Prepared for

SOUTH FLORIDA WATER MANAGEMENT DISTRICT

ANALYSIS OF EVERGLADES AGRICULTURAL AREA (EAA) FARM DATA TO IDENTIFY AREAS OF OPPORTUNITY FOR WATER QUALITY IMPROVEMENT THROUGH SOURCE CONTROLS

Contract No: PC P501742

FINAL ANALYSIS REPORT

STATISTICAL ANALYSIS ON EAA FARM WATER QUALITY DATA

(Deliverable 4.1)

Prepared by

Stanley Consultants Inc.
1601 Belvedere Road, Suite 400
West Palm Beach, FL 33406

ZFI Engineering & Construction Inc. (formerly Z-Facilitators Inc.)
3251 Progress Dr. Suite A
Orlando, FL 32826

March, 2006

Client

SOUTH FLORIDA WATER MANAGEMENT DISTRICT

Project

ANALYSIS OF EVERGLADES AGRICULTURAL AREA (EAA) FARM DATA TO IDENTIFY AREAS OF OPPORTUNITY FOR WATER QUALITY IMPROVEMENT THROUGH SOURCE CONTROLS

Contract No.	*ZFI Project No.*
PC P501742	**PJ0423 - 03**

Date of Issue

March 17, 2006

Report Title

Final Analysis Report

Statistical Analysis on EAA-Farm Water Quality Data

(deliverable 4.1)

Consultants

Stanley Consultants Inc.

ZFI Engineering & Construction Inc.

Distribution

Copy No	*Issued to*
Copy 1-3	South Florida Water Management District
Copy 4	Stanley Consultants Inc.
Copy 5	ZFI Engineering & Construction Inc.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The purpose of this report is to respond to directives of the Everglades Forever Act (EFA) and to public inquiries on regulatory data. During the 2003 legislative session, the 1994 EFA was amended to include reference to the March 17, 2003, Conceptual Plan for Achieving Long-term Water Quality Goals (Long-Term Plan). Although the Long-Term Plan for the Everglades Agricultural Area (EAA) recognized that the combined performance of the EAA BMP regulatory program and the STAs has exceeded expectations, supplemental adaptive management measures were identified to ultimately achieve water quality goals in the EPA. Additionally, in recent years, the District has received multiple inquiries from the public on what the 10 years of regulatory farm level data says about differences in phosphorus concentration and load between farms and between EAA sub-basins. The intended use of this data was not to make comparisons between farms or EAA sub-basins. Until now, a collective analysis of Everglades Agricultural Area (EAA) farm level regulatory data had not been conducted.

In response to the Long-term Plan requirements and in an effort to respond to the public's concerns, this exploratory statistical analysis on the existing farm level regulatory data was initiated. The primary objective of the analysis was to glean as much information as possible from the data relative to identifying high-level relationships and "screening tools" for BMP Program optimization opportunities. The District commissioned Stanley Consultants, Inc., in association with Z-Facilitators Inc. to perform this analysis. The analysis was then reviewed by EAA agricultural BMP experts to determine practical application of results (quoted comments in Appendix C). This Executive Summary consolidates the statistical analysis findings and the agricultural/BMP considerations as noted by the experts.

The scope of work required statistical analysis of specified data and interpretation of its results towards responding to specific, and somewhat ambitious, questions designed to cull out relationships that may exist. Within each section, assumptions and limitations associated with the dataset and scope were pointed out. The questions are consolidated into six major topics:

- The Seasonality analysis (section 3) depicts the seasonal and long term variations of farm P loads and concentrations per EAA sub-basin and also by characteristic groups of land use, soil type, or water detention levels within each sub-basin;
- The Location analysis (section 4) looks for phosphorus load and concentration relationships between adjacent structures along a canal;
- The Potential Hydraulic Interconnection analysis (section 5) provides for a preliminary screening of farms that could establish hydraulic connections based on proximity and discharge characteristics;
- The Inflow analysis (section 6) focuses on whether statistical correlations could identify relationships between inflows from Lake Okeechobee and EAA farm discharges;
- The Spikes analysis (section 7) identifies high load discharge events and evaluates

their potential impact on total discharges per EAA sub-basin and also by characteristic groups of land use, soil type, or water detention levels within each sub-basin; and

- The Detention analysis (section 8) explores relationships between rainfall and runoff and between rainfall and P loads in farm discharges for multiple characteristic groups of land use, soil type, and water detention levels.

The analysis area consists of sub-basins S3 and S8 along the Miami Canal, and sub-basin S5A along the West Palm Beach Canal. The basic dataset consists of the permittee-reported farm level data recorded from Water Year (WY) 1995 to WY2003 for the farms located in the analysis area, and farm information documented in District files. The effect of farm characteristics was evaluated based on four different categories: land use, soil type, EAA sub-basin location, and water detention level. Within each category, characteristics included three sub-basin locations, seven land uses, five soil types, and three water detention levels. The individual structure datasets (e.g., daily for flow and up to 3–week composites for concentration) were aggregated as necessary to meet the scope objectives for each of the analyses. Changes in the farm characteristic groupings were tracked on an annual basis.

There were various limitations associated with the use of categories for the analysis. In contrast with the farm level water quality data which is routinely and systematically collected, the site specific farm information is a compilation of best available information from District files. Some of this information is collected on an as-needed basis. However, despite these limitations, these categories were included in the exploratory analysis on the basis that pointing out the inherent differences among farms —based on best available information and noting the limitations of the sources— offers better grounds for understanding and interpreting the data than not considering them at all.

The normality of the datasets was tested prior to selecting statistical methods. Test results indicated the significant skewness of most of the time series of P concentration and load. Therefore, non-parametric methods were used for the analyses. Table 2-3 describes the specific tests used.

As anticipated, analysis results provide general high-level “screening tool” type findings. There are several constraints, mainly due to data availability and precision that limit the usefulness of the results. The specific conclusions indicated below verify assumptions associated with farm level discharges (seasonality analysis results), serve as a high level screening tool for more detailed analysis (potential interconnection and spikes analysis results), or identify supplemental data or refined analyses that would provide better results (location, inflow and detention analyses):

Seasonality

- The decomposition analysis indicates that farm discharges in the S3 and S8 sub-basins as a whole show either moderate reduction or stabilization in their P concentrations and unit loads. S5A, however, shows a significant decrease in concentration and a slight increase in load, suggesting an increase in flow. The long-term trend indicates that S5A concentrations at the beginning of the series were significantly higher than in the other sub-basins. The long-term trends for the characteristic groupings (soils, land uses and water detention levels) show moderate variation over years consistent with the sub-basin results.

- There is qualitative existence and consistency (in term of cyclic variation phases) of the seasonality changes in the time series. The seasonal variations in P load and concentration were found to depend on multiple characteristics.
- Analysis of flow trends as a function of land use, soil type, or water detention levels, was not included in the scope of work for this analysis. Its evaluation, however, may serve to better understand the separate effect of flow and concentration on P load trends. Use of multiple characteristic groupings may serve to understand the combined effect of various characteristics. Use of stepwise functions can be applied to the dataset, when it is suspected that sudden changes may have affected P loads and flows (e.g., removal of land to construct a stormwater treatment area or STA).

Location

- The location analysis did not find meaningful correlations between the discharge loads and concentrations of adjacent structures. These concentrations, and the loads derived from them, only represent the water quality conditions at the outlet, on the upstream side of the structure. A major constraint for the analysis is assumed to be the lack of farm irrigation information and canal water quality data (irrigation schedules, flows, water quality). This information is not collected and, thus, it is not available for analysis. The "no correlation" finding cannot be assumed as definite since the information available to perform the analysis does not represent the complete hydrology and phosphorus transport processes that would explain the potential correlation.
- Evaluation through a site-specific analysis is likely to be more appropriate to address this section of the scope of work.

Potential Hydraulic Interconnection

- The practice of interconnecting drainage systems of multiple farms is a practice that already exists in the EAA where farms are typically same-owner farms. The first level screening-type analysis identified farms that may be candidates for hydraulic interconnection.
- Analysis of the many practical constraints for connecting was outside the analysis scope. Site specific evaluation of critical factors such as ownership, crop type and water retention capacity to determine the feasibility of these arrangements is

needed. Site-specific studies may also be necessary to assess how runoff diversion and intermixing could serve to improve water quality.

- An assessment of land uses and water management practices among adjacent farms would benefit this analysis.

Lake Okeechobee Inflows

- Limited farm irrigation information and canal water quality data (irrigation schedules, flows, water quality), and reduced Lake Okeechobee inflow information in comparison to Sub-area discharge data were attributed as the major reasons for the lack of correlation between Lake Okeechobee inflows and discharges from farms downstream. Use of daily averages for each month and data aggregated by sub-areas may also have simplified the analysis such that correlations could not be found.

- Different aggregation and lag techniques to test the correlation of data may improve the analysis method.

Spikes

- Days with high P load and concentration discharges (defined as spikes for this analysis) were identified for all characteristic groups. These spike events can be associated with higher discharge pumping, higher P concentration or both. The available data does not provide the level of detail to identify the specific causes behind these discharges or whether they are avoidable. An estimate of the phosphorus contribution of spike events indicated that they represent approximately half of the phosphorus loads discharged in a year, and that these events occur within relatively short time spans (one to two days). Spikes occurred both during the dry and wet season. Spikes occurring during the dry season have a larger contribution to the total seasonal loads, while wet season spikes are higher.

- It is recommended that farm level spike information be provided to permittees for self assessment and understanding of the effect of spike events on annual loadings. This information will be useful for the permittee to identify the causes for these events and adjust farming practices to optimize BMPs as applicable.

Detention

- Rainfall-runoff and rainfall-P load curves were derived using a simplified standard rationale based on rainfall events (i.e., consecutive rainfall days). Correlations between rainfall and farm discharge flows were identified, thus providing a general relationship between the criteria defined rainfall-event and the resulting runoff for multiple characteristic groupings. However, no direct relationship between rainfall and P loads could be verified through this analysis. Therefore, the specific effect of rainfall detention on phosphorus concentration and load could not be evaluated to address the scope of work.

- A site-specific evaluation including detailed field data collection may be necessary. Use of modeling was also suggested as a tool to more specifically understand the effects of detaining runoff on flows and loads at the farm level. Information on water table levels was not available for the analysis and is necessary for the type of assessment described by the scope.

Analysis findings support expectations associated with EAA farm discharges that had previously only been assumed and not based on actual review of data. The specific causes behind the P load and concentration relationships presented in this analysis, however, cannot be identified with certainty because of the lack of information on the many factors affecting the characteristics of farm discharges. Use of more refined datasets is necessary. It is recommended that suggestions for more refined analyses be provided to the University of Florida IFAS for consideration in their BMP research efforts and data analysis (using BMP research data collected by IFAS at the farm level), or for additional field investigation. The analyses may fit within the objectives of the EFA-mandated BMP farm research that is conducted by IFAS under the Master Research Permit (issued as a requirement of Chapter 40E-63, F.A.C.) to the EAA-EPD.

1.0 INTRODUCTION

1.1 Study Background

Florida's 1994 Everglades Forever Act (EFA), F.S. 373.4592, established long-term water quality goals designed to restore and protect the Everglades Protection Area (EPA). As defined in the Act, the EPA includes Water Conservation Areas (WCAs) 1, 2A, 2B, 3A, 3B, the Arthur R. Marshall Loxahatchee National Wildlife Refuge, and the Everglades National Park. A primary component of the EFA is the Everglades Construction Project which includes a combination of phosphorus source control programs using mandatory best management practices (BMPs) and downstream treatment within manmade stormwater treatment areas (STAs).

During the 2003 legislative session, the 1994 EFA was amended to include reference to the March 17, 2003, Conceptual Plan for Achieving Long-term Water Quality Goals (Long-Term Plan). Although the Long-Term Plan for the Everglades Agricultural Area (EAA) recognized that the combined performance of the EAA BMP regulatory program and the STAs has exceeded expectations, supplemental adaptive management measures were identified to ultimately achieve water quality goals in the EPA. Accordingly, the Process Development and Engineering (PDE) component of the Long-Term Plan's overall water quality improvement strategy directed activities and funds towards "Identifying opportunities to maintain and improve upon the performance of source controls (BMPs) in reducing overall pollutant loads[1]" discharging from specific basins, including the EAA.

Under the EAA BMP regulatory program, permittees are required to monitor individual discharges to District canals for phosphorus concentrations and loads. This regulatory data is intended to be used in a secondary compliance methodology should the EAA as a whole be determined to be out of compliance. Because the EAA has been in compliance and has in fact exceeded the water quality goals for the BMP program since its inception, the secondary farm level regulatory data and compliance methods have not been utilized. Both the EAA basin-level (primary) and the individual permit-level (secondary) data are published annually in the South Florida Environmental Report (SFER).

In recent years, in response to the published data, the District has received public inquiries on why differences in phosphorus concentration and load exist between farms in the EAA and between sub-basins within the EAA. In the SFER[2], the District has reported on the limited usefulness of the permit-level regulatory data for making such comparisons. There are many complexities associated with water quality discharged from the farms to EAA canals and recycled back onto farms for irrigation. These complexities encompass physical, chemical and biological processes. Soil characteristics, timing and distribution of rainfall, cropping patterns, geographic location, and hydrology are also factors that vary throughout the area and affect the characteristics of farm

[1] Long-Term Plan page 5-2.
[2] 2006 South Florida Environmental Report, Chapter 3, EAA Permit-Level Monitoring Results

discharges. Specific information on these factors is not part of the individual permit-level data set. For these reasons, the regulatory permit-level data set has been most useful for making relative comparisons between water years for the same farm to provide advice to the permittee on BMP optimization opportunities for their particular farm. Until this analysis was commissioned, the data set as a whole had not been analyzed.

Therefore, in recognition of the Long-Term Plan requirement to identify additional opportunities for water quality improvement and to further address public inquiries, this exploratory statistical analysis of the existing permit-level data was conducted. The South Florida Water Management District (District) commissioned Stanley Consultants, Inc. in association with Z-Facilitators Inc. to perform statistical analysis of the EAA permit (farm) level regulatory data for the purpose of gleaning as much information from the data set as possible relative to optimization of the existing regulatory program. The statistical assessment was followed by a technical review by EAA agricultural BMP experts to determine practical application of results. This document constitutes the final report summarizing the statistical analysis findings and the agricultural/BMP considerations noted by the experts.

1.2 Analysis Objectives and Scope

It is important to emphasize that the objective is to complete an exploratory statistical analysis based on the limited permit level regulatory data and farm information collected as part of the regulatory program. The analysis is not attempting to use the regulatory data to do a research-type analysis or compliance assessment. Because of the distinct limitations of the regulatory data set and the complex nature of the drainage system, several assumptions had to be made. These assumptions were considered appropriate for this analysis because the primary goal of the scope was to take the existing unrefined data set and identify high-level relationships, if they exist. The analysis does not intend to identify new BMPs or to evaluate specific occurrences at individual farms. Additionally, because of constraints on time and funds, the analysis data set only includes the farms located within the EAA sub-basins S-3, S-8 and S-5A.

It is the intent of this scope to perform an exploratory analysis to identify relationships, if they exist, and to utilize the results as a high-level “screening tool” to identify opportunities for further improvement of the source control program in the EAA. Expectations on how the information in this report will be used shall be made clear up front: It is not possible to make a compliance determination with the limited scope of the data set and the EFA establishes that permittees within the EAA that are in full compliance with Chapter 40E-63, F.A.C., requirements —as EAA permittees are— shall not be *required* to implement additional water quality improvement measures. It is also not possible to perform a research-type analysis with the limited scope of the dataset. The University of Florida Institute of Food and Agricultural Services (IFAS) is currently the lead agency in conducting BMP research in the EAA.

The scope of work required statistical analysis of specified data and interpretation of its results towards responding to specific and somewhat ambitious questions designed to cull out relationships that may exist. The questions were consolidated into six major topics:

- The Seasonality analysis (section 3) depicts the seasonal and long term variations of farm P loads and concentrations per EAA sub-basin and also by characteristic groups of land use, soil type, or water detention levels within each sub-basin;
- The Location analysis (section 4) looks for phosphorus load and concentration relationships between adjacent structures along a canal;
- The Potential Hydraulic Interconnection analysis (section 5) provides for a preliminary screening of farms that could establish hydraulic connections based on proximity and discharge characteristics;
- The Inflow analysis (section 6) focuses on whether statistical correlations could identify relationships between inflows from Lake Okeechobee and EAA farm discharges;
- The Spikes analysis (section 7) identifies high load discharge events and evaluates their potential impact on total discharges per EAA sub-basin and also by characteristic groups of land use, soil type, or water detention levels within each sub-basin; and
- The Detention analysis (section 8) explores relationships between rainfall and runoff and between rainfall and P loads in farm discharges for multiple characteristic groups of land use, soil type, and water detention levels.
- The Report includes a discussion of each major topic, a description of the datasets, data preparation and statistical methods, and conclusions and recommendations. The graphs and tables supporting the analysis are included in Appendix A, a detailed discussion of the statistical methods is presented in Appendix B, and the quoted comments from the technical reviewers are included in Appendix C.

2.0 THE PROJECT DATASETS AND STATISTICAL ANALYSIS APPROACH

2.1 The Analysis Area

The analysis area consists of sub-basins S3 and S8 along the Miami Canal, and sub-basin S5A along the West Palm Beach Canal. Flow through the Miami Canal is controlled upstream by the EAA structure S-354 (named S3IN in this analysis) and controlled downstream by S-8 (S8OUT). Flow through the West Palm Beach Canal is controlled upstream by Station S-352 (S5AIN) and controlled downstream by Station S-5A (S5AOUT). A map depicting the analysis area, major District works, and basin wide monitoring locations is shown in Figure 2-1.

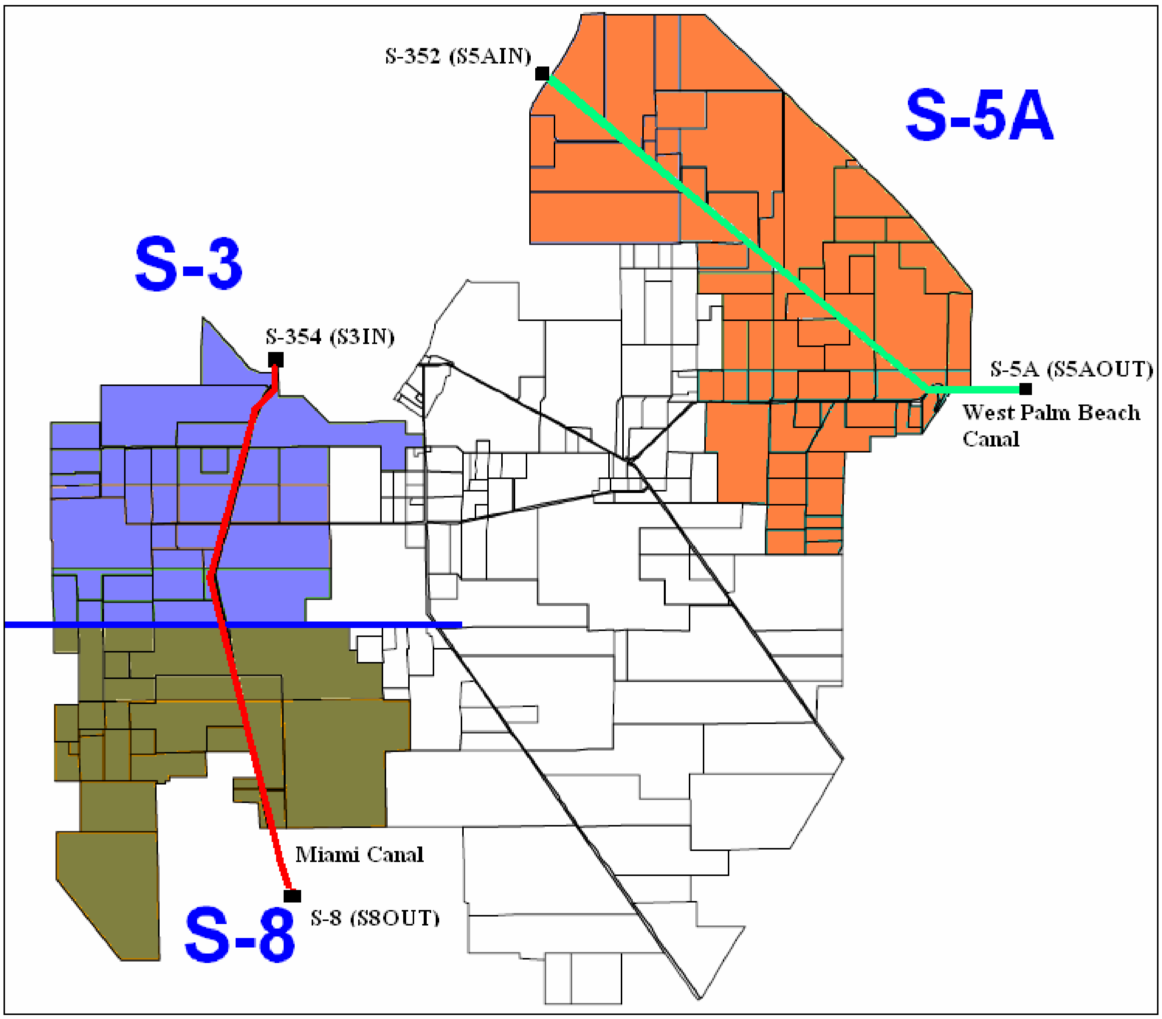


Figure 2- 1: Analysis area, major District works, and basin wide monitoring locations

2.2 The Project Datasets

The scope of work specifies the data, period of analysis, and reference documents that should be used for the development of this project. The scope of work does not include collection of additional data. The purpose of the analysis was to conduct statistical analysis of the permit level regulatory data. This report will note when the scope-specified dataset places constraints on the analysis or findings and will make recommendations on how the data set or analysis might be expanded or refined. The scope of work, for instance, concentrates on the analysis of phosphorus loads and concentrations. Analysis of flow was not included in the scope; however, its analysis could serve to separate the effects of flow and concentration from loads. This is an opportunity for refinement. Nevertheless, all results presented in this report are only based on the specified regulatory datasets as directed in the scope.

The basic dataset consists of the permittee-reported farm level data recorded from Water Year (WY) 1995 to WY2003 for the farms located in the analysis area, and farm information documented in permit files (characteristics of sub-basin location, land use, soil type and water detention levels). The period of data analyzed reflects the complete WY data that were available at the time this analysis was initiated. For the analysis on the potential effects of Lake Okeechobee inflows, water quality and flow data at Lake Okeechobee inflow structures to S3 and S5A, and EAA basin discharge structures from these sub-basins to the Everglades were also used. Table 2-1 provides a description of each data set, and the analyses for which they were used.

Table 2-1: Project Dataset

Datasets	Format	Coverage[1]	Application	Data Source
Farm-level daily discharge flows	Time series	per structure	Used for all analyses: Seasonality Location Potential Interconnection Spikes Inflow Detention	Data are reported by permittees, and reviewed by the SFWMD
Farm-level composite P concentrations	Time series[2]	per monitored structure		
Farm-level daily P loads calculated from the daily discharge and P composites	Time series	per structure		
Farm-level daily rain gage readings	Time series	per structure	Detention	
Daily flow and P concentration at EAA basin wide monitoring structures	Time series	1993-2003 at EAA structures S352, S354, S5A, and S8	Inflow	DBHydro Database reported by the SFWMD
Farm structure locations, and boundaries of farm and EAA sub-basins	GIS shape files	1996-2003	Location and Inflow analysis Structure aggregation	SFWMD
Discharge structures and the water quality monitoring stations per farm[3]	Access Database	per farm	Structure aggregation, Loading computation	SFWMD
Soil types, land uses, and BMPs[4]	Spreadsheet	per farm	All analyses	SFWMD

1. Coverage is the entire analysis area (EAA sub-basins S3, S8, and S5A) during the period WY1995 to WY2003, unless indicated above.
2. Observed composite periods ranged between 6 and 21 days. Composite concentrations were assumed representative of each day of the composite period for purposes of this analysis (composites are based on flow-weighted or time-weighted collection methods).
3. The structure and water quality monitoring arrangement for each farm may change from year to year. These changes were tracked and incorporated in the analysis.
4. Land use data were manually compiled through hardcopy file review of annual reports submitted by permittees. Changes in land use and BMPs during the period were incorporated in the analysis as described in section 2.3.3

2.2.1 Hierarchy of datasets provided by the District

Daily flow and composite P concentrations are measured at farm structures. Based on hydrology, land ownership, and farming activity, these structures are grouped into farms, unit areas, and sub-basins. The hierarchy of data structure is described by the following:

- Sub-basin – denotes drainage basins within the EAA region that are associated with individual storm water treatment areas (STAs). Example: Sub-basin S5A;
- UAID (Unit Area ID) – denotes one or more farms utilizing the same structure or structures to represent the water quality data associated with the farm Unit Area ID. Example: UAID 113 which is located in S5A;
- GISID (or Farm ID) – identifies a parcel of land typically owned or operated by one farm. One farm may contain several structures. The Unit Areas are delineated based on the hydrological conditions. Example: GISID=50-064-01, which is part of UAID 113.
- Structure ID - used to identify every discharging structure with fixed geographic coordinates in the field. Structures may be turned on and off, and may be aggregated in data processing. Individual farms may contain several structures. Example: structure IDs OC04.1TS-K and OC04.1TS-L, which belong to GISID 50-064-01.

2.3 Data Set Aggregation and Numeric Methods

The scope of work poses several questions associated with structure level discharges, farm discharges, and of discharges from groups of farms by characteristic categories. The scope of work also asks for comparisons of discharges from different months and seasons. Responses to these inquiries are based on the analysis of general trends from groups of farms. Evaluating specific occurrences or individual farm tendencies in the farm datasets is not within the scope of the analysis. General relationships at the structure and farm levels were analyzed if specifically required by the scope (e.g., Location and Potential Interconnection analyses, respectively). Because the main datasets are available at the farm structure level for relatively short periods (e.g., daily for flow or 3–week composites for concentration), the datasets were aggregated to conduct the group-level statistical analysis at the requested time frequencies (monthly, annually, seasonally, etc.). The basis for farm aggregation and numeric methods used for the different analyses is presented below.

2.3.1 Structure – Farm aggregation

Flow and phosphorus concentration data were provided by the District for 125 water quality monitoring structures located within the analysis area. These 125 structures were used by 102 farms discharging runoff to EAA canals. The farm-structure aggregation arrangement changed from year to year during the analysis period. For example all 125 structures in sub-basins S3, S8 and S5A belonged to 98 farms in WY1998, while the same number of structures belonged to 84 farms in WY2000. These changes to the structure-farm arrangement are incorporated into the analysis. The aggregated farm-level time series (discharge, phosphorus loads and concentrations) is the basic dataset for the analysis.

2.3.2 Area-normalized P loads

Flow and load data are normalized by farm size in order to compute comparable unit values that can be aggregated among farms. The original datasets included some negative flow values. These invalid values were excluded from the analysis.

$$Load_{farm,date} = \sum_{i=1}^{N} structure\_load_{i,date} \quad (2\text{-}1a)$$

$$Flow_{farm,date} = \sum_{i=1}^{N} structure\_flow_{i,date} \quad (2\text{-}1b)$$

$$Concentration = \begin{cases} 0 & \text{if flow}_{\text{farm}} = 0 \text{ for each data} \\ \dfrac{Load_{farm}}{Flow_{farm}} \times 119.8 & \text{if flow}_{\text{farm}} > 0 \text{ for each data} \end{cases} \quad (2\text{-}1c)$$

$$Unit\_load = Load / Farm\_Area$$
$$Unit\_flow = Flow / Farm\_Area \quad (2\text{-}1d)$$

Where: N = the number of structures of farm.
Load = lbs/day
Flow = mgd
Concentration = ppb
119.8 = unit conversion factor
Unit_load = lbs/day-acre
Unit_flow = mgd/acre
Farm-area = acres

2.3.3 Farm categorization

The effect of farm characteristics was evaluated based on four different categories: land use, soil type, location, and water detention level. These categories reflect the effect of specific land features and farming activities on farm discharges. Within each category, characteristics included three sub-basin locations, seven land uses, five soil types, and three water detention levels. A characteristic from each of these categories was assigned to each farm within the analysis area. The farm time series (daily flow, phosphorus loads and concentrations) were grouped into single or multiple characteristic groups based on the analysis. Characteristic time series were compared within the categories to identify whether differences in P loads and concentration (as required by the analyses) existed and how they relate to each other.

Table 2-2: Farm Categories and Characteristic Groups Assigned to Farms

Category	Characteristic Groups
Location	S3, S8, or S5A
Soil Type	Sands, Dania, Lauderhill, Pahokee, or Terra ceia
Water Detention Level	0.5", 1.0", or $\geq$ 1.5"
Land uses	A: Sugarcane with minimal rotation or other uses B: Sugarcane in rotation with corn, or some parcels dedicated to corn C: If the farm is either: (1) a sugarcane field in rotation with rice or fallow flooded parcels, (2) an area containing sugarcane fields and urban land, or (3) an area containing sugarcane and pasture lands. D: Areas with sugarcane and sod, or primarily sod E: Citrus groves, tree farms, fruit tree, or combinations appear to prevail F: Sugarcane and vegetables, vegetables, or areas where vegetables appear to prevail G: Agricultural lands with industrial uses in it (e.g., mills)

As an example, the spatial distributions of land use, soil type and water detention levels across farms in the analysis area during WY1998 are shown in Figures 2-2, 2-3, and 2-4, respectively.

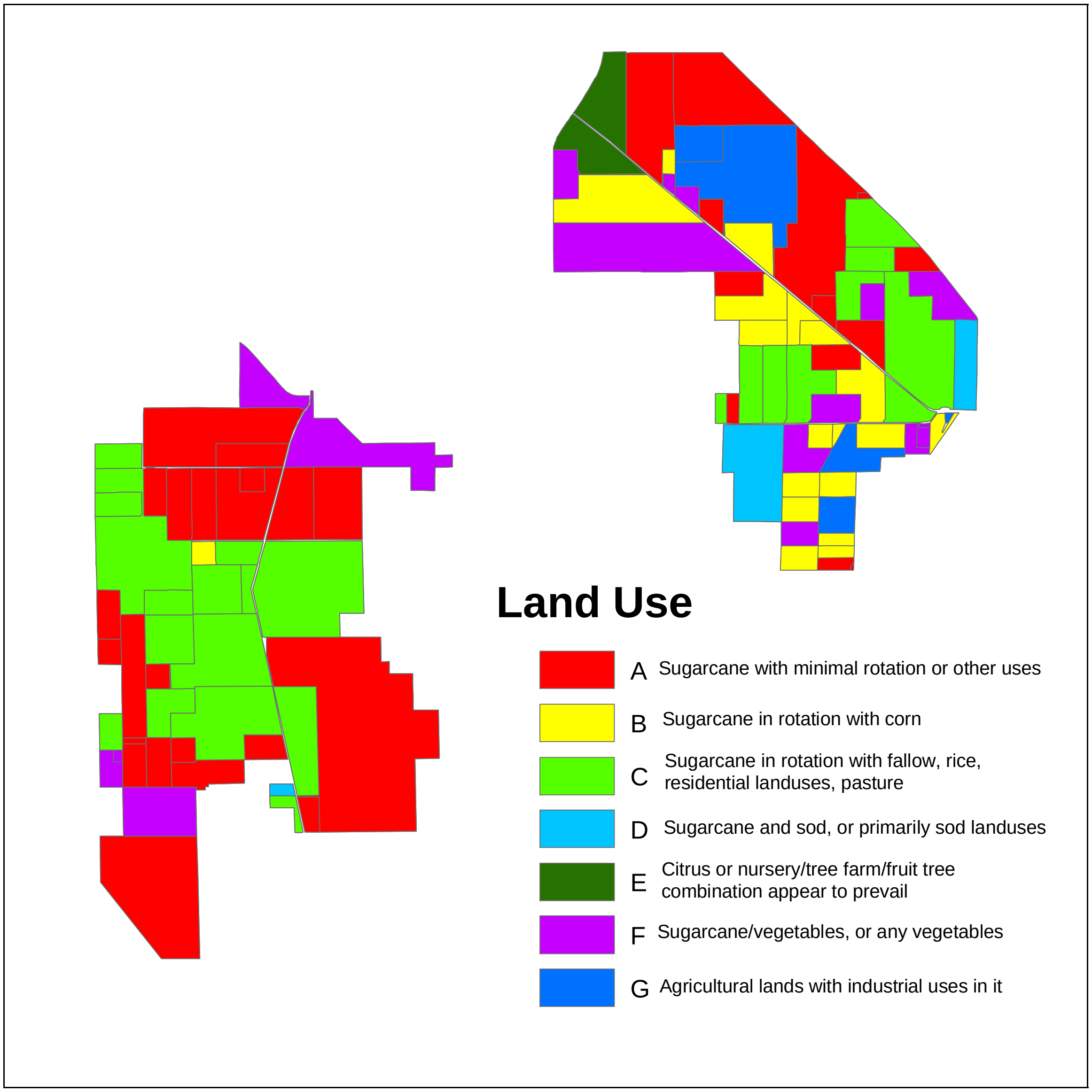


Figure 2-2: Land use distribution over the analysis area (WY1998)

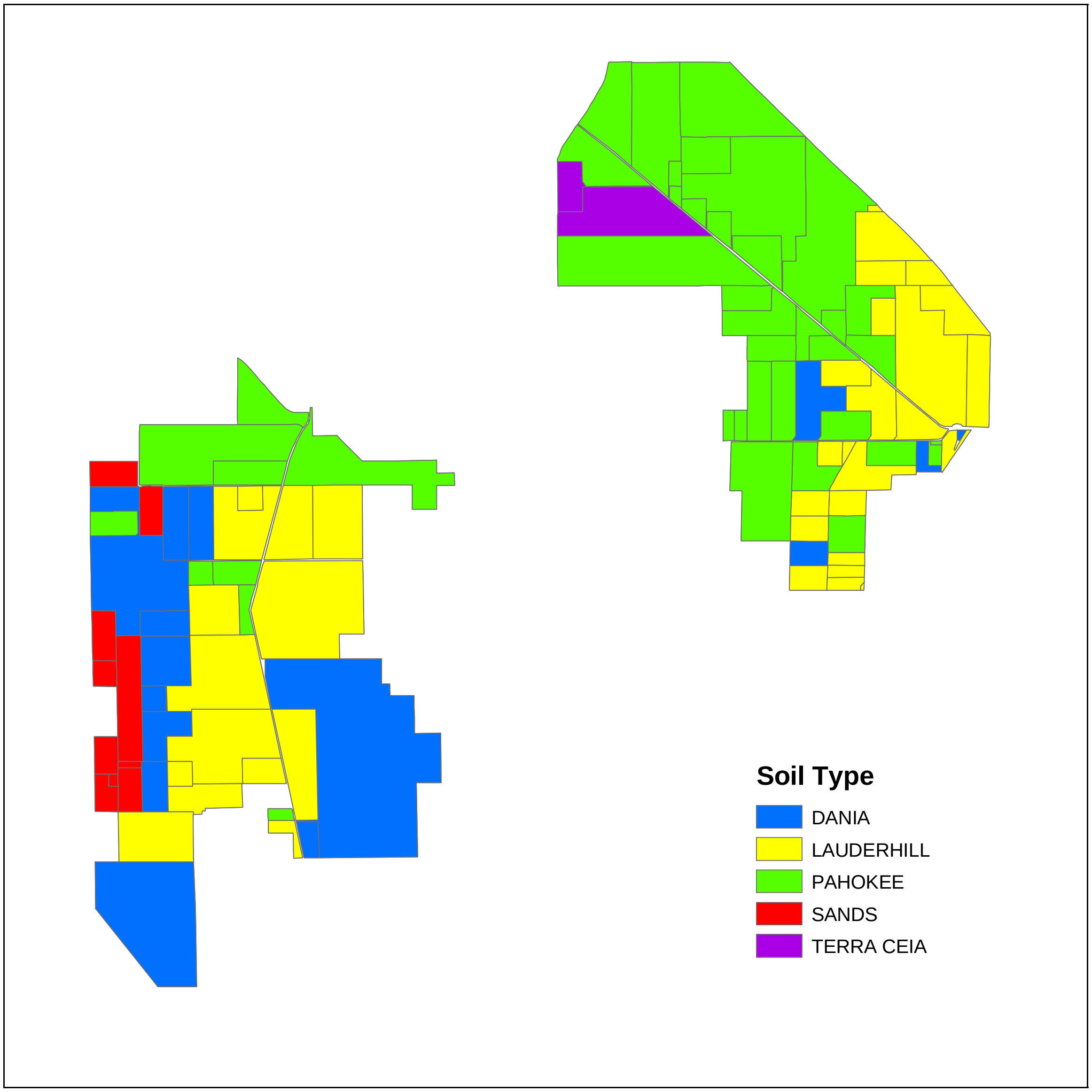


Figure 2-3: Soil type distribution over the analysis area (WY1998)

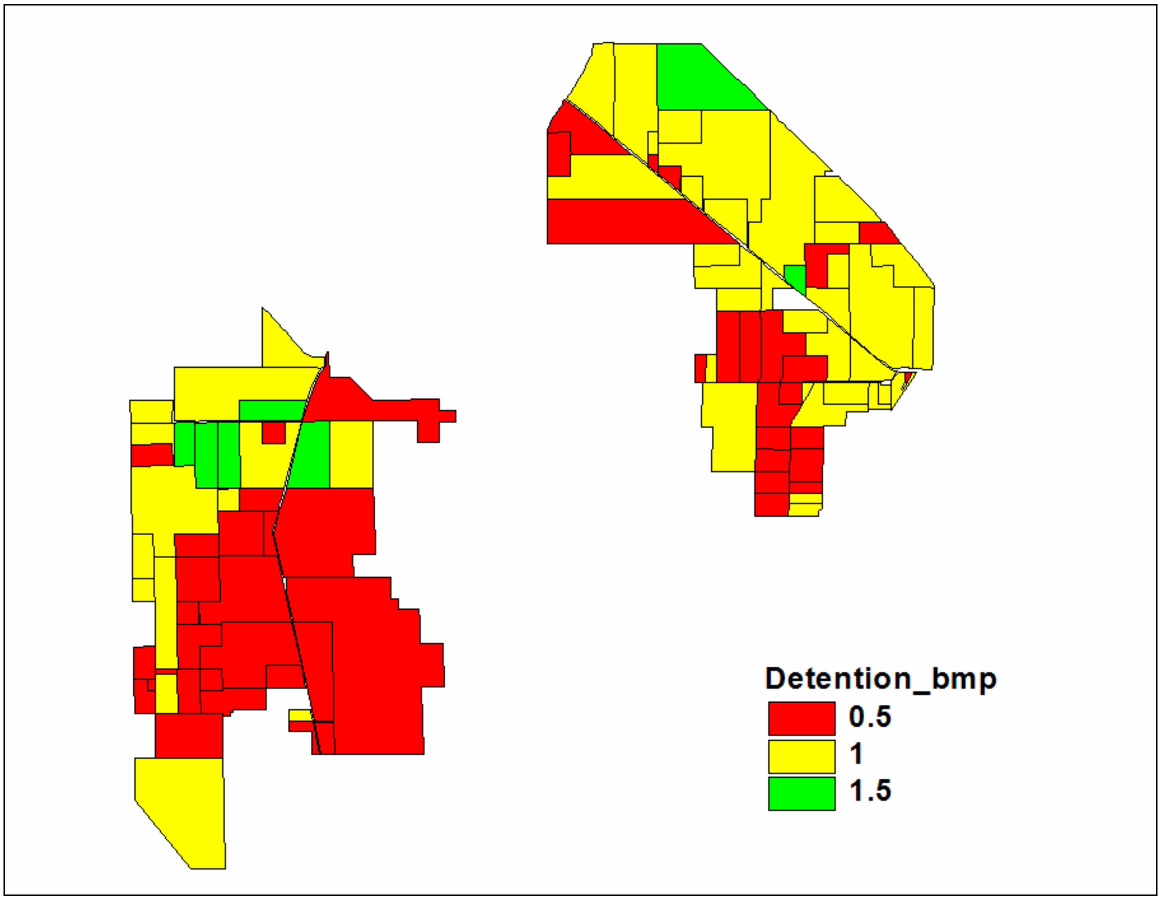


Figure 2-4: Water detention level distribution over the analysis area (WY1998)

The majority of the farms appeared to maintain similar BMPs and land uses during the analysis period; however, changes in these categories were tracked on an annual basis. Farms were regrouped when changes occurred. Table 2-3 shows an example of how changes were tracked for farm 50-007-02 during the analysis period. This farm would be analyzed in Land use B from WY1999 to WY2000 and in Land use D during the periods WY1995 to WY1998, and WY2001 to WY2003.

Table 2-3: Example of land use change and regrouping

Sub-basin	GISID	WY	LANDUSE	NOTE
S5A	50-007-02	1995	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	1996	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	1997	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	1998	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	1999	B	Sugarcane in rotation with corn
S5A	50-007-02	2000	B	Sugarcane in rotation with corn
S5A	50-007-02	2001	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	2002	D	Sugarcane and sod, or primarily sod land uses
S5A	50-007-02	2003	D	Sugarcane and sod, or primarily sod land uses

There are various limitations associated with the use of categories for the analysis. In contrast with the farm level water quality data which is routinely and systematically collected, the farm information is a compilation of best available information from farm permit files. Some of this information is collected on an as-needed basis. Indicated below are specific considerations that should be pointed out to the reader when reviewing the category analyses. However, despite these limitations, these categories were included in the exploratory analysis on the basis that pointing out the inherent differences among farms —based on best available information and noting the limitations of the sources— offers better grounds for understanding and interpreting the data than ignoring these characteristics. The results of the analysis shed light on how categories and characteristic groups can be refined. Based on the experience with this analysis, new or revised categories (e.g., farm area, flow, or multiple groupings based on water quality) are recommended for future analysis. Input or verification from individual permittees on the characteristic groupings and the characteristics assigned per farm may also serve to refine grouping arrangements. The assigned characteristics are based on review of available historic records for the analysis period.

2.3.3.1 Land uses

Results associated with the land use category should be viewed with particular caution. As discussed with District staff, there is some level of uncertainty associated with the land use characteristics. Crop variations across the years may cause significant effects in water quality, and the acreage and cropping patterns associated with each crop are not documented or systematically tracked for all farms as the permit-required water quality, flow, and rainfall data are. The level of detail in the information provided for crop types also vary from farm to farm. Crop type information was obtained by manual review of farmers' reports or permit information, based on availability. When permittees provided additional detail on their crops, this information was incorporated. For instance, a farmer reporting sugarcane and corn was classified in category B (sugarcane in rotation with corn), while a different farmer reporting sugarcane and vegetables, without providing the detail that the vegetable was corn, would have been classified in category F (Sugarcane/vegetables or any vegetable).

The District-provided data set included general land uses per year and per farm. The characteristics of all farms within the analysis area are shown in appendix Table A-Sec2-tab.1 to Table A-Sec2-tab3. Land use information provided by the District was incomplete for some of the farms. For example, Farm 26-003-01 only has land use records for periods WY1995-WY1996 and WY1998-WY2001. The available land use information was applied to the analysis.

2.3.3.2 Water Detention Levels
Each farm within the EAA is mandated to implement a BMP plan consisting of water management, nutrient, and sediment control BMPs equivalent to 25 credit points. The only BMP category used for this analysis was the water management BMP based on detention levels, and it was divided into characteristic groupings of 0.5", 1" and ≥ 1.5" detention. The characteristic levels do not imply that some farms have lower requirements than others. Those farms claiming the lower water detention levels (0.5") have increased levels of sediment controls or nutrient management practices. The BMP plan is assumed to define an equivalent level of effort among permittees and reflect BMPs that are feasible to implement based on farm specific conditions. The characteristic levels only reflect the emphasis in the selection of the alternative BMP activities (i.e., greater or lower detention strategies). BMPs may change during the years due to owner or crop changes (e.g. conversions from sugarcane to sod). Variations in the water detention level category were tracked and incorporated in the analysis from year to year.

2.3.3.3 Soil Types
The District provided the soil type assigned to each farm and this is the soil found to be the most representative for each farm. The District assumed that changes in soil type across the years (e.g., due to subsidence) would not significantly affect analysis results for the nine-year period evaluated. Nevertheless, it should be noted that subsidence is known to vary between 0.5 inch and 1.0 inch per year for the EAA.

2.4 General Statistical Methods

The analysis evaluates the temporal variation and spatial distribution of the time series datasets of phosphorus load and concentration for groups of farms in the categories described above. Indicated below is a description of the statistical methods used in the analysis. A detailed discussion of these statistical methods is provided in Appendix B.

2.4.1 Normality Test

Time series of phosphorus load and concentration collected at individual farms were tested for their normality by using the Shapiro-Wilkes Test and the Kolmogorov-Smirnov Test. The normality test of time series verifies whether or not a sampling dataset exhibits a normal distribution. If the test verifies the normality, many well-established statistical methods suitable for normal distributions can be used for data analysis. If the results are negative, non-parametric methods need to be applied to analysis. The normal distribution applied for the normality test is expressed as:

$$f(x) = e^{-(1/2)[x-\mu)/\sigma]^2} /(\sigma\sqrt{2\pi}) \qquad (2\text{-}4)$$

Results of these tests on most of the time series of P concentration and load indicate the skewness of the data is significant and probability distributions are considerably off from the normal distributions. Therefore it is concluded that the majority of P series may not be treated as normal distributions and that use of non-parametric methods is needed. In the analysis, the non-parametric Wilcoxon Ranks Test (Appendix B) method is applied to test large non-normal datasets.

2.4.2 Applied non-parametric statistical methods

In addition to the statistic testing methods described, the following methods or models were applied to depict and evaluate temporal variations in farm water quality data. The analyses were conducted with the farms grouped by single or multiple characteristics under the different categories.

Table 2-4: Statistical methods per Analysis

Analyses	Statistical Method
Seasonality	The Time-series Decomposition Model was used to depict the seasonal variations and multiple-year trends of P loads and concentrations.
Location	The Scatter Diagram Method and the Pearson Correlation Coefficient were used to depict the relationship between discharges from adjacent farm structures and statistically quantify the degree of correlation
Potential Interconnection	Farm discharge mean concentration and load values, and probability distributions for the wet and the dry seasons were compared among adjacent farms using the Wilcoxon Rank Testing Method
Inflow	The Scatter Diagram Method and the Pearson Correlation Coefficient were used to depict the relationship between Lake Okeechobee inflow structures and farm discharges, and statistically quantify the degree of correlation. Multiple regression equations were derived describing the relationship between the inflow structure and farm discharges. The Wilcoxon Rank Testing Method was used to compare the level of dependence described by the regression.
Spikes	The Box Plot Method was used to identify extremely rare high load and concentration values for the wet and the dry seasons. The contribution of these rare events was calculated and compared to the rest of the dataset.
Detention	Rainfall-runoff and Rainfall-P load curves were derived by using Scatter Diagram Methods and fitting trend curves based on lower order polynomials (linear or quadratic).

3.0 ANALYSIS OF SEASONALITY

This analysis intended to address the following questions presented in the scope of work:

- *Are phosphorus discharges distributed homogenously throughout the year or are they concentrated during specific months or agricultural activities?*
- *Calculate and compare monthly or seasonal phosphorus loads and concentrations for the EAA sub-basins and by categories.*
- *What are the characteristics of farms that have more consistent discharge levels?*

Approach: The analysis evaluates P load and concentration as a function of time of the year, EAA basin location (sub-basin), land use, soil type, and water detention levels. The method assumes that within each sub-basin the characteristic land use, soil type, or water detention levels can affect load and concentrations independently of each other. This assumption simplifies the analysis, but presents limitations to the analysis results. This analysis was performed using statistically-derived farm level P load and concentration trends. Trends are based on farm level load and concentration values that were calculated by aggregating the structure level datasets described in Section 2. The analysis was performed for each of three sub-basin locations (S3, S8, or S5A) by comparing loads and concentrations between: 1) the seven land use characteristics, 2) the five soil type characteristics, and 3) the three water detention level characteristics. For purposes of this analysis the term category means farms with at least one known characteristic that is the same. For example, for all farms within the S3 basin, average area-normalized P load and average concentration data were plotted for each of the four land use characteristics (A, B, C, and F) that were identified during the nine-year period. Long-term trends and seasonal variation curves were developed, and the curves compared among land use characteristics. This analysis was repeated for basins S8 and S5A. The same analysis was performed for each sub-basin independently for the soil type characteristics, for the water detention level characteristics, and for the sub-basins as a whole. Analysis of flow trends as a function of land use, soil type, or water detention levels, was not included in the scope of work for the analysis, however, its evaluation would serve to better understand their effect on P load trends. The supporting data for this analysis is presented in Appendix A.

3.1 Analysis Results

Are phosphorus discharges distributed homogenously throughout the year or are they concentrated during specific months or agricultural activities?

The analysis results depict qualitative existence and consistency (in term of cyclic variation phases) of the seasonality in the time series. Variations in P discharges throughout the year are generally observed in the time series of the farm category datasets. However, some categories of farms exhibit stronger and more consistent variation patterns. For instance loads from S3 farms characterized as having Lauderhill soils (Appendix A Figure A-Sec3-Soil Type-fig.d) show stronger and more consistent variation than S5A farms characterized as Dania soils (Appendix A Figure A-Sec3-Soil Type-fig.c). The seasonal variation of P load for the farms in S5A with Dania soils is stronger from 1994 to 1999 than from 2000 to 2002. Analysis results indicate that most P series collected from individual farms and aggregated over different categories have identifiable periodic motion. Figure A-Sec3-Landuse-fig.a to Figure A-Sec3-Subbasin-fig.c of Appendix A depict how P loads and concentrations were distributed during the year for each category of farms. The decomposition method verifies expectations that P load peaks occur during the wet season (from June to October) while P loads are relatively low during the dry season months. Using the same abscissa scales for all plots, we observe the following facts from the multiyear time series plots (see Table 2.2 for a description of land uses):

- Land use characteristics A and B have higher unit-area loads in sub-basin S5A than in S3 or S8 (Appendix A Figures A-Sec3-Landuse-fig.a, b, c, d, e, f). Land use characteristics A and B in S5A increased to high loading levels starting in WY 1997, while farms of the land use F were consistently high during the entire analysis period in all sub-basins (Appendix A Figures A-Sec3-Landuse-fig.c, f, n, o, p).
- Evaluation of soil classification characteristics within S3, S8 and S5A does not show clear distribution patterns. Both Dania and Lauderhill soils have higher P loads in S5A (Appendix A Figures A-Sec3-Soil Type-fig.c and f) when compared to the same soil characteristics in sub-basins S3 and S8 (Appendix A Figures A-Sec3-Soil Type-fig.a, d for S3, and A-Sec3-Soil Type-fig.b, e for S8 respectively). Other soil characteristics have relatively low-level P loads in S5A. In sub-basins S3 and S8 along the Miami Canal, the effect of soils on phosphorus levels is not clearly differentiated by this analysis.
- Detention level characteristics may affect P loads and concentration as shown in the data plots, but those effects are not predominant. Instead, the water detention level characteristic needs to be combined with other factors such as soil and land uses. The 0.5" detention characteristic is associated with lower P loads in the S3 sub-basin, but is associated with higher P loads in the S8 and S5A sub-basins (Appendix A Figures A-Sec3-BMP-fig.a, b, and c). The 1.0" detention characteristic does not show significantly different P loads than the 0.5" detention category in the S3 sub-basin. However, the 1.0" detention category is associated with lower P loads than the 0.5" category in sub-basins S5A and S8.

What are the characteristics of farms that have more consistent discharge levels?

Based on the observations indicated above, seasonal variations in P load and concentration depend on multiple characteristics including sub-basin location, land uses, soil types, BMPs, and others. As explained above, this analysis assumes these characteristics are independent of each other. A detailed analysis that evaluates P loads and concentrations as a function of multiple characteristics is necessary to more precisely address this part of the scope.

Although the analysis performed cannot answer this specific question, it does show interesting long term trends for the categories. Identifying the specific causes for these variations, however, is complex and extends beyond the scope of this statistical analysis.

The long-term trends indicate that farm discharges in the S3 and S8 basins as a whole show either moderate reduction or stabilization in their P concentrations and unit loads as shown in Figure 3-1a, Figure 3-1b and Figure 3-1c. However, S5A basin shows a significant decrease in concentration and a slight increase in load basin wide, suggesting an increase in flow. The long-term trend indicates that S5A concentrations at the beginning of the series were significantly higher than in the other sub-basins. Similar to the sub-basin results, long term trends associated with most farm categories (soils, land uses and water detention levels) show moderate variation over the years. However, considerable increase or decrease can be observed in the time series trends of a small number of farm categories. For example, a downward trend in P load can be observed for S-8 Land use F farms (Appendix A Figure A-Sec3-Landuse-fig.o). It is known that in 1999, farm basin 50-011-05 (classified as a land use F farm) was taken out of production and used for STA-5. This type of background information is critical in explaining variations in the time series.

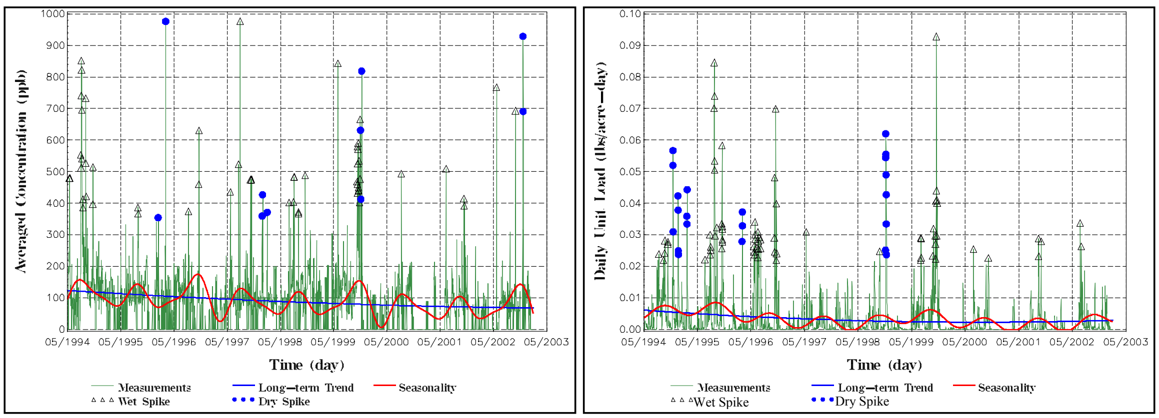


Figure 3-1a: P Concentrations & Loads for S3 Basin - Long term trend, seasonality and spikes

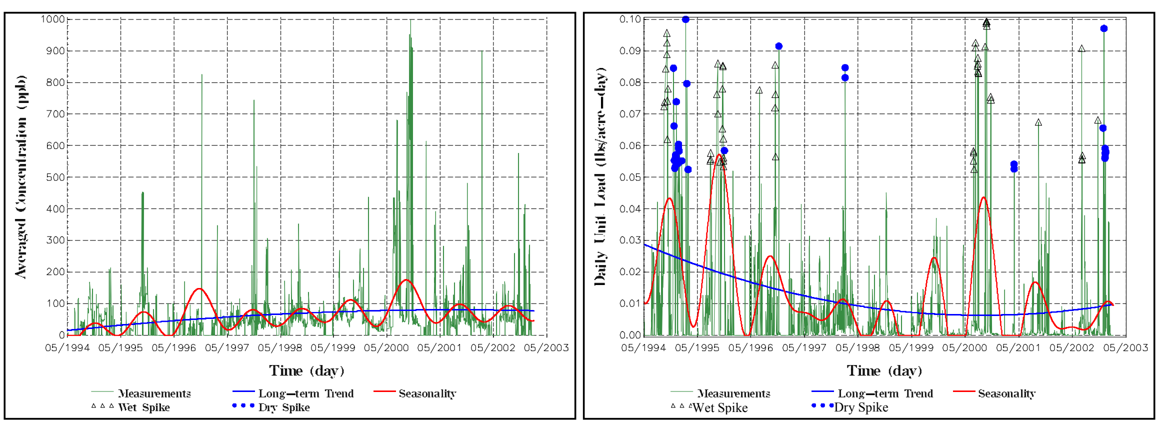


Figure 3-1b: P Concentrations & Loads for S8 Basin - Long term trend, seasonality and spikes

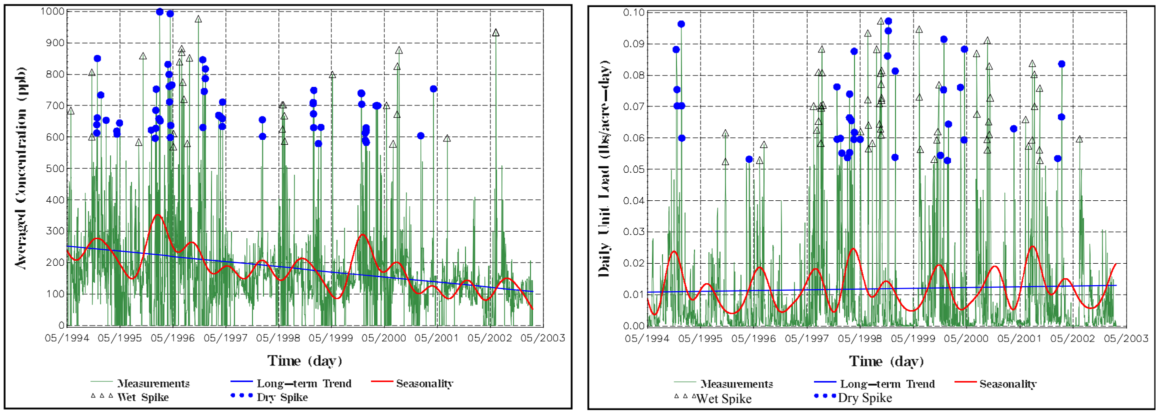


Figure 3-1c: P Concentrations & Loads for S5A Basin - Long term trend, seasonality and spikes

Calculate and compare monthly or seasonal phosphorus loads and concentrations for the basin as a whole and by categories.

The quantitative analyses initially conducted to address this section of the scope provided average monthly loads and concentrations based on the total nine years of data. These results were based on "yearly averaged daily" and "yearly averaged monthly" load and concentration datasets (i.e., for an "average year" based on the total nine years of data). A statistical analysis of similarity of the raw phosphorus concentration and load datasets on which these average values were based indicated that similarities among the years are not statistically significant, thus, the average loads and concentrations derived may not be representative of all years. Therefore, monthly P loads and concentrations for farm categories are not presented in this report to prevent misunderstanding that these average values can be applied to any single year without regard to year-specific hydrologic or farming conditions. This portion of the scope cannot be addressed with the existing analysis because a quantitative analysis would need to be conducted for single or representative years to ensure that expected variations from year to year are incorporated in the results and ensure accurate interpretation.

3.2 Statistical Analysis Methodology

3.2.1 Assumptions

The method utilized assumes that the characteristics of land use, soil type or water detention level, act independently of each other relative to loads and concentrations within each sub-basin. The classical decomposition model was applied to reveal temporal variations for this time-series analysis. The residuals were assumed to be white noise (W(t)). No other assumptions on white noise were made as the qualitative feature component (long-term trend and seasonality) was of primary interest.

Fourth order harmonics were used to fit seasonal variations, as they appear sufficient and stable to describe cyclic variation in most time series. The goodness of fit was obtained mainly through observation. The long-term trends were fitted with 2nd or 3rd order polynomials. The numbers of order determination were based on visual examination.

The long-term trend for each series considered in the decomposition model was assumed to be one continuous function over the study period. A stepwise function may need to be applied to certain series if changes with long-term effects are suspected to have occurred to the farms during the study period.

3.2.2 Seasonality Decomposition and Analysis

3.2.2.1 Approach of Decomposition

Like in many other hydrological time series, the multi-year P series discharged from EAA farms may contain three basic variation patterns:

- Long-term trend variation caused by gradual land use changes, ecosystem variation, long-term changes of hydrological conditions and long-term variation of farming practices (e.g., improved BMP implementation methods);
- Seasonal variation likely related to hydrological seasons and seasonal farming activities;
- Irregular load or concentration variation generated by infrequent, local, and short-term events and activities.

The long-term trend and the seasonality embedded in the P series may be decomposed using a gradual variation function (e.g. linear or quadratic polynomials) and a harmonic function respectively. The procedure of decomposing a time series includes:

- Assuming a long-term fitting model, normally a linear or quadratic polynomial;
- Assuming a harmonic model, normally a harmonic function;
- Using the method of least square estimate to determine the optimal fitting function parameters; and
- Assessing the long term variation and seasonality based on the representative modeling results.

If the harmonic function does not fit the time series in terms of occurrence and phases of seasonal motion, the seasonality may not exist or may be weak in the considered series.

3.2.2.2 Applied decomposition in the seasonality analysis

For trend development, the most typical method is to regress the time series with linear and lower level polynomials depending on the dataset. The seasonality component is usually represented by periodical polynomials. In this study, we have modeled the long-term trends by a quadratic regression, and modeled cyclic variations by a 4^{th} order harmonic regression model.

Thus we consider the decomposition:

$$P_{NL} \text{ or } P_C = \sum_{i=0}^{3} c_i t^i + \left(a_0 + \sum_{i=1}^{4} \left[a_i \sin((i-1)wt) + b_i \cos((i-1)wt) + W(t) \right] \right) \quad (3\text{-}1)$$

where t is time measured as day numbers from May 1, 1994 (i.e. t=1 for 05/01/1994 and t=2 for 05/02/1994, etc.); a_i, b_i and w are coefficients to be determined during fitting. The first term represents the component of long-term variation, the second term estimates the seasonal change, and $W(t)$ is the residual noise term representing the irregular and local P load and concentration fluctuations.

The method of decomposition is applied to the following P concentration and P load time series (P series) and presented in Appendix A:

- Figure A-Sec3-Landuse-fig.x: Averaged P series over farms of every land use (land uses A-G) within each EAA sub-basin (S3, S8 and S5A);
- Figure A-Sec3-Soil Type-fig.x: Averaged P series over farms of every soil type (a total of 5 soil types) within each EAA sub-basin;
- Figure A-Sec.3-BMP-fig.x: Averaged P series over farms of different water detention levels (detention 0.5", 1.0" and >1.5") within each EAA sub-basin; and
- Figure A-Sec3-Subbasin-fig.x: Averaged P series over each EAA sub-basin S3, S8 and S5A.

("fig.x" stands for the sequence of figures, e.g., fig.x=a, b ,c… etc.)

3.2.2.3 Example – Long-term trend and seasonal variation

As an example, which shows relatively clear long-term variation and seasonality, we consider the decomposition of the averaged area-normalized P load (P_{NL}) and concentration (P_C) series over all farms of land use category A within the S3 sub-basin. The concentration and load series are shown in Figure 3-2.

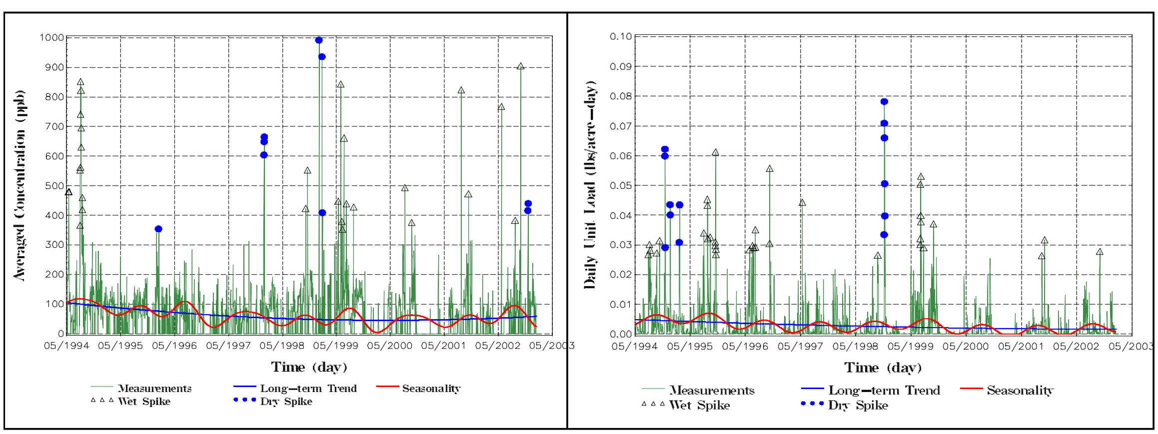


Figure 3-2: Averaged P concentration and load series for farms of land use A over the S3 sub-basin

Both P_{NL} and P_C series show a gradually decreasing long-term variation from WY1994 to WY2003 and seasonal variation within most water years. Concentration peaks (i.e.,

statistical outliers as defined in Section 7.0 of this report) occurred during WY97-98 and WY99-00 and load peaks in WY96 to WY97 and WY99 to WY00 should not be considered as having significant impact on the long-term variation.
Programmed in SAS, and applying the Least Square Estimate (LSE), we can determine coefficients included in Equation 3-1 as:

For P_C :	For P_{NL} :
c_1=8.18E-6, c_2=0.00637, c_3=98.1283;	c_1=1.789E-9, c_2=-8.82E-6, c_3=0.0125;
a_0 =-2.2021,	a_0 =-0.00012,
a_1=4.5286, a_2=-6.4849, a_3=-3.3965	a_1=-0.00076, a_2=-6.4849, a_3=-0.00090
a_4 =-38.9667,	a_4 =-0.00272,
b_1=29.0862, b_2=-22.7279, b_3=19.3882	b_1=0.00161, b_2=-0.00015, b_3=0.000426
b_4=36.2646;	b_4=0.00189;
w=0.00431	w=0.00437

In this example, the trend function shows that both loads and concentrations reduce gradually during the study period, which is consistent to the direct observation on the concentration time series. The seasonal function indicates that higher values of P loads and concentrations would likely occur during June to October for land use A (Primarily sugarcane with minimal rotation or other uses) over the S3 sub-basin. Peak loads and concentrations most likely occurred in August during the period of study.

4.0 LOCATION EFFECT

This analysis intended to address the following questions presented in the scope of work:

- *Is the water quality from structures located adjacent to each other similar?*
- *Does the water quality from structures sharing same canal waters follow similar trends?*

Approach: The original intent of this analysis was for comparison of water quality at the farm structure level; however, with the available regulatory data, the analysis could not conclusively answer the questions posed by the scope. A more detailed analysis and dataset, which is outside the scope of this project, would be required. A preliminary analysis was conducted by graphically and statistically comparing water quality collected at structures adjacent to each other along the same canal. The relationship pattern between composite daily P concentrations from two adjacent structures was examined graphically by using the scatter diagram method. The existence of correlation would prove the "cause and effect" or the "common cause and response" relationships between discharge concentrations from two adjacent structures. To statistically compare the correlation, the Spearman Correlation Coefficient, which is the numerical index of quantifying the degree of correlation, was computed.
Results of the preliminary analysis show that most structures had weak or no correlation between them. However, "no correlation" can not exclude the actual interactive effect between two adjacent structures since the discharge concentrations available to the study may not represent the complete hydrology and phosphorus transport processes. Evaluation through a site-specific analysis is likely to be more appropriate to address this section of the scope of work. Additional data including tailwater canal water quality, irrigation volumes and schedules, and specific activities or crops tributary to the structures may also be necessary for an effective analysis that isolates the effect of adjacent structures.

4.1 Analysis Results

Is the water quality from structures located adjacent to each other similar? Does the water quality from structures sharing same canal waters follow similar trends?

The preliminary analysis results based on the available data did not reveal the similarity of water quality at adjacent structures sharing the same canal section. Most scatter diagrams show no-correlation or weak correlation of water quality between two adjacent structures. For example, Figure 4-1 shows "no-correlation" between the adjacent structure pair MC21.5TW and MC23.0TE, and the pair MC18.8TW and MC18.8TE. The intersecting points are widely distributed, indicating that the water quality variations at these two structures are not related to each other.

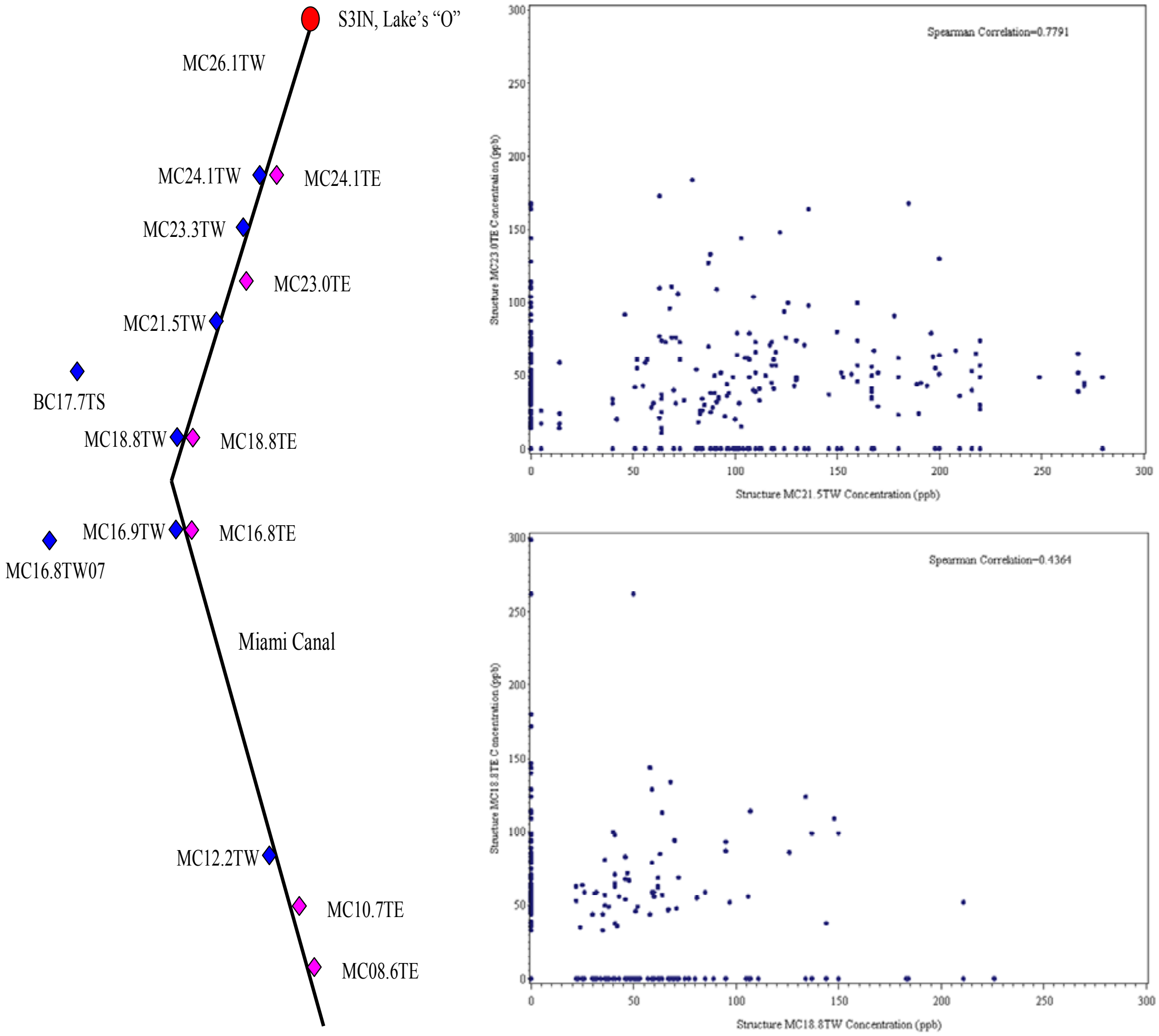


Figure 4- 1: Scatter diagram of scatter diagram of adjacent structures

Scatter diagrams based on the composite daily P concentrations at adjacent structures are shown in figures of Appendix A Figure A-Sec4-Miami-fig.1 to Figure A-Sec4-Miami-fig.15, and Figure A-Sec4-WPB-fig.1 to Figure A-Sec4-WPB-fig.33.

These inconclusive results may be explained by the following:

- The water quality variation of individual structures depends primarily on land use, farming activities and hydrological conditions. Although a discharge structure of a farm may share a common canal section with an adjacent structure of another farm, water quality at the structure is predominantly influenced by agricultural and hydrological features inside the farm upstream of the discharging point.
- Canal water may affect the discharge water quality at a structure through water intake and groundwater exchange. However, the P concentration data available for the analysis only represent the water quality conditions at the outlet on the upstream side of the structure.

A more detailed analysis utilizing additional datasets is recommended for a conclusive correlation between structures. Canal water quality and irrigation volumes and schedules, and specific activities or crops tributary to the structures is necessary for a sound analysis that isolates the effect of adjacent structures. Evaluation through a site-specific analysis is more appropriate to address this section of the scope of work.

4.2 Statistical Analysis Methodology

4.2.1 Assumptions

For the structure comparison, relationships between sets of three adjacent structures along a canal section were studied. A total of 9 discharge structures were selected for analysis along the Miami Canal and 18 structures along the West Palm Beach canal. Structures in tributary canals were not studied. It was assumed that the effect of groundwater quality and canal water quality should be limited to the near field. Figure 4-2 depicts the rationale used.

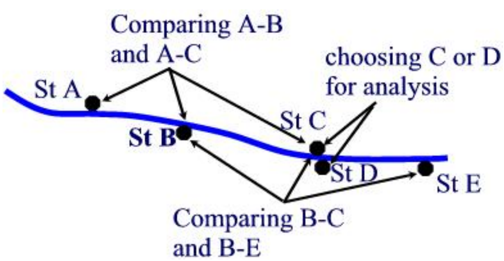


Figure 4-2: Selection of structures for near field water quality analysis

The flow-weighted daily P concentration series are applied. P loads at adjacent structures were not compared, as the load magnitude depends directly on farm sizes and the number of discharge structures of selected farms.

4.2.2 Comparison between structures

The reported composite daily P concentrations at two adjacent structures were compared using a scatter diagram. The scatter diagram is a useful tool to visually explore correlations of large data sets without regard to time. If P concentrations from two structures are related or share a common cause for variation, a scatter plot would show that most points are located in a narrow band. Otherwise, points would spread over the graph. The Spearman Correlation Coefficient, which is the numerical index of quantifying the degree of correlation for non-nominal datasets, was computed to evaluate location correlations.

- The composite daily P concentration series were applied for the preliminary analysis. Daily data series during the study period from WY94 to WY03 were used to construct the diagram. It is noted that some structures do not have complete data coverage over the study period, thus their sample sizes were reduced.

5.0 POTENTIAL HYDRAULIC INTERCONNECTION

This analysis intended to address the following question of the scope of work:

- *Analyze flows, loads and concentrations of adjacent farms to identify farms that may be candidates to "team up" (get connected, e.g., culverts with risers that could be opened or closed as needed by either party). Water quality would improve by increasing the combined retention capacity.*

Approach: The analysis identified farms that could establish hydraulic connections based on proximity, and on whether the mean values and probability distributions of their phosphorus loads and concentrations were statistically distinct. The farms could interconnect providing for a larger area for more flexibility with water management practices (recycling runoff to prevent or minimize combined discharge). The intent of this simplified approach is to develop a first level screening tool and assumes that farms identified will need follow-up on consideration of practical constraints for connecting such as ownership, runoff retention capacity, and crop type. The results of this analysis are limited to use as a first level screening tool. Mean seasonal loads and concentrations for the nine year period were calculated for individual farms in the S3, S8 and S5A basins, farms were grouped into 17 sub-areas based on geographic location denominated "potential interconnection groups". Mean phosphorus load and concentration for farms within each potential interconnection group were compared for the wet and dry seasons, and potential candidates for interconnecting within each sub-area were identified as those farms with the most statistically different values. The practice of interconnecting drainage systems of multiple farms is a practice that already exists in the EAA where farms are typically same-owner farms. This analysis acts as a first level screening tool to identify other potential hydraulic connection opportunities based on water quality and location. A follow-up in depth analysis of ownership, retention levels and crop type is necessary to further evaluate the potential for interconnection and is outside the scope of this analysis.

5.1 Analysis Results

Analyze flows, loads and concentrations of adjacent farms (with emphasis in higher load/concentration farms) to identify farms that may be candidates to "team up".

The evaluation acts as a first level screening tool to identify farms that may be candidates for hydraulic interconnection based on location and water quality. Site specific evaluation would be necessary to further evaluate critical factors such as ownership, crop type and water retention capacity to determine the feasibility of establishing hydraulic connections. Site-specific studies are necessary to assess how runoff diversion and intermixing could serve to improve water quality and are outside the scope of this analysis. As a follow-up to the first level screening, identified farms must consider issues associated with engineering feasibility, hydrology, cost, farming practices, and political aspects to understand potential constraints. Because these factors are beyond the scope of this analysis and were not analyzed, the statistical approach yields limited screening level information. Supplemental evaluation of the identified farms is required to determine the applicability of the arrangements found in the statistical analysis.

5.2 Statistical Analysis Methodology

5.2.1 Assumptions

Potential interconnection groups were introduced to support this analysis. The groups were delineated based on farm boundaries. It was assumed that within each potential interconnection group, all farms would have equal opportunity to be connected with others without consideration of ownership, crop type, or water retention capacity. This study only provides the statistical significance of potential interconnections based on the monitored discharges and P series. Constraints and site specific conditions would require further evaluation to identify farms that are potential candidates to hydraulically interconnect.

5.2.2 Potential Interconnection System Analysis

Farms were arranged into 17 groups based on the following two criteria:

- Farms should be adjacent to each other
- District canals should not divide farms within each potential interconnection group

As a result, farms within each group are geographically connected to each other and groups are generally bounded by tributary canals, as potential interconnection could not occur across District canals. Some farms located in the lower area of sub-basin S3 are not included in potential interconnection groups as those farms are either separated by District canals, or have tributaries cutting through the individual preventing potential interconnection. The delineation of the 17 potential interconnection groups is presented in Figure 5-1.

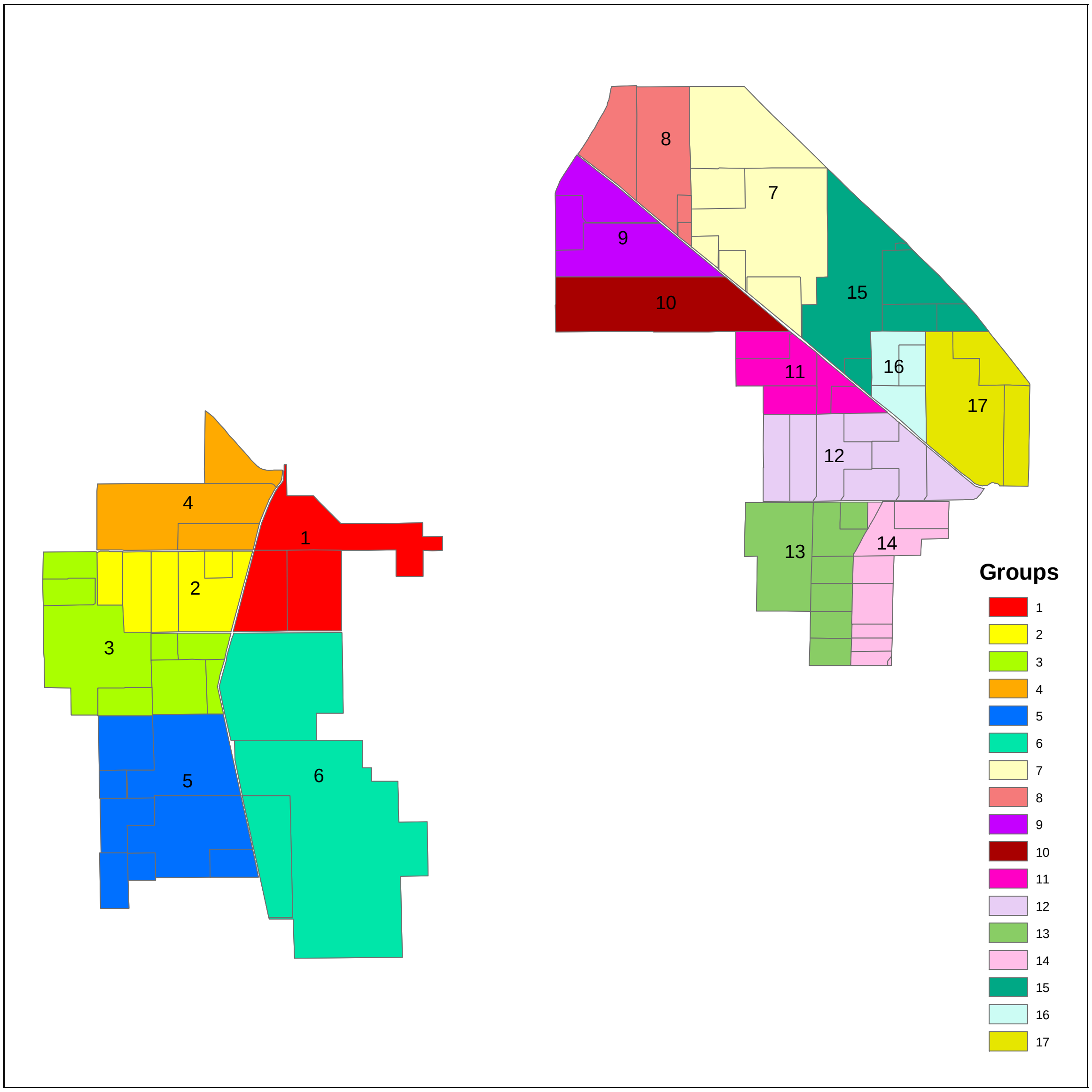


Figure 5-1: Delineation of Potential Interconnection Groups

Potential interconnections are identified for farms within each group (if they share common boundaries). For example, as shown in Figure 5-2, farms in the dark color are assigned to be potential interconnection group 3. There may be many combinations of potential interconnections within the group as indicated by the arrows.

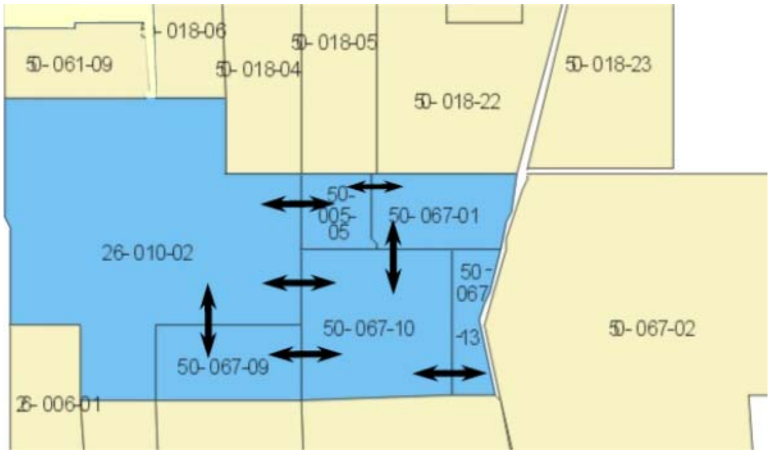


Figure 5-2: Potential Interconnection Groups

The rationale consisted of selecting the most statistically appropriate connection between all potential candidates by comparing the farm phosphorus series and identifying the most distinct farms in terms of both mean values and probability distributions. Due to the non-normality of phosphorus loads and concentrations, a non-parametric analysis method, the Wilcoxon Rank Testing was used.

Data collected during the wet season (May 1 to October 31 of each water year) and the dry season (November 1 to April 30) are analyzed separately, as farm discharges may vary significantly between hydrologic seasons. It is observed that phosphorus loads and concentrations often result in different potential interconnection arrangements due to their individual distribution of means. Both phosphorus loads and concentrations are assessed in the analysis. Results of the preliminary analysis, used as a first level screening tool, indicate that there are some potential candidates. A more detailed analysis evaluating ownership, land use and existing BMP retention levels would be required to further refine the results. The supporting data for the preliminary analysis is presented in Appendix A Table A-Sec5-tab.1 for wet season and Table A-Sec5-tab.2 for dry season.

6.0 ANALYSIS OF INFLOW IMPACT

This analysis intended to address the following questions presented in the scope of work:

- *Is there a relationship between the EAA Basin water quality, District basins water quality, or individual farms water quality and pass-through waters (Lake Okeechobee releases/298 District Diversion Projects)?*

Approach: The analysis was limited to determining whether statistical correlations could be identified between Lake Okeechobee and groups of downstream farms (sub-areas) based on phosphorus load and concentration data from specific farm discharge structures and Lake Okeechobee inflow structures S-354 (to Miami Canal) and S-352 (to West Palm Beach Canal). There were various limitations inherent in this analysis because of the limited dataset and associated assumptions so that statistically meaningful correlations between Lake Okeechobee inflows and discharges could not be identified. Lake Okeechobee structures discharged to the EAA less frequently relative to farm sub-area discharges to EAA canals. Additionally, the dataset does not include farm irrigation or District canal data (irrigation schedules, flows, and P concentrations) because they are not available for the EAA. As such, the cascading effect of the limited inflow events on the EAA canals, of those canals on the quality of farm irrigation waters, and of the irrigation water on the discharge events could not be incorporated. Use of an average daily phosphorus load and concentration assigned to each month and farms grouped into sub-areas instead of individual farms may have also simplified the analysis such that correlations could not be found.

As a first step to the simplified analysis performed, scatter graphs were created based on average daily phosphorus loads and concentrations for each month for groups of farms discharging directly to or through a tributary canal to the same associated segment of the Miami or the West Palm Beach canals. For instance, sub-basin S3 was divided into 3 sub-area groupings (e.g., S-31, S-32 and S-33), and sub-basins S8 and S5A into 2 and 4 sub-areas, respectively. The scatter points paired the concentration or loads from the inflow structures with those of the downstream sub-areas. Pearson Correlation Coefficients were derived for the graphs. As a second step, a multivariable regression model was derived to compare the relative correlation of Lake Okeechobee inflows on sub-areas, and of the sub-areas among each other. The supporting data for this preliminary analysis is in Appendix A.

6.1 Analysis Results

Is there a relationship between the EAA Basin water quality, District basins water quality, or individual farms water quality and pass-through waters (Lake Okeechobee releases/298 District Diversion Projects)?

The scatter diagrams showed that average daily inflows at structures S-354 and S-352 did not have strong correlation with discharges from sub-area groupings of farms. The multivariable linear regression model indicated that among these weak correlations only relatively stronger effects from Lake Okeechobee inflows (S-354) were observed on sub-area S-32, that is, the southern portion of the S-3 basin. These results, however, are not definitive. Because of data limitations, assumptions and the methods used, other approaches should be explored before reaching a conclusion. As indicated above, only EAA basin inflow and farm discharge data were available. Farm irrigation practice and District canal data (irrigation schedules, flows, water quality) are necessary to understand water cycling. The statistical approach may also need refinement to identify meaningful correlations. Use of sub-areas groupings of farms and average daily loads and concentrations for each month may not provide sufficient precision to capture the effect of Lake Okeechobee inflows on individual farms. Different farm aggregation techniques and, most likely, evaluation of farms on an individual basis is needed. Use of lagging techniques, instead of daily averages for each month, may also be more appropriate to understand the effects of Lake Okeechobee. Additional data and analysis is necessary to verify and enhance the findings from this preliminary analysis. The type of supplemental analysis that is necessary is complex and extends beyond the objectives of this preliminary statistical analysis.

6.2 Statistical Analysis Methodology

6.2.1 Assumptions

The major assumption involved in the analysis is associated with not having farm inflow water quality, quantity, and irrigation schedule information. Conducting the analysis without this information assumes that Lake Okeechobee inflows to the EAA and farm discharges will alone be sufficient to establish statistical relationships between Lake Okeechobee inflows and farm discharges. This lack of information impedes understanding water recycling between District canals and farms. Lake waters flow into EAA canals, from which farmers irrigate their lands supplemental to rainfall. Excess irrigation water and rainfall runoff are discharged back to District canals which in turn can be recycled onto downstream farms based on crop needs and farming operations. Farming activities may contribute to or remove phosphorus from these discharges.

The frequency at which data are collected at the Lake inflow locations may also contribute to the difficulty in deriving meaningful correlations. The historical data for the Lake Okeechobee inflows are typically collected on a monthly frequency. As such, discharge values from sub-areas were calculated by averaging daily farm values for each month. For scatter diagrams at S-352, the number of points used are the number of months that the inflow structure and the sub-area have valid monthly data, and should be less than 9 water years x 12 months = 108. It was assumed that these datasets would be sufficiently precise to identify statistical relationships.

Data from two District stations, S-354 (named as S3IN in the analysis) and S-352 (named as S5AIN), respectively, form the inflow datasets (discharge, phosphorus load and concentration) for the analysis. The period of inflow datasets is consistent with the farm datasets from WY1995 to WY2003. In addition, the outflow datasets taken at S8

(S8OUT) and Station S5A (S5AOUT) are also included in the analysis as the most downstream points of the District canals and outflow locations from the basin.

The sub-area grouping concept was used for this analysis as a first step to evaluate general P distributions in a spatial scale without excessive statistical computation of farm level data. Use of sub-area groupings, also, would serve to more clearly depict correlations by smoothing the significant noise of some of the individual farm series. The delineation method assumes that similarities among farms within the same geographic sub-area are more significant than potential differences, that is, results for a sub-area are reflective of what the results would have been for the individual farms within the sub-area. Nevertheless, time series associated with sub-areas aggregate farms without consideration of whether they have different land uses, soils, water detention levels, or water quality characteristics. Responses on a farm level, thus, may differ from each other based on their differences. EAA sub-basin S3 was divided into 3 sub-areas (S-31, S-32 and S-33), sub-basin S8 was divided into 2 sub-areas (S-81 and S-82) and sub-basin S5A into 4 sub-areas (S-5A1, S-5A2, S-5A3 and S-5A4). Sub-areas were delineated based on the considerations indicated below and are shown in Figure 6-1. Sub-area delineation was not based on a detailed hydrology and farm characteristic review, thus, the simplified criteria limit the applicability of the findings. Results from the Sub-area comparison are limited to conceptually understanding the general spatial response of discharge P loads and concentrations under varying Lake Okeechobee inflow conditions.

- Sub-areas are located along the Miami Canal and the West Palm Beach Canal which links them to a major drainage system;
- Each sub-area contains at least one drainage tributary which collects discharge from all farms within the sub-area and directly discharges to the Miami Canal or the West Palm Beach Canal;
- Sub-areas are delineated along the farm boundaries so that each farm is completely included in one sub-area.

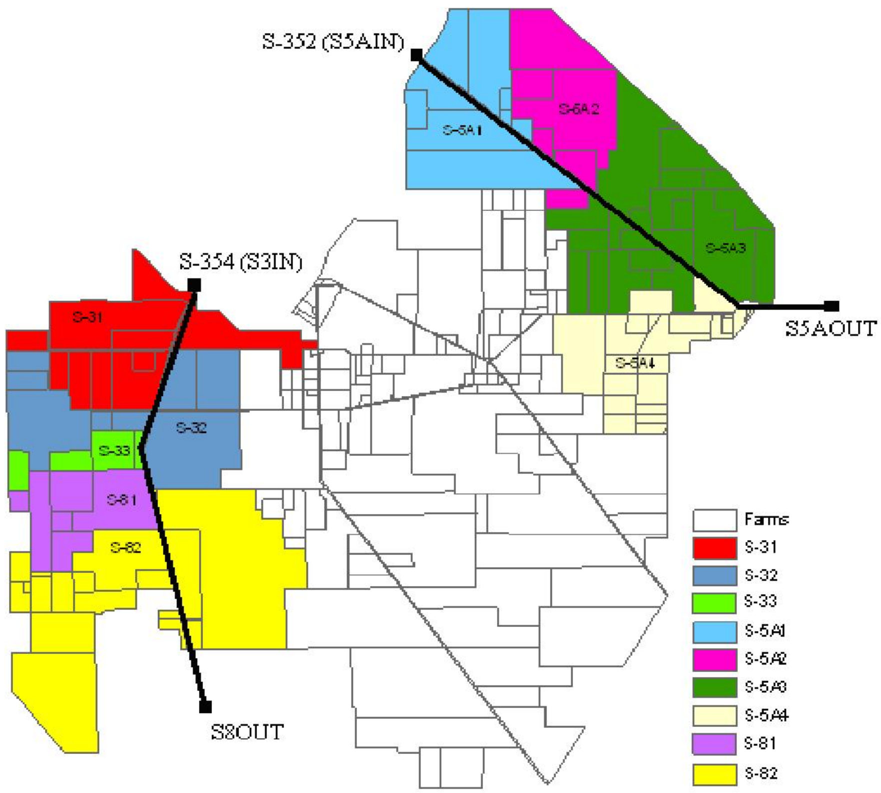


Figure 6-1: Sub-area Delineation

6.2.2 Correlation of Inflow and Sub-Area Discharges

The method of Scatter Diagram was first used to evaluate potential correlations between the inflow and individual downstream sub-area discharges. The scatter diagram is a useful tool to visually explore correlations of large datasets without regard to time. If phosphorus concentrations from inflows and individual sub-area discharges are related, a plot showing points where the concentrations from inflows are the abscissas and the sub-area discharges are the ordinates would show clusters of points. Otherwise, points would spread over the graph. In addition, the Pearson Correlation Coefficient, which is the numerical index of quantifying the degree of correlation, was computed to describe the relationship of the series. Data points plotted in the scatter diagram are the mean monthly values of P concentrations and loads. For example, if within a month, inflow has a total of 10 daily records of P concentration, and these individual concentrations are flow weighted concentrations, that data point used for the scatter diagram would be the arithmetic average of these 10 values. The number of points in scatter diagrams should be the number of months that both inflow and individual sub-area discharge series have valid monthly data and should be less than 9 water year x 12 months = 108. The scatter diagrams indicate that there is no correlation between Lake Okeechobee inflows at both S-354 and S-352 and their respective downstream sub-areas.

The generated scatter diagrams are presented in the Appendix A as Figures A-Sec6-Miami-fig.1 to A-Sec6-Miami-fig.6 for the relation between the S3IN and individual downstream sub areas, and Figures A-Sec6-WPB-fig.1 to A-Sec6-WPB-fig.5 for the relation between S5AIN and individual sub area series. In these scatter diagrams, both daily P concentrations and P loads are plotted to support the analysis.

The following scatter diagrams are examples describing the relationship between the inflow and sub area discharges:

- Figure 6-2 shows the lack of correlation between the inflow to Miami Canal and discharge from the adjacent S-31; and
- Figure 6-3 shows the lack of correlation between the inflow to WPB Canal and discharge from the adjacent S-5A1.

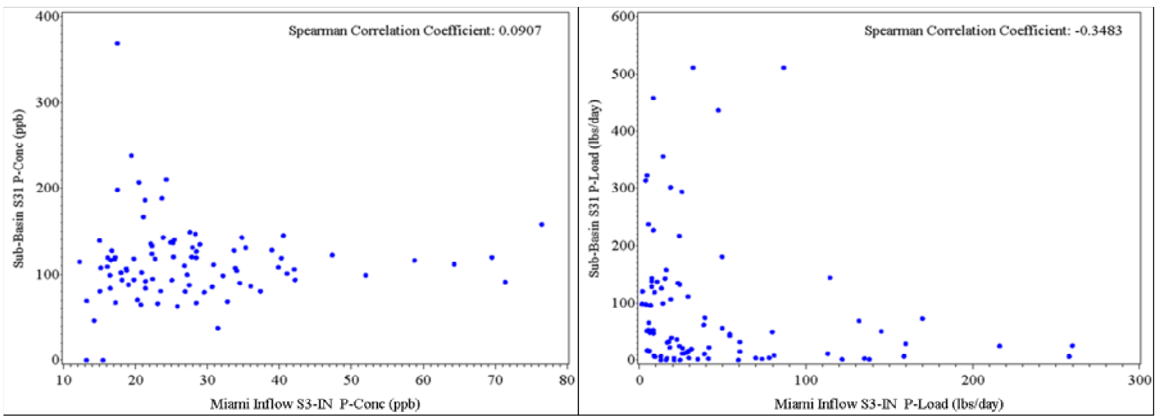


Figure 6-2: Correlation between Inflow to Miami Canal and Discharge from sub-area S31 – P concentration and P load.

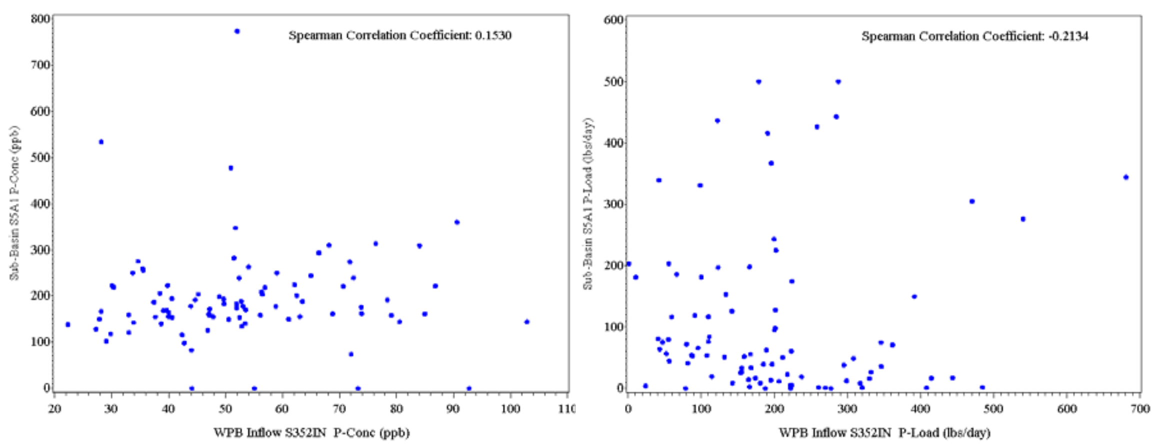


Figure 6-3: Correlation between Inflow to WPB Canal and Discharge from sub-area S-5A1 – P concentration and P load.

From both figures, the upstream inflows (S3IN and S5AIN) have limited correlation on the adjacent farm discharges, even though the selected sub areas S-31 and S-5A1 are immediately adjacent to the EAA Basin inflow point. The correlation coefficients for the considered situations are at or below the value of 0.2, which is too low to show any meaningful correlation.

6.2.3 Evaluation of inflow impact by multiple-regression

The multiple regression analysis identified, for each sub-area, which upstream sub-area discharges had stronger impact than others. Although, the correlation between two or more individual sub area series could be weak, one or more upstream series could have stronger general impact to the downstream series than others.

The multiple-regression method is applied to quantitatively evaluate the dependency of discharge P series of the target downstream sub-area on inflow and other upstream sub-areas). For example, assuming the phosphorus of a downstream area would be affected by phosphorus released from upstream discharges, then

$$S_{31} = f_{31}(S_{3IN});$$
$$S_{32} = f_{32}(S_{3IN}, S_{31})$$
$$...............$$
$$S_{82} = f_{82}(S_{3IN}, S_{31}, ...S_{81})$$

Considering the phosphorus load, we use the multiple regression models to build up the potential relations:

$$P_{NL,S31} = a_0 + a_1 P_{NL,S3IN} + \varepsilon_{S31}$$
$$P_{NL,S32} = b_0 + b_1 P_{NL,S3IN} + b_2 P_{NL,S31} + \varepsilon_{S32}$$
$$P_{NL,S33} = c_0 + c_1 P_{NL,S3IN} + c_2 P_{NL,S31} + c_3 P_{NL,S32} + \varepsilon_{S33} \qquad (6\text{-}1)$$
$$\cdots\cdots\cdots$$
$$P_{NL,S82} = g_0 + g_1 P_{NL,S3IN} + g_2 P_{NL,S31} + \cdots\cdots + g_5 P_{NL,S81} + \varepsilon_{S82}$$

The error term associated with each regression model maintains the relations to be probabilistic, rather than deterministic. Coefficients are calculated using the least square method, and the results are shown in Table 6-1 for the Miami Canal and Table 6-2 for the WPB Canal, respectively.

The dependence is determined by performing the non-parametric testing (e.g., Wilcoxon Rank Test) on deviation of individual independent variables from the modeled regression curve, and computing p-values which describe the hypothesis that "the considered coefficient is equal to zero". If the testing result indicates that the considered coefficient is similar to zero (or $p > 0.05$), the independent variable associated with the tested coefficient would have little impact to the modeled value. Or otherwise, if $p<0.05$, it means that the tested coefficient is far from zero, and the considered independent value would have impact on the dependent variable.

The computed p-Values associated with each coefficient of the multiple regression models are listed together with the coefficient in Table 6-1 and Table 6-2.

Table 6-1: Coefficients and p-Values of the multiple regression models applied to evaluate the impact of upstream P loads on the P loads of considered downstream sub-areas in Miami Canal

Load							
Parameter Estimation							
Item	Constant	S3-IN	S-31	S-32	S-33	S-81	S-82
S-31	108.5513	-0.4290					
S-32	64.0257	-0.4538	0.5677				
S-33	161.2134	-0.1824	-0.4321	0.7206			
S-81	6.0059	-0.0407	0.0356	0.1539	0.0062		
S-82	44.8682	-0.0929	0.1810	0.1496	0.1240	1.7781	
S-8OUT	3.8723	0.1152	-0.1560	1.0398	0.0209	0.6593	0.1455
p-Value Listing							
Item	Constant	S3-IN	S-31	S-32	S-33	S-81	S-82
S-31	<0.0001	0.0552					
S-32	<0.0001	0.0078	<0.0001				
S-33	0.0106	0.7862	0.2618	0.0836			
S-81	0.1726	0.3755	0.1784	<0.0001	0.3987		
S-82	0.0237	0.6478	0.1258	0.3073	0.0002	0.0003	
S-8OUT	0.8556	0.5922	0.2172	<0.0001	0.5738	0.2277	0.2076
Concentration							
Parameter Estimation							
Item	Constant	S3-IN	S-31	S-32	S-33	S-81	S-82
S-31	111.9910	0.0574					
S-32	95.0879	0.1034	0.2574				
S-33	-21.1346	2.3512	0.2571	0.2358			
S-81	37.4436	0.3322	0.0192	0.0457	0.0577		
S-82	50.3370	0.2545	0.0421	0.0972	0.0772	-0.2173	
S-8OUT	5.3103	0.0484	0.0132	0.0016	0.0194	0.0726	0.0307
p-Value Listing							
Item	Constant	S3-IN	S-31	S-32	S-33	S-81	S-82
S-31	<0.001	0.8826					
S-32	<0.001	0.8233	0.0444				
S-33	0.6411	0.0133	0.3262	0.2755			
S-81	0.0377	0.3822	0.8521	0.5916	0.1737		
S-82	0.0332	0.6016	0.7491	0.3734	0.1590	0.1186	
S-8OUT	0.1481	0.5149	0.5375	0.9248	0.3363	<0.0001	<0.001

Table 6-2: Coefficients and p-Values of the multiple regression models applied to evaluate the impact of upstream P loads on the P loads of considered downstream sub-areas of West Palm Beach Canal

Load						
Parameter Estimation						
Item	Constant	S352-IN	S-5A1	S-5A2	S-5A3	S-5A4
S-5A1	98.1999	0.0511				
S-5A2	25.9153	-0.1007	1.1273			
S-5A3	37.0979	-0.0949	0.4382	0.2296		
S-5A4	297.1031	-0.6852	-1.2347	1.6045	1.2456	
S-5AOUT	52.6036	0.2843	0.5755	-0.1237	0.1992	0.3949
p-Value Listing						
Item	Constant	S352-IN	S-5A1	S-5A2	S-5A3	S-5A4
S-5A1	0.0002	0.6413				
S-5A2	0.0804	0.0865	<.0001			
S-5A3	0.0211	0.1330	0.0026	0.0463		
S-5A4	<.0001	0.0035	0.0238	0.0002	0.0019	
S-5AOUT	0.2257	0.0694	0.1099	0.6726	0.4538	<.0001
Concentration						
Parameter Estimation						
Item	Constant	S352-IN	S-5A1	S-5A2	S-5A3	S-5A4
S-5A1	187.6687	0.1362				
S-5A2	97.2627	0.0442	0.2426			
S-5A3	169.1697	1.9323	-0.1017	-0.3269		
S-5A4	235.2742	-1.1113	0.0209	-0.1632	0.1441	
S-5AOUT	28.0408	0.0595	0.0051	0.0569	-0.0211	0.0272
p-Value Listing						
Item	Constant	S352-IN	S-5A1	S-5A2	S-5A3	S-5A4
S-5A1	<.0001	0.8330				
S-5A2	0.0010	0.9215	0.0015			
S-5A3	0.0400	0.1106	0.6297	0.2581		
S-5A4	<.0001	0.0719	0.8432	0.2636	0.0092	
S-5AOUT	0.0008	0.5637	0.7698	0.0207	0.0254	0.1329

As shown in the table of parameter estimation, the multiple regression models representing P load at S-32 would be:

$$P_{NL,S32} = 64.0257 - 0.4538 P_{NL,S3IN} + 0.5677 P_{NL,S31} + \varepsilon \qquad (6\text{-}2)$$

From the test on the null hypothesis of "the computed coefficient is equal to zero", the computed p-Value for the coefficient of S3-IN is 0.008 and for that of S-31 is <0.0001. Since all p-values are small than 0.05, it means that both coefficients for S3-IN and S-31 are close to the estimated value. Therefore, both $P_{NL,S3IN}$ and $P_{NL,S31}$ have impact on the target $P_{NL,S32}$.

Also considering the regression on S-81, from the above table, the regression model is:

$$P_{NL,S81} = 6.0059 - 0.04P_{NL,S3IN} + 0.04P_{NL,S31} + 0.15P_{NL,S32} + 0.01P_{NL,S33} + \varepsilon \qquad (6\text{-}3)$$

The p-Value for $P_{NL,S32}$ is nearly zero, indicating this coefficient is significantly unlike "0". The p-values for other coefficients are larger than 0.05, meaning those coefficients are similar to zero, and thus those variables may have weak influence in the model. Therefore, we may conclude that $P_{NL,S81}$ would be significantly affected by $P_{NL,S32}$ from S-32 adjacent to S-82, but less affected by S3-IN, S-31, and S-33.

Comparing the results from the scatter diagram (Section 6.2.2) and the multiple regressions modeling (Section 6.2.3), we observe that the P load at S3IN has the meaningful impact on the P load from sub area S-32 (p-Value=0.008) that is augmented by the impact from S-31 (p-Value<0.0001). This result is consistent with the scatter diagram (Figure 6-4), which shows the correlation coefficient is -0.3390, much larger than other scatter plot values. The absolute of the correlation coefficient (0.3390) is greater than 0.2 which is the threshold above which a correlation can be considered meaningful.

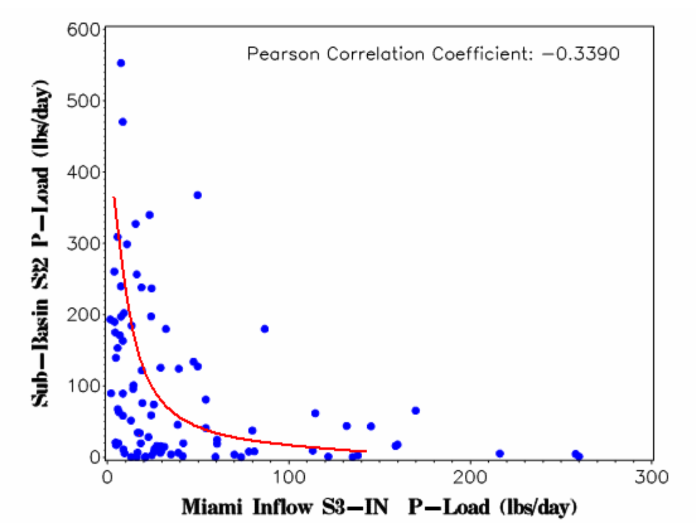


Figure 6-4: Scatter diagram for the P Load between S3IN and S-32

As indicated in the correlation section (6.3.2), no correlation was found between Lake Okeechobee inflows and sub-areas by observing the scatter plots and the resulting Pearson Correlation Coefficients. The Wilcoxon Rank Test results indicate that only structure S-354 (Miami Canal) shows strong impact on a downstream sub-area (S-32). Tables 6-1 and 6-2 show some dependence among sub-areas, however, it cannot be necessarily associated to downstream effect. Results of the preliminary analysis are inconclusive.

7.0 ANALYSIS OF PHOSPHORUS DISCHARGE SPIKES

This analysis intended to address the following questions of the scope of work:

- *What is the contribution of incidental spikes on water quality?*
- *Are these spikes more common in a specific type of farm or condition?*
- *Are high phosphorus discharge farms also farms that are subject to spikes?*
- *Are there water years (WY) or seasons when spikes are more frequent?*
- *Determine the impact of incidental exceedances and their effect on annual WY calculations.*

Approach: For purposes of this analysis the term spikes refers to high load discharge events based on averaged daily loads and concentrations for groupings of farms. The term spike does not apply to short term high concentrations at a specific location, as it is typically used when conducting a water quality field investigation at an individual site. The farm level datasets available for the analysis do not offer this level of precision (permittee reported data represent up to 21-day composite periods). Spikes loads and concentrations can result from higher pumping, higher concentration or both. The purpose of this analysis is to determine if there are consistent patterns associated with high load discharge events and to quantify their annual contribution to the total P load discharged during the year, not to establish cause-effect relationships. Site-specific-short term evaluations are outside the scope of this analysis.

In order to identify the extreme spike discharge from the other data points, it is assumed that the spikes are equivalent to statistical outliers involved in the time series. Based on this assumption, the averaged daily time series for load and concentration prepared for the seasonality analysis were evaluated using the Box Plot method to identify spikes for groupings of farms. The Box Plot method defines a statistical upper range for a set of data. Spikes are statistically defined outliers above this threshold. The threshold for the analysis was set at the 75th percentile value plus three times the difference between the 25th and the 75th percentile values of each data set. Outliers should not be considered necessarily invalid data, but single observations that are significantly different from the rest of the data. Outliers can indicate problems in sampling or data collection, or they can also represent unusual responses to a situation that calls for additional investigation. Once daily spikes for the farm groups were identified, average wet and dry season phosphorus loads were calculated with and without the spike loads.

7.1 Analysis Results

What is the contribution of spike discharge events on water quality? Determine the impact of incidental exceedances and their effect on annual WY calculations.

The analysis indicates that phosphorus loads from spike days represented approximately half of the phosphorus loads discharged during a year. Although most spikes occur occasionally over a short time span (one to two days), their high magnitude contributes considerably to the total annual load. The specific causes for spike discharges (flow, concentration, or both), or whether they could be associated with unrepresentative data (contaminated samples, etc.) were not within the scope of this analysis.

Are spikes more common in a specific type of farms or conditions? Are high phosphorus discharge farms also farm that are subject to spike?

Comparisons of spike load contribution to the total phosphorus load are listed in Table 7-1, Table 7-2, and Table 7-3 for different land uses, soil types and water detention levels. The analysis focused on farm groupings. No analysis was completed for individual farms. These are general observations worth noting for the groups:

- Associated with land uses, the average increase of the time series mean due to existence of spikes is computed to be 75% for the dry season, 54% for the wet season, and 60% for the annual series respectively. This contribution appears to be significant due to the extremely high magnitude of the spikes. Sharp differences in phosphorus loads and concentrations were observed between land uses A and C (see Table 2-2 for description of land uses) in comparison with other land uses as shown in Table 7-1.
- The spikes for the different soil types averaged 59%, 55% and 56%, for the dry season, wet season and annual series, respectively (Table 7-2). These values are lower than those for the land uses, conveying that the soil group time series have smoother distributions. A smoother distribution is perhaps due to the fewer categories within this group. For example, all farms are gathered into 7 groups based on their land uses, and only in 5 groups based on their soil type.
- Based on the water detention levels, spike contribution to the phosphorus loads varies from 54% for the dry season, to 44% for the wet season and to 48% for the annual values (Table 7-3). Since three groups are subdivided based on assigned water detention levels (0.5", 1.0" and 1.5"), the averaged spike contribution for the detention category appears to be less than for the land use and soil categories.
- In terms of seasons, the analysis indicates that spikes occurring during the dry season have a larger contribution to the total seasonal loads (from 54% to 75% based on the grouping used) than the contribution of wet season spikes to the total seasonal loads (from 44% to 55% based on the grouping used). While in terms of the overall total phosphorus loads, the wet season spikes have larger total value than the dry season spikes. For example, the means with spikes during the dry season were at or below 3 lbs/acre-season (year), while 4 (four) means with spikes exceeded this level during the wet season.

Are there water years (WY) or seasons when spikes are more frequent?

The number of spikes recorded for the different categories and from season to season varies from year to year. Daily spikes along the years can be observed in the Seasonality graphs:

- Land uses[3] (A-G) within each EAA Sub-Basin S3, S8 or S5A (Appendix A Figure A-Sec3-Landuse-fig.a to Figure A-Sec3-Landuse-fig.q);
- Soil types within each EAA Sub-Basin S3, S8 and S5A (Appendix A Figure A-Sec3-Soil Type-fig.a to A-Sec3-Soil Type-fig.l);
- Water management detention level groups (0.5", 1.0" and >1.5") within each EAA Sub-basin S3, S5, and S8 (Appendix A Figure A-Sec3-BMP-fig.a to A-Sec3-BMP-fig.h); and
- EAA Sub-basins S3, S8 and S5A (Appendix A Figure A-Sec3-Subbasin-fig.a to A-Sec3-Subbasin-fig.c).

Table 7-1: Impact of spike discharge on the total phosphorus loads for different land uses (lbs/acre-season(year))

P load series	Means w/ spikes-dry season	Means w/o spikes – dry season	Relative difference %	Means w/ spikes-wet season	Means w/o spikes-wet season	Relative difference %	Means w/ spikes-annual	Means w/o spikes-annual	Relative difference %
Land use A	1.07	0.49	119%	1.90	0.96	98%	2.97	1.45	105%
Land use B	2.29	1.52	51%	4.00	2.18	84%	6.29	3.70	70%
Land use C	0.36	0.14	151%	1.16	0.58	98%	1.51	0.72	109%
Land use D	0.62	0.47	31%	1.28	1.09	17%	1.90	1.57	22%
Land use E	0.57	0.47	21%	0.89	0.77	15%	1.45	1.24	17%
Land use F	1.34	0.82	64%	2.26	1.56	44%	3.60	2.38	51%
Land use G	3.04	1.64	85%	3.62	2.99	21%	6.66	4.63	44%
Avg. P Contribution			75%			54%			60%

Table 7-2: Impact of spike discharge on the total phosphorus loads for different soil types (lbs/acre-season(year))

P load series	Means w/ spikes-dry	Means w/o spikes - dry	Relative difference %	Means w/ spikes-wet	Means w/o spikes-wet	Relative difference %	Means w/ spikes-annual	Means w/o spikes-annual	Relative difference %
DANIA	0.84	0.59	43%	1.45	1.01	45%	2.29	1.59	44%
LAUDERHILL	1.94	1.08	79%	3.13	1.71	83%	5.07	2.79	82%
PAHOKEE	0.65	0.35	85%	1.14	0.84	35%	1.78	1.19	49%
SANDS	1.48	0.90	65%	3.09	1.48	108%	4.57	2.38	92%
TERRA_CEIA	0.62	0.50	25%	1.26	1.19	5%	1.88	1.69	11%
Avg. P Contribution			59%			55%			56%

[3] See Table 2.2 for a description of land uses.

Table 7- 3: Impact of spike discharge on the total phosphorus loads for different Water Detention Levels (lbs/acre-season(year))

P load series	Means w/ spikes-dry	Means w/o spikes - dry	Relative difference %	Means w/ spikes-wet	Means w/o spikes-wet	Relative difference %	Means w/ spikes-annual	Means w/o spikes-annual	Relative difference %
0.5 inch	1.37	0.84	63%	2.52	1.63	55%	3.90	2.47	58%
1.0 inch	1.35	0.83	64%	2.01	1.45	39%	3.36	2.27	48%
>1.5 inch	0.50	0.37	34%	0.89	0.64	39%	1.39	1.02	37%
Avg. P Contribution			54%			44%			48%

7.2 Statistical Analysis Methodology

7.2.1 Assumptions

Spikes were defined as accurate/valid data that were significantly high in comparison with other data in the datasets. Permittee collected data are routinely submitted and reviewed by District staff. A year end summary is provided back to the permittee each water year, prior to calculation of annual farm level unit loads and concentrations. This provides the permittee an additional opportunity to review their data. In order to identify the extreme spike discharge from the other data points, it is assumed that the spikes are equivalent to statistical outliers involved in the time series. Based on this assumption, the method of upper outer fence (or the Box-Plot method) is used to identify the spikes from the time series.

7.2.2 Identification of Spikes

The Box Plot method, which examines the quartiles of a time series, is used to identify those statistic extreme values representing relatively rare occurrences. In this method, the statistically abnormal values of a series are identified as:

$$P_{spike} \geq Q_u + 3.0 \times (Q_u - Q_L) \quad (7\text{-}1)$$

where Q_u, the upper quartile is the 75th percentile of the data set.

Q_L, the lower quartile, is the 25th percentile of the data set.

The computation of spike values is shown in Figure 7-1.

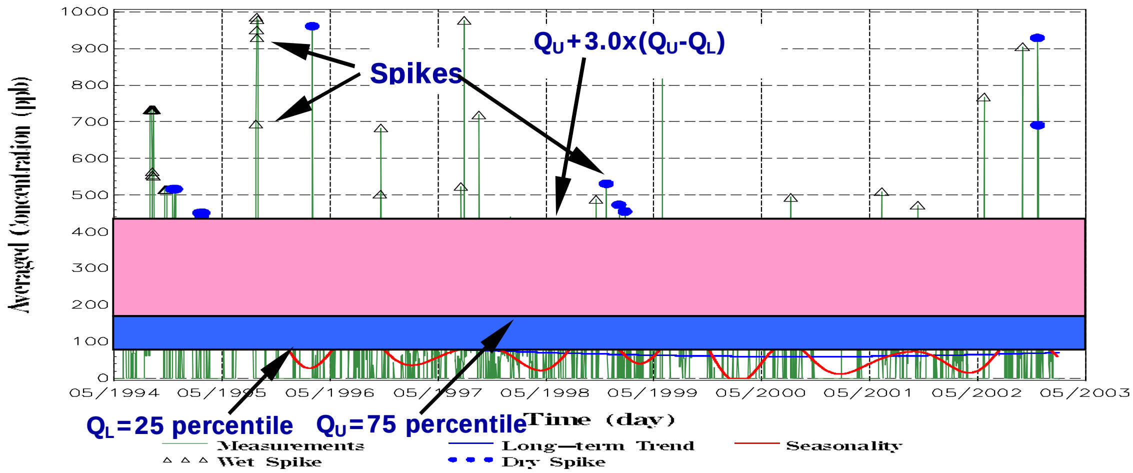


Figure 7- 1: Identification of spikes – Box Plot Method

Selection of the parameter 3.0 in Eq. (7-1), depends on the time series characteristics and the spike value distributions. It is chosen based on the opinion of the professional conducting the statistical evaluation. Some technical references use a value of 1.5, and others suggest using 3.0 or even higher. In this study, the parameter of 3.0 was selected for filtering out the outliers as many high individual discharge values are found in the time series. Typically a parameter of 1.5 or 3.0 is used to identify whether high values are "rare" events. When using the higher parameter, the number of events selected as rare events is less. For demonstration, a normally distributed dataset (the farm data for this analysis is not normally distributed), the 75th percentile stands for 1.53σ (one tail) one standard deviation from the mean value (the mean equals the median in a normal distribution). The difference between the 75th and 25th percentile stands for 3.06 σ. If using the parameter 3.0 with the difference (i.e. 3.06 σ), the spike values identified would have a deviation of 9.18 σ, and the parameter 1.5 would have a deviation of 4.6 σ. For a normal distribution, 3.0 σ covers 99.7% of total events. Thus using the parameter 3.0 at 9.18 σ would identify rare events, and relatively much rarer than using the parameter of 1.5.

To evaluate the 75th percentile, Q_u, and the 25th percentile, Q_L, those time series values less or equal to zero were eliminated, as they are not contributing to the upper level of spike values. It was observed that high values occurred frequently in the time series. Assuming all values in time series are valid to present field conditions, it was necessary to define a statistical measure to filter out all "normal" values and to identify those abnormal spikes.

Spikes of phosphorus discharge are identified for the following conditions, and all spike results are plotted together with the seasonality data (Appendix A):

- Figure A-Sec3-Landuse-fig.x: Averaged P series over farms of every land use (land uses A-G) within each EAA sub-basin (S3, S8 and S5A);
- Figure A-Sec3-Soil Type-fig.x: Averaged P series over farms of every soil type

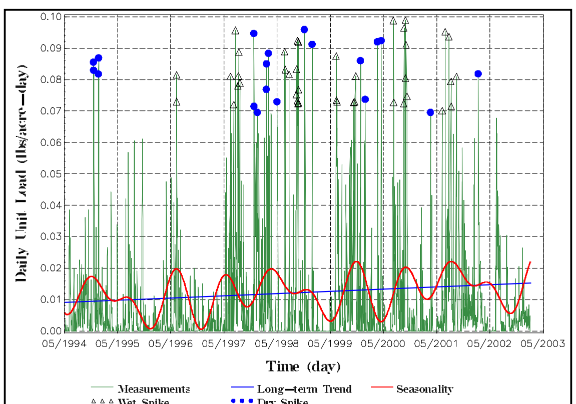
(a total of 5 soil types) within each EAA sub-basin;

- Figure A-Sec.3-BMP-fig.x: Averaged P series over farms of different water detention levels (detention 0.5", 1.0" and >1.5") within each EAA sub-basin; and
- Figure A-Sec3-Subbasin-fig.x: Averaged P series over each EAA sub-basin S3, S8 and S5A.

("fig.x" stands for the sequence of figures, e.g., fig.x=a, b ,c… etc.)

Box Plot upper limits are determined for individual time series. A phosphorus value may be defined as a spike in one P series, but not considered as a spike in another P series. For example, the upper limit of box plot for P load series of 0.5" water detention level averaged over S5A Sub-basin is about 0.07 lb/acre-day, while the upper limit level for 1.0" detention averaged over S3 Sub-basin is 0.032 lb/acre-day. Observed from Figure 7-2a and Figure 7-2b, spikes associated with these two time series are identified based on different datasets. The different upper threshold level depends upon the magnitude and distribution of the two datasets. Figure 7-2a shows a large number of high P loads occurred in the 0.5" detention level in S5A sub-basin dataset, which results in higher values for the 25th and 75th percentiles, and Figure 7-2b shows a limited numbers of high P loads in the 1.0" detention in the S3 sub-basin dataset, which resulted in lower values for the 25th and 75th percentiles.

Although the Box-Plot upper limits are established over multiple years for considered phosphorus series, spike values are marked separately as dry-season spikes and wet-season spikes indicating their seasonal occurrence (Figure 7-2a and Figure 7-2b for 0.5" and 1.0" detention, respectively).



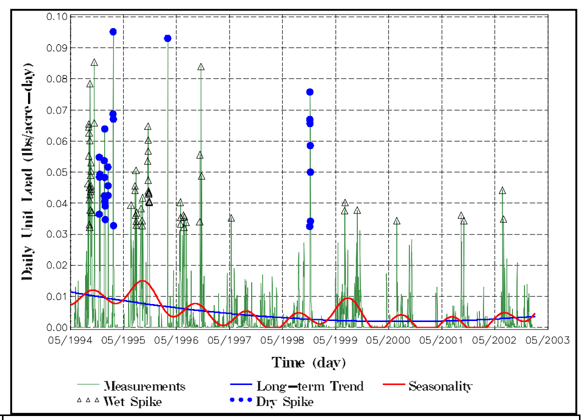


Figure 7-2 a: P load for 0.5in-BMP averaged over Sub-basin S5A

Figure 7-2 b: P load for 1.0in-BMP averaged over Sub-basin S3

The contribution of daily spikes to the total load of the various farm groupings is evaluated by comparing the total phosphorus discharge with and without the identified daily spike values. Calculation is based on the daily time series aggregated by land uses, soil, and detention level characteristics. If on one day the phosphorus load is counted as a spike, the total daily value of this day will be counted as a spike value.

Comparisons of spike load contribution to the total phosphorus load are listed in Table 7-

1, Table 7-2 and Table 7-3 for different land uses, soil types and BMP detention volumes respectively. Aggregating all farmlands within the same category over the Study Area generates the multiple year time series. For example, P load series for Land use A is constructed by aggregating all time series of land use A during WY1995 to WY2003.

8.0 DETENTION

The scope of work asked various questions associated with detention levels. These questions are summarized as follows:

- *What is the effect of rainfall detention on phosphorus concentration and load?*
- *How much do farms typically detain based on reported data?*
- *How often do farms deviate from the assigned detention level category and are there any trends that can be devised?*

Approach: This analysis explores relationships between rainfall and runoff, and rainfall and phosphorus load based on multiple characteristic groupings of land use, soil type and water detention levels. Answers to the questions in the scope of work were explored by developing rainfall-runoff and rainfall-phosphorus loading regression curves for groups of farms with the same characteristic land uses, soil types, and water detention levels. The curves representing the groups were compared to identify differences and similarities and how the contrasting characteristic may affect the rainfall-runoff and rainfall-phosphorus load relationships. The curves were derived from scattered points grouped into "rainfall events" (i.e., consecutive rainfall days). The simplified criteria to define rainfall events provides a standardized rationale for this screening analysis, however, it does not cover all possible field scenarios. This is a limitation of the analysis.

A review of historic delayed pumping practices was also conducted to understand deviations from the assigned detention levels, that is, the relationship between the point at which pumping begins and the measured rainfall. The analysis evaluated delayed discharges based on reported daily flow and daily rain gage readings. A simplified criterion to define a rainfall event was established for the analysis. Rainfall detention levels of 0.5, 1.0 and 1.5-inch were assigned to each farm and three categories of rainfall events were specifically created for the purpose of simplifying the analysis. The delayed pumping analysis accounted for reported instances when discharge occurred before or after assigned water detention levels based on rain gage readings. However, there are other factors, such as water table levels, that play a significant role in the decision making process to control farm discharge. Water table levels were not part of the analysis dataset because they are not included in the permittees routinely submitted data. The District concluded that without this information, the questions as posed in the scope could not be adequately addressed. Consequently, this portion of the scope was removed from the analysis. A separate farm-by-farm evaluation is necessary to accomplish the scope of work objectives.

The supporting data for this analysis is presented in Appendix A.

8.1 Analysis Results

What is the effect of rainfall detention on phosphorus concentration and load?

This question cannot be addressed based on the data available and the criteria used for the analysis. Regression curves with adequate correlation levels were found to establish rainfall-runoff relationships for most categories, however, they were not found to assess rainfall versus load. Comparison of the regression curves indicates that land use and soils have considerable impact on the rainfall-runoff correlation while assigned water management detention levels were found to have limited impact.

How much do farms typically detain based on reported data?

The rainfall runoff curves provide a general relationship between the criteria-defined rainfall event and the resulting runoff for the multiple characteristic groups. These graphs are presented in Appendix A Figures A-Sec8-fig.a to A-Sec8-fig.al. How much farms typically detain, however, cannot be addressed with the reported data for each farm. Detailed site specific information (e.g., surface water management, configuration, monitoring, control water levels) and modeling would be necessary to calculate rainfall volume detained for specific rainfall events at individual farms.

Most of the rainfall-runoff regression curves show that the event-based correlation between rainfall and runoff are significant and relatively stable. Rainfall-runoff relations are considered to be stable as large event-rainfall volumes result in asymptotically limited runoff volumes (below a line at 45 degree in the first quadrant). The following rainfall-runoff variation characteristics can be observed from the regression analysis:

- For Land uses A and B (see Table 2-2 for a description of land uses), the soil types Lauderhill and Pahokee yield significantly correlated rainfall-runoff graphs at different rainfall detention levels. Soil types Dania and Sands for these land uses, however, have a relatively weak correlation at different rainfall detention levels. For example Figure 8-1a and Figure 8-1b show the comparison of strong Land A-Lauderhill-1.5” BMP and weak Land A-Dania-1.5” BMP.
- For Land uses C and D, soils Dania and Pahokee have stronger correlation of rainfall-discharge than Lauderhill.
- Terra Ceia soils have limited distribution across the area and are only associated with Land uses B and F. Many rainfall events are identified for Soil Terra Ceia in land use F, but only few events in Land use B. The rainfall-runoff correlation for Soil Terra Ceia in Land use F appears to be significant.

Some combinations of land use, soil and BMP detention level had few reported rainfall events and were statistically insufficient to generate regression curves.

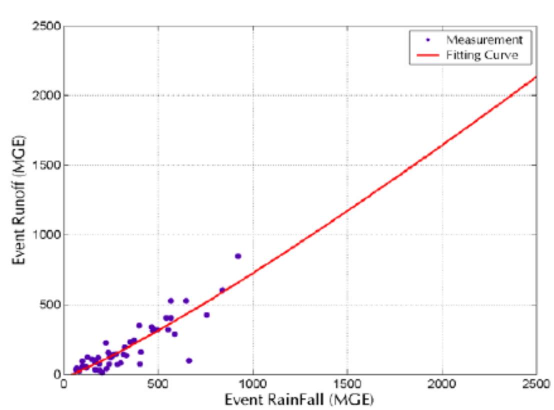


Figure 8-1a: Rainfall-runoff correlation for Land A, soil type Lauderhill, BMP 1.5".

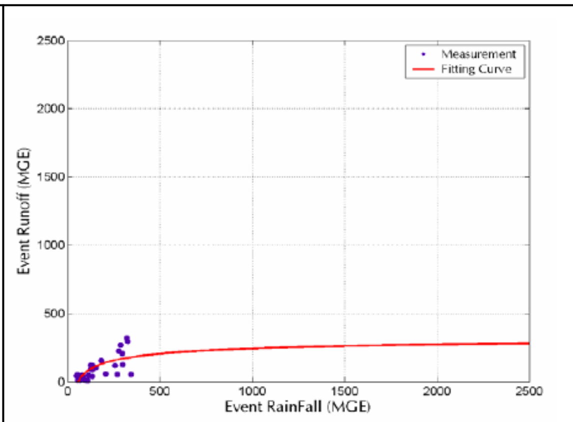


Figure 8-1b: Rainfall-runoff correlation for Land A, soil type Dania, BMP 1.5 ".

How often do farms deviate from the assigned detention level category and are there any trends that can be devised?

The scope dataset is not adequate for an assessment of historic delayed pumping practices. A screening based on the simplified event criteria and on rain gage readings may not provide an accurate assessment and could be misinterpreted as a compliance assessment. This type of analysis requires a farm-by-farm assessment based on review of detailed site-specific data and water management rationale at each farm.

8.2 Statistical Analysis Methodology

8.2.1 Assumptions

For the development of rainfall runoff and rainfall P loading regression curves, rainfall events were defined using the criteria described below:

- A rainfall event is preceded by at least two dry days before the rain starts and covers at least two consecutive raining days;
- Single-day rain events were not included. It was estimated that South Florida showers of short duration would likely be absorbed through soil absorption, evaporation or retained onsite with minimal effect on farm detention capacity. Because single day events may not have the same effects on the discharge as multiple day events, and the multiple-day data set was sufficiently large, it was not considered necessary to include single-day events for the rainfall-runoff assessment.
- A rainfall event stops if the continuous daily record (e.g., flow) breaks.

Table 8- provides an example of how rainfall events were derived based on the criteria. The "duration" column indicates the consecutive days of a rainfall event, and each event

is separated by at least two dry days. The rainfall volume and flow volume columns are computed based on the unit load and farms acreage. The P load is computed based on the runoff and composite concentration.

Schemes of structure aggregation to farms are changing slightly from water year to water year. When making the farm level rainfall time series, all variations involved in structure-farm aggregation are included.

Table 8-1: Rainfall event selection and flow/P loads calculation

Sub-basin-S8; GISID50-061-10; Land use A, Soil Dania, BMP 0.5, Acreage 22879 ft2							
Date	Event No	Duration (day)	Rain (in/day)	Unit Flow (mgd/acre)	Unit Load (lbs/acre-day)	Rainfall volume (mgd)	Flow volume (mgd)
8/15/1998	2	8	0.15	0.00189	0.001039	93.2	43.2
8/16/1998	2	8	0.37	0.00052	0.000289	229.9	12.0
8/17/1998	2	8	0.44	0.00000	0.000000	273.4	0.0
8/18/1998	2	8	0.98	0.00000	0.000000	608.8	0.0
8/19/1998	2	8	0.77	0.00195	0.001122	478.4	44.6
8/20/1998	2	8	0.11	0.00443	0.002548	68.3	101.3
8/21/1998	2	8	0.2	0.00462	0.002656	124.3	105.6
8/22/1998	2	8	0.15	0.00377	0.002168	93.2	86.2
9/16/1998	3	6	0.1	0.00000	0.000000	62.1	0.0
9/17/1998	3	6	1	0.00000	0.000000	621.3	0.0
9/18/1998	3	6	0.34	0.00167	0.000722	211.2	38.1
9/19/1998	3	6	0.11	0.00293	0.001272	68.3	67.1
9/20/1998	3	6	0.94	0.00389	0.001689	584.0	89.1
9/21/1998	3	6	1.39	0.00656	0.002847	863.6	150.2
12/28/1998	4	4	0.15	0.00000	0.000000	93.2	0.0
12/29/1998	4	4	0.025	0.00000	0.000000	15.5	0.0
12/30/1998	4	4	0.3	0.00000	0.000000	186.4	0.0
12/31/1998	4	4	0.015	0.00000	0.000000	9.3	0.0
4/27/1999	5	4	0.65	0.00000	0.000000	403.8	0.0
4/28/1999	5	4	0.825	0.00296	0.001013	512.5	67.8
4/29/1999	5	4	0.105	0.00465	0.001592	65.2	106.5
4/30/1999	5	4	0.19	0.00340	0.001161	118.0	77.7

* The shaded area indicates the first day during the event on which the farm started to discharge

8.2.2 Impact of Rainfall on Discharge and P Load Analysis

8.2.2.1 Rainfall series aggregation

The generated farm level rainfall events and their corresponding runoff and phosphorus volumes are aggregated based on multiple characteristic groupings of land use, soil and water detention levels. This type of grouping assumes that runoff volume and phosphorus release from an individual farm depend primarily on these characteristics. Grouping farms of the same land use, soil and water management detention levels provides a large enough sample for a viable statistical analysis.

As such, over a group of farms of the same land use, soil and water detention level, we can reduce the farm-level time series of rainfall, discharge flow and phosphorus load into a set of event-based rainfall volume, discharge volume and total phosphorus load which are all independent of time of occurrence.

8.2.2.2 Analysis approach

Relations of rainfall versus runoff and rainfall versus phosphorus load are analyzed by regression methods that represent rainfall event values by fitting curves. Given a set of individual rainfall event volumes and the corresponding flow and loads, the regression curve would demonstrate the correlation between the independent variables (rainfall) and the dependent variables (runoff volume and phosphorus loads). All combinations of land use, soil and water management detention level are investigated by the regression method.

Figure 8-2a and Figure 8-2b show examples of scatter diagrams of rainfall and runoff and rainfall and phosphorus loads. The dataset gathers all rainfall events and the corresponding flow (a) and load (b) over farms of land use A, BMP 1.0" and soil Dania. It is observed that event rainfall and event runoff have relatively close correlation for this combination of categorization. Event flow volume increases with respect to rainfall volume. The P loads, however, show a relatively weak correlation with rainfall. This inconsistency may be attributed to various reasons, including the fact that the relationship between rainfall and P load may not be described by a single regression curve. Seasonal variation due to farming activities could significantly affect the rainfall phosphorus load relation. More complex categorization (farm area, number of structures, etc.) or detailed data are necessary. The fact that daily and event phosphorus loads are derived from concentration composite values can also affect the correlations. Also, if farms behave very differently within a category, variations within the group may considerably disperse the scatter points group. More detailed analysis on the correlation between rainfall events and the event phosphorus loading outside the scope of work is necessary to address the inconsistency.

The regression curves applied in the analysis are lower order polynomials (linear or quadratic).

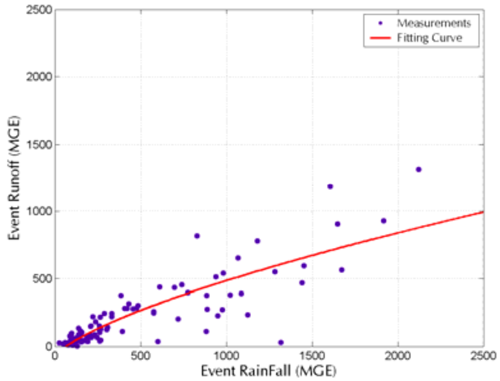


Figure 8-2a: Event rainfall and discharge correlation (Land use A, Soil Dania, and BMP1.0)

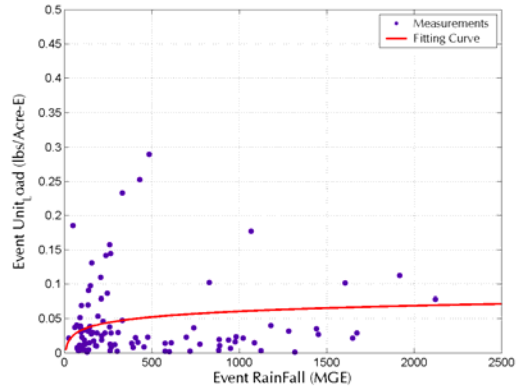


Figure 8-2b: Event rainfall and P-Load correlation (Land use A, Soil Dania and BMP 1.0)

9.0 CONCLUSIONS AND RECOMMENDATIONS

This analysis provides general findings based on the statistical analysis of farm level regulatory data collected under the requirements of Chapter 40E-63, F.A.C. These data represent an extensive set of flow, phosphorus concentration and load data at the farm level. However, since the datasets were not collected for this type of analysis, there are several constraints associated with the results. The specific conclusions indicated below verify assumptions associated with farm level discharges (seasonality analysis results), serve as a high level screening tool for more detailed analysis (potential interconnection and spikes analysis results), or identify supplemental data or refined analyses that would be needed (location, inflow and detention analyses). Indicated below is a brief summary of the results for the individual sections and recommendations:

Seasonality

- The decomposition analysis indicates that farm discharges in the S3 and S8 sub-basins as a whole show either moderate reduction or stabilization in their P concentrations and unit loads. S5A, however, shows a significant decrease in concentration and a slight increase in load, suggesting an increase in flow. The long-term trend indicates that S5A concentrations at the beginning of the series were significantly higher than in the other sub-basins. The long-term trends for the characteristic groupings (soils, land uses and water detention levels) show moderate variation over years consistent with the sub-basin results.

- There is qualitative existence and consistency (in term of cyclic variation phases) of the seasonality changes in the time series. The seasonal variations in P load and concentration were found to depend on multiple characteristics.

- Analysis of flow trends as a function of land use, soil type, or water detention levels, was not included in the scope of work for this analysis. Its evaluation, however, may serve to better understand the separate effect of flow and concentration on P load trends. Use of multiple characteristic groupings may serve to understand the combined effect of various characteristics. Use of stepwise functions can be applied to the dataset, when it is suspected that sudden changes may have affected P loads and flows (e.g., removal of land to construct a STA).

Location

- The location analysis did not find meaningful correlations between the discharge loads and concentrations of adjacent structures. These concentrations, and the loads derived from them, only represent the water quality conditions at the outlet, on the upstream side of the structure. A major constraint for the analysis is assumed to be the lack of farm irrigation information and canal water quality data (irrigation schedules, flows, water quality). This information is not collected and, thus, it is not available for analysis. The "no correlation" finding cannot be assumed as definite since the information available to perform the analysis does not represent the complete hydrology and phosphorus transport processes that would explain the potential correlation.

- Evaluation through a site-specific analysis is likely to be more appropriate to address this section of the scope of work.

Potential Hydraulic Interconnection

- The practice of interconnecting drainage systems of multiple farms is a practice that already exists in the EAA where farms are typically same-owner farms. The first level screening-type analysis identified farms that may be candidates for hydraulic interconnection.

- Analysis of the many practical constraints for connecting was outside the analysis scope. Site specific evaluation of critical factors such as ownership, crop type and water retention capacity to determine the feasibility of these arrangements is needed. Site-specific studies may also be necessary to assess how runoff diversion and intermixing could serve to improve water quality.

- An assessment of land uses and water management practices among adjacent farms would benefit this analysis.

Lake Okeechobee Inflows

- Limited farm irrigation information and canal water quality data (irrigation schedules, flows, water quality), and reduced Lake Okeechobee inflow information in comparison to Sub-area discharge data were attributed as the major reasons for the lack of correlation between Lake Okeechobee inflows and discharges from farms downstream. Use of daily averages for each month and data aggregated by sub-areas may also have simplified the analysis such that correlations could not be found.

- Different aggregation and lag techniques to test the correlation of data may improve the analysis method.

Spikes

- Days with high P load and concentration discharges (defined as spikes for this analysis) were identified for all characteristic groups. These spike events can be associated with higher discharge pumping, higher P concentration or both. The available data does not provide the level of detail to identify the specific causes behind these discharges or whether they are avoidable. An estimate of the phosphorus contribution of spike events indicated that they represent approximately half of the phosphorus loads discharged in a year, and that these events occur within relatively short time spans (one to two days). Spikes occurred both during the dry and wet season. Spikes occurring during the dry season have a larger contribution to the total seasonal loads, while wet season spikes are higher.

- It is recommended that farm level spike information be provided to permittees for self assessment and understanding of the effect of spike events on annual loadings. This information will be useful for the permittee to identify the causes for these events and adjust farming practices to optimize BMPs as applicable.

Detention

- Rainfall-runoff and rainfall-P load curves were derived using a simplified standard rationale based on rainfall events (i.e., consecutive rainfall days). Correlations between rainfall and farm discharge flows were identified, thus providing a general relationship between the criteria defined rainfall-event and the resulting runoff for multiple characteristic groupings. However, no direct relationship between rainfall and P loads could be verified through this analysis. Therefore, the specific effect of rainfall detention on phosphorus concentration and load could not be evaluated to address the scope of work.

- A site-specific evaluation including detailed field data collection may be necessary. Use of modeling was also suggested as a tool to more specifically understand the effects of detaining runoff on flows and loads at the farm level. Information on water table levels was not available for the analysis and is necessary for the type of assessment described by the scope.

Summary

Analysis findings support expectations associated with EAA farm discharges that had previously only been assumed, and not based on actual review of data. The specific causes behind the P load and concentration relationships presented in this analysis, however, cannot be identified with certainty because of the lack of information on the many factors affecting the characteristics of farm discharges. Use of more refined datasets is necessary. It is recommended that suggestions for more refined analyses be provided to the University of Florida IFAS for consideration in their BMP research efforts and data analysis (using BMP research data collected by IFAS at the farm level), or for additional field investigation. The analyses may fit within the objectives of the EFA-mandated BMP farm research that is conducted by IFAS under the Master Research Permit (issued as a requirement of Chapter 40E-63, F.A.C.) to the EAA-EPD.

10.0 REFERENCES

1) South Florida Water Management District *Statistical Analysis of Everglades BMP Program On-Farm Total Phosphorus Concentration and Load Data* 1995

2). University of Florida Institute of Food and Agricultural Sciences *(UF/IFAS) Research Report* 1995

3). South Florida Water Management District *FINAL REPORTEVERGLADES PROTECTION AREA TRIBUTARY BASINS* 2003

4) South Florida Water Management District: *Everglades Program-Best Management Practices Annual Report* 1995-2003

5). James T. McClave, Terry Sincich *Statistics*, 8th edition, Prentice Hall, c2000

6). David G. Kleinbaum, Lawrence L. Kupper, Keith E. Muller *Applied regression analysis and other multivariable methods* Boston, Mass. : PWS-Kent Pub. Co., c1988

7). Brockwell, Peter J. *Introduction to Time Series and Forecasting* 2nd edition New York : Springer, c2002

8). Alvin C. Rencher *Methods of multivariate analysis* New York : J. Wiley, 2002

9) SAS Institute Inc. *SAS Online Documents*. 2000

10) D'Agostino, R.B., and Stephens, M. (1986) *Goodness of Fit* New York Marcel Dekker