26.7
L2-LNMP [a]	RK/SMLR	Spherical	0.2493	3.0703	1,195	8.1
L3-MP	LNK	Spherical	1.6420	5.7110	19,973	28.8
L3-LNMP [a]	RK/SMLR	Spherical	3.1220	4.6380	15,786	49.2
L4-MP	LNK	Spherical	1.3030	6.7040	36,883	19.4
L4-LNMP [a]	RK/SMLR	Spherical	3.8280	4.6560	18,785	82.2

L1-LNTP

Interpolation method	Fitted model	Coregionalization matrix Range (m)		Nugget		Sill		Nugget/ sill (%)	
				L1-LNTP	SLOPCME33	L1-LNTP	SLOPCME33	L1-LNTP	31.9
CK	Cubic	38,954	L1-LNTP	0.5352	0.1196	1.6792	1.3336	SLOPCME333	45.8
			SLOPCME33	0.1196	1.314	1.3336	2.868	L1-LNTP and SLOPCME33	9.0

Abbreviations: CK = cokriging; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; LNK = lognormal kriging; MP = Mehlich phosphorus; RK/SMLR = regression kriging with stepwise multiple linear regression; SOLPEME33 = mean slope (percent) within a 3×3 pixel moving window; TP = total phosphorus;.

[a] Parameters of semivariogram were derived from residuals of global trend model with stepwise multiple linear regression.

Table 4-7. Error statistics of the best prediction methods for Mehlich phosphorus (MP) in µg/g and total phosphorus (TP) in µg/g across the Santa Fe River Watershed.

Property	$RMSE_v$ (µg/g)			ME_v (µg/g)			RPD		
	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK
L1-MP	97.8	*41.3*	41.7	-49.8	*8.2*	12.5	0.37	*0.89*	0.88
L2-MP	274.0	*24.9*	160.6	-114.8	*-0.1*	-62.9	0.07	*0.75*	0.12
L3-MP	221.8	*32.3*	603.9	-99.3	*6.7*	191.6	0.13	*0.92*	0.05
L4-MP	468.4	*48.3*	338.9	-191.21	*10.4*	-114.8	0.09	*0.86*	0.12
L1-TP	420.7	416.1	*404.9*	12.5	*7.5*	-29.8	1.22	1.23	*1.26*

Property	$RMSE_c$ (µg/g)			ME_c (µg/g)			RPD		
	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK
L1-MP	72.9	*30.8*	36.9	-37.6	*3.1*	8.9	0.56	*1.32*	1.10
L2-MP	294.9	*55.8*	183.1	-137.8	*-11.2*	-80.4	0.10	*0.55*	0.17
L3-MP	219.7	*47.6*	*583.9*	-101.7	*16.6*	-183.6	0.23	*1.04*	0.08
L4-MP	288.6	*55.5*	175.4	-133.0	*22.7*	-66.5	0.19	*1.00*	0.32
L1-TP	339.1	*274.8*	341.*4*	*-8.0*	57.3	-43.3	1.53	*1.89*	1.52

Abbreviations: CK = cokriging; L1-MP = MP at 0-30cm; L2-MP = MP at 30-60cm; L3-MP = MP at 60-120cm; L4-MP = MP at 120-180cm; L1-TP = TP at 0-30cm; LNK = lognormal kriging; ME_v = mean error calculated on validation set; MP = Mehlich phosphorus; $RMSE_c$ = root mean square error calculated on calibration set; $RMSE_v$ = root mean square error calculated on validation set; ME_c = mean error calculated on calibration set; RK/SMLR = regression kriging with stepwise multiple linear regression; RPD = residual prediction deviation; TP = total phosphorus.
The best RMSE of models with all validation sets are shown in italics.

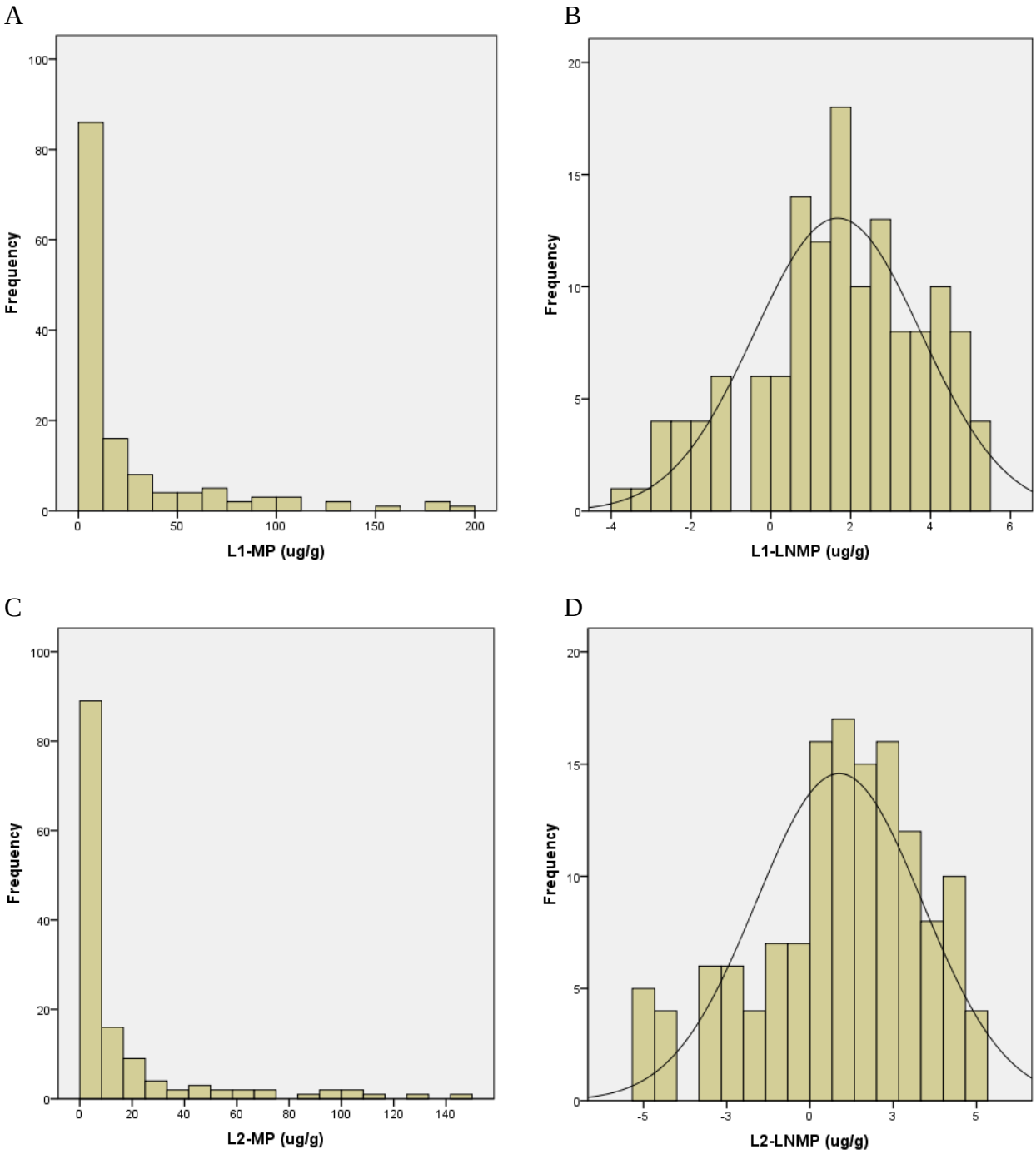


Figure 4-1. Histograms of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g datasets in the Santa Fe River Watershed: (A) MP at 0-30cm; (B) Log_e(MP) at 0-30cm; (C) MP at 30-60cm; (D) Log_e(MP) at 30-60cm; (E) MP at 60-120cm; (F) Log_e(MP) at 60-120cm; (G) MP at 120-180cm; (H) Log_e(MP) at 120-180cm; (I) TP at 0-30cm; (J) Log_e(TP) at 0-30cm.

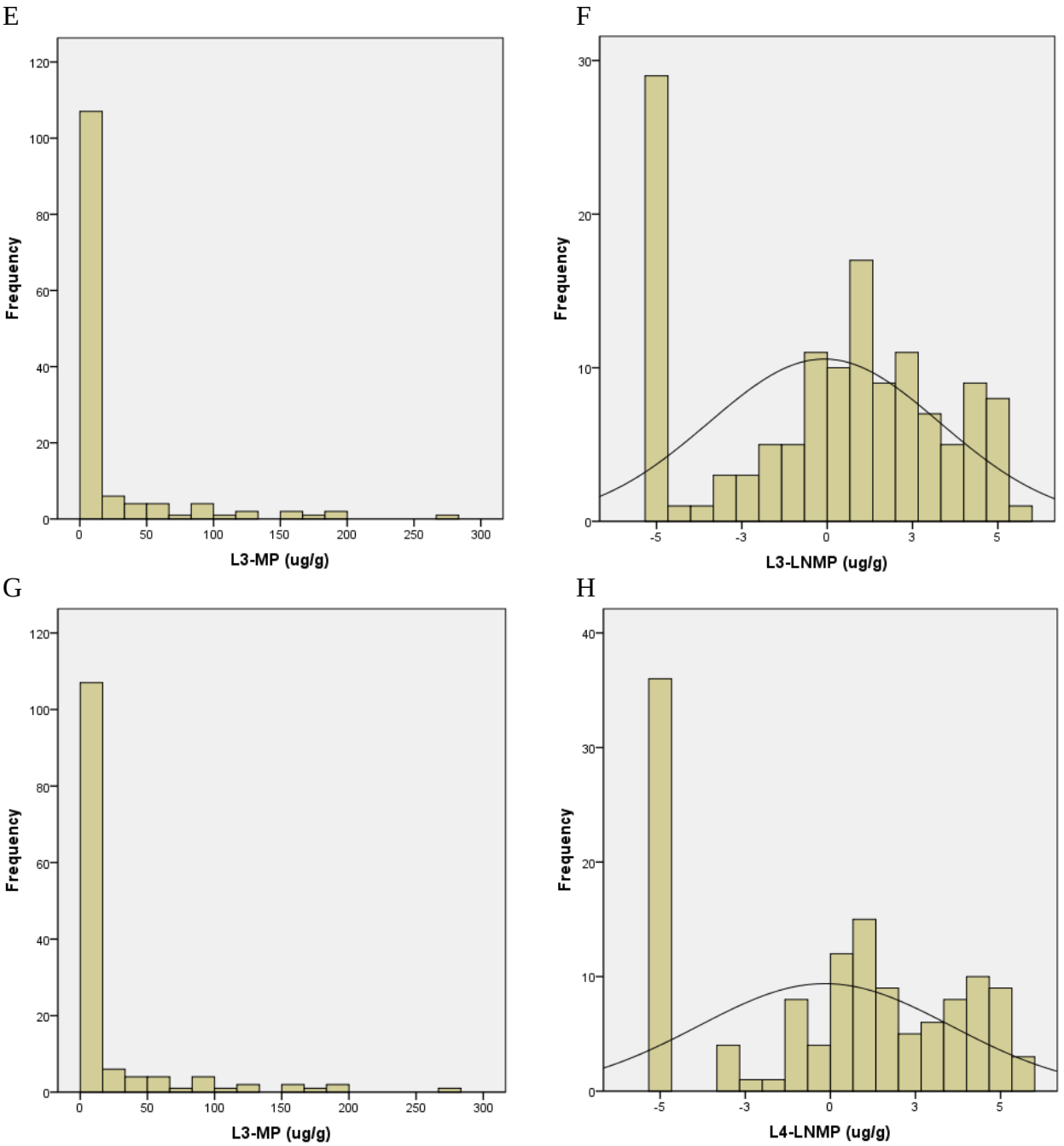
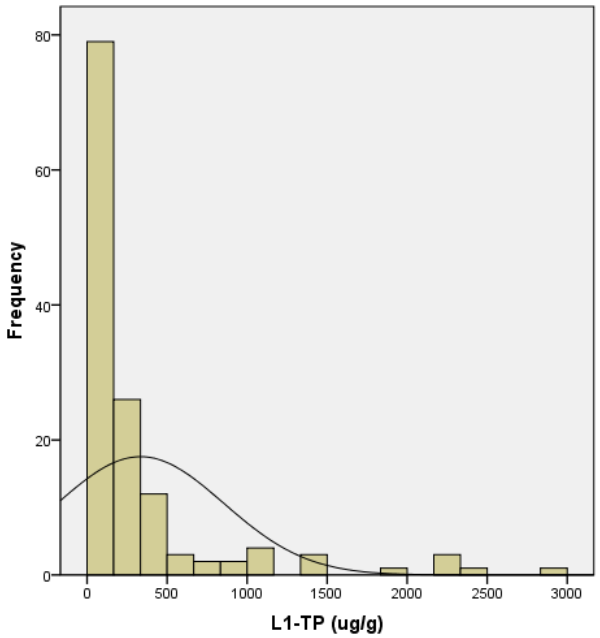


Figure 4-1. Continued.

I



J

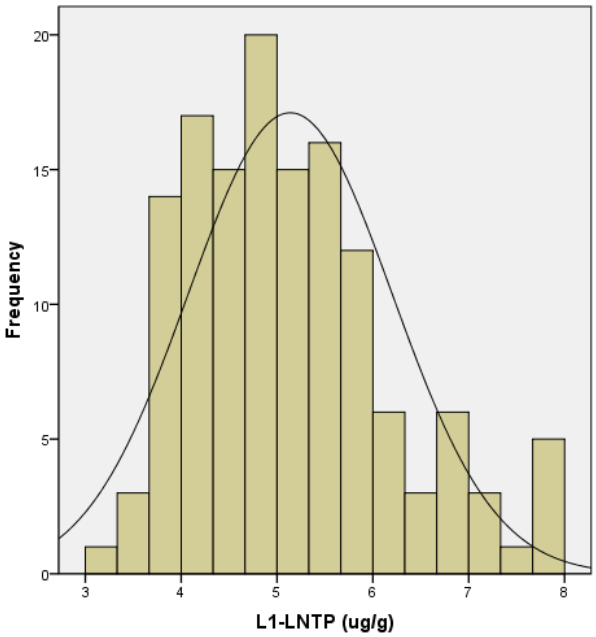


Figure 4-1. Continued.

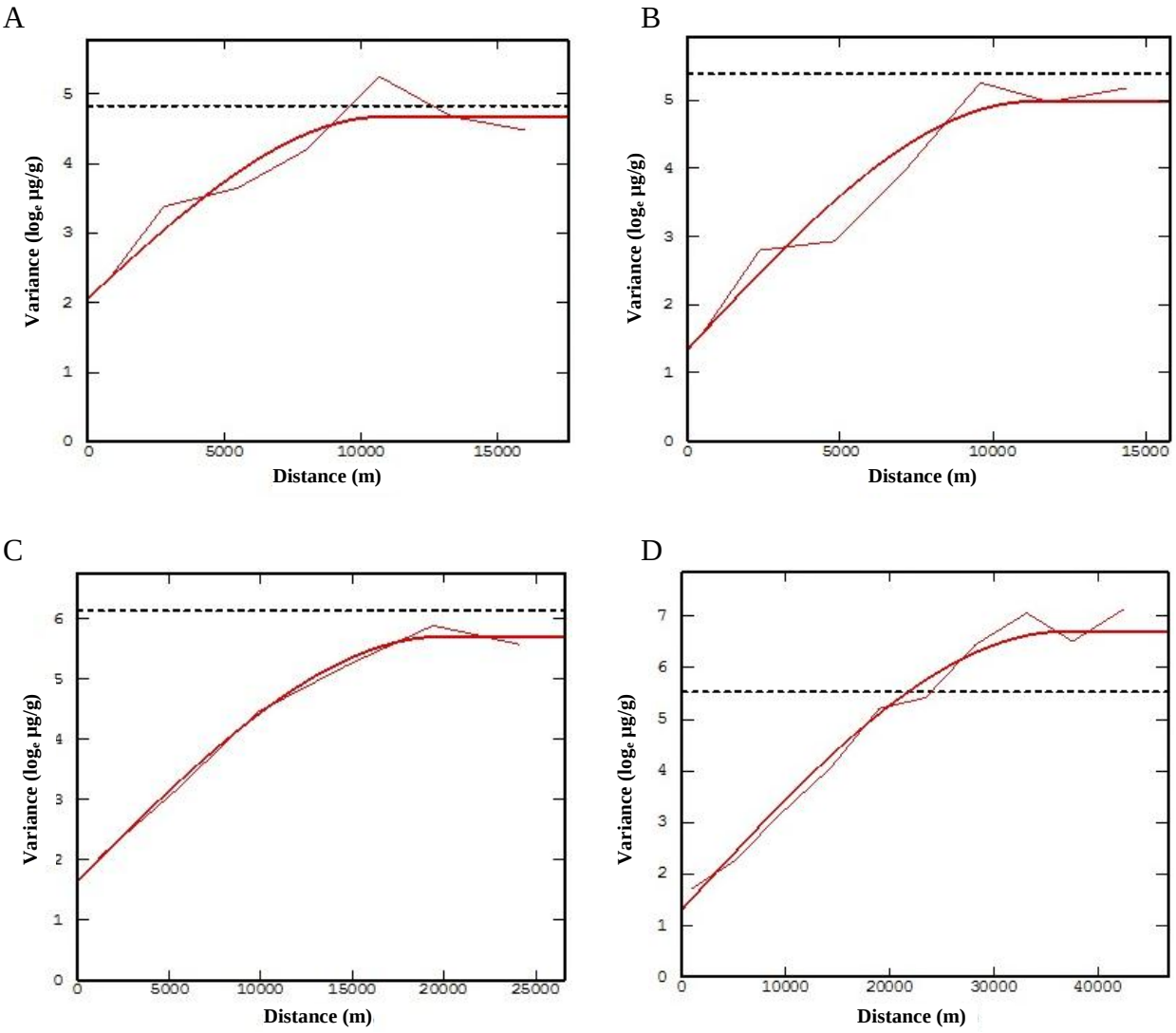
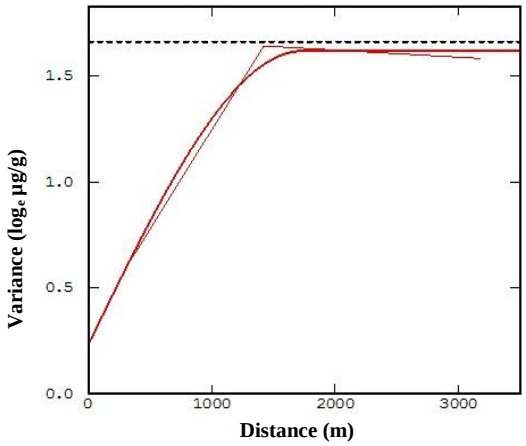
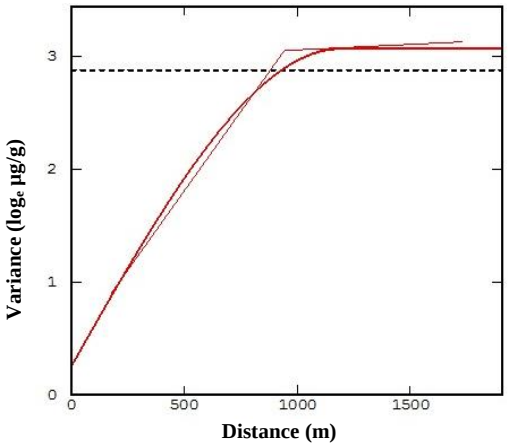


Figure 4-2. Semivariograms for the best spatial prediction method for log$_e$-transformed Mehlich phosphorus (MP) in µg/g at different four layers and a cross-semivariogram for total phosphorus (TP) in µg/g at L1 in different layers across the Santa Fe River Watershed.: (A) Lognormal kriging (LNK) of LNMP at L1(0-30 cm); (B) LNK of LNMP at L2 (30-60 cm); (C) LNK of LNMP at L3 (60-120 cm); (D) LNK of LNMP at L4 (120-180 cm); (E) Kriging on residuals of LNMP at L1; (F) Kriging on residuals of LNMP at L2; (G) Kriging on residuals of LNMP at L3; (H) Kriging on residuals of LNMP at L4; (I) samivariograms of LNTP and mean slope in percent within a 3×3 pixel moving window (SLOPC33) and a cross-variogram between LNTP and SLOPC33 at L1.

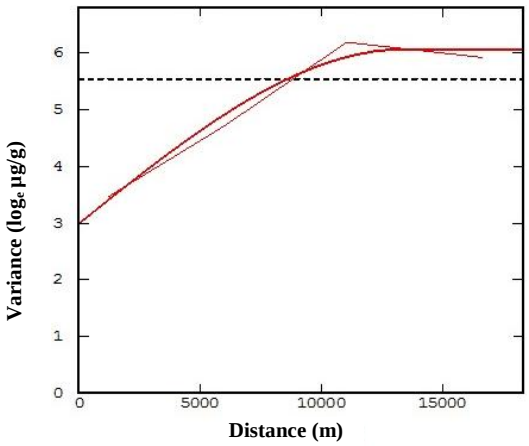
E



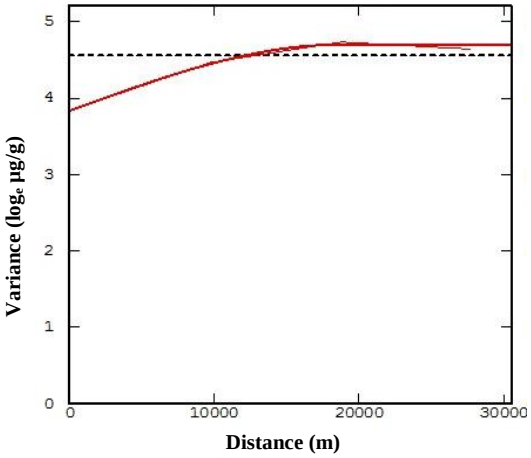
F



G



H



I

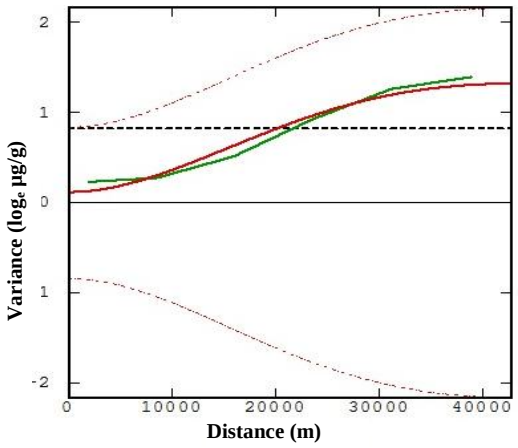
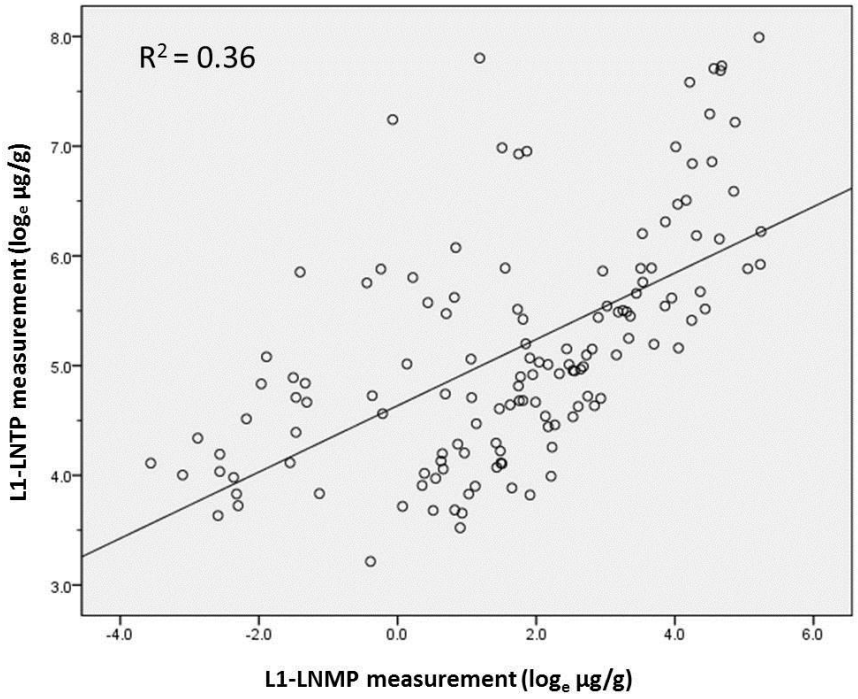


Figure 4-2. Continued.

A



B

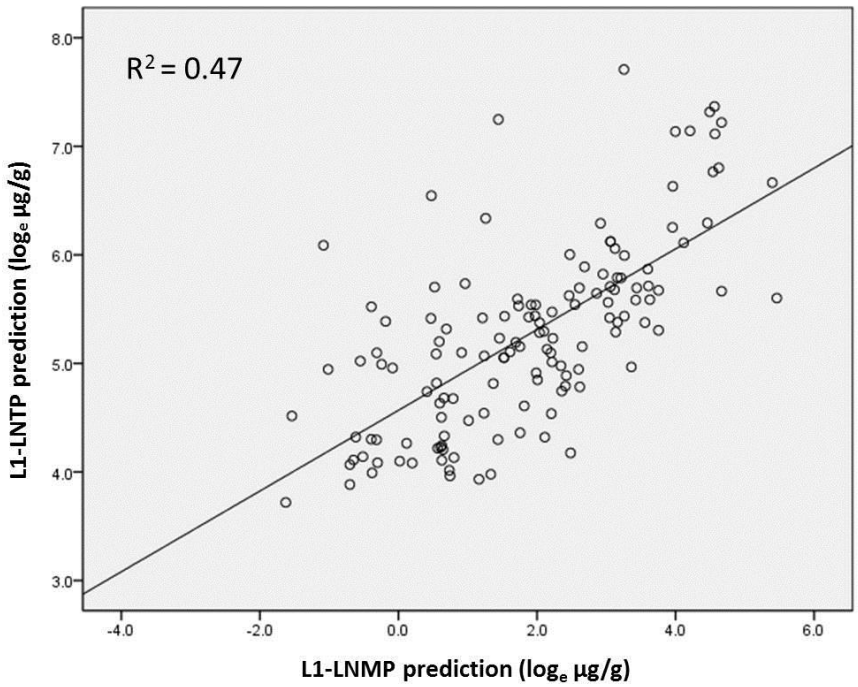


Figure 4-3. Scatter plots of measurements and predictions of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g derived by regression kriging with stepwise multiple linear regression in the Santa Fe River Watershed. (A) MP measurements at L1 (0-30 cm) and TP measurements at L1; (B) MP predictions at L1 and TP predictions at L1.

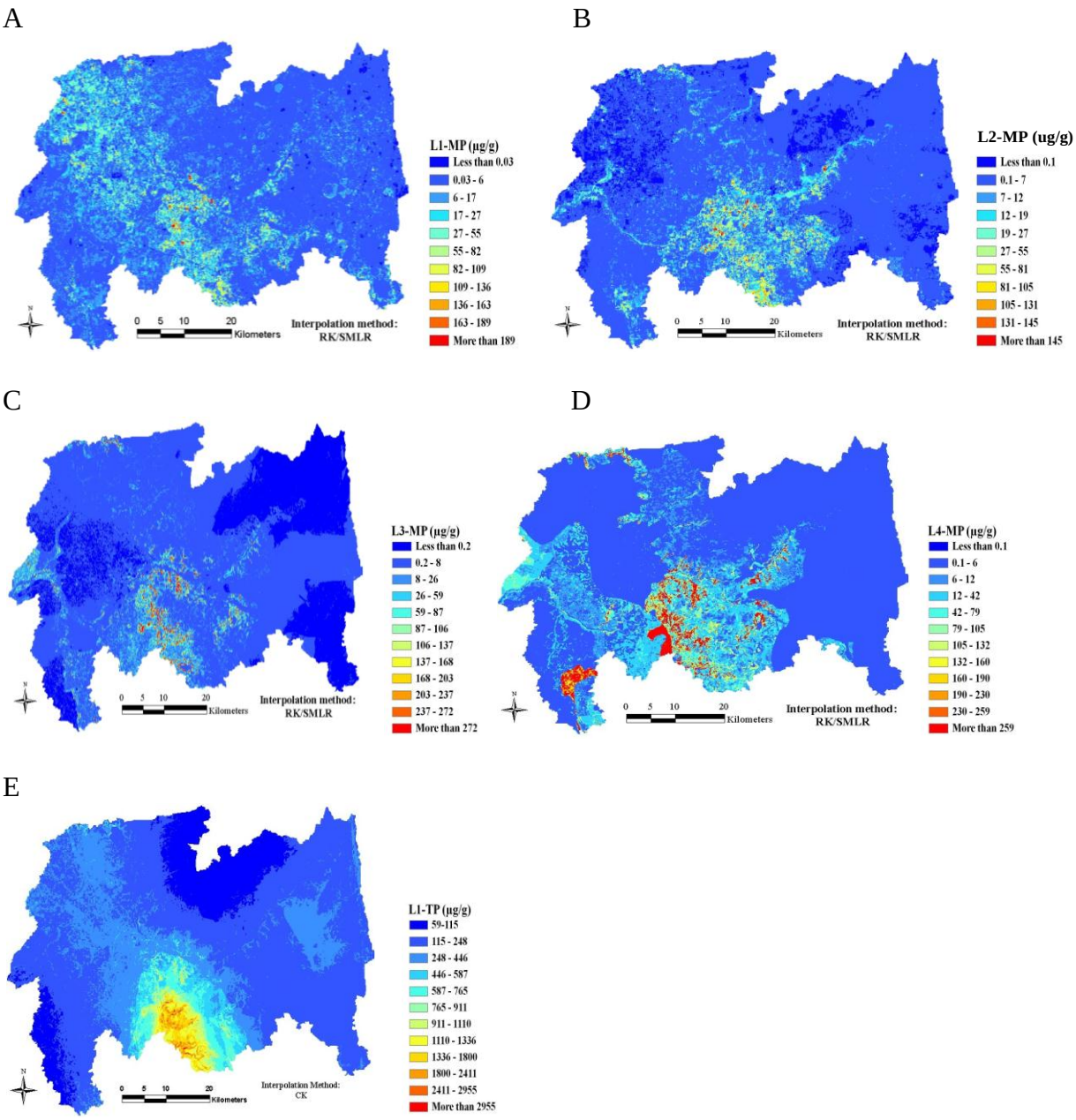


Figure 4-4. Final predictions in the Santa Fe River Watershed.: (A) Mehlich phosphorus (MP) in µg/g at L1 (0-30 cm) by regression kriging using stepwise multiple linear regression to model the global trend (RK/SMLR); (B) MP at L2 (30-60 cm) by RK/SMLR; (C) MP at L3 (60-120 cm) by RK/SMLR; (D) MP at L4 (120-180 cm) by RK/SMLR; (E) total phosphorus (TP) in µg/g at L1 by cokriging.

CHAPTER 5
MODELING OF SOIL PHOSPHORUS USING SATELLITE IMAGERY AND ANCILLARY SPATIAL ENVIRONMENTAL DATA IN THE SANTA FE RIVER RANCH BEEF UNIT

Because phosphorus (P) is attached to soil materials, erosion largely determines the particulate P (PP) movement in the landscape. Sources of PP in streams include eroding surface soil, stream banks, channel beds, and plant material (e.g. leaves) and these PP from soil erosion and dissolved P from runoff are the main factors affecting the P transport to underground/surface waters (Randall et al., 2002). These excessive P in water accelerate eutrophication, which causes abnormal production of algae and aquatic plants and lack of dissolved oxygen and may lead to poor water quality. Mander et al. (1998) stated that P runoff has always been reported to positively correlate with stream runoff; thus, it is critical to understand the behavior and spatial distribution of P to optimize the benefits from agronomical prospective (e.g., maximized plant growth or harvest) as well minimize environmental impact through conservation management. Agricultural soils are considered as the main diffuse source of P reaching freshwaters due to imports of fertilizer and livestock feeds (Foy and Withers, 1995; Haygarth et al., 1998).

The State of Florida supplies about 25% of the world's phosphate needs and 75% of US domestic needs including mainly fertilizer for agricultural purpose, vitamins, soft drinks, and toothpaste (Greenhalgh, 2005). Phosphate mining in the State of Florida provides one of the leading export commodities, however the risk of environmental damage from mining activities is high.

Remote sensing techniques have been used for identifying soil properties through direct sensing and indirect inference from biophysical features and phenomena measured on the earth's surface (Campling et al., 2002; Peng et al., 2003; Lillesand and Kiefer, 2008; Sridhar et al., 2009; Eldeiry and Garcia, 2010). Stochastic methods, such as regression models, have been used to quantify relationships between soil properties and spectral data/indices derived from satellite

imagery (add reference). The spectral data allow inference on land cover, land use, specific vegetation properties such as chlorophyll content, crop stress, and more (add reference). These spectral data can be fused with ancillary environmental properties to build complex models to predict soil properties, such as soil P. The SCORPAN model (McBratney et al., 2003) and STEP-AWBH model (Grunwald et al., 2011) provide conceptual frameworks for the implementation of soil predictions.

Despite of limitations of remote sensors (e.g. cloud cover and/or atmospheric scatter) multispectral scanners such as Landsat, Satellite Pour l'Observation de la Terre (SPOT) and IKONOS have been successfully used to support mapping of soil properties (Hengl et al., 2007; Rivero et al., 2007a; Wu et al., 2009). For example, McKenzie and Ryan (1999) used quantitative environmental variables including climate, parent material, normalized difference vegetation index (NDVI), etc. derived from a digital elevation model (DEM), Landsat TM, and airborne gamma radiometric remote sensing data to model TP and total carbon in Australia. They concluded that it is feasible to derive environmental correlations from extensive, fine grain remote sensing data to predict soil properties that was not possible with traditional methods (i.e., polygon-based methods). Wu et al. (2009) conducted and compared the prediction of soil organic matter (SOM) using geostatistical methods (Ordinary Kriging (OK) and Cokriging (CK)) with remote sensing data (Landsat ETM) for a 367 km^2 study area in China. In their research, CK outperformed OK indicating that remote sensing data have the potential to improve the accuracy and reliability of SOM prediction.

This chapter is focused on the regional assessment of Mehlich P (MP) and total phosphorus (TP) spatial distributions at four different depths in the Santa Fe River Ranch Beef Unit (SFRRBU). The overall objective was to map the spatial distribution of MP and TP from

site-specific soil samples using a mixed modeling approach that utilized SCORPAN and geostatistical methods with available exhaustive environmental datasets including land uses, soils, hydrological data, satellite images, and geological information.

In this chapter the following hypotheses were investigated: (a) multivariate or hybrid geospatial models (RK and CK) that incorporate exhaustive environmental datasets improve the prediction of MP and TP when compared to a univariate method (OK) that only uses soil property measurements. (b) A finer spatial resolution of remote sensing data (e.g. IKONOS and LIDAR) improves the accuracy of predicting MP and TP across the landscape.

With these hypotheses the specific objectives were to: (a) identify relationships between MP and TP, respectively, and environmental landscape properties in the SFRRBU. (b) investigate the correlation of MP and TP among different soil layers (0-30, 30-60, 60-120, and 120-180 cm) in the SFRRBU. (c) assess the efficacy of different spatial resolutions of remotely sensed data for modeling of MP and TP spatial distributions in the SFRRBU. (d) characterize the spatial patterns of MP and TP across the SFRRBU using a variety of geostatistical methods (OK, RK, and CK). (e) validate each geospatial interpolation method using an independent validation set and identify the best method that predicts MP and TP in the SFRRBU.

Materials and Methods

Soil Sampling, Laboratory Analysis, and Exhaustive Spatial Environmental Data

Soil sampling, the extraction procedures of MP and TP in this chapter were the same as in Chapter 4 and discussed in detail in Chapter 2. A comprehensive list of the environmental data used in the study is summarized in Table 2-3, Chapter 2.

Comparison of Soil Phosphorus at Four Depths

To analyze correlations among four depths in soil profiles Pearson's correlations were calculated on natural log-transformed MP data. The Mann-Whitney test was performed to

investigate how the median concentrations of MP and TP vary according to the different soil depths.

Supervised Land Use Classification from Landsat ETM+ and IKONOS Image

2003 Landsat ETM+ and 2006 IKONOS satellite images were used to derive land use classification information with assumption of unchanged land use over 10 years (Driver, 2007, personal communication). The IKONOS image was taken on April 28, 2006 and the image was geometrically corrected using 34 ground control points collected with the Trimble® GPS Pathfinder® Pro XR (Trimble Navigation Limited, Sunnyvale, CA). The image was re-projected to Albers Equal Area Conic map projection. Then the SFRRBU boundary was used to clip out the study area. Supervised classification was performed on the clipped image and 181 reference points were used to evaluate the accuracy of the land use classification with the image.

In case of the 2003 Landsat ETM+ imagery the image was taken on February 11, 2003 and the study area was clipped out similar to the IKONOS image. Since classification was done by the FFWCC in 2004 without accuracy assessment (Stys, 2009, personal communication) due to budget cut, the same 181 reference points used for the 2006 IKONOS image were used to evaluate the accuracy of classification on the 2003 Landsat ETM+ image. Producer's accuracy, user's accuracy, overall classification accuracy, and overall kappa statistics were calculated to compare the accuracy of classification of the two images.

Evaluation of the benefits of different spatial resolutions on prediction map of soil properties

To assess the efficacy of different spatial resolutions of remotely sensed data for modeling of MP and TP spatial distributions in the SFRRBU, various environmental variables were assembled. They included land use/land cover, topographic attributes (e.g., slope, aspect, catchment area and CTI), and spectral-derived information (e.g., vegetation indices, tasseled cap

indices, and principal components of the reflectance bands) and were divided into two resolution groups (i.e., the coarse-resolution group and the fine-resolution group). Variables named coarse-resolution (30 m) data were derived from the 2003 Landsat ETM+ image and DEM from NED, while variables called fine-resolution (4 m) data were derived from the 2006 IKONOS image and DEM from LIDAR. Coarse resolution datasets used in the SFRRBU were clipped from GIS layers used in the SFRW (compare previous chapter) because they were 30 m resolution and the SFRRBU is nested within the SFRW. Fine resolution datasets were processed at the GIS laboratory, Soil and Water Science Department at the University of Florida. These environmental variables were then used in RK/SMLR to model spatial distributions of MP and TP across the landscape and validated with an independent validation data to assess the effectiveness of fine and coarse resolution environmental data to predict MP and TP, respectively.

Quantify Relationships between Soil Phosphorus and Environmental Variables

To identify relationships between MP and TP and various environmental variables SLMR was performed. The measurements of soil MP and TP were considered the dependent variables and various environmental variables including information from satellite images (e.g., land use/land cover, NDVI, and reflectance from both 2003 Landsat ETM+ and 2006 IKONOS image), topographic indices (e.g., slope and aspect), soil information (e.g., soil order and soil drainage class), depth to water table, etc., were considered as predictor variables in SLMR. Table 2-3 shows a detailed list of the data used in the SFRRBU.

To test the significance of targeted predictors (i.e., land use, soil drainage, soil order, and geologic unit), one-way analysis of variance (ANOVA) was used in case of equal variance groups based on Levene's test ($p > 0.05$) or either one-way Welch's ANOVA (if the number of samples in each group was equal or greater than 10) or the Brown-Forsythe's ANOVA (if the

number of samples in each group was less than 10) (Myers and Well, 2003; Vasques et al., 2010). Classification of homogenous groups in land use and soil drainage was conducted by either Tukey's test for equal variance groups or Dunnett's T3 test for unequal variance groups (Dunnett, 1980).

Geostatistical Modeling of Soil Properties

Univariate (OK) and multivariate methods (RK and CK) were employed to model the spatial distribution of MP and TP across the SFRRBU. Due to the positively skewed distribution of sample data they were transformed to logarithm (base-e) and ordinary kriging (OK) was performed to predict $\log_e$MP (LNMP) at four depths (i.e., L1, L2, L3, and L4) and $\log_e$TP (LNTP) at L1 (Webster and Oliver, 2007). These LNMP and LNTP at each depth were back-transformed to the original unit (μg/g) after kriging. For RK SLMR was used as a regression model to identify the relationships between MP and various environmental variables and to map the global trend of MP across the landscape (Wang et al, 2009; Vasques et al., 2010). The residuals that were produced by subtracting the global trend from raw measurements were kriged by OK and added back to the global trend model to produce the final maps (Odeh et al., 1995). To predict soil TP the same RK-SLMR approach was utilized. Collocated CK was used since MP and TP were sparsely measured at some locations and other environmental variables, including remote sensing data and DEM derived data, were available everywhere within the study area (Goovaerts, 1997, 2000; Wackernagel, 2003; Rivero et al., 2009).

Note that OK used only MP and TP measurements with their spatial autocorrelations, while RK and CK used not only MP and TP measurements but also various auxiliary environmental variables to map the spatial distribution of MP and TP across the landscape.

Various software packages were used for this research including ArcGIS® 9.3 (Environmental Systems Research Institute, Redlands, CA) for displaying the final continuous

spatial distribution of MP and TP and SPSS® 17.0 (SPSS Inc., Chicago, IL) for statistical analysis including exploratory data analysis, correlation and significant tests, and SMLR. All geostatistial analysis was conducted in ISATIS® 8.02 (Geovariances, Avon, France). ERDAS IMAGINE® 9.1 (Leica Geosystems Geospatial Imaging, LLC, Norcross, GA) was used for focal analysis of raster data which is based on the attributes in a neighborhood (e.g. 3 x 3 or 7 x 7 moving window) across the landscape. Chapter 2 has more detailed description of geostatistical modeling.

Model Validation

For prediction and validation of each MP and TP map by LNK, RK and CK in the SFRRBU, the original sample data were randomly selected and divided into two groups: Calibration dataset (about 70%) and validation dataset (about 30%).

The ME, RMSE, and RPD were used to evaluate the performance of each prediction model and the equations are shown below in Equation 5-1, 5-2 and 5-3, respectively.

$$\text{ME} = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n} \tag{5-1}$$

$$\text{RMSE} = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)}{n}} \tag{5-2}$$

$$\text{RPD} = \frac{SD}{RMSE\sqrt{\frac{n}{(n-1)}}} \tag{5-3}$$

where y_i = measured values; $\hat{y}_i$ = predicted values; n = number of measured values with i = 1, 2,…., n.; SD = standard deviation.

Results and Discussion

Exploratory Data Analysis

Table 5-1 shows the descriptive statistics of raw and $\log_e$-transformed MP and TP data including mean, median, maximum, minimum, skewness, etc. Like many other environmental datasets (Kravchenko and Bullock, 1999; Goovaerts, 2001; Hengl et al., 2004 and 2007; Regalado and Ritter, 2006), the datasets at all depths had some very low and high values representing a positively skewed distribution (Figure 5-1). The logarithm transformation ($\log_e$) was used to approximate a normal distribution and stabilize variances. Total dataset of MP was ranging from a minimum of 0.75 µg/g at L3 to maximum of 937.16 µg/g at L4. Total phosphorus at L1 showed minimum of 69.38 µg/g and maximum of 3886.58 µg/g. Overall, MP measurements ranged from 3.14 µg/g to 754.34 µg/g, with a mean of 97.33 µg/g, and a median of 73.06 µg/g at L1, from 0.75 µg/g to 847.88 µg/g, with a mean of 81.28 µg/g, and a median of 57.75 µg/g at L2. The values of MP at L3 varied from 0.93 µg/g to 721.83 µg/g, with a mean of 99.71 µg/g, and a median of 70.54 µg/g, while those of MP at L4 ranged from 6.58 to 937.16 µg/g, with a mean of 122.72 µg/g, and a median of 76.39 µg/g. Lastly, TP dataset showed a range from 69.38 µg/g to 3886.58 µg/g, with a mean of 844.31 µg/g and a median of 651.27 µg/g at L1.

The highest mean MP was found at L4 (L4-MP; 122.72 µg/g) followed by L3 (L3-MP; 99.71 µg/g), L1 (L1-MP; 97.33 µg/g) and L2 (L2-MP; 81.28 µg/g). The mean MP among four depths showed a decrease from L1 to L2 and an increase from L2 to L3 and L4. Many soil samples that had low MP values were missing due to field condition (e.g., high water table) in the wetland forest at L4 and this might have caused the highest mean and median of MP in that layer. For all measured MP at four depths and TP at L1, the means of the dataset were much

greater than the medians of the dataset. This might reveal that a few very high values in the sample data caused the inflation of the mean.

Table 5.2 shows that LNMP was significantly correlated among all depths (L1 – L4) and between L1-LNMP and L1-LNTP at the 0.05 significance level. The highest correlation existed in between L1 and L2 (0.80) and the lowest correlation could be found in between L2 and L4 (0.40). The overall pattern of correlations at four depths showed that adjacent layers (i.e., L1 – L2, L2 – L3, and L3 – L4) had relatively higher correlations (0.80, 0.74 and 0.66, respectively) while the layers further from each other (i.e., L1 – L3, L1 – L4, and L2 – L4) had relatively lower correlations (0.67, 0.45 and 0.40, respectively). This indicated that vertical spatial distributions of MP at the top layer might influence more MP at the adjacent layer than layers further down the profile. This was consistent with the results in the SFRW from chapter 4.

However, despite the significant correlations found between MP in layers at all depths the median difference of MP among layers, except between L2 and L4, were not all significant at the 0.05 significance level by the Mann-Whitney test and showed a decreasing trend in layers at deeper depth (i.e., from L1 to L2 (0.24), to L3 (0.04) and to L4 (-0.04) and from L2 to L3 (-0.20), and to L4 (-0.28)).

Land Use Classification from Landsat and IKONOS Image

Table 5-3 shows the error matrix including reference and classified data used in each land use classification category, and the result of accuracy assessment for 2003 Landsat ETM+ supervised classification. According to producer's accuracy water had the highest accuracy (90%) followed by agriculture (87.0%), improved pasture (83.0%), rangeland (72.0 %), pineland forest (50%), upland forest (44.4%), urban (33.3%), and wetland forest (13.3%). On the other hand, user's accuracy was the highest in both water and urban (100%) and decreased in the following order: agriculture (80%), improved pasture (75.9%), upland forest (54.5%), rangeland

(50%), pineland (31.3%), and wetland forest (22.2%). The overall accuracy of forest classes (i.e., pineland forest, wetland forest and upland forest) was relatively poor (36.5%; 19 out of 52) compared to that (75.2%; 97 out of 129) of non-forest classes (i.e., agriculture, rangeland, urban, water, and improved pasture). This could be explained by the coarse resolution of Landsat ETM+ that could not distinguish specific forest classes within pixels and non-forest classes showed a relatively homogenous distribution within the study area (See Figure 2-10). For producer's accuracy, only 5 out of 10 (50%), 2 out of 15 (13.3%), and 12 out of 27 (44.4%) were accurately classified as pineland forest, wetland forest and upland forest, respectively, and these were lower than the overall classification accuracy (64.1%) of the 2003 Landsat ETM+ classification. This may be due to the fact that those forest classes had similar spectral signatures among them and 30 m resolution of the Landsat ETM+ image was too coarse to differentiate forest types in the study area. This could be also supported by the fact that one water reference point was classified as improved pasture. Despite of different spectral signatures between the two classes, spectral signatures from water were merged into those from improved pasture, because the extent of water was less than 30 meter and was close to improved pasture. In case of urban land use, producer's accuracy was only 33.3% implying that separating urban areas from agriculture, rangeland, or improved pasture was rather difficult because of spectral pattern from bare soil within those classes. Reese et al. (2002) indicated these classifications were quite difficult with a single image. More extensive ground truth data and multi-seasonal images could be used to improve accuracy of these land use classification in the study area. However, this was not the main objective of this research.

Table 5-4 is a summary of error matrix and accuracy assessment of the 2006 IKONOS supervised classification. Producer's accuracy of rangeland was the lowest (44.0%) and urban

(88.9%) and water (100%) were the highest, while user's accuracy of wetland forest was the lowest (55.6%). The overall forest classes achieved accuracy of 67.3% which was a significant improvement compared to the one (36.5%) from the 2003 Landsat ETM+ image, while non-forest classes achieved relatively similar accuracies (79.1% from the 2006 IKONOS image and 75.2% from the 2003 Landsat ETM+ image).

Table 5-5 shows the comparisons of areal extents of each land use from the 2003 Landsat ETM+ and the 2006 IKONOS images and land use classification were divided into two groups; one was 'forest' which included upland forest, wetland forest, and pine trees and the other was 'non-forest' which includes improved pasture, rangeland, agriculture, urban and barren soil, and water.

Based on areal extents classified from the 2003 Landsat ETM+ compared to the ones from the 2006 IKONOS image, upland forest, urban and barren soil, and agriculture area increased by 3.6 times (from 9.4 to 22.1%), 2.3 times (from 9.4 to 22.1%), and 1.8 times (from 7.8 to 13.7%), respectively, while pine trees, rangeland, and wetland forest decreased by 80% (from 5.0 to 1.0 %), 80% (from 12.7 to 2.9%), and 20% (from 24.6 to 18.5%), respectively. Only improved pasture and water area remained relatively unchanged (i.e., from 39.1 to 38.8 % and from 0.5 to 0.3%, respectively).

Even though variations were detected for each land use classified, the overall areal extents of forest and non-forest category classification between the 2003 Landsat ETM+ and the 2006 IKONOS images were rather similar by 6% increase (from 39.0 to 41.5%) and 4% decrease (from 61.0 to 58.5%), respectively.

The overall classification accuracy of the forest group from the 2003 Landsat ETM+ image was only 36.5% (19 correct out of 52 reference points) which was relatively poor compared to

67.3% (35 correct out of 52 reference points) from the 2006 IKONOS image, while the non-forest group had a comparable accuracy with 75.2% (97 correct out of 129 reference points) from the 2003 Landsat ETM+ image compared to the accuracy of 79.1% (102 correct out of 129 reference points) from the 2006 IKONOS image. One may conclude that because of the finer spatial resolution the 2006 IKONOS image could be useful to differentiate the different forest groups (upland forest, wetland forest, and pine trees) despite of the similarity of their spectral responses but did not show significant improvement on classifying non-forest classes (improved pasture, rangeland, agriculture, and urban and barren soil) due to their distinct spectral differences.

Hyperspectral remote sensing data that provide more detailed spectral information and higher horizontal accuracy of reference points in the forest may be desired to increase accuracy of forest classification. Two or more (high-resolution) satellite images of different seasons, if possible, may also reduce seasonal variations so that they may help to increase accuracy of forest or other land uses classification.

The user's accuracy showed significantly improved classification using the 2006 IKONOS image by as high as 3.2 times (from 31.1 to 100 %), 2.5 times (from 22.2 to 55.6%), and 1.6 times (from 50.0 to 78.6%) inferring on pineland, wetland forest, and rangeland, respectively, when compared to the 2003 Landsat ETM+ image. In contrast, all other land uses (upland forest, agriculture, urban, improved pasture, and water) showed less than 30% differences in user's accuracy among the two images.

For producer's accuracy the 2006 IKONOS image showed significant improvement to classify wetland forest, urban, and upland forest as high as 5 times (from 13.3 to 66.7%), 2.7 times (from 33.3 to 88.9%), and 1.6 times (from 44.4 to 70.4%), respectively, when compared to

the 2003 Landsat ETM+. In cases of wetland forest, urban, and upland forest, finer resolution from the 2006 IKONOS image was small enough to differentiate objects within pixels to capture different spectral responses from them, while the resolution of the 2003 Landsat ETM+ image was so coarse that different spectral responses from those objects within the pixels were smoothed out and no differences were captured.

On the other hand, the producer's accuracy of rangeland from the 2003 Landsat ETM+ image was decreased by 39% (from 72 to 44%) compared to the producer's accuracy of rangeland from the 2006 IKONOS classification. This may be due to the fact that the 2006 IKONOS image taken in the spring season (April 26th) allowed capturing large amounts of grass growth in rangeland. In contrast, the 2003 Landsat ETM+ was taken February 11th causing more reference points of rangeland to be classified as improved pasture due to the reduced ability to spectrally discriminate during the winter season when grass growth is limited.

To summarize, the classification accuracy for the 2003 Landsat ETM+ was poorer, in terms of overall classification accuracy and kappa statistics, when compared to the classification for the 2006 IKONOS image (Table 5-3 and Table 5-4). It showed that there were distinct differences between the two images to discriminate forest types based on poorer producer's and user's accuracy.

Evaluation of the benefits to use different spatial resolutions for soil maps

Regression kriging with SMLR with land use, topographic, and spectral-derived information from two different spatial resolutions (coarse and fine data) were performed to compare their soil P prediction performances in the SFRRBU. Table 5-6 shows goodness-of-fit statistics of the SMLR and error statistics including RMSE, ME, and RPD of RK/SMLR derived from coarse and fine datasets. The global trend models of SMLR based on the calibration dataset for coarse data explained 36, 29, 20, and 10% of the variability of LNMP at L1, L2, L3,

and L4, respectively, and 32% of L1-LNTP. According to the R^2 from the calibration dataset for fine data, 34% (L1-LNMP), 27% (L2-LNMP), 12% (L3-LNMP), and 13% (L4-LNMP) of the variability were explained by the global trend models of SMLR. The higher R^2 values obtained in SMLR from the validation datasets at all depths using in the coarse data indicate that coarse data captured more variability of soil P compared to the fine data. The RMSEs were 134.2, 119.1, 103.3, 128.0, and 634.3 µg/g on coarse sets of validation data and 142.2, 131.9, 112.4, 128.4, 644.8 µg/g were obtained with fine data for L1-MP, L2-MP, L3-MP, L4-MP and L1-TP, respectively. These statistics indicate that fine resolution data did not provide an advantage to predict the distribution of soil P in the study area. Fine data contained variation in soil and environmental properties (e.g. spectral data) which the model could not utilize to improve predictions of soil P in the SFRRBU. Instead, the model probably perceived this fine-resolution variation (e.g. tree crowns, shades, and little areas of exposed soils) as random noise. In effect, land use in the study area were the important sources of spatial variation in soil P, and the fine resolution sensors were too detailed to represent this P variation at the proper scale that it occurred. On the other hand, coarse data better captured the landscape variation at a scale compatible to soil P patterns, averaging out within-pixel variation, otherwise perceived as random noise at finer resolution. In other words, coarse data had a smoothing effect within the pixel that better captured soil P global trends within the study area, where a relatively homogenous distribution of land uses exists.

Semivariograms on the residuals of the global trend derived by SMLR were produced to assess the spatial structures of soil MP and TP in the study area and various parameters including best fitted models, nugget, sill, and range are summarized in Table 5-7 and depicted in Figure 5-2. Spherical models were the best fitted models for all semivariograms of coarse and fine data.

Compared to nuggets at each depth in LNK, relatively large nuggets were found in all residual soil MP and TP semivariograms, except for L1-LNTP in coarse data. These findings are not unusual since large nuggets are common in residual semivariograms of soil-environmental properties (Hengel et al., 2004; Rivero et al., 2007a; Vasques et al., 2010).

Compared to sills from LNK (Table 5-8), sills between coarse/fine data of residuals derived from SMLR (Table 5-7) showed a decreased pattern at L1-LNMP and L2-LNMP in both coarse and fine data and L4-LNMP in fine data, while an increased trend was found at L3-LNMP and L1-LNTP in both data and L4-LNMP in coarse data. A decreased sill could be due to the fact that variances were reduced in the residuals at those layers. Nugget/sill ratios of soil P residuals varied between 36.8 and 58.0 for the coarse data and between 44.3 and 64.2 for the fine data (L1-LNMP to L4-LNMP). At L1-LNTP the ratio was very low with 19.3 (coarse data) and 47.3 (fine data). The spatial autocorrelation (range) in soil P residuals varied between 279 and 770 m (coarse data) and 378 and 1318 m (fine data) (Table 5-7), which were much shorter compared to the ranges derived from soil P observations with lengths between 561 and 1689 m (Table 5-8). These findings suggest that the more complex model (RK) which incorporates modeling of the non-spatial (SMLR) and spatial (LNK of residuals) structures has superior ability to describe soil variation, when compared to simple LNK. Compared with ranges of LNK (Table 5-8), all ranges of residuals, except for L3-LNMP, were decreased after the trend removed (Table 5-7).

Figure 5-3 depicts prediction maps of soil properties derived by RK/SMLR with coarse and fine data in the SFRRBU. Overall, the main factors in coarse and fine data influencing the spatial distributions of soil P by SMLR were topographic (e.g., aspect and slope), spectral (e.g., reflectance of band 5 from IKONOS) and land use information (e.g., upland forest and improved

pasture) (Table 5-9). Even though these different environmental variables were selected as predictor variables in SMLR both maps produced by RK/SMLR with coarse and fine datasets showed a similar overall pattern of spatial distributions of soil properties. Lower concentrations of soil P were found in the northern part of the study area (mixture of wetland forest and upland forest) along with upland forest in the southeast, while higher soil P concentrations occurred along with higher slope areas. This may suggest that fine data were not necessary to understand the overall patterns of soil P distributions in the SFRRBU as coarse data also showed similar spatial patterns.

Other research on the effect of spatial resolutions on modeling of soil properties has provided similar findings. For instance, Wu et al. (2008) verified the effect of different resolutions of DEM (4, 6, 8, 10, 20, and 30 m) on the relationships between soil properties (e.g., pH, OM, Ca, P, etc.) and terrain attributes (e.g., elevation, slope, aspect, and topographic wetness index (TWI)) in hilly landscapes (100 ha in size) in southwestern China. They pointed out that the most accurate prediction model is not always created by the highest resolution DEM. They stated that as DEM cell size increased important local landscape features were lost and this might lead to inaccurate results in a given area, and thus, an appropriate DEM resolution should be wisely chosen for a particular landscape. Thompson et al. (2001) compared four soil-landscape models to predict A horizon depth (AHD) with 143 soil samples in west central Minnesota. Each model used variables derived from 10 or 30 m DEMs with 0.1 or 1 m vertical precisions from either USGS or field survey and had similar capabilities in predicting AHD (adjusted R^2 of models ranged from 0.42 to 0.59). They suggested that higher resolution DEMs may not be necessary for generating useful soil-landscape models in their study area.

The results of higher $RMSE_v$ by fine data and a similar overall pattern of predicted maps derived by coarse and fine data may suggest that fine data did not provide an advantage for soil properties mapping in the SFRRBU because it may generally take more resources to obtain and process finer data. However, this conclusion may not be assumed for other areas since this is a site-, soil attribute-, and sample design- specific result. Li (2010) also supported this conclusion by stating that "the influence of the spatial resolution of RS (remote sensing) data on successful RK applications is unclear or is mixed."

Relationships between Soil Properties and Environmental Variables

Significances of land use, soil order, and soil drainage on LNMP and LNTP were tested and homogeneous group with their respective means in each category are presented in Table 5-10. Note that the influence of geological unit was not tested because the entire study area is located on the same geological unit (Coosawhatchie formation). Leven's test showed that L1-LNMP and L1-LNTP in land use (p-value = 0.000) and L1-LNMP in soil drainage (p-value = 0.000) had unequal variances, so Brown-Forsythe's ANOVA (rangeland and excessively drained class < 10) and Dunnett's T3 test were used to differentiate homogenous groups according to group mean at the 0.05 significance level. However, L1-LNMP and L1-LNTP based on soil order were not significantly different at the 0.05 significance level.

Note that pineland forest and the urban class land uses had only two samples in each category and may not truly represent the concentrations of MP and TP in the study area, and thus, they were removed before the test was run. Dunnett's T3 test on land use indicated that agriculture had the highest mean L1-LNMP (4.55 $\log_e$ µg/g) and decreased in the following order: improved pasture (4.42 $\log_e$ µg/g), rangeland (4.40 $\log_e$ µg/g), upland forest (3.60 $\log_e$ µg/g), and wetland forest (3.30 $\log_e$ µg/g). Based on Table 5-8, L1-LNMP in agriculture and improved pasture was statistically higher than L1-LNMP in upland forest and wetland forest.

The highest mean of L1-LNTP was found in agriculture (6.72 $\log_e$ µg/g) followed by improved pasture (6.67 $\log_e$ µg/g), rangeland (6.59 $\log_e$ µg/g), upland forest (6.11 $\log_e$ µg/g), and wetland forest (5.95 $\log_e$ µg/g). It showed that only wetland forest had significantly lower L1-LNTP compared to all other land use categories in the study area. Excessive usage of fertilizer and animal waste could be responsible for higher concentrations of MP and TP in the SFRBBU and this supports the conclusion that land use is an important factor to control the spatial distributions of MP and TP as in the SFRW from Chapter 4.

Influence of soil drainage on L1-LNMP and L1-LNTP was also tested. The highest group mean for L1-LNMP occurred in excessively drained soils (4.62 $\log_e$ µg/g) and decreased in the following order: moderately well drained (4.53 $\log_e$ µg/g), somewhat poorly drained (4.43 $\log_e$ µg/g), well drained (4.42 $\log_e$ µg/g), and poorly drained (3.19 $\log_e$ µg/g), while L1-LNTP showed the highest group mean in moderately well drained (6.75 $\log_e$ µg/g) followed by well drained (6.73 $\log_e$ µg/g), somewhat poorly drained (6.62 $\log_e$ µg/g), excessively drained (6.42 $\log_e$ µg/g), and poorly drained (5.79 $\log_e$ µg/g) soils. However, poorly drained soils showed statistically lower group means of LNMP and LNTP compared to other soil drainage classes.

Geostatistical Modeling of Soil Properties and Model Validation

Stepwise multiple linear regression was performed with various environmental variables including soil order, soil drainage, satellite derived information, etc. to model the global trend. Several environmental variables were selected by SMLR as important factors to relate to soil P (MP and TP) in the study area and results are summarized in Table 5-11. The global trend models derived by SMLR with coarse data including land use, soil order, soil drainage, etc. explained 63, 52, 55, and 40% of the variability of LNMP at L1, L2, L3 and L4, respectively, and 58% in case of L1-LNTP in model calibration mode.

The DRASTIC index was selected at all layers of LNMP including L1-LNTP and negatively impacted the prediction of soil P. The DRASTIC Index, the empirical index created by the USEPA, is an acronym based on depth to water, (net) recharge, aquifer media, soil media, topography (slope), impact of the vadose zone media, and hydraulic conductivity of the aquifer to standardize evaluation on groundwater pollution potential based on hydrogeological setting (Babiker et al., 2005). However, it may not be straightforward to relate this index to soil properties since seven parameters contributed to the index to represent vulnerability of potential groundwater pollution. But SMLR indicates that the DRASTIC index is a controlling factor for P distribution.

Land use/land cover (specifically rangeland and upland forest classes) at all depths, except for L3, influenced the prediction of LNMP either negatively or positively. In contrast, Ultisols influenced positively the soil P models at all depths. This suggests that land use and soil order are also important factors to explain the spatial distribution of soil P in the SFRRBU, confirming the findings presented in Chapter 4 in the SFRW. The distributions of MP and TP can be clearly related to geomorphological structures (e.g. slope, soil order, soil drainage, geological formation) as well as to anthropological factors (e.g. land use).

As seen in Chapter 4, several variables selected as impacting factors were related to soil wet/dry condition. For example, soil available water capacity, sand content in soil, bulk density, slope percent, hydrologic groups B and D, and the well drained soil class could be related to soil drainage condition.

Table 5-12 shows the most highly correlated environmental variables with LNMP at each depth and L1-LNTP at the 0.01 significance level. They are DTW (Spearman correlation coefficient = 0.33 at L1-LNMP, 0.28 at L2-LNMP, 0.25 at L3-LNMP, and 0.35 at L1-LNTP)

and slope derived from NED had a correlation coefficient of 0.31 at L4-LNMP. These variables were used as secondary variables to predict soil P at each depth by CK.

Error statistics such as RMSE, ME, and RPD calculated on calibration and validation data for LNK, RK/SMLR, and CK are summarized in Table 5-13. Based on values of $RMSE_v$ and RPD LNK outperformed RK/SMLR and CK at all depths for both MP and TP in the SFRRBU, while CK at L1-MP and L4-MP outperformed other methods based on ME_v. The fact that LNK outperformed RK/SMLR and CK in the study area somewhat contradicts previous research (McBratney et al. 2000; Mueller and Pierce, 2003; Hengl et al., 2004; Verfaillie et al., 2006; Vasques et al., 2010) as well as Chapter 4 but several possible explanations can be offered.

Firstly, spatial dependence (autocorrelation) of MP and TP in the model of LNK could better represent the distributions of these soil properties than models derived from ancillary environmental variables that link to soil properties in RK/SMLR and CK. In other words, the environmental covariance that related to soil MP and TP in RK/SMLR and CK across the landscape was expected to help explaining better the variability of MP and TP. But it turned out that the spatial dependence structures of soil MP and TP resulted in stronger models (LNK) than the ones derived through environmental correlation (RK/SMLR and CK).

Secondly, the backtransformation from logarithmic unit to original unit after SMLR modeling and kriging of the residuals might have expanded errors in the process of RK/SMLR and produced larger $RMSE_v$ and ME_v.

Thirdly, sensors that capture the scale (i.e., the grain resolution) at which environmental variables in the model were presented did not match the scale at which MP and TP behaved. In other words, the pixel resolution of investigated fine and coarse scale environmental datasets did not allow developing models with lower errors or larger predictive capabilities. The scales of

environmental variables were either too fine or too coarse to capture the behavior or variability of soil properties across the landscape indicating that the optimal resolution of environmental variables to predict MP and TP lies in between the selected coarse and fine resolution data or a much coarser or finer resolution is necessary to develop the best prediction models.

Fourthly, the quality of environmental variables used in SMLR and CK for modeling may not be accurate enough to capture the variability of soil properties in the study area. For example, the overall accuracy of land use classification derived from 2003 Landsat ETM+ was only 64.1% in the SFRRBU.

Lastly, the positional accuracy of fine and coarse scale environmental data may have contributed to shifts when superimposing them which may have led to mismatches between spatially explicit soil P observations and spatial environmental data. To assess the error introduced through positional displacement is beyond this research because of the mixed data sources incorporated in models which included satellite imagery, DEMs, and environmental grids created by various agencies.

Other research studies investigating scale effects on modeling of soil properties are rare. For example, Penížek and Borůvka (2006) investigated the advantages of methods incorporating terrain attributes into the prediction of soil depth with 603 soil profiles in an area of 1,327 km^2 in Southern Bohemia. The terrain attributes included elevation, slope, and aspect and the methods compared were OK, CK, RK, and linear regression (RE).

They found that OK and CK performed the best (110 and 110.3 cm, respectively) and RK and RE performed the worst (108.4 and 79 cm, respectively) to predict the mean observed soil depths (110 cm) within the study area. The poorer performance by RE and RK was caused by weak correlations between the soil depth data and terrain attributes used in the model (R^2 = 0.186

with slope, 0.246 with aspect, and 0.102 with elevation). Another reason could be the fact that if areas in steep slopes were overlooked during sampling, extreme biased values might have shown in the prediction by RK. Nour et al. (2006) also compared various kriging methods for the estimation of daily rainfall in the Canadian Boreal forest to suggest potential sites for the installation of further weather stations. Geostatistical methods in their research included simple kriging (SK), OK, and LNK as univariate methods and kriging with external drift (KED) using elevation and slope (KED-ELEV and KED-SLOPE, respectively) as multivariate methods. They achieved RMSE of 4.90, 4.78, 4.81, 4.96, 5.74 mm for SK, OK, LNK, KED-ELEV and KED-SLOPE, respectively, and concluded that SK, OK and LNK were comparative methods with respect to prediction performance and adding secondary variables did not improve predictions in their research. Relatively mild slope of the study area and weak correlation coefficients of rainfall/elevation and rainfall/slope (< 0.2) could be offered to explain this result. Favorable sites were proposed based on the areas with high kriging variance, accessibility, land use/land cover, and road networking. Li (2010) performed OK, universal kriging (UK), and RK to predict SOM and evaluated RK's capability for improving SOM prediction in China using external auxiliary information including soil map (SOIL), vegetation indices (VI) derived from a Landsat 5 TM image, and various topographic information (elevation, slope, wetness indices, etc.). The results showed that OK and UK produced equal or better prediction maps (RMSE: 0.61 and 0.62, respectively) than RK did (RMSE: 0.62~0.67 depending on auxiliary data added in the model). By adding SOIL and VI information RK significantly deteriorated the SOM prediction (RMSE of 0.67) indicating that inclusion of more information does not always significantly improve prediction performance. The author claimed that RK technique might not work in the area where soil samples bear very strong spatial autocorrelations or the linear regression model assumption

in RK is not appropriate. If this was the case, a non-linear regression method (e.g., regression tree) might have revealed better interrelationships between SOM and external auxiliary information, and thus, the model might be improved.

Lognormal kriging produced generally smoother maps compared to RK/SMLR and CK in the study area. This was an expected result since LNK minimized the error variance by least squares and the variations of soil P values were smoothed out. These smoothing effects were presented in the predicted maps at all depths (Figure 5-4). Overall, lower predicted concentrations of MP and TP in soils were found in the northern area where a mixture of upland and wetland forest exists, while higher predicted concentrations of MP and TP were found in the northwestern and central part where cattle grazing and resting under shadows in improved pasture predominates. In addition, there is a nursery located near by the main entrance of the west side of the study area and leaching or debris of fertilization from the nursery could be transported by surface runoff into the stream and carried into the SFR on the northern side of the SFRRBU. All predicted maps show the higher concentrations of MP and TP transgressing from west to the central portion, and somewhat in the southern part of the study area. These hotspots high in soil P may impair groundwater if proper attention is not placed.

Due to the higher correlation (Pearson correlation coefficient = 0.89) between L1-LNMP and L1-LNTP in the SFRRBU compared to the correlation coefficient of 0.60 in the SWFR, overall distributions of these soil properties showed somewhat similar patterns even though magnitudes of concentrations were different. This indicates that the map of L1-LNMP may resemble the general distribution pattern of L1-LNTP so that more time- and labor- consuming extraction of TP might be avoided for an estimation of rough trends for TP in the study area.

As conclusions of this chapter, 1) pair-wise comparisons of LNMP among the four depths and L1-LNMP and L1-LNTP, 2) comparison of land use classification with Landsat ETM+ and IKONOS images, 3) relationships between MP/TP and selected environmental variables (land use, soil drainage, and soil order), 4) benefits and limitations to use different spatial resolution data for soil mapping, and 5) geostatistical modeling of MP at four depths and TP at L1 in the SFRRBU were investigated.

LNMP was significantly correlated among all depths and between L1-LNMP and L1-LNTP at the 0.05 significance level. The overall pattern of correlations at four depths showed that adjacent layers (i.e., L1- L2, L2- L3 and L3-L4) had relatively higher correlations (0.80, 0.74, and 0.66, respectively) and the layers further from each other (i.e., L1-L3, L1-L4, and L2-L4) had relatively lower correlations (0.67, 0.45, and 0.40, respectively). This suggests that vertical transport process of MP influences the spatial distribution from upper layer to adjacent layers.

The overall classification accuracy was 64.1% and kappa statistic was 0.47 for the 2003 Landsat ETM+ image, while 75.7% and 0.71 were achieved for the overall classification and kappa statistic, respectively, from the 2006 IKONOS image. Each land use classified showed some variations of classification accuracy; however, the overall areal extents of forest and non-forest category classification between the two images were somewhat similar (39.0 vs. 41.5% and 61.0 vs. 58.5%, respectively).

Dunnett's T3 test on land use indicated that agriculture had the highest mean L1-LNMP (4.55 $\log_e$ µg/g) and decreased in the following order: improved pasture (4.42 $\log_e$ µg/g), rangeland (4.40 $\log_e$ µg/g), upland forest (3.60 $\log_e$ µg/g), and wetland forest (3.30 $\log_e$ µg/g). Influence of soil drainage on L1-LNMP and L1-LNTP also tested. The highest group mean for

L1-LNMP existed in excessively drained soils (4.62 $\log_e$ µg/g) and decreased in the following order: moderately well drained (4.53 $\log_e$ µg/g), somewhat poorly drained (4.43 $\log_e$ µg/g), well drained (4.42 $\log_e$ µg/g) and poorly drained soils (3.19 $\log_e$ µg/g), while L1-LNTP showed the highest group mean in moderately well drained soils (6.75 $\log_e$ µg/g) followed by well drained (6.73 $\log_e$ µg/g), somewhat poorly drained (6.62 $\log_e$ µg/g), excessively drained (6.42 $\log_e$ µg/g), and poorly drained (5.79 $\log_e$ µg/g) soils. These indicate that anthropological and geomorphological factors impacted the distributions of MP and TP.

The results of higher $RMSE_v$ from fine data and similar overall pattern of predicted maps derived from coarse and fine data suggest that fine data did not provide an advantage for soil properties mapping in the SFRRBU because it may generally take more resources to obtain and process finer data. This result rejects the hypothesis that a finer spatial resolution of remote sensing data (e.g. IKONOS and LIDAR) improves the accuracy of predicting MP and TP across the landscape.

Based on $RMSE_v$ and RPD the best prediction methods of MP and TP were LNK at all depths in the study area and rejecting the hypothesis that multivariate or hybrid geospatial models (RK and CK) that incorporate exhaustive environmental datasets improve the prediction of MP and TP when compared to a univariate method (LNK) that only uses soil property measurements. This specifies that incorporating the spatial covariation between soil properties and auxiliary environmental properties did not provide more accurate predictions than raw measurements of soil samples alone in the SFRRBU.

As Parry (1998) stated that soil contamination may also lead to underground/surface water contamination as pollutants from soils and surface water runoff leach into an aquifer and/or well field. To address the quality and quantity of water various monitoring programs and regulations

have been established for optimizing land use activities and minimizing environmental impact. The EPA water quality criteria state that P should not exceed 0.05 mg/L if streams discharge into lakes or reservoirs, 0.025 mg/L within a lake or reservoir, and 0.1 mg/L in streams or flowing waters not discharging into lakes or reservoirs to control alga growth (USEPA, 1986). Stoner (2011) mentioned that N and P pollution has the potential to become one of the costliest and the most challenging environmental problems based on the facts that 50% of U.S. streams have medium to high levels of N and P and 78% of assessed coastal waters exhibit eutrophication. The United States has adopted many management practices to response the concern like Stoner's by reducing P loss from agriculture. For example, within the Okeechobee Basin, Florida, agricultural BMPs are implemented for over 30 years and significant amount of P imports from fertilizer and cattle feed supplements were decreased (Dunne et al., 2007). Anderson and Flaig (1995) and Evans and Corrandini (2001) suggested other BMPs used to reduce and mitigate P loss include cover crops, crop residue management, crop rotations, improved nutrient/waste management systems, pasture land management, and the use of constructed and natural wetlands.

The results achieved from this research can be used as a preliminary assessment tool to evaluate the effectiveness of these BMPs if water quality data and soil distribution maps are linked together.

Table 5-1. Descriptive statistics of Mehlich phosphorus and total phosphorus data at each layer in the Santa Fe River Ranch Beef Unit.

Statistics	Total set	Calibration	Validation	Total set	Calibration	Validation
	L1-MP (µg/g) 0 – 30 cm			L1-LNMP (log_e µg/g) 0 – 30 cm		
Observations	145	101	44	145	101	44
Mean	97.33	91.50	110.69	4.11	4.10	4.13
Median	73.06	71.73	79.96	4.29	4.27	4.39
Std. deviation	102.32	84.06	135.51	1.10	1.05	1.21
Kurtosis	14.89	10.20	12.20	0.50	0.64	0.33
Skewness	3.23	2.54	3.17	-0.77	-0.88	-0.63
Minimum	3.14	3.43	3.14	1.14	1.23	1.14
Maximum	754.34	570.66	754.34	6.63	6.35	6.63
Range	751.20	567.23	751.20	5.49	5.12	5.49
	L2-MP (µg/g) 30 – 60 cm			L2-LNMP (log_e µg/g) 30 – 60 cm		
Observations	145	101	44	145	101	44
Mean	81.28	78.94	86.65	3.89	3.92	3.82
Median	57.75	58.19	52.30	4.06	4.06	3.96
Std. deviation	93.31	71.13	131.71	1.18	1.16	1.25
Kurtosis	32.16	6.14	26.77	2.59	3.38	1.64
Skewness	4.58	2.16	4.73	-1.25	-1.53	-0.73
Minimum	0.75	0.75	0.96	-0.29	-0.29	-0.04
Maximum	847.88	410.34	847.88	6.74	6.02	6.74
Range	847.13	409.59	846.92	7.03	6.31	6.78
	L3-MP (µg/g) 60 – 120 cm			L3-LNMP (log_e µg/g) 60 – 120 cm		
Observations	132	95	37	132	95	37
Mean	99.71	92.88	117.25	4.10	4.06	4.20
Median	70.54	71.15	63.42	4.26	4.27	4.15
Std. deviation	112.18	95.51	146.75	1.08	1.08	1.08
Kurtosis	9.40	7.23	7.72	1.07	1.57	-0.35
Skewness	2.72	2.41	2.59	-0.51	-0.76	0.14
Minimum	0.93	0.93	7.15	-0.07	-0.07	1.97
Maximum	721.83	507.09	721.83	6.58	6.23	6.58
Range	720.90	506.16	714.68	6.65	6.30	4.61
	L4-MP (µg/g) 120 – 180 cm			L4-LNMP (log_e µg/g) 120 – 180 cm		
Observations	114	80	34	114	80	34
Mean	122.72	117.12	135.92	4.34	4.38	4.24
Median	76.39	76.71	74.57	4.34	4.34	4.32
Std. deviation	143.80	122.20	186.58	1.02	0.93	1.23
Kurtosis	14.82	17.85	10.38	0.35	0.69	-0.22

Table 5-1. Continued.

Statistics	Total set	Calibration	Validation	Total set	Calibration	Validation
	L4-MP (µg/g) 120 – 180 cm			L4-LNMP (log_e µg/g) 120 – 180 cm		
Skewness	3.40	3.52	3.00	-0.36	-0.45	-0.17
Minimum	6.58	7.36	6.58	1.88	2.00	1.88
Maximum	937.16	869.43	937.16	6.84	6.77	6.84
Range	937.16	869.43	930.58	4.96	4.77	4.96
	L1-TP (µg/g) 0 – 30 cm			L1-LNTP (log_e µg/g) 0 – 30 cm		
Observations	145	101	44	145	101	44
Mean	844.31	819.64	900.93	6.44	6.43	6.46
Median	651.27	586.49	784.73	6.49	6.49	0.89
Std. deviation	656.95	657.64	622.09	0.83	0.80	6.44
Kurtosis	2.98	0.33	4.18	-0.07	-0.20	0.20
Skewness	1.51	1.03	1.86	-0.42	-0.46	-0.37
Minimum	69.38	79.27	69.38	4.24	4.37	4.24
Maximum	3886.58	2531.69	3886.58	8.27	7.84	8.27
Range	3817.20	2452.42	3817.20	4.03	3.47	4.03

Abbreviations: L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); TP = total phosphorus (µg/g).

Table 5-2. Pair-wise comparison of log_e-transformed Mehlich phosphorus (MP) in µg/g among the different four depths and total phosphorus (TP) in µg/g at top layer in the Santa Fe River Ranch Beef Unit.

Pair	Number of samples	Pearson Correlation	Mann-Whitney Test Median difference (log_e µg/g)
L1-LNMP – L2-LNMP	145	0.80*	0.24
L1-LNMP – L3-LNMP	128	0.67*	0.04
L1-LNMP – L4-LNMP	110	0.45*	-0.04
L2-LNMP – L3-LNMP	128	0.74*	-0.20
L2-LNMP – L4-LNMP	110	0.40*	-0.28*
L3-LNMP – L4-LNMP	111	0.66*	-0.08
L1-LNMP – L1-LNTP	145	0.89*	-2.20*

Abbreviations: L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); TP = total phosphorous (µg/g).

* Statistically significant at the 0.05 significance level.

Table 5-3. Error matrix and accuracy assessment of the 2003 Landsat ETM+ supervised classification in the Santa Fe River Ranch Beef Unit.

ERROR MATRIX

Classified Data	Reference Data								
	Pineland forest	Wetland forest	Upland forest	Agriculture	Range land	Urban	Water	Improved pasture	Row Total
Pineland forest	5	3	3	0	2	1	0	2	16
Wetland forest	0	2	6	0	1	0	0	0	9
Upland forest	0	5	12	0	2	0	0	3	22
Agriculture	0	0	0	20	0	3	0	2	25
Range land	5	3	5	0	18	3	0	2	36
Urban	0	0	0	0	0	6	0	0	6
Water	0	0	0	0	0	0	9	0	9
Improved pasture	0	2	1	3	2	5	1	44	58
Column Total	10	15	27	23	25	18	10	53	181

ACCURACY ASSESSMENT

Class name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy (%)	User's Accuracy (%)
Pineland	10	16	5	50.0	31.3
Wetland forest	15	9	2	13.3	22.2
Upland forest	27	22	12	44.4	54.5
Agriculture	23	25	20	87.0	80.0
Rangeland	25	36	18	72.0	50.0
Urban	18	6	6	33.3	100.0
Water	10	9	9	90.0	100.0
Improved pasture	53	58	44	83.0	75.9
Total	181	181	116		

Overall Classification Accuracy = 64.1%

Overall kappa statistics = 0.47

Table 5-4. Error matrix and accuracy assessment of the 2006 IKONOS supervised classification in the Santa Fe River Ranch Beef Unit.

ERROR MATRIX

Classified Data	Reference Data								
	Pineland forest	Wetland forest	Upland forest	Agriculture	Range land	Urban	Water	Improved pasture	Row Total
Pineland forest	6	0	0	0	0	0	0	0	6
Wetland forest	1	10	5	0	2	0	0	0	18
Upland forest	1	4	19	1	1	0	0	0	26
Agriculture	0	0	0	18	4	0	0	3	25
Range land	0	0	1	0	11	0	0	2	14
Urban	1	0	0	1	0	16	0	1	19
Water	0	0	0	0	0	0	10	0	10
Improved pasture	1	1	2	3	7	2	0	47	63
Column Total	10	15	27	23	25	18	10	53	181

ACCURACY ASSESSMENT

Class name	Reference Totals	Classified Totals	Number Correct	Producer's Accuracy (%)	User's Accuracy (%)
Pineland forest	10	6	6	60.0	100
Wetland forest	15	18	10	66.7	55.6
Upland forest	27	26	19	70.4	73.1
Agriculture	23	25	18	78.3	72.0
Rangeland	25	19	11	44.0	57.9
Urban	18	19	16	88.9	84.2
Water	10	10	10	100.0	100.0
Improved pasture	53	63	47	88.7	74.6
Total	181	181	137		

Overall Classification Accuracy = 75.7%

Overall kappa statistics = 0.71

Table 5-5. Comparisons of areal extent based on the 2003 Landsat ETM+ and the 2006 IKONOS supervised classification in the Santa Fe River Ranch Beef Unit.

Land use		Area (%) from Landsat	Area (%) from IKONOS	Total Forest area (%) from Landsat	Total Non-Forest area (%) from IKONOS
Forest	Upland	9.4	22.1	39.0	41.5
	Wetland	24.6	18.5		
	Pineland	5.0	1.0		
Non-Forest	Improved pasture	39.1	38.8	61.0	58.5
	Rangeland	12.7	2.9		
	Agriculture	7.8	13.7		
	Urban and barren soil	0.8	2.8		
	Water	0.5	0.3		
Total		100	100	100	100

Table 5-6. Goodness-of-fit statistics of the stepwise multiple linear regression models (SMLR) and error statistics of regression kriging with SMLR using coarse (30 m) and fine resolution (4 m) dataset in the Santa Fe River Ranch Beef Unit.

		Coarse resolution [1]					Fine resolution [2]				
SMLR*		L1-LNMP	L2-LNMP	L3-LNMP	L4-LNMP	L1-LNTP	L1-LNMP	L2-LNMP	L3-LNMP	L4-LNMP	L1-LNTP
Cal.	R^2	0.36	0.29	0.20	0.10	0.32	0.34	0.27	0.12	0.13	0.38
Val.		*0.46*	*0.42*	*0.28*	*0.17*	*0.45*	0.17	0.40	0.01	0.13	0.07

		Coarse resolution					Fine resolution				
RK/SMLR		L1-MP	L2-MP	L3-MP	L4-MP	L1-TP	L1-MP	L2-MP	L3-MP	L4-MP	L1-TP
	RMSE	*80.7*	73.7	*100.4*	*127.9*	*536.3*	84.8	*70.6*	100.9	128.9	559.3
Cal.	ME	20.7	19.8	28.9	*32.7*	138.4	*19.6*	25.5	37.5	46.5	127.6
	RPD	1.04	0.98	0.99	1.09	1.23	0.99	1.00	0.94	0.94	1.17
	RMSE	*134.2*	*119.1*	*103.3*	*128.0*	*634.3*	142.2	131.9	110.9	128.4	644.8
Val.	ME	44.4	*25.3*	*33.8*	*38.6*	227.0	*42.3*	30.7	46.5	48.0	*200.0*
	RPD	1.00	1.09	1.40	1.44	0.97	0.94	0.99	1.29	1.43	0.95

Abbreviations: Cal. = Calibration dataset; L1-LNMP = Log_e of Mehlich soil P at 0 – 30cm; L2-LNMP = Log_e of Mehlich soil P at 30 – 60cm; L3-LNMP = Log_e of Mehlich soil P at 60 – 120cm; L4-LNMP = Log_e of Mehlich soil P at 120 – 180cm; L1-LNTP = Log_e of total P at 0 – 30cm; ME = mean error; RMSE = root mean square error; RK/SMLR = regression kriging with stepwise multiple linear regression; RPD = residual prediction deviation; R^2 = coefficient of determination; SMLR = stepwise multiple linear regression; Val. = Validation dataset;

[1] Coarse resolution dataset were derived from Landsat ETM+ and digital elevation model (DEM) from national elevation dataset.

[2] Fine resolution dataset derived from were derived from IKONOS and DEM from light detection and ranging (LIDAR).

* Statistically significant at the 0.10 significance level.

Table 5-7. Summary of semivariogram parameters derived from residuals of global trend model with stepwise multiple linear regression for Mehlich phosphorus (MP) in µg/g and total phosphorus (TP) in µg/g in the Santa Fe River Ranch Beef Unit.

Resolution	Layer	Fitted model	Nugget	Sill	Range (m)	Nugget/Sill (%)
COARSE	L1-LNMP	Spherical	0.3070	0.7211	424	42.6
	L2-LNMP	Spherical	0.5763	0.9932	279	58.0
	L3-LNMP	Spherical	0.4307	1.1669	664	36.8
	L4-LNMP	Spherical	0.4581	0.9126	770	50.2
	L1-LNTP	Spherical	0.0894	0.4634	473	19.3
FINE	L1-LNMP	Spherical	0.4148	0.7246	468	57.3
	L2-LNMP	Spherical	0.6843	1.0656	1318	64.2
	L3-LNMP	Spherical	0.5371	1.2124	784	44.3
	L4-LNMP	Spherical	0.4291	0.8360	658	51.3
	L1-LNTP	Spherical	0.1733	0.3668	378	47.3

Abbreviations: COARSE = coarse resolution dataset (30 m); FINE = fine resolution dataset (4 m); L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus in µg/g; TP = total phosphorus in µg/g.

Table 5-8. Summary of semivariogram parameters by the lognormal kriging interpolation method for Mehlich phosphorus at different four depths in the Santa Fe River Ranch Beef Unit.

Property	Best interpolation	Fitted model	Nugget	Sill	Range (m)	Nugget/Sill (%)
L1-LNMP		Spherical	0.1675	1.2220	1689	12.05
L2-LNMP		Spherical	0.4870	1.1830	1531	29.16
L3-LNMP	LNK	Spherical	0.2372	0.8710	601	21.40
L4-LNMP		Spherical	0.4265	0.8427	818	33.60
L1-LNTP		Spherical	0.1057	0.3525	561	23.07

Abbreviations: LNK = lognormal kriging; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); TP = total phosphorus (µg/g).

Table 5-9. Global trend model of log_e-transformed Mehlich phosphorus (LNMP) in µg/g at four different layers and total phosphorus (LNTP) in µg/g at L1 (0-30cm) represented by the stepwise multiple linear regression models (SMLR) with land use, topographic, and spectral information from coarse and fine datasets in the Santa Fe River Ranch Beef Unit.

SMLR with coarse data (30 m) derived from Landsat ETM+ and NED

Property	Regression equation*
L1-LNMP	8.559 + 0.068 * [PC3ME33] + 0.096 * [CSLOPEPCT] - 0.428 * [PAST]
L2-LNMP	8.004 + 0.068 * [PC3ME33] + 0.118 * [CSLOPEPCT] + 0.604 * [CUPFOR]
L3-LNMP	6.670 + 0.038 * [PC3ME33] + 0.098 * [CSLOPEPCT] - 0.434 * [PAST] - 0.410 * [ASPECTNO]
L4-LNMP	3.583 + 0.601 * [CASPECTW] +0.149 * [CSLOPEPCT]
L1-LNTP	9.390 + 0.049 * [PC3ME33] + 0.086 * [CSLOPEPCT] + 0.366 * [CUPFOR]

SMLR with fine data (4 m) derived from IKONOS and LIDAR

Property	Regression equation*
L1-LNMP	4.526 - 1.407 * [ASPECTS] - 1.067 * [WETF] - 0.401 * [FUPFOR]
L2-LNMP	6.272 - 1.125 * [ASPECTS] - 1.360 * [WETF] - 0.004 * [IK06B5]
L3-LNMP	4.278 - 0.939 * [ASPECTS] - 0.839 * [WETF]
L4-LNMP	4.361 - 1.452 * [ASPECTS] + 0.603 * [FASPECTW]
L1-LNTP	4.331 - 0.842 * [ASPECTS] + 0.638 * [FASPECTW] - 0.376 * [IK06RATIO] + 0.049 * [FSLOPEPCT] - 0.017 * [IK06TC3] + 0.418 * [RANGE]

Abbreviations: ASPECTNO = north-facing slope derived from NED data; ASPECTS = south-facing slope derived from LIDAR data; CSLOPEPCT = slope (percent) derived from National Elevation Dataset (NED); CUPFOR = upland forest land use classified from Landsat ETM+; ETM+ = Enhanced Thematic Mapper Plus (ETM+); FSLOPEPCT = slope (percent) derived from LIDAR data; FUPFOR = upland forest land use classified from IKONOS; LIDAR = LIght Detection and Ranging; IK06B5 = reflectance of band 5 derived from IKONOS; IK06RATIO = infrared / red band from IKONOS; IK06TC3 = tasseled cap 3 derived from IKONOS; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); PAST = improved pasture classified from Landsat ETM+; PC3ME33 = mean principal component 3 derived from Landsat ETM+ averaged within a 3×3 pixel window; RANGE = rangeland land use classified from IKONOS; TP = total phosphorus (µg/g); WETF = wetland forest land use classified from IKONOS.
* Statistically significant at the 0.10 significance level.

Table 5-10. Analysis of variance (ANOVA) and homogeneous groups of loge-transformed Mehlich phosphorus at 0-30cm in the Santa Fe River Ranch Beef Unit according to the Dunnett's T3 test at the 0.05 significance level.

Variable and factor	Levene	ANOVA	Welch	B-F	Homogeneous Groups [1]	Group mean (log_e µg/g)
L1-LNMP by land use	14.8*	8.6*	5.6*	*8.2**	Agriculture [a]	4.55
					Improved pasture [a]	4.42
					Rangeland [a b]	4.40
					Upland forest [b]	3.60
					Wetland forest [b]	3.30
L1-LNTP by land use	7.6*	5.8*	4.4*	*6.1**	Agriculture [a]	6.72
					Improved pasture [a]	6.67
					Rangeland [a]	6.59
					Upland forest [a]	6.11
					Wetland forest [b]	5.95
L1-LNMP by soil drainage	11.3*	14.6*	11.0*	*22.0**	Excessively drain [a]	4.62
					Moderately well drain [a]	4.53
					Somewhat poorly drain [a]	4.43
					Well drain [a]	4.42
					Poorly drain [b]	3.19
L1-LNTP by soil drainage	3.1*	12.7*	7.7*	*17.4**	Moderately well drain [a]	6.75
					Well drain [a]	6.73
					Somewhat poorly drain [a]	6.62
					Excessively drain [a b]	6.42
					Poorly drain [b]	5.79

Abbreviations: B-F = Brown-Forsythe's ANOVA test statistic; L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; Levene = Levene's test statistic; MP = Mehlich phosphorus in µg/g; TP = total phosphorous in µg/g; Welch = Welch's ANOVA test statistic;

[1] Columns with different letters have significantly different means at the 0.05 significance level, based on Dunnett's T3 test.

The preferred test based on the number of samples and equality of variance is shown in italics.

Table 5-11. Global trend model of log_e-transformed Mehlich phosphorus (LNMP) in µg/g at four different layers and total phosphorus (LNTP) in µg/g at L1 represented by the stepwise multiple linear regression models with all coarse data and regression coefficients in the Santa Fe River Ranch Beef Unit.

Property	R^2 Calib.	Valid.	Adj.	Regression equation *
L1-LNMP	0.63	0.48	0.62	10.427 - 0.047 * [DRASTIDX] + 0.484 * [ULT] - 0.458 * [RANGE]
L2-LNMP	0.52	0.39	0.50	10.199 - 0.043 * [DRASTIDX] + 0.920 * [ULT] + 0.493 * [UPFOR] - 0.013 * [SANDPCT]
L3-LNMP	0.55	0.17	0.49	24.691 - 0.043 * [DRASTIDX] + 0.049 * [CLAYPCT] + 0.614 * [ULT] - 0.656 * [RANGE] - 12.963 * [AWC] - 0.116 * [ETMB1] + 0.078 * [ETMB2] - 6.463 * [BULK] - 0.600 * [HYDROGRPD] - 0.849 * [DRAINWD] + 0.782 * [HYDROGRPB]
L4-LNMP	0.40	0.20	0.33	12.593 + 0.087 * [SLOPEPCT] - 0.032 * [ETMB1] + 0.844 * [ULT] - 0.026 * [DRASTIDX] - 0.034 * [SANDPCT] - 0.480 * [EVEGOCLAY] + 0.464 * [UPFOR] + 1.604 * [SPOD]
L1-LNTP	0.58	0.43	0.56	12.189 - 0.029 * [DRASTIDX] - 0.764 * [HYDROGRPD] - 0.021 * [ETMB1] + 0.034 * [SILTPCT] - 0.374 * [RANGE]

Abbreviations: Adj. = adjusted on calibration data; AWC = soil available water capacity (cm/cm); BULK = bulk density (g/cm^3); Calb. = calibration dataset; CLAYPCT = clay content percent in soil; DRAINWD = well drained in soil drainage class; DRASTIDX = depth to water table, net recharge, aquifer media, soil media, topography, impact of the vadose zone, hydraulic conductivity of the aquifer (0-255); ETMB1 = reflectance of band1 derived from Landsat Enhanced Thematic Mapper Plus (ETM+) (0-255); ETMB2 = reflectance of band2 derived from Landsat ETM+ (0-255); EVEGOCLAY = clayey sand in hydrogeologic class; HYDGRPB = moderate infiltrations rate in soil hydrologic group; HYDGRPD = very slow infiltrations rate in soil hydrologic group;L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); R^2 = coefficient of determination; RANGE = rangeland classified from Landsat ETM+; SANDPCT = sand content percent in soil; SILTPCT = silt content percent in soil; SLOPEPCT = slope (percent) derived from National Elevation Dataset (NED); SPOD = Spodosols in soil order class; TP = total phosphorus (µg/g); ULT = Ultisols in soil order; UPFOR = upland forest land use classified from Landsat ETM+; Valid. = validation dataset
* Statistically significant at the 0.10 significance level.

Table 5-12. Spearman correlation of selected variables for cokriging in the Santa Fe River Ranch Beef Unit.

Property	Selected variable	Spearman correlation*
L1-LNMP	DTW	0.33
L2-LNMP	DTW	0.28
L3-LNMP	DTW	0.25
L4-LNMP	NEDSLOPE	0.31
L1-LNTP	DTW	0.35

Abbreviations: DTW = depth to water table (cm); L1-LNMP = Log_e(L1-MP) at 0-30cm; L2-LNMP = Log_e(L2-MP) at 30-60cm; L3-LNMP = Log_e(L3-MP) at 60-120cm; L4-LNMP = Log_e(L4-MP) at 120-180cm; L1-LNTP = Log_e(L1-TP) at 0-30cm; MP = Mehlich phosphorus (µg/g); NEDSLOPE = slope (percent) derived from National Elevation Data (NED) digital elevation model (DEM) dataset; TP = total phosphorus (µg/g).
* Statistically significant at the 0.01 significance level.

Table 5-13. Error statistics of the best prediction methods for Mehlich phosphorus and total phosphorus across the Santa Fe River Ranch Beef Unit data.

Property	$RMSE_v$ (µg/g)			ME_v (µg/g)			RPD		
	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK
L1-MP	*105.7*	146.0	107.1	14.3	60.8	*-9.6*	*1.28*	0.92	1.25
L2-MP	*117.0*	139.7	119.3	*-9.0*	26.5	-27.7	*1.10*	0.92	1.08
L3-MP	*87.1*	101.9	124.7	*11.0*	15.0	-89.3	*1.21*	1.03	0.85
L4-MP	*92.9*	128.6	101.4	*-7.8*	10.5	-5.5	*1.31*	0.95	1.20
L1-TP	*481.0*	838.1	871.5	*9.7*	592.4	-560.0	*1.32*	0.76	0.73

Property	$RMSE_c$ (µg/g)			ME_c (µg/g)			RPD		
	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK	LNK	RK/ SMLR	CK
L1-MP	54.2	85.8	67.4	*-7.6*	43.4	-31.5	*0.68*	0.43	0.55
L2-MP	*50.7*	75.3	61.9	*-14.3*	31.1	-33.0	*0.37*	0.25	0.31
L3-MP	*55.1*	96.4	119.5	*-10.1*	12.8	-82.5	*0.55*	0.31	0.25
L4-MP	*97.3*	114.3	988.5	*-32.2*	7.3	67.7	*0.43*	0.37	0.04
L1-TP	*244.3*	789.5	774.7	*-41.5*	555.9	-588.5	*2.11*	0.65	0.67

Abbreviations: CK = cokriging; L1-MP = MP at 0-30cm; L2-MP = MP at 30-60cm; L3-MP = MP at 60-120cm; L4-MP = MP at 120-180cm; L1-TP = TP at 0-30cm; ME_c = mean error calculated on calibration set; ME_v = mean error calculated on validation set; MP = Mehlich phosphorus (µg/g); $RMSE_c$ = root mean square error calculated on calibration set; $RMSE_v$ = root mean square error calculated on validation set; RK/SMLR = regression kriging with stepwise multiple linear regression; TP = total phosphorus (µg/g);
The best RMSE, ME, and RPD of each model were shown in italics.

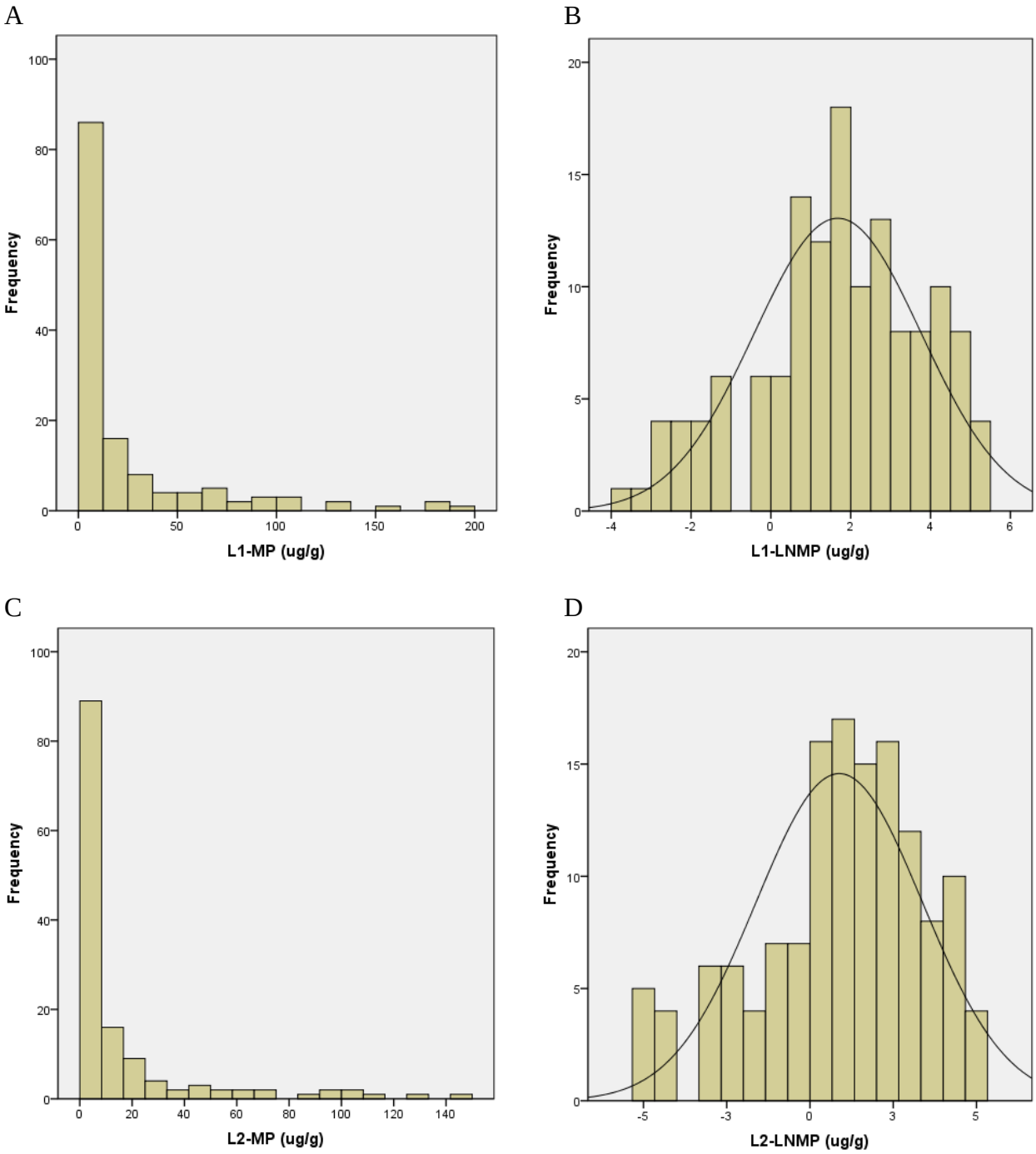


Figure 5-1. Histograms of Mehlich phosphorus (MP) in $\log_e$ µg/g and total phosphorus (TP) in $\log_e$ µg/g datasets in the Santa Fe River Ranch Beef Unit: (A) MP at 0-30cm; (B) Log_e(MP) at 0-30cm; (C) MP at 30-60cm; (D) Log_e(MP) at 30-60cm; (E) MP at 60-120cm; (F) Log_e(MP) at 60-120cm; (G) MP at 120-180cm; (H) Log_e(MP) at 120-180cm; (I) TP at 0-30cm; (J) Log_e(TP) at 0-30cm.

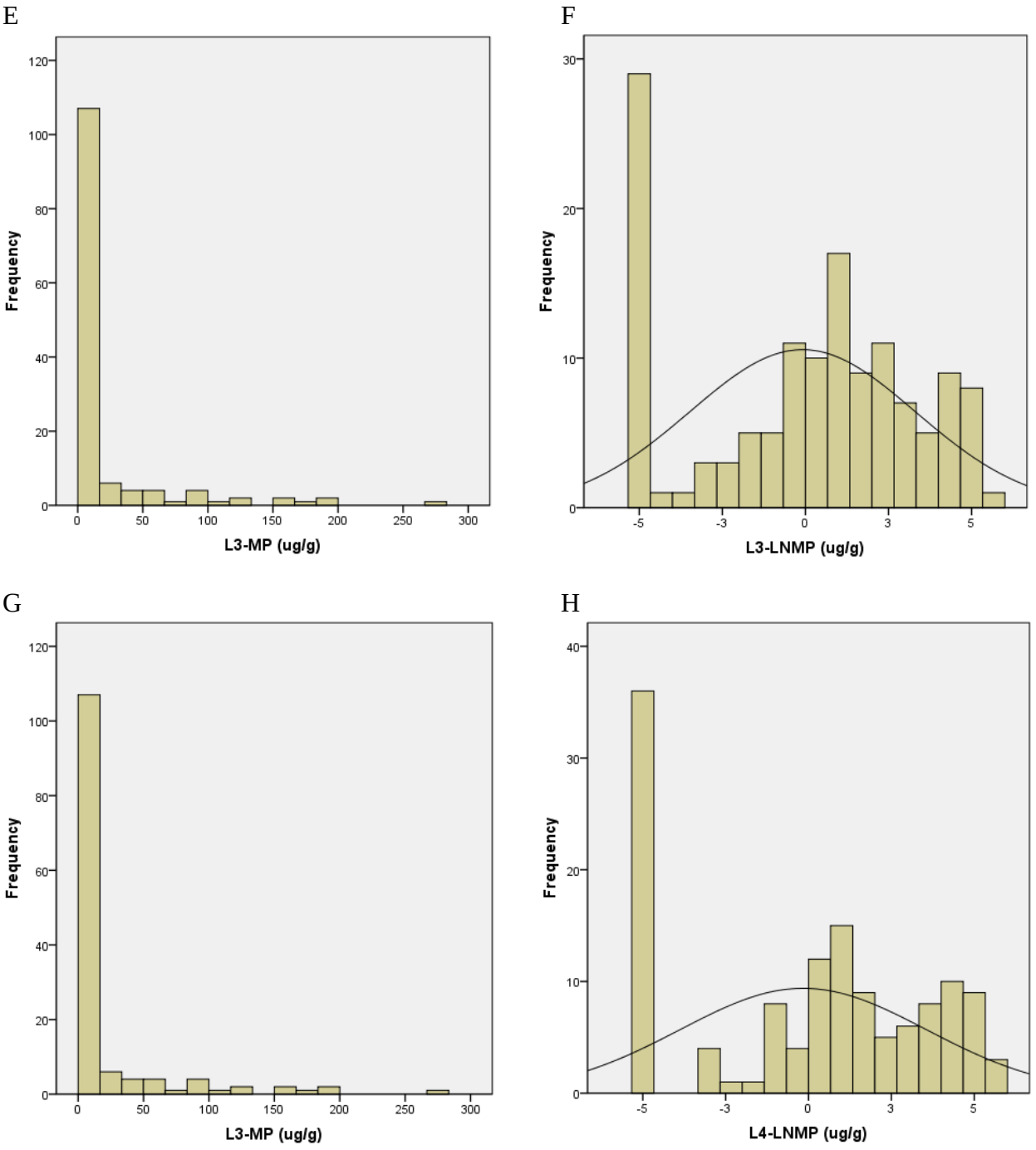
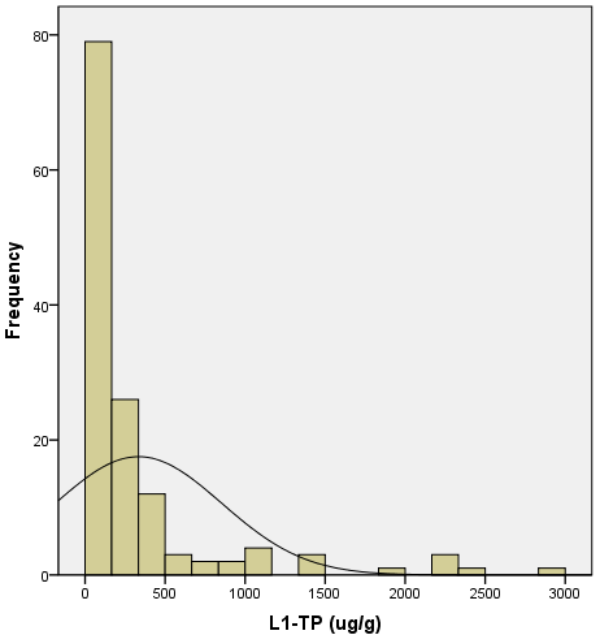


Figure 5-1. Continued.

I



J

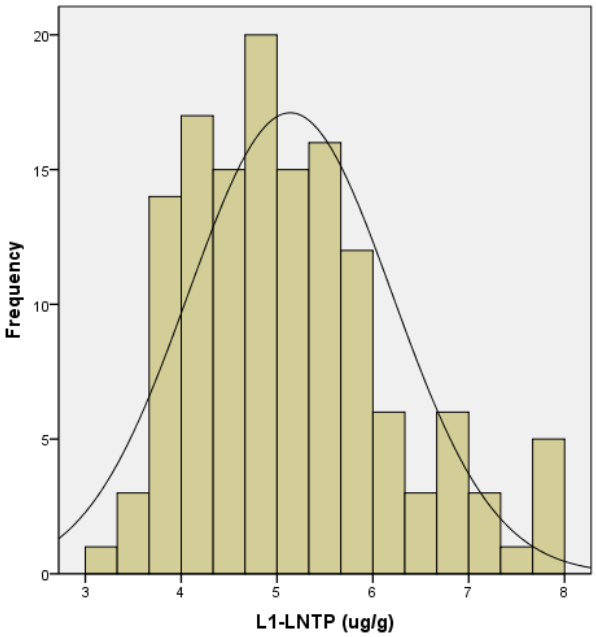


Figure 5-1. Continued.

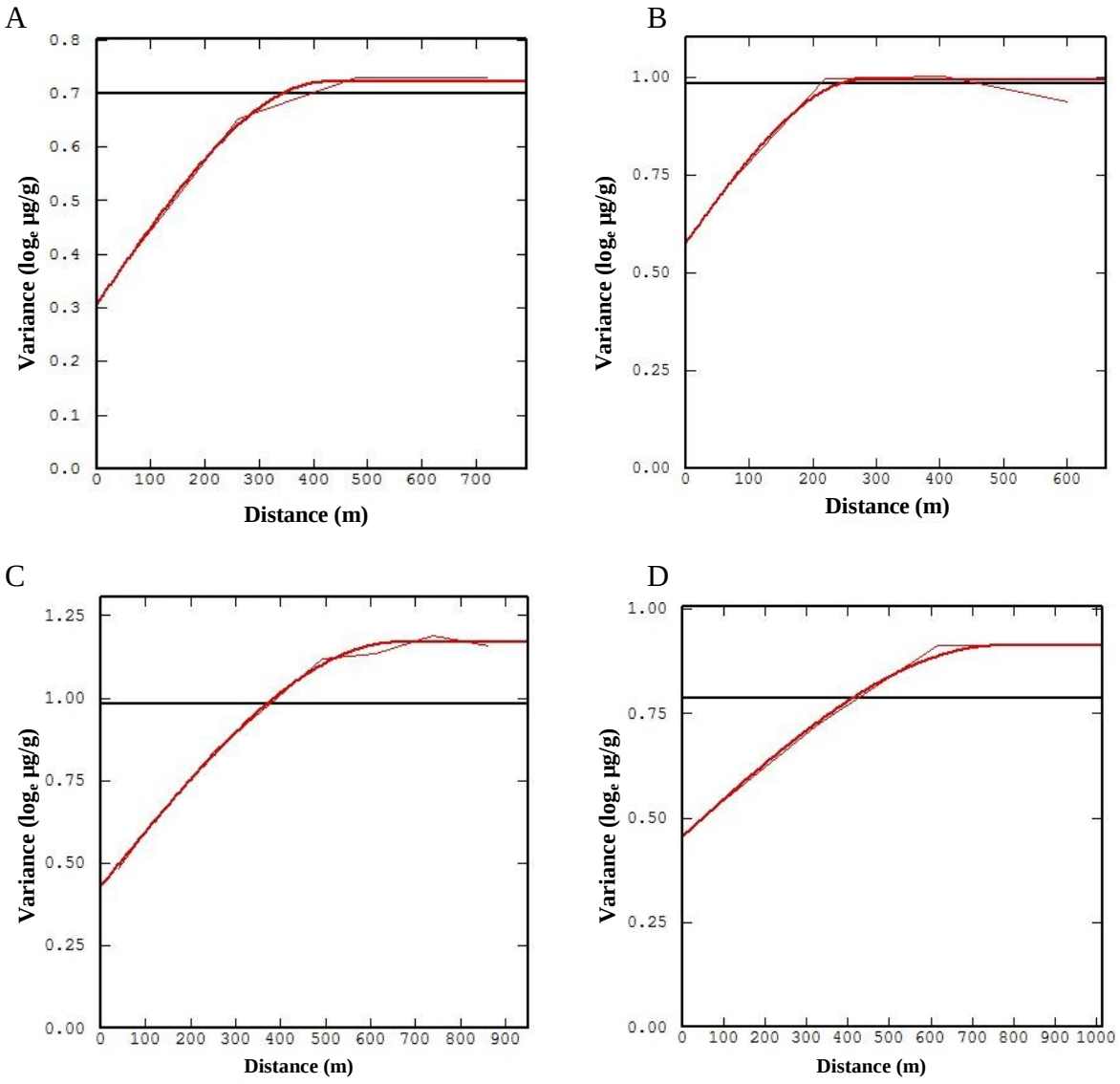


Figure 5-2. Semivariograms for kriging on residuals with coarse and fine data across the Santa Fe River Ranch Beef Unit.: (A) Kriging of Mehlich phosphorus (MP) residuals with coarse data at L1 (0-30 cm); (B) Kriging of MP residuals with coarse data at L2 (30-60 cm); (C) Kriging of MP residuals with coarse data at L3 (60-120 cm); (D) Kriging of MP residuals with coarse data at L4 (120-180 cm); (E) Kriging of total phosphorus (TP) residuals with coarse data at L1; (F) Kriging of MP residuals with fine data at L1; (G) Kriging of MP residuals with fine data at L2; (H) Kriging of MP residuals with fine data at L3; (I) Kriging of MP residuals with fine data at L4; (J) Kriging of TP residuals with fine data at L1.

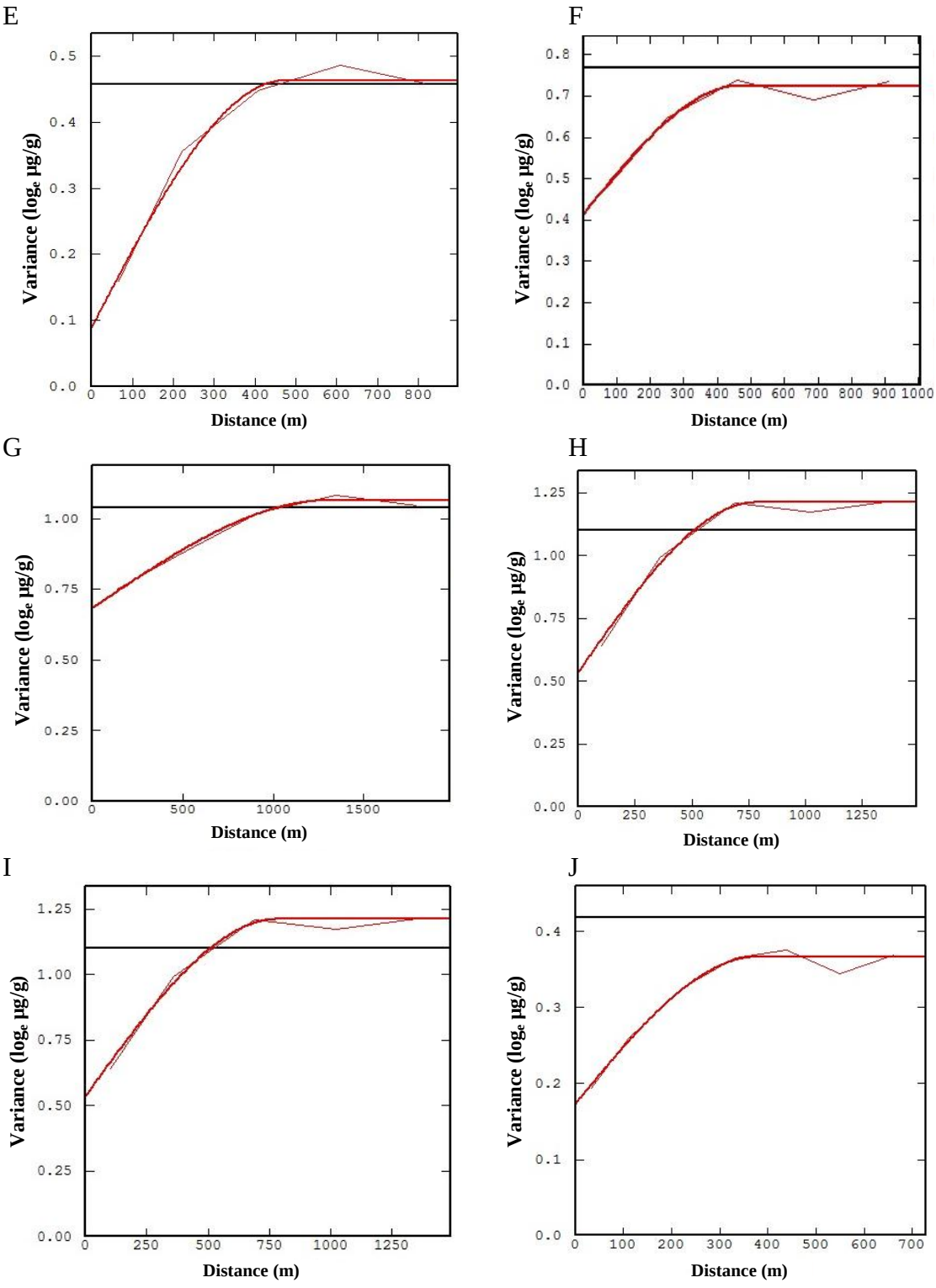


Figure 5-2. Continued.

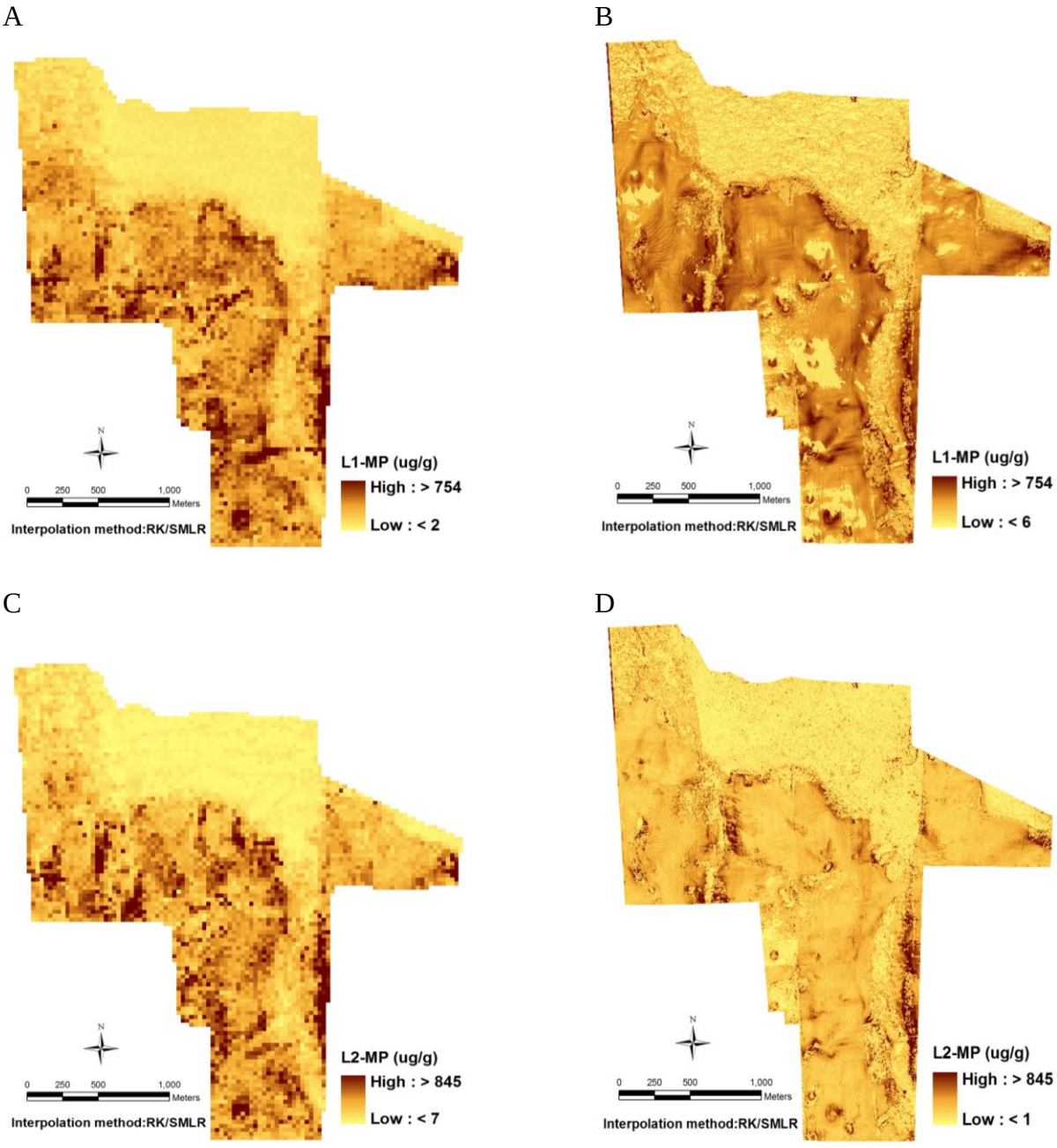
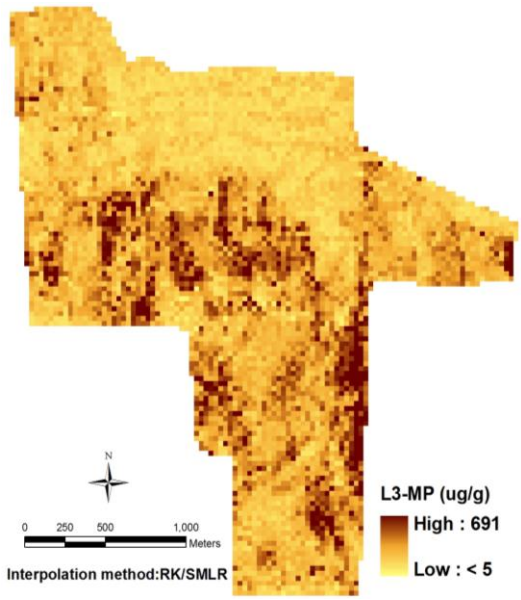
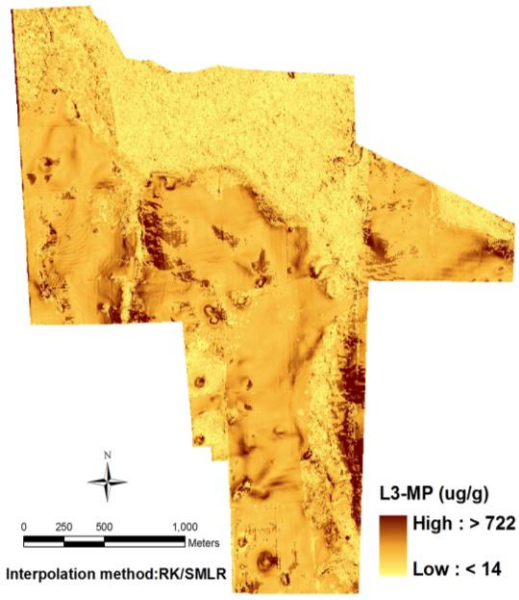


Figure 5-3. Prediction maps of Mehlich phosphorus (MP) and total phosphorous (TP) by regression kriging with stepwise multiple linear regression (RK/SMLR) with coarse (COARSE) and fine (FINE) resolution dataset in the Santa Fe River Ranch Beef Unit: (A) MP at L1 (0-30 cm) by RK/SMLR with COARSE; (B) MP at L1by RK/SMLR with FINE; (C) MP at L2 (30-60 cm) by RK/SMLR with COARSE; (D) MP at L2 by RK/SMLR with FINE; (E) MP at L3 (30-60 cm) by RK/SMLR with COARSE; (F) MP at L3 by RK/SMLR with COARSE; (G) MP at L3 by RK/SMLR with FINE; (H) MP at L4 (120-180 cm) by RK/SMLR with COARSE; (I) TP at L1 by RK/SMLR with COARSE; (J) TP at L1 by RK/SMLR with FINE.

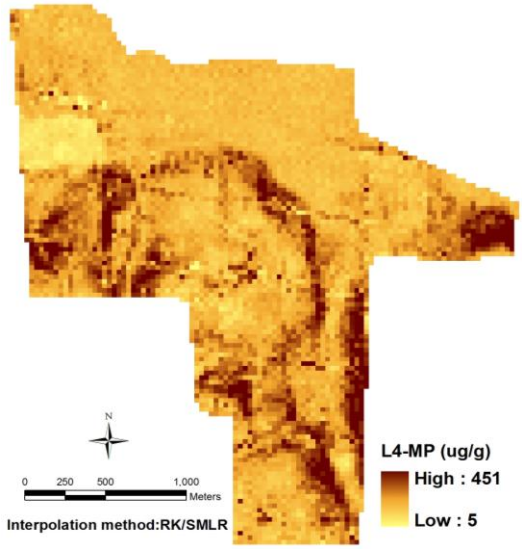
E



F



G



H

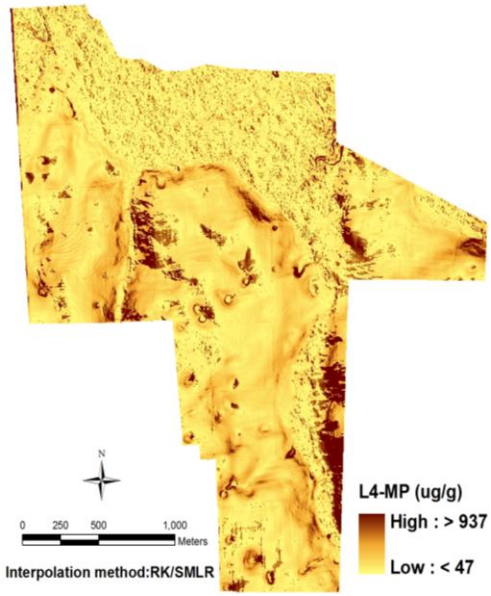
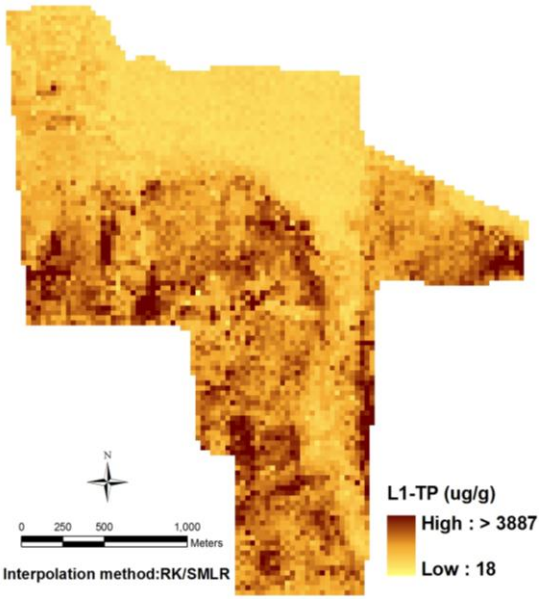


Figure 5-3. Continued.

I



J

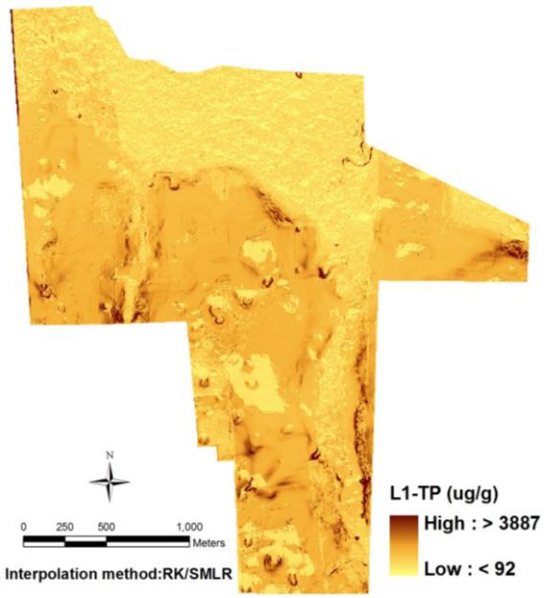


Figure 5-3. Continued.

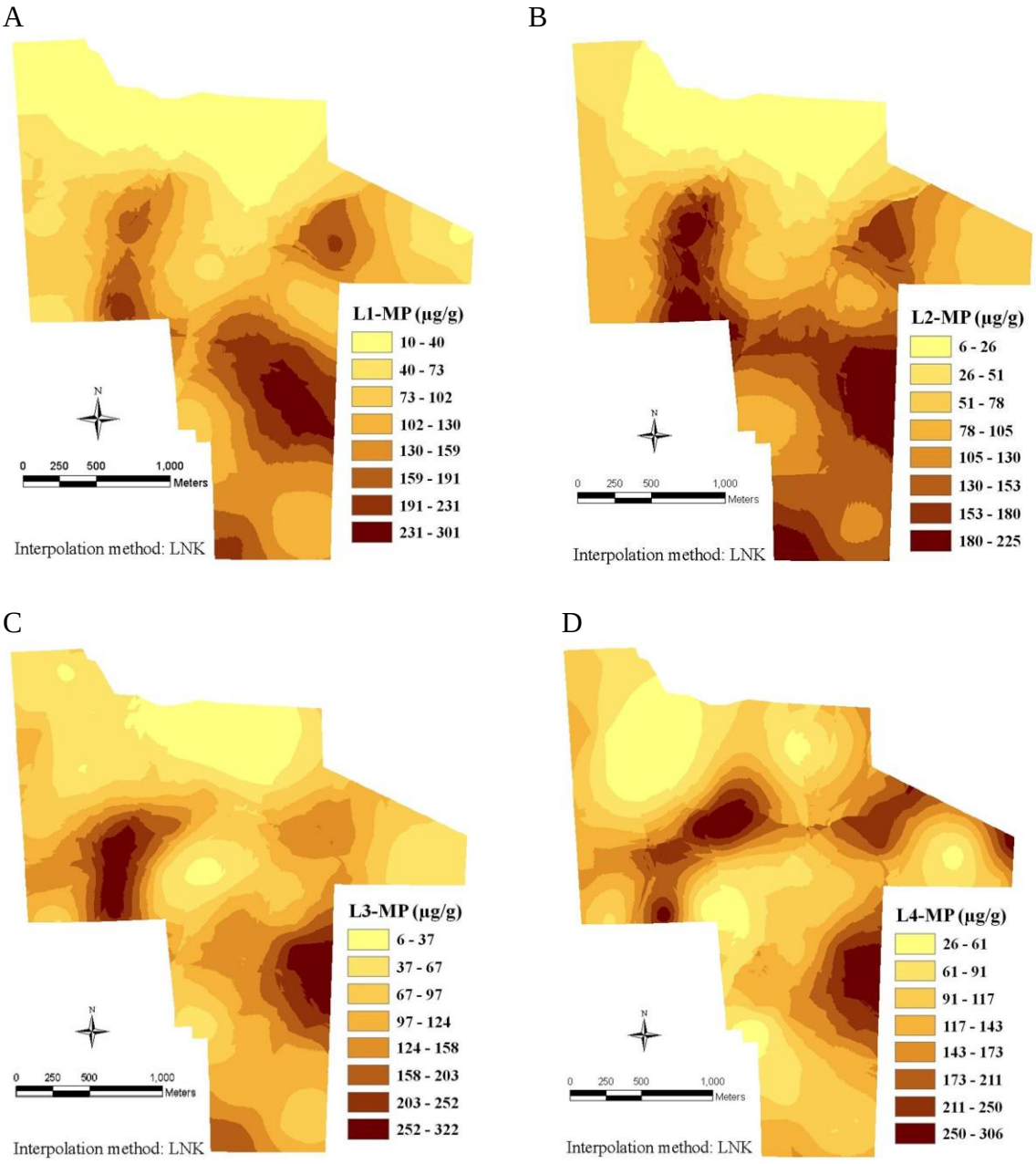


Figure 5-4. Final maps of prediction in the Santa Fe River Ranch Beef Unit: (A) Mehlich phosphorus (MP) at L1 (0-30 cm) by lognormal kriging (LNK); (B) MP at L2 (30-60 cm) by LNK; (C) MP at L3 (60-120 cm) by LNK; (D) MP at L4 (120-180 cm) by LNK; (E) total phosphorus (TP) at L1 by LNK.

E

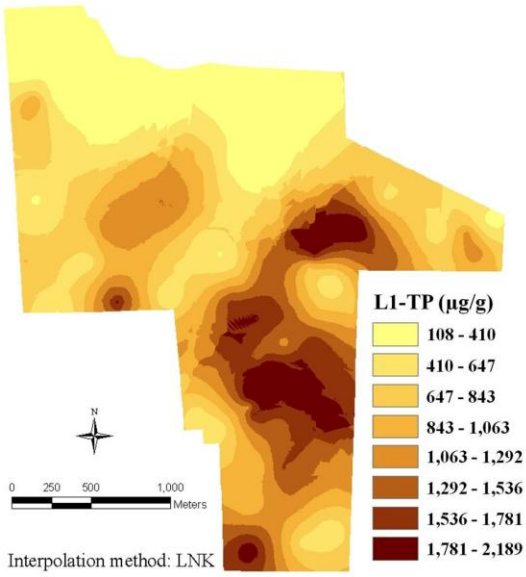


Figure 5-4. Continued.

CHAPTER 6
SUMMARY AND SYNTHESIS

In conjunction with air and water, soil is one of most precious and diminishing resources on the earth because of its dynamic, physical, chemical and biological functions and soil contamination typically can be caused by various sources.

Even though P is essential to all forms of life on earth the major environmental issues related to soil P are land degradation and accelerated eutrophication and water quality monitoring over many years showed that TP loads transported into the Gulf of Mexico from the SFR has increased.

This research was motivated by the fact that it is essential to identify factors controlling soil MP and TP contents and characterize the spatial distribution of MP and TP before proper land use management or assessment of soil ecosystem services can be established.

The results showed that MP among different layers were significantly correlated both in the SFRW and the SFRRBU indicating that the spatial distribution of MP at the upper layer might have influenced the adjacent layer beneath it through vertical transport processes.

The results also indicated that land use, soil drainage, soil order, topographic information, and geologic unit in the SFRW were statistically significant factors controlling the variation of MP. Findings support the hypothesis that land use at the top soil and geology at the subsoil rank highest to predict MP.

Error statistics in the SFRW showed that multivariate methods (RK/SMLR and CK) outperformed the univariate method (LNK) demonstrating that incorporating the spatial covariation between soil and auxiliary environmental properties are more accurate than the univariate method (LNK). This can be very beneficial for modeling large areas, such as a watershed, because soil sampling can be constrained by limited resources such as labor, time,

finance, where auxiliary data may be relatively easily achieved or updated to generate better prediction maps.

Based on the overall accuracy and kappa statistics from two remote sensing images for the SFRRBU, the 2006 IKONOS provided better classification accuracy when compared to the classification of the 2003 Landsat ETM+, mainly due to higher spatial resolution from the IKONOS image. Even though the accuracy of discriminating different forest types (e.g., upland, wetland, and pineland forest) from the Landsat ETM+ were poor, the overall accuracy of forest and non-forest were similar for both images. This can be interpreted that the coarse resolution satellite image (Landsat ETM+) can still provide similar overall patterns of land use compared to fine resolution image (IKONOS) in the study area.

According to $RMSE_v$ and RPD from RK/SMLR with coarse (derived from IKONOS and LIDAR) and fine (derived from Landsat ETM+ and DEM from NED) data, fine resolution data did not have superior predictive capabilities to infer on soil P when compared to coarse resolution data. This may be explained by several sources of data errors (e.g. positional accuracy error, accuracy of soil attribute measurements, etc.) or that the finer scale environmental co-variates did not have the capabilities to improve predictions.

Based on $RMSE_v$ and RPD, the best prediction methods of MP and TP were LNK at all depths in the SFRRBU and this rejects the hypothesis that multivariate or hybrid geospatial models (RK and CK) that incorporate exhaustive environmental datasets improve the prediction of soil MP and TP when compared to a univariate method (OK) that only uses soil property measurements. This specifies that incorporating the spatial covariation between soil properties and auxiliary environmental properties does not necessarily improve the accuracy of soil predictions. This may be coupled to the scale (extent) of the relatively small study area

(SFRRBU-5.58km^2) which captured sufficiently the fine and medium spatial autocorrelation of soil P from measurements. Thus, the incorporation of ancillary environmental properties did not further improve predictions of MP and TP. In other words, overall the spatial autocorrelation of MP and TP alone in LNK explained their variability better than the combination of global and local trends so auxiliary variables in RK/SMLR and CK did not provide better explanatory power.

Our findings from this research improved our understanding of the spatial distribution of soil MP and TP and how they relate to other environmental properties (e.g., land use/land cover, soil drainage, and/or soil type). We also identified interrelationships between soil MP and TP and various controlling environmental variables and mapped their distributions in the SFRW and the SFRRBU.

LIST OF REFERENCES

Abuzar, M., Ryan, S., 2001. Integrating multi-source remote sensing data for soil mapping in Victoria. Geoscience and remote sensing symposium. International Geoscience and Remote Sensing Symposium (IGARSS), Sydney, Australia, July 9-13, 2001. The Institute of Electrical and Electronics Engineers (IEEE). 6, 2495-2497.

Alachua County Environmental Protection Department (ACEPD), 2009. Fertilizer standards and management practices code (Ordinance 09-06). The Board of County commissioners, Alachua County, Florida.

Alva, A.K., Huang, B., Paramasivam, S., Sajwan, K.S., 2002. Evaluation of root growth limiting factors in spodic horizons of spodosols. J. of plant nutrition. 25, 2001-2014.

Anderson, D.L., Flaig, E.G., 1995. Agricultural best management practices and surface water improvement and management. Water Science and Technology 31, 109-121.

Babiker, I.S., Mohamed, M.A.A., Hiyama, T., Kato, K., 2005. A GIS-based DRASTIC model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan. Science of the Total Environment. 345, 127-140.

Bailloeul, T., Duan, J., Prinet, V., Serra, B., 2003. Urban digital map updating from satellite high resolution images using GIS data as a priori knowledge. In: Proceedings of the 2nd Geoscience and Remote Sensing Symposium (GRSS)/ International Society for Photogrammetry and Remote Sensing (ISPRS) Joint Workshop, Berlin, Germany, May 22-23, 2003. pp. 283–287.

Barnes, E.M., Baker, M.G., 2000. Multispectral data for mapping soil texture: possibilities and limitations, Applied Engineering in Agriculture. 16, 731-741.

Berkelaar, E., 2001. The effect of aluminum in acidic soils on plant growth. Educational concerns for hunger organization, Development notes 71, 1-3.

Birr, A.S., Mulla, D.J., 2001. Evaluation of the phosphorus index in watersheds at the regional scale. J. Environ. Qual. 30, 2018-2025.

Birk, R..J., Stanley, T., Snyder, G. I., Henning, T. A., Fladeland, M. M., Policelli, F., 2003. Government programs for research and operational uses of commercial remote sensing data. Remote Sens. Environ. 88, 3-16.

Bishop, T.F.A., Minasny, B., 2005. Digital soil-terrain modeling: The predictive potential and uncertainty. In: Grunwald, S.(Ed.), Environmental soil-landscape modeling – geographic information technologies and pedometrics. CRC Press, Boca Raton, Florida. pp. 185-213.

Bochove, E.V., Theriault, G., Dechmi, F., Rousseau, A.N., Quilbe, R., Leclerc, M.L., Goussard, N., 2006. Indicator of risk of water contamination by phosphorus from Canadian agricultural land. Water Sci. and Tech. 53, 303-310.

Bochove, E.V., Theriault, G., Dechmi, F., Leclerc, M.L., Goussard, N., 2007. Indicator of risk of water contamination by phosphorus: temporal trends for the Province of Quebec from 1981 to 2001. Canadian J. Soil Sci. 87, 121-128.

Borůvka, L., Mládková, L., Penižek, V., Drábek, O., Vasat, R., 2007. Forest soil acidification assessment using principal component analysis and geostatistics. Geoderma 140, 374-382.

Bouma, J., 1997. Precision agriculture: Introduction to the spatial and temporal variability of environmental quality. In: Lake, J.V., Bock, G.R., Goode, J.A. (Ed.s.), Precision agriculture: spatial and temporal variability of environmental quality., 210. John Wiley and Sons, Ciba Foundation Symposium, Wageningen, The Netherlands. pp.5-17.

Brady, N.C., Well, R.R., 2002. The nature and properties of soils (13th edition). Macmillan Publishing. Englewood Cliffs, New Jersey.

Bui, E.N., 2004. Soil survey as a knowledge system. Geoderma 120, 17-26.

Bruland, G.L., Grunwald, S., Osborne, T.Z., Reddy, K.R., Newman, S., 2006. Spatial distribution of soil properties in Water Conservation Area 3 of the Everglades. Soil Sci. Soc. Am. J. 70, 1662-1676.

Burrough, P.A., McDonnell.,R.A., 1998. Principles of Geographic Information Systems: Spatial Information Systems and Geostatistics (2nd edition). Oxford University Press, New York.

Cahoon, L.B., Ensign, S.H., 2004. Excessive soil phosphorus levels in eastern North Carolina: temporal and spatial distributions and relationships to agriculture and other land uses. Nutrient Cycl. in Agro. 69, 111-125.

Campling, P., Gobin, A., Feyen, J., 2002. Logistic modeling to spatially predict the probability of soil drainage classes. Soil Sci. Soc. Am. J. 66, 1390-1401.

Canfield Jr., D.E., Hoyer, M.V., 1988. Regional geology and the chemical and trophics state characteristics of Florida lakes. Lake and Reservoir Management 4, 21-31.

Collins, M.E., 2009. Key to soil orders in Florida. Institute of Food and Agricultural Sciences Extension, University of Florida, Gainesville, FL. SL43.

Corstanje, R., Grunwald, S., Reddy, K.R., Osborn, T.Z. and Newman, S., 2006. Assessment of the spatial distribution of soil properties in a Northern Everglades. J. Environ. Qual. 35, 938-949.

Cox, J., Kautz, R. MacLaughlin, M, Gilbert, T., 1994. Closing the gaps in Florid's wildlife habitat conservation system. Florida Game and Fresh Water Commission, Tallahassee, Florida.

Daverede, I.C., Kravchenko, A. N., Hoeft, R.G., Nafziger, E.D., Bullock, D. G., Warren, J. J., Gonzini, L. C., 2004. Phosphorus runoff from incorporated and surface-applied liquid swine manure and phosphorus fertilizer. J. Environ. Qual. 33, 1535-1544.

Davis, M.R., 1994. Topsoil properties under tussock grassland and adjoining pine forest in Otago, New Zealand. New Zealand J. Agricultural Research 37, 465-469.

DeBusk, W.F., Newman, S., Reddy, K.R., 1999. Spatial and temporal patterns of soil phosphorus enrichment in Everglades Water Conservation Area 2A. J. Environ. Qual. 30, 1438-1446.

De Gryze, S., Six, J., Bossuyt, H., Van Oost, K., Merckx, R., 2008. The relationship between landform and the distribution of soil C, N and P under conventional and minimum tillage, Geoderma 144, 180-188.

Department of animal science, The University of Florida. Available at: http://www.animal.ufl.edu/facilities/sbru/. Last verified: Jun. 8, 2011.

Di, K., Ma, R., Wang, J., R. Li., 2003. Coastal mapping and change detection using high-resolution IKONOS satellite imagery. In: Proceedings of the NSF DG 2003 Conference, Boston, MA., pp.343-346.

Dobos, E., Micheli, E., Baumgardner, M.F., Biehl, L., Helt, T., 2000. Use of combined digital elevation model and satellite radiometric data for regional soil mapping. Geoderma 97, 367-391.

Dong, A., Simsiman, G.V., and Chesters, G., 1983. Particle-size distribution and phosphorus levels in soil, sediment, and urban dust and dirt samples from the Menomonee River Watershed, Wisconsin, U.S.A., Water Research. 17, 569-577.

Dubberly, D.W., 2007. Urban turf fertilizer rule. Florida Department of Agriculture and Consumer Services. Fact Sheet, August, 17.

Dunne, E.J., McKee, K.A., Clark, M.W., Grunwald, S., Reddy, K.R., 2007. Phosphorus in agricultural ditch soil and potential implications for water quality. Jurnal of soil and water conservation 62, 244-252.

Dunnett, C.W., 1980. Pair-wise multiple comparisons in the unequal variance case. J. Am. Stat. Assoc. 75, 796-800.

Elbir, T., 2003. Comparison of model predictions with the data of an urban air quality monitoring network in Izmir, Turkey. Atmospheric Environment 37, 2149-2157.

Eldeiry, A.A., Garcia, L.A., 2010. Comparison of ordinary kriging, regression kriging, and Cokriging techniques to estimate soil salinity using LANDSAT images. J. irrigation and Drainage Engineering 136, 355-364.

Evans, B.M., Corradini, K.J., 2001. BMP pollution reduction guidance document. Environmentla Resources Research Institute, Pennsylvania State University.

Fageria, N.K., Baligar, V.C., Bailey, B.A. 2005. Role of cover crops in improving soil and row crops productivity. Communications in Soil Science and Plant Analysis. 36:2733-2757.

Fernald, E.A., Purdum, E.D. (Eds), 1998. Water resource atlas of Florida, Institute of Science and Public Affairs, Florida State University, Tallahassee, Florida. 312 p.

Florida Department of Environmental Protection (FDEP). 2001. Basin status report: Suwannee. Tallahassee, FL

Florida Department of Environmental Protection (FDEP). 2006. TMDL protocol-version 6.0. Task Assignment. 003.03/05-003.

Florida Fish and Wildlife Conservation Commission (FFWCC), 2004. Florida vegetation and land cover data for Florida from 2003 Landsat ETM+ imagery. Raster layer. Spatial resolution: 30 m. FFWCC, Tallahassee, Florida.

Follett, R.F., Kimble, J.M., Lal, R. (Eds.) 2000. The potential of U.S. grazing lands to sequester carbon and mitigate the greenhouse effect. CRC Press, Boca Raton, Florida.

Foy, R.H., and Withers, P.J.A., 1995. The contribution of agricultural phosphorus to eutrophication. Ferilizer Soc. Proc. 365, 1-32.

Friesen, D.K., Rao, I.M., Thomas, R.J., Obcrson, A., Sanz, J.I., 1997. Phosphorus acquisition and cyclig in crop and pasture system in low fertility tropical soils. Plant and soil 196, 289-294.

Gallo, K., Ji, L., Reed, B., Eidenshink, J., Dwyer, J., 2005. Multi-platform comparisons of MODIS and AVHRR normalized difference vegetation index data. Remote Sens. Environ. 99, 221-231.

Garcia, A.R., Iorio, A.F. de., 2003. Phosphorus distribution in sediments of Morales Stream (tributary of the Matanza-Riachuelo River, Argentina). The influence of organic point source contamination. Hydrobiologia 492, 129-138

Geoeye Inc. Available at: http://www.geoeye.com/CorpSite/products-and-services/imagery-sources/Default.aspx. Last verified: Jun. 8, 2011.

Giri, C., Zhu, Z., Reed, B., 2005. A comparative analysis of the global land cover 2000 and MODIS land cover datasets. Remote Sens. Environ. 94, 123-132.

Gitelson, A.A., Vina, A., Masek, J.G., Verma, S.B., Suyker, A.E., 2008. Synoptic monitoring of gross primary productivity of maize using Landsat data. Geoscience and remote sensing letters 5, 133-137.

Goodchild, M.F., 1994. Integrating GIS and remote sensing for vegetation analysis and modeling - methodological issues. J. Vege. Sci. 5, 615-626.

Goovaerts, P., 1997. Geostatistics for natural resources evaluation. Oxford University Press, New York, New York.

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydro., 228, 113-129.

Goovaerts, P., 2001 Geostatistical modeling of uncertainty in soil science. Geoderma 103, 3-26.

Gray, P.N., Farrell, C.J., Kraus, M.L., and Gromnicki A.H., 2005. Lake Okeechobee: A synthesis of information and recommendations for its restoration. Aububon of Florida, Miami, FL. 106 pp. + appendices.

Greenhalgh, T., Ladner, J., Rupert, F., 2005. Geology of Florida. Biennial report No. 23. Florida Geological Survey, Tallahassee, FL.

Grimm, R., Behrens, T., Märker, M., Elsenbeer, H., 2008. Soil organic carbon concentrations and stocks on Barro Colorado Island - digital soil mapping using Random Forest analysis. Geoderma 146, 102–113.

Grunwald, S., 2005. What do we really know about the space-time continuum of soil-landscapes? In: Grunwald, S. (Ed.), Environmental soil-landscape modeling–geographic information technologies and pedometrics. CRC Press, Boca Raton, Florida, pp.3-36.

Grunwald, S., Goovaerts, P., Bliss, C.M., Comerford, N.B., Lamsal, S., 2006. Incorporation of auxiliary information in the geostatistical simulation of soil nitrate-nitrogen. Vadose Zone J. 5, 391-404

Grunwald, S., Rivero, R.G., Reddy, K.R., 2007. Understanding spatial variability and its application to biogeochemistry analysis. In: Sarkar D., Datta R. and R. Hannigan (Eds.), Concepts and Applications in Environmental Geochemistry. Elsevier, Amsterdam, The Netherlands, pp. 435-462.

Grunwald S. 2008. Reconstruction and scientific visualization of earthscapes considering trends and spatial dependence structures. New J. Phys. 10, No. 125011, 15 pp.

Grunwald, S., Reddy, K.R., 2008. Spatial behavior of phosphorus and nitrogen in a subtropical wetland. Soil Sci. Soc. Am. J. 72, 1174-1183.

Grunwald, S., 2009. Multi-criteria characterization of recent digital soil mapping and modeling approaches. Geoderma 152, 195-207.

Grunwald, S., Thompson, J.A., Boettinger, J.L., 2011. Digital soil mapping and modeling at continental scales – finding solutions for global issues. Soil Sci. Soc. Am. J. (Anniversary Special Paper). (in press)

Hallas, J.F., Magley, W., 2008. Nutrient and dissolved oxygen TMDL for the Suwannee River, Santa Fe River, Manatee Springs, Fanning Springs, Branford Spring, Ruth Spring, Troy Spring, Royal Spring, and Falmouth Spring. TMDL Report. Florida Department of Environmental Protection. 117p.

Haygarth, P.M., Heathwaite, A.L., Jarvis, S.C., Harrod, T.R., 1999. Hydrological factors for phosphorus transfer from agricultural soils. Advances in Agronomy 69, 153-178.

Haygarth, P.M., Hepworth, L., Jarvis, S.C., 1998. Forms of phosphorus transfer in hydrological pathways from soil under grazed grassland. Eur. J. Soil Sci. 49, 65-72.

He, Z.L., Zhang, M.K., Stoffella, P.J., Yang, X.E., Banks, D.J., 2006. Phosphorus concentrations and loads in runoff water under crop production. Soil Sci. Soc. Am. J. 70, 1807-1816.

Hengl, T., Heuvelink, G.B.M., Stein, A., 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma 120, 75-93.

Hengl, T., Huevelink, G.B.M., Rossiter, D.G., 2007 About regression-kriging: From equations to case studies. Computers and Geosciences 33, 1301-1315.

Hewitt, A.E., 1993. Predictive modeling in soil survey. Soil and Fertilizers 56, 305-314.

Heuvelink, G.B.M., Webster, R., 2001. Modelling soil variation: past, present, and future. Geoderma 100, 269-301.

Hinsinger, P., 2001. Bioavailability of soil inorganic P in the rhizosphere as affected by rootinduced chemical changes: a review. Plant and Soil 237, 173-195.

Holmberg, M., Forsius, M, Starr, M., Huttunen, M., 2006. An application of artificial neural networks to carbon, nitrogen and phosphorus concentrations in three boreal streams and impacts of climate change. Ecol. Modeling 195, 51-60.

Hopkins, W.G., Huner, N.P. A., 2004. Introduction to plant physiology. John Wiley and Sons, Inc., New York, NY.

Hornsby, D., 2007. Influences on the distribution and occurrence of nitrate-nigrogen and total phosphorus in the water resources of the Suwannee river water management district. Ph.D dissertation, University of Florida, Gainesville, Florida, USA.

Hudson, B.D., 1992. The soil survey as paradigm-based science. Soil Sci. Soc. Am. J. 56, 836-841.

The International Development Research Center (IDRC), 1999. Available at: http://www.idrc.ca/en/ev-83275-201_871023-1-IDRC_ADM_INFO.html. Last verified: Jun. 8, 2011.

Jenny, H., 1941. Factors of soil formation. McGraw-Hill, NY.

Khalid, R.A., Patrick, W.H. Jr., DeLaune, R.D., 1977. Phosphorus sorption characteristics of flooded soils. Soil Sci. Soc. Am. J. 41, 305-310.

Kravchenko, A., Bullock, D.G., 1999. A comparative study of interpolation methods for mapping soil properties. Agron. J. 91, 393-400.

Kuo, S., Mikkelsen, D.S., 1979. Distribution of iron and phosphorus in flooded and unflooded soil profiles and their relation to phosphorus adsorption. Soil Science, 127,18-25.

Kuo, S., Baker, A.S., 1982. The Effect of soil drainage on phosphorus status and availability to corn in long-term manure-amended soils. Soil Sci. Soc. Am. J. 46:744-747.

Kustas, W.P., Li, F., Jackson, T.J., Prueger, J.H., MacPherson, J.I., Wolde, M., 2004. Effects of remote sensing pixel resolution on modeled energy flux variability of croplands in Iowa. Remote Sens. Environ. 92, 535-547

Lamsal, S., Grunwald, S. Bruland, G.L., Bliss, C.M., Comerford, N.B., 2006a. Regional hybrid geospatial modeling of soil nitrate-nitrogen in the Santa Fe River Watershed. Geoderma 135, 233-247.

Latimer, J.C., Filippelli, G.M., 2002. Eocene to Miocene terrigenous inputs and export production: geochemical evidence from ODP Leg 177, Site 1090. Palaecogeography, Palaeoclimatlogy, Palaeoecology 182, 151-164.

Levene, H., 1960. Robust tests for equality of variances. In: Olkin, I. (Ed.), Contributions to probability and statistics: essays in honor of Harold Hotelling. Stanford University Press, Palo Alto, California, pp. 278-292.

Li, Y., 2010. Can the spatial prediction of soil organic matter contents at various sampling scales be improved by using regression kriging with auxiliary information?. Geoderma 159, 63-75.

Lim, H.S., MatJafri, M.Z., Abdullah, K., 2009. Chlorophyll measurement from Landsat TM imagery. Oceans 1-4.

Liu, X., Zhao, K., Xu, j., Zhang, M., Si, B., Wang, F., 2008. Spatial variability of soil organic matter and nutrients in paddy fields at various scales in southeast China. Environ. Geology 53, 1139-1147.

López-Granados, F., Jurado-Exposito, M., Pena-Barragan, J.M., Garcia-Torres, L., 2005. Using geostatistical and remote sensing approaches for mapping soil properties. Euro. J. Agron. 23, 279-289.

Leone, A.P., Wright, G.G., Corves, C., 1995. The application of satellite remote sensing for soil studies in upland areas of Southern Italy. Int. J. Remote Sens. 16, 1087-1105.

Lillesand, T.M., Kiefer, R.W., 2008. Remote sensing and image interpretation (6^{rd} edition). John Wiley and Sons, NY.

Mafongoya, P.L., Giller, K.E., Odee, D., Gathumbi, S.K., Nduga, S.M., Sitompul, S. M., 2004. Benefiting from N_2-Fixation and Managing Rhizobia. In: Noordwijk, M., Cadisch, G., Ong, C.K., (Eds), Below-ground interactions in tropical agroecosystems: concepts and models with multiple plant components. CABI Publishing, Oxfordshire, U.K., pp 227-242.

Mahasandana, S., Tripathi, N.K., Honda, K., 2009. Sea surface multispectral index model for estimating chlorophyll *a* concentration of productive coastal waters in Thailand. Canadian J. Remote Sens. 35, 287-296.

Mander, Ü., Kulla, A., Tammb, V., Kuusemetsa, V., Karjus, R., 1998. Impact of climatic fluctuations and land use change on runoff and nutrient losses in rural landscapes. Landscape and urban planning, 41, 229-238.

Marceau, D.J., Hay, G.J., 1999. Remote sensing contributions to the scale issue. Canadian J. Remote Sens. 25, 357-366.

Marella, R.L., 2009. Water withdrawals, use, and trend in Florida, 2005 and trends 1950-2005: scientific investigations report 2009-5125, U.S. Geological Survey.

Masuoka, P.M., Claborn, D.M., Andre, R.G., Nigro, J., Gordon, S.W, Klein, T.A., Kim, H.C., 2003. Use of IKONOS and Landsat for malaria control in the Republic of Korea. Remote Sens. Environ. 88, 187-194.

McBratney, A.B., De Gruijter, J.J., 1992. A continuum approach to soil classification by modified fuzzy k-means with extragrades. J. Soil Sci. 43, 159-175.

McBratney, A.B., Odeh, I.O.A, Bishop, T.F.A., Dunbar, M.S., Shatar, T.M., 2000. An overview of pedometric techniques for use in soil survey. Geoderma 97, 293-327.

McBratney, A.B., Santos, Mendonça, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3-52.

McKenzie, N.J., and Ryan., P.J., 1999. Spatial prediction of soil properties using environmental correlation. Geoderma 89, 67-94.

Mermut, A.R., Eswaran, H., 2001. Some major developments in soil science since the mid-1960s. Geoderma 100, 403-426.

Mishra, U., Lal, R., Liu, D., Meirvenne, M.V., Predicting the Spatial Variation of the Soil Organic Carbon Pool at a Regional Scale. Soil Sci. Soc., Am. J., 74, 906-914.

Mora-Vallejo, A., Claessens, L., Stoorvogel, J., Heuvelink, G.M.B. Small scale digital soil mapping in southeastern Kenya. Catena 76, 44-53.

Mueller, T.G., Pierce, F.J., 2003. Soil carbon maps: enhancing spatial estimates with simple terrain at tributes at multiple scales. Soil Sci. Soc. Am. J. 67, 258-267.

Mumby, P.J., Edwards, A.J., 2002. Mapping marine environments with IKONOS imagery: enhanced spatial resolution can deliver greater thematic accuracy. Remote Sens. Environ. 82, 248-257.

Murphy, J., J.P Riley., 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chim. Acta 27, 31-36.

Myers, J.L., Well, A.D., 2003. Research design and statistical analysis. Lawrence Erlbaum Associatges, Mahwah, NJ.

Nair, V.D., Portier, K.M., Graetz, D.A., Walker, M.L., 2004. An environment threshold for degree of phosphorus saturation in sandy soils. J. Environ. Qual. 33, 107-113.

National Aeronautics and Space Administration (NASA). Available at: http://geo.arc.nasa.gov/sge/health/sensor/sensors/ikonos.html. Last verified: Jun. 8, 2010.

National Climatic Data Center (NCDC), National Oceanic and Atmospheric Administration (NOAA)., 2008. Monthly surface data. NCDC, Asheville, North Carolina. Available at: http://www.ncdc.noaa.gov/oa/ncdc.html. Last verified: Jun. 26, 2011.

National Elevation Data (NED), United States Geological Survey (USGS)., 1999. Raster layer. Original Scale: 1:24,000. Spatial resolution: 30 m. USGS, Sioux Falls, South Dakota.

Natural Resources Conservation Service (NRCS), U.S. Department of Agriculture (USDA)., 2006. Soil Survey Geographic (SSURGO) database. Original scale: 1:24,000. NRCS, Fort Worth, Texas.

Ng Kee Kwong, K.F, Bholah., A., Volcy. L., Pynee, K., 2002. Nitrogen and phosphorus transport by surface runoff from a silty clay loam soil under sugarcane in the humid tropical environment of Mauritius. Agri., Ecosys. and Environ. J. 91, 147-157

Nour, M.H., Smit, D.W., Gamal El-Din., 2006. Geostatistical mapping of precipitation: implications for rain gauge network design. Water Science and Technology 53, 101-110.

Odeh, I.O.A., McBratney, A.B., Chittleborough, D.J., 1995. Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. Geoderma 67, 215–226.

Oorts, K., Garnier, P., Findeling, A., Mary, B., Richard, G., Nicolardot, B., 2007. Modeling soil carbon and nitrogen dynamics in no-till and conventional tillage using PASTIS model. Soil Sci. Soc. Am. J. 71, 336-346.

Parry, R., 1998. Agricultural phosphorus and water quality: A U.S. Environmental Protection Agency Perspective. J. Environ. Qual. 27, 258-261.

Peng, W., Wheeler, D.B., Bell, J.C., Krusemark, M.G., 2003. Delineating patterns of soil drainage class on bare soils using remote sensing analyses. Geodema 115, 261-279.

Penížek, V., Borůvka, L., 2006. Soil depth prediction supported by primary terrain attributes: a comparison of methods. Plant Soil Environ. 52, 424-430.

Peters, J.M., Basta, N.T., 1996. Reduction of excessive bioavailable phosphorus in soils by using municipal and industrial wastes. J. Environ. Qual. 25, 1236-1241.

Phupaibul, P., Chitbuntanorm, C., Chinoim, N., Kangyawongha, P., Matoh, T., 2004. Phosphorus accumulation in soils and nitrate contamination in underground water under export-oriented asparagus farming in Nong Ngu Lauem village, Nakhon Pathom province, Thailand. Soil Sci. and Plant Nutri. 50, 385-393.

Pierzynski, G.M., Sims, J.T., Vance, G. F., 1994. Soil and environmental quality. CRC Press Inc, Boca Raton, Florida.

Ramnarine, R., 2003. Predicting phosphatic soil distribution in Alachua County, Florida. M.S. Thesis, University of Florida, Gainesville, Florida, USA.

Randazzo, A., Jones, D., 1997. The geology of Florida. University Press of Florida, Gainesville, Florida.

Randall, G., Mulla, D., Rehm, G., Busman, L., Lamb, J., Schmitt, M., 2002. Phosphorus transport and availability in surface waters. University of Minnesota extension service. Pub. FO-6796. Available at: http://www.extension.umn.edu/distribution/cropsystems/DC6796.html. Last verified: Jun. 8, 2011.

Raulston, M., Johnson, C., Webster, K., Purdy, C., Ceryak, R., 1998. Suwannee river water management District. Chapter 10. In: Fernald E.A. and Purdum, E.D (ed). Water Resources Atlas of Florida, Florida State University. Tallahassee, Florida. pp. 194-213.

Raven, P.H., Berg., L.R., 2001. Environment (3rd edition). Wiley. 2001

Rechcigl, J., 1997. Phosphorus – Natural versus pollution levels: Lake Okeechobee case study. In: Proceedings of the 47th annual beef cattle short course. University of Florida. Gainesville, FL. May 7-9, 1997, FL. pp. 1-9.

Regalado, C.M., Ritter, A., 2006. Geostatistical tools for characterizing the spatial variability of soil water repellency parameters in a laurel forest watershed. Soil Sci. Soc. Am. J. 70, 1071-1081.

Rivero, R.G., Grunwald, S., Binford, M.W., Osborne, T.Z., 2009. Integrating spectral indices into prediction models of soil phosphorus in a subtropical wetland. Remote Sens. Environ. 113, 2389-2402.

Rivero, R.G., Grunwald, S., Bruland, G.L., 2007a. Incorporation of spectral data into multivariate geostatistical models to map soil total phosphorus variability in a Florida wetland. Geoderma (Special Issue Pedometrics) 140, 428-433.

Rivero, R.G., Grunwald, S., Osborne, T.Z., Reddy, K.R., Newman, S., 2007b. Characterization of the spatial distribution of soil properties in Water Conservation Area-2A, Everglades, Florida. Soil Sci. 172, 149-166.

Rogan, J., Chen, D., 2004. Remote sensing technology for mapping and monitoring land-cover and land-use change. Progress in Planning 61, 301-325.

Sabesan, A., 2004. Geo-spatial assessment of the impact of land cover dynamics and distribution of land resources on soil and water quality in the Santa Fe river watershed. M.S. Thesis, University of Florida, Gainesville, Florida, USA.

Sandoval-Pérez, A.L., Gavito, M.E., García-Oliva, F.G., Jaramillo, V.J., 2009. Carbon, nitrogen, phosphorus and enzymatic activity under different land uses in a tropical, dry ecosystem. Soil Use and Management 2009, 419-426.

Sawaya, K.E., Olmanson, L.G., Heinert, N..J., Brezonik, P.L., Bauer, M.E., 2003. Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. Remote Sens. Environ, 88, 144-156.

Schaefer, K., Einax, J.W., Simeonov, V., Tsakovski, S., 2010. Geostatistical and multivariate statistical analysis of heavily and manifoldly contaminated soil samples. Analytical and bioanalytical chemistry 396, 2675-2683.

Sharpley, A., B. Foy, P. Withers. 2000. Practical and innovative measures for the control of agricultural phosphorus losses to water: An overview. J. Environ. Qual. 29, 1–9.

Sharpley, A.N., Tiessen, H., Cole, C.V., 1987. Soil phosphorus forms extracted by soil tests as a function of pedogenesis. Soil Sci.Soc. Am. J. 51, 362-365.

Shober, A.L., 2008. Soil and fertilizers for master gardeners: Phosphorus in the home landscape. Institute of Food and Agricultural Sciences Extension, University of Florida, Gainesville, FL. SL 261.

Shonk, J.L., Gaultney, L.D., Schulze, D.G., Van Scoyoc, G.E., 1991. Spectroscopic sensing of soil organic matter content. Trans. ASAE 34, 1978-1984.

Silgram, M., Jackson, D.R., Bailey, A., Quinton, J., Stevens, C., 2010. Hillslope scale surface runoff, sediment and nutrient losses associated with tramline wheelings. Earth Surf. Process. Landforms 35, 699-706.

Soil Science Society of America (SSSA), 1997. Glossary of soil science terms. Madison, Wisconsin. Available at: http://soils.usda.gov/education/facts/soil.html. Last verified: Jun. 8, 2010.

Soil Taxonomy, 1999. A Basic System of Soil classification for Making and Interpreting Soil Surveys (2nd edition). US Department of Agriculture (USDA). Washington, D.C.

Solaimani, K., Shokrian, F., Tamartash, R., Banihashemi, M., 2011. Landsat ETM+ based assessment of vegetation indices in highland environment. J. Advances in Developmental Research 2, 5-13.

Solim, S.U., Wanganeo, A., 2008. Excessive phosphorus loading to Dal Lake, India: Implications for managing shallow eutrophic lakes in urbanized watersheds. Int. Review of Hydrobiology 93, 148-166.

Somenahally, A. Weindorf, D.C., Darilek, L., Muir, J.P. Wittie, R. Thompson, C., Morgan, C.L.S., 2009. Spatial variability of soil test phosphorus in manure-amended soils on three dairy farms in North Central Texas. J. Soil and Water Conserv. 64, 89-97.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effect?. Remote Sens. Environ. 75, 230-244

Sridhar, B.B.M., Vincent, R.K., Witter, J.S., Spongberg, A.L., 2009. Mapping the total phosphorus concentration of biosolid amended surface soils using LANDSAT TM data. Science of the Total Environment 407, 2894-2899.

Stefanov, W.L., Netzband, M., 2005. Assessment of ASTER land cover and MODIS NDVI data at multiple scales for ecological characterization of an arid urban center. Remote Sens. Environ. 99, 31-43.

Stoner, N.K., 2011. Working in partnership with States to address phosphorus and nitrogen pollution through use of a framework for state nutrient reductions. Memorandum. United States Environmental Protection Agency. Washington, D.C.

Stys, B., Kautz, R., Reed, D., Kertis, M., Kawula, R., Keller, C., Davis, A., 2004. Florida vegetation and land cover data derived from 2003 Landsat ETM+ imagery. Florida Fish and Wildlife Conservation Commission, Tallahassee, Florida

Sullivan, D.G., Shaw, J.N., Rickman, D., 2005. IKONOS imagery to estimate surface soil property variability in two Alabama physiographies. Soil. Sci. Soc. Am. J. 69, 1789-1798.

Suwannee River Water Management District (SRWMD), 2002a. Surface water quality and biological annual report. Department of water resources. WR-01-02-04.

Suwannee River Water Management District (SRWMD), 2002b. WARN water monitoring atlas. Department of water resources. WR-02-02.

Suwannee River Water Management District (SRWMD), 2002c. Nitrate profile of the lower Santa Fe River. Department of water resources. WR-02-03.

Suwannee River Water Management District (SRWMD), 2003a. Ground water quality annual report. Department of water resources. WR-02/03-02.

Suwannee River Water Management District (SRWMD), 2003b. Surface water quality and biological annual report. Department of water resources. WR-02/03-03.

Suwannee River Water Management District (SRWMD), 2004. Surface water quality and biological annual report. Department of water resources. WR-04/05-02.

Suzumura, M., Ueda, S., Sumi, E., 2000. Control of phosphate concentration through adsorption and desorption processes in groundwater and seawater mixing at sandy beaches in Tokyo Bay, Japan. J. of Oceanography. 56, 667-673.

Tian, Y., Woodcock, C.E., Wang, Y., Privette, J.L., Shabanov, N.V., Zhou, L., Zhang, Y., Buermann, W., Dong, J., Veikkanen, B., Hame, T., Andersson, K., Ozdogan, M., Knyazikhin, Y., Myneni, R.B., 2002. Multiscale analysis and validation of MODIS LAI product sampling strategy. Remote Sens. Environ. 83, 431-441.

Triantafilis, J., Odeh, I.O.A., McBratney, A.B., 2001. Five geostatistical models to predict soil salinity from electromagnetic induction data across irrigated cotton. Soil Sci. Soc. Am. J. 65, 869-878.

Tyrell, T., 1999. The relative influence of nitrogen and phosphorus on oceanic primary production. Nature 400, 525-531.

United States Environmental Protection Agency (USEPA), 1986. Quality criteria for water. EPA-440/5-86-001, Office of water regulations and standards, Washington DC.

United States Environmental Protection Agency (USEPA), 2005. Nutrient water quality requirements. Available at: http://www.epa.gov/waterscience/criteria/nutrient/ecoregions. Last verified: Jun. 8, 2011.

United States Geological Survey (USGS), 1999. National Elevation Dataset (NED). Raster layer. Original Scale: 1:24,000. Spatial resolution: 30 m. USGS, Sioux Falls, South Dakota.

Vaithiyanathan, P., Correll, D.L., 1992. The Rhode river watershed: Phosphorus distribution and export in forest and agricultural soils. J. Environ. Qual., 21, 280-288.

Vasques, G.M., Grunwald, S., Sickman, J.O., 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma 146, 14-25.

Vasques, G.M., Grunwald, S., Sickman, J.O., 2009. Visible/near-infrared spectroscopy modeling of dynamic soil carbon fractions. Soil Sci. Soc. Am. J. 73, 176-184.

Vasques, G.M., Grunwald, S., Comerford, N.B., Sickman, J.O., 2010. Regional modeling of soil carbon at multiple depths within a subtropical watershed. Geoderma 156:326-336.

Verfaillie, E., Lancker, V.V., Meirvenne, V., 2006. Multivariate geostatistics for the predictive modeling of surficial sand distribution in shelf seas. Continuental shelf research. 26, 2454-2468.

Villapando, R.R., Graetz, D.A., 2001. Phosphorus sorption and desorption properties of the spodic horizon from selected Florida spodosols. Soil Sci. Soc. Am. J. 65, 331-339

Wackernagel, H., 2003. Multivariate geostatistics: An introduction with applications (3rd edition). Springer-Verlag, Berlin, Germany.

Wang, Y., Zhnag, X., Huang, C., 2009. Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the Loess Plateau, China. Geoderma 150, 141-149.

Webster, R., Oliver, M.A., 2005. Modeling spatial variation of soil as random functions., In:Grunwald S.(Ed.), Environmental soil-landscape modeling –geographic information technologies and pedometrics, CRC Press, Boca Raton, Florida. pp.241-287.

Webster, R., Oliver, M.A., 2007. Geostatistics for environmental scientists (2nd edition). John Wiley and Sons, Chichester, United Kingdom.

Weintraub, M.N., Scott-Denton, L.E., Schmidt, S.K., and Monson, R. K., 2007. The effects of tree rhizodeposition on soil exoenyme activitiy, dissolved organic carbon, and nutrient availability in a subalpine forest ecosystem. Oecologia 154, 327-338.

White, P.J., Hammond, J.P., 2006. Updating the estimates of the sources of phosphorus in UK waters. Defra final report for project WT0701CSF. Defra, London. U.K.

White, R.E., 2006. Principles and practice of soil science: the soil as a natural resource. Balckwell Science Ltd., Malden, MA, USA.

Williams, P.C., 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In: Norris, K., Williams, P. (Eds.), Near-infrared Technology in the Agricultural and Food Industries. American Association of Cereal Chemists, St. Paul, MN, 143–167.

Wu, C., Wu, J., Luo, Y., Zhang, L., DeGloria, S.D., 2009. Spatial prediction of soil organic matter content using cokriging with remotely sensed data. Soil Sci. Soc. Am. J. 73, 1202-1208.

Wu, W., Fan, Y., Wang, Z., Liu, H., 2008. Assessing effects of digital elevation model resolutions on soil-landscape correlations in a hilly area. Agriculture, Ecosystems and Environment 126, 209-216.

Yaalon, D.H., 2000. Down to earth: why soil- and soil science- matters. Millennium essay, Nature 407, 301p.

Zhang, Y., Zhao, Y.C., Shi, X.Z., Lu, X.X., Yu, D.S., Wang, H.J., Sun, W.X., Darilek, J.L., 2008. Variation of soil organic carbon estimates in mountain regions: a case study from Southwest China. Geoderma 146, 449-456.

BIOGRAPHICAL SKETCH

Jinseok Hong was born and raised in Daegu, South Korea. He had an opportunity to explore new culture for one year each at the University of California at Davis in 1995 as well as the University of Nebraska in Lincoln, Nebraska in 1998. He earned a Bachelor of Engineering degree in 1999 from Keimyung University in Daegu, South Korea and a Master of Science degree in 2003 from the University of Florida in Gainesville, Florida. He worked for the Center for Watershed Management as an Engineer Intern in Sarasota County, Florida in 2005. He continued to pursue a Doctor of Philosophy degree at the University of Florida while he worked in the GIS laboratory, Soil and Water Science Department at the University of Florida. He married Eunkyung Lee in 2005 and has a daughter, Ellen L. Hong. He is currently teaching courses related to Geomatics as an Assistant Professor at the Department of Engineering Technology, Surveying, and Digital Media at East Tennessee State University, Johnson City, Tennessee.