Appendix 1-2: Peer-Review Panel and Public Comments on Draft Volume I

During September–November 2014, comments from the peer-review panel were provided on the 2015 SFER WebBoard (www.sfwmd.gov/webboards); public comments were also received during the review period. Panel comments were prepared under separate Purchase Orders with the South Florida Water Management District. This content was not edited by the SFER production staff and appears verbatim as posted on the WebBoard.

Note: Comments and recommendations from the panel were sought on, but not limited to, the following general questions targeted for two levels of review:

1. ***Accountability Review.*** *Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate? Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report? Are findings linked to management goals and objectives? The panel may also provide constructive guidance for the District's large-scale programs, particularly as related to water quality assessment and control across the agency.*
2. ***Technical Review.*** *Are the findings and conclusions supported by "best available information," or are there gaps or flaws in the information presented in the document? Are there other interpretations of the data or other available information that should be considered by the authors and presented to decision makers? If so, the panel shall identify specific studies that should be addressed or available data to support alternative findings.*

2015 Draft South Florida Environmental Report (SFER)

Chapter 3A – Water Quality in the Everglades Protection Area

"AA" Review Prepared by Vladimir Novotny

Accountability Review. This review targets progress and achievements of expectations in District programs and projects that are generally descriptive or standardized, and may deal with cross-cutting themes or content.

This chapter (1) provides an assessment of water quality within the Everglades Protection Area (EPA) during Water Year 2014 (WY2014) (May 1, 2013–April 30, 2014), (2) fulfills numerous reporting requirements of the Everglades Forever Act (EFA), (3) provides a preliminary assessment of total phosphorus (TP) criterion achievement, and (4) provides an annual update of the comprehensive overview of nitrogen and phosphorus concentrations and loads throughout the EPA. The SF Water Management District (District) evaluated in this chapter for WY 2014 62 pesticides and 15 water quality constituents. In total throughout the previous years the total number of parameters was 109 but some parameters, for example, metals, were found as not threatening the water quality of the Everglades system.

The reporting goals of the chapter were:

1. To summarize areas and times where water quality criteria are not being met and indicate in excursions over space and time.
2. Discuss factors contributing to excursions from water quality criteria and provide evaluations of natural background conditions where existing standards may not be appropriate.
3. Present an updated review of pesticide and priority pollutant data made available during WY 2014
4. Present a preliminary TP criterion achievement assessment for different areas within the EPA for the most recent five year period (WY2010-WY2014).
5. Summarize phosphorus and nitrogen concentrations measured in surface waters within different portions of the EPA.
6. Summarize the flow and phosphorus loads entering different portions of the EPA and describe spatial and temporal trends observed.
7. Describe and discuss factors contributing to any spatial and temporal trends observed.

The team of authors over the years developed a good writing format and protocol along with graphical presentations which make the chapter very uderstandable and easy to compare to previous years. They are consistent with units, using mostly SI unit system, and provide conversion to the other systems when needed. In an overal assessment it can be stated that the chapter as presented fulfiled the above goals.

Desciption of the system and its water quality issues and problems

The entire Everglades Protection Area (EPA) encompasses 10,020 km^2 (3,877 sq miles or 2.5 million acres) of the original Everglades marsh which is currently divided into several large separate distinct shallow impoundments. The most important is the Everglades National Park

(ENP). In addition to rainfall inputs, which is significant when compared to surface sources, the surface water inflows from agricultural tributaries, such as the Everglades Agricultural Area (EAA) to the north and the C-139 basin to the west, feed the EPA. The EPA also receives surface water inflows originating from Lake Okeechobee to the north and from predominantly urbanized areas to the east. Today, surface runoff passes through several large Stormwater Treatment Areas (STAs) providing treatment and removal of phosphorus from runoff (see Chapter 5b) . The STAs are shallow marshes covered by submerged and emerged aquatic vegetation.

The EPA consists, from north to south, of Water Conservation Area 1 (WCA-1)/Arthur R. Marshall Loxahatchee National Wildlife Refuge (Refuge), WCA -2A, WCA-2B, WCA-3A, WCA-3B and ENP. The main purpose of the chapter was to analyze and interpret the historic trends which considered four phases: (1) This historic Baseline Period of WY1979-WY1993; (2) Phase I intermediate period of WY 1994-WY2004 which includes initial but limited implementation of Best Management Practices (BPMs) on agricultural and some small urban lands and construction of STAs; (3) Phase II of WY-2005 to WY2013 which includes further optimization and enhancement of performance of BMPs and STAs and adding more restoration projects under the Everglades Protection Act; and (4) the current Water Year 2014.

The first and overwhelming impression anyone reviewing this chapter will have is the fact that the BPMs, STAs and other restoration measures undertaken over the last fifteen years are working and in many instances, highlighted in this review, the water quality goals are being met. This WY report presents also positive trends and progress in water quality improvement and the District and other Florida agencies should receive praise for their work. However, great challenges still remain and work must continue and intensify. It can also be pointed out that due to ongoing long term climatic changes the challenges will be increasing and may even include land losses as the oceans will rise and disruptions by more frequent extreme weather conditions such as hurricanes.

Comment on water quality assessment: On page 3a-11 the writers define the criteria for assessing the impaired water quality based on the frequency of exceedances whereby exceedances between 0 to 5% of samples would be classified as minimal concern, 5 to 10 % exceedances would be classified as potential concern, and 10% exceedances would be classified as concern. They stated that such classification is based on Section 303(d) which is a section of the Clean Water Act dealing with impaired waters and TMDL. However, under this section the allowed exceedances are far more stringent. The more lenient above exceedances have been used by states and EPA for the CWA Section 305 assessment for reporting water quality to Congress by the states. For example, exceedance of 5 % of Dissolved Oxygen (DO) samples would not be acceptable by the Section 303(d) water quality regulations and even frequency of DO exceedance on the level of 1% would imply, according to the federal Water Quality Criteria, a violation.

With respect to pesticides, these lenient exceedances cannot be used for priority pollutants and apparently this year the Florida Department of Natural Resources adopted site specific (for Everglades) Class III water quality standard.

DO in the Everglades was in the last years reports evaluated according to a site specific standard that calculated the allowable DO based on the time of the year and temperature, evidently mimicking the "natural" DO concentrations in the system. This standard was discussed extensively in the previous WY reviews and the discussion will not be repeated herein; however, the arguments against the protective effectiveness of this harmonic standard allowing DO concentrations to drop to very low levels remain.

Pesticides were evaluated under the assumption that the Class III criteria values represent instantaneous maximum concentrations for which any exceedance constitutes a non-attainment of designated use. Pesticides were categorized based on whether the parameter was detected at concentrations above the MDL (potential concern) or at concentrations exceeding Class III criteria or chronic toxicity values (concerns). MDL was not defined in the text of the chapter and most likely it means Minimum Detectable Limit.

The criterion for assessment of phosphorus is discussed on page 3a-12 without actually stating what the criterion is. This standard (10 µg TP/L) is revealed later in the chapter in the section dealing with phosphorus.

Unit system. The authors have used predominantly SI system which is the most preferable unit system for reports aimed at an extensive national and international audience. When US system is used the chapter provides conversion factors. There is one odd unit invented by authors, the Kacre-ft, first introduced on page 3A-15 which is an awkward mix of metric prefix with an archaic US unit. Should we also have units like Kinch for 25.4 meters or mlb for 0.45 grams? Certainly not. If the authors insist on using acre-ft (why?) then 14.6×10^3 acre-ft (18×10^6 m^3) would be more appropriate. For larger volumes, 99% of world hydrologists would use Km^3 or 14.6 "kacre-ft" = 0.018 Km^3.

Hurricane Andrea. Table 3A-1 on page 3A-15 reports water quality data collected at G-300 and G-301 diversion structures. Although the path of the hurricane missed southern Florida, a large rainfall impacted hydrology and water quality inside EPA. Of note also were low DO concentration 2 to 3 mg/L that could have adversely affect fish population..

Excursion Analysis. Problems with exceedances of Florida Class III criteria were identified and summarized in Table 3A-2. Four water quality parameters were identified: Dissolved Oxygen, Alkalinity, pH, alkalinity, and specific conductance. These parameters were evaluated for inflow, interior, and outflow for each water body of the EPA. It should be pointed out that under the present scenario the DO standard is different for the inflow and outflow (Florida Class III) and interior (site specific alternate criterion – SSAC). This leads to a questionable situation where interior water quality meets the site specific more lenient standard but the same water quality would fail the inflow and outflow standard which is more stringent. The responsible agencies

should again revisit the standard and, by a Use Attainability Analysis, develop a unifying side specific standard that would protect indigenous biota. During WY2014 seven interior stations violated the SSAC criterion. The DO fluctuations and exceedances in the EPA are related to the nutrient concentration promoting algal growth that by respiration depresses oxygen during cloudy days and darkness. Many wetlands are naturally dystrophic which is exhibited by a low DO.

The water quality excursions related to pH and alkalinity may not be a concern. The apparent violations are (1) cross-correlated (low alkalinity is often accompanied by low pH); (2) related to geology of the Everglades, (3) natural acidity related to formation of humic acids in wetlands; (4) the chemistry of the interior areas is dominated by rainfall that has naturally very low alkalinity and low pH that would violate the surface water criteria. Typically, natural unpolluted rainfall has pH around 5.7 and has very low alkalinity. This again may call for an UAA type analysis. Alkalinity problem was detected only in the inflow into Refuge area.

Specific conductance was evaluated by the Florida Class III specific conductance which was quoted in the chapter as : *Class III criterion allows for a 50% increase above background conditions in specific conductance or 1227 µS/cm, whichever is* ***greater***. The reviewer believes that it criterion should end with ….**smaller** not …greater. For example if the natural specific conductance is 400 µS/cm, typical for EPA then the limit should be 600 µS/cm and not 1227 µS/cm. The latter would damage fish and other organisms due to salinity shock. This could result in more criteria excursions than that reported in the chapter.

Pesticides and metals. The report has not reported problems of concern with the concentrations of pesticides and metals. Metals were not measured during WY 2014. On line 596 identify or correct …including III, …

Phosphorus. The second half of the chapter deals with the phosphorus and nitrogen content of waters inside of EPA. For these waters, the Florida Department of Environmental Protection established a relatively stringent numeric criterion of 10 µg/L for Total phosphorus (TP) applied to a long term geometric average of the concentrations. For nitrogen only narrative criterion is available; however, studies have proven that the EPA system is phosphorus limited. In addition to TP, the District study also measured orthophosphate and soluble reactive phosphorus (SRP). The study presents also spatial and temporal trends which are satisfactorily shown in tables, charts and maps.

In the WY 2014, for the first time in history, the geometric mean in all interior areas (Refuge, WCA2, WCA3, and ENP) were below the long term standard of 10 µg/L. Specifically, the interior concentrations inside EPA was only 3.6 µg/L. This is very good news. The spatial gradient, as one would expect, is from the highest in the Refuge area in the north to the lowest concentration in the south in the ENP. Refuge areas receive agricultural phosphorus rich runoff. The most dramatic decreases in TP concentrations occurred in the water conservation areas WC2 and WC3. The STA study in this WY2014 report (Chapter 5b) documents that once the TP concentrations become very low, e.g., near the criterion limit of 10, further decline does not

occur, as a matter of fact in this particular study at very low concentrations (1 to 10 µg/L) the removals were negative.

Since the removals of TP in very low concentrations by known mechanisms attenuating P in Everglades marshes cannot be identified or proven and the information on P in wet atmospheric deposition over ENP and WCAs were not provided, there is no information available to speculate that dilution by rainfall would drop TP concentrations below 10 µg/L. Based on the information provided in Chapter 10, WY2014 was one of the wettest years which could cause the very low TP concentrations throughout the EPA system; however, Figure 3A-13 show that flows through the EPA have not significantly changed during the last five years.

Referring to lines 800-802, define the four part test of meeting the P standard for Everglades.

The flow and mass balance flow charts in Figures 3A-11 and 12 on pages 3A 34-35 are very illustrative; however, the atmospheric contribution appears not to be included. It could be as much as 30% of the load. Please, do something with the kacre-ft unit on these figures, in Table 3A-7 and also in the text of the chaopter. For example, on line 806, (1,096 kacre ft) would be 1.3 km^3.

The decreases are even more significant for orthophosphate (page 3A-39) wherein the interior concentrations are around 1 µg/L. What is the minimum detection limit?

Nitrogen. Apparently, the EPA is phosphorus limited in most parts; therefore, nitrogen concentrations may not be as important as phosphorus but historically N concentrations have appeared to be on a higher side. Some marine water bodies in the everglades system such as estuaries draining Lake Okeechobee (see Chapter 10) were found to be N limited. Line 991 states that "When nitrogen is limited biota (algae and plants) can offset this limitation by fixation of atmospheric N_2." This statement should be checked. Very few algae if any can fix N_2, only cyanobacteria microorganisms (e.g., Anabaena) can fix N_2 but cyanobacteria are not algae. Legume plants can fix N_2 but can they be found in large quantities in the EPA? Cyanobacteria were found years ago and maybe even today in Lake Okeechobee and many other Florida lakes and marshes.

Line 996 in parentheses should read ..(TKN= ammonia + organic nitrogen)… Recent investigations revealed that both P and N must be controlled if cyanobacteria infestation reaches very high hyper-eutrophic levels (see, Paerl et al, 2010) which is not the case of the EPA and Lake Okeechobee today.

The data on N statistics in Table 3A-7 show significant improvement since the Baseline period. Further improvement can be expected by improving the performance of agricultural BMP, STAs, by improving mileage of automobiles and in the near future even by switching to hybrid and electric cars. Automobile traffic is a significant source of NO_x emissions. However, Figure 3A-17 shows that most of nitrogen present in the EPA system is organic N. Unlike nitrates or ammonia , organic nitrogen is not readily available for algal growth,

Concluding Assessment

Chapter 3A is well written and organized and provides a wealth of information. Accordingly, based on the information, it can be concluded that the status of the EPA has been significantly improving owing to the implementation of STAs, agricultural and some urban BMPs, flow management and also due to more favorable meteorological conditions with less catastrophic disruptions. Since the major STAs have been operated only few years after finishing and optimization and expansion of BMPs is ongoing, further improvements can be expected.

The chapter again pointed out obvious long lasting problems and controversies with the DO concentrations in the Everglades which are partially natural, partially caused by anthropogenic sources of nutrients in the system.

The authors have accomplished all goals that they identified for this chapter.

The authors made an effort to avoid confusion with units which was endemic to the past reports and even today to some other chapters. Today at the beginning of the new millennium it may be necessary to adhere to both unit system with conversions and not try to invent new units.

Literature:

Paerl, H.W., H. Xu, M. McCarthy, G. Zhu, B. Qin, Y. Li, and W. Gardner (2010) Controlling harmful cyanobacteria blooms in a hyper-eutrophic lake (Lake Taihu, China): The need for a dual nutrient (N & P) management strategy, *Water Research* doi:101016/j.waters.2010.09.018

REVIEW CLOSURE BY VLADIMIR NOVOTNY (AA)
DRAFT 2015 SFER – VOLUME 1, CHAPTER 3A

This chapter describes the progress achieved in water quality control in the Everglades Protection Area (EPA) consisting of the National Wildlife Refuge (WCA-1), Water Conservation Areas 2 and 3 and Everglades National Park (ENP). The trend analyses presented throughout the chapter persuasively show the decreasing trend of concentrations of phosphorus and several other key pollutants. This is a result of the last fifteen years of implementing and in the last few years fine also tuning the Best Management Practices (BMPs) in the agricultural areas and storm water treatment practices in STA protecting the system. Again, this chapter was well written and presented good statistical analyses of concentrations, trends and sources of several key pollutants.

In this period, after years of planning and implementation several milestones outlined in the Everglades Protection Area have been achieved such as meeting the very stringent phosphorus standard (goal) in most areas of the EPA, yet the efforts must continue, many of them in perpetuity, to keep the ENP healthy and thriving.

In their response, the authors adequately addressed the reviewer's comments and suggestions. If the Comment is not mentioned in this closure, no change or further clarifications were not needed. Specifically:

In *Comments 2* the authors and the reviewer in their comments continued the discussion on adequacy and appropriateness of magnitude and frequency of allowable excursion used in the analysis. The authors used the assessment frequency commonly used in the so called Clean Water Act - Section 305 reports that allows 5% and 10% of samples not to meet the standard before concern with the water quality would be triggered. The reviewer pointed out that such "lenient" frequency generally used by states in the mandatory state water quality reports to the Congress may not be protective enough and would be too soft to identify impaired water body. This is the reviewer's scientific opinion. However, it is recognized that the writers of the report must follow the state guidelines. The authors pointed out that the State Impaired Waters Rule used as the basis for the report excursion analysis has been approved by the US EPA for reporting and the allowed percent of exceedance actually is a percent exceedance at the 90% confidence interval which makes the assessment exceedance criterion somewhat more stringent.

Similarly, *in Comment* 3 authors and the reviewer again continued their discussion on the adequacy of the dissolved oxygen (DO) standard for the internal areas of the EPA which was in the current 2015 report further complicated by adding another and vastly different standard for

the inflow and outflow zones of the areas comprising the EPA. This question will be revisited in the discussion of Comment 8.

In *Comments 5 and 6* the authors responded to the editorials by the reviewer about the definition of the phosphorus standard and mixed atypical unit for annual flow which will be corrected by the authors.

Comment 8 focused on the conflicting assessment of dissolved oxygen (DO) of the interiors and input/output zones of the areas of the EPA. The point in question is the fact that the site specific DO standard for interior areas calculated by a harmonic model where the input is temperature and time of the year, allows very low DO concentrations that normally would be of great concern if, for example, federal DO criteria would have been applied. However, all agree that the low DOs in the system, considering the already established fact that the phosphorus concentrations in the interim areas are already at or below the very stringent P standard, are most likely of natural origin, characteristic of subtropical wetlands. Furthermore, the authors are again bound to report the DO in the context of the state standard.

Nevertheless, the reviewer again is using his prerogative to point out a dichotomy between the standards of the interior and the narrow input/output zones wherein the yardstick is a far more stringent Class III DO standard. This leads to an odd situation of meeting the site specific DO criterion in the interior on one side and extensive noncompliance with Class III DO standard in the input/output zones, yet, there is almost nothing available to bring about the compliance short of installing dozens of energy demanding aeration devices to remedy something that may be natural to the system. And the authors apparently agreed that the low DOs may be natural as they are in the interior sites.

Comment 9 deals with the excursions of pH and alkalinity from the state standards. Both the reviewer and the authors agree that low pH and alkalinity (these two parameters are cross-correlated) are due to natural precipitation inputs into the system which dominate especially in the EPA. Normally, a Site Specific criterion should have been developed under these circumstances. The reviewer suggests that Use Attainability Analysis is a proper vehicle to develop a site specific standard because, according to the US EPA UAA guidelines (40 CFR 131.10(g)), the first of the six key conditions under which the standard can be modified is the situation wherein natural water quality does not meet the standard by which the use is defined. Otherwise the current and future reports have to report the noncompliance and be subjected to critique and calls for action to remedy the violations. However, similarly to the last years report, the district has put a low priority on solving this problem.

Comment 10 deals with the semantic of the Florida specific conductance standard. The reviewer pointed out illogical definition of the upper limit in the Florida narrative for the standard (“smaller” instead “greater”). But it is recognized that the objective of the SF Environmental Report is not to dispute the phrasing of the standard but the Florida DEP should take a note. Alleged problems with specific conductance were satisfactorily explained by the authors.

Comment 12 is an extensive discussion of phosphorus standard and trends of TP concentrations. Again, the most important finding of the chapter is the fact that the phosphorus loads to the system have been decreasing and the stringent geometric mean limit of 10 μ P/L is generally met in the ENP and in most of the other conservation areas. The reviewer mentioned in his comment on Chapter 3A soluble reactive phosphate (SRP) which was not included in this chapter; however, it was included in Chapters 3B and 5B dealing with the sulfur –mercury relationship throughout the EPA and performance of Stormwater Treatment Areas, respectively, which he also critiqued. STAs are adjoining the EPA. So the SRP data were taken and are available but not mentioned in Chapter 3A. The authors satisfactorily explained the reviewer's concern with alleged incomplete information on the atmospheric inputs and pointed out where this information was included and is available.

Comment 13 is an extensive discussion of the nitrogen impact on the system. The EPA system has been found phosphorus limited and most of the nitrogen in the system is particulate organic N which is less available to algae. However, in their response the authors revealed that the Everglades ecosystem is known for having some areas where cyanobacteria may constitute a large portion of periphyton and be a source of nitrogen. This may be a point of concern because cyanobacteria (they are not algae) may under favorable conditions develop as bloom and overpopulate both the water column and sediment where, unlike algae, may survive and reappear years later as another bloom.

KEY FINDINGS AND RECOMMENDATIONS BY VLADIMIR NOVOTNY (AA) DRAFT 2015 SFER – VOLUME 1, CHAPTER 3A

- Chapter 3A was well written and provided a good assessment of water quality and its trends in the Everglades Protection Area consisting of the National Wildlife Refuge Area (WCA-1), Water Conservation Areas 2 and 3 and the Everglades National Park.
- The chapter fulfilled the reporting requirements, and provided assessment of Total Phosphorus (TP) criterion in the EPA system. The stringent protective goal of 10 µg/L of TP was met in WY 2014 in all interior areas of the EPA system and also in the inflow to the ENP. The trend in TP concentrations over the years is decreasing.
- The majority of other water quality parameters, including pesticides meet the standing Florida water quality standards.
- Four parameters (dissolved oxygen, pH, alkalinity, and specific conductance) deviated from the standard; however, the exceedances for pH and alkalinity are most likely caused by concentrations that are of natural origin, namely, due to naturally low pH and alkalinity in the precipitation dominating the inputs into the EPA system. Developing a Site Specific standards for pH and alkalinity may be needed otherwise the repeating reporting of "violations" in every annual report may be confusing to the public. This issue is a repeating item because it was also introduced in the 2014 SFER critique.
- Frequent DO concentration excursions were classified as a concern. However, there is dichotomy with having two vastly different DO standards applied to the internal waters and the rim and outflow areas. In the interior the Site Specific Standard mimic the natural occasionally low DO concentrations characteristic for subtropical wetlands. For the inflow/outflow a far more stringent Class III freshwater standard was used which led to extensive violations. Physically, outflow quality is very close to that in the interior. This dual assessment of the same quality, one classifying it as acceptable the other as extensive violation of the standard, is also confusing, inappropriate and should be addressed. This issue was also extensively discussed in the 2014 SFER.
- The disclosure of development of cyanobacteria in some areas of the EPA system should be addressed in the next year SFER. Cyanobacteria blooms if they develop are difficult to control and represent a serous ecological nuisance and produce toxins. However, it appears that the average TP concentrations in the entire EPA system are at the level not favoring Cyanobacteria blooms but the threat may increase as a result of global warming. Furthermore such blooms have developed in other parts of the Everglade system such as the Lake Okeechobee, which feeds the EPA, and the estuaries.

Peer Review of Draft 2015 SFER Chapter 3B: Regional Mercury and Sulfur Monitoring

Authors: Otto Stein (AA) and Vladimir Novotny (A)
Level of Panel Review: Technical

Technical Questions:

1. Are the findings and conclusions supported by "best available information," or are there gaps or flaws in the information presented in the document?

2. Are there other interpretations of the data and other available information that should be considered by the authors and presented to decision makers? If so, the panel shall identify specific studies that should be addressed or available data to support alternative findings.

The review is divided in to several sections. First are several broad questions and comments regarding the interpretation of reported results. These reflect the general impressions of the reviewers after reading and reflection of the entire document. These are followed by more focused comments and concerns that in many cases back up the first section, but in other cases are merely requests for clarification or better presentation. In general, these were created on a section by section analysis. This is followed by comments on figures and tables, editorial suggestions

Broad questions and comments to be addressed:

The Chapter 3b in the 2015 report provides an assessment of sulfur and mercury (Hg) status within the Everglades Protection Area (EPA), and surrounding areas during the Water Year 2014 (WY2014) (May 1, 2013–April 30, 2014) reporting period. The information provided in this chapter is an update to Chapter 3B of the *2014 South Florida Environmental Report* (SFER) – *Volume I*. This chapter updates the status of mercury and sulfur monitoring in the Everglades and deals with the serious problem of high and unacceptable methyl mercury contamination of higher trophic level organisms, specifically largemouth bass (LMB)and fish eating birds and exceedances of the safety criteria for human consumption as well as data on lower trophic fish.

The formation of methyl mercury (MeHg) is related to the biological activity of sulfur reducing bacteria (SRB), the same bacteria that, by reducing sulfate to sulfide, can release immobile iron and aluminum bound phosphorus into pore water as phosphate. Therefore this process may also be driving eutrophication, increasing dissolved organic matter, which was found in the 2013 report as one of the factors affecting formation of MeHg. The sources of sulfate include: (1) inputs from atmospheric deposition, (2) inputs of connate seawater (connate water= water entrapped in sedimentary rocks during the time of their formation) into canals, and (3) export of sulfate from the Everglades Agricultural Area (EAA).

Unlike the 2013 report (Chapter 3b) two years ago which, for the first time, presented an extensive synthesis on the interactions and effects of sulfates on formation of methyl mercury and contamination of fish and higher trophic organisms, the previous Chapter 3b in the 2014

report was mostly descriptive of data taken by the South Florida Water Management District staff and generally refrained from any synthesis of the findings. The reviewers of this chapter also reviewed the 2013 and 2014 reports and found that the 2013 Chapter 3b effort identified the causes and solutions of the methyl mercury contamination, the probable causes, interactions among the parameters, impacts on biota, and possible remedies (if any) and therefore represented the most comprehensive scientific endeavor since the mercury problem was identified more than forty years ago.

In the 2013 report Chapter 3b, the research by the authors and affiliated scientists produced evidence that the sulfate effect on MeHg accumulation in largemouth bass (LMB) was most pronounced when the water sulfate concentration was between 1 to 5 mg/L. This was explained by the fact, supported also by the literature, that the higher, possibly toxic levels of sulfide concentrations produced by the SRBs could have retarding effects on the formation of MeHg, which led to a conclusion, expressed by the "A" reviewer in the 2013 report review, that reducing sulfate levels below a certain low level, e.g. 1 mg/L, advocated as a criterion for the sulfate management by the authorities, may be counterproductive. This was confirmed in several places in the 2013 report which outlined a discovery of the SFWMD research team that the formation of MeHg is maximal in the "Goldilocks' window" of sulfate concentration between 1 to 5-10 mg/L and also depended on the concentration of *aromatic* dissolved organic. These low $SO_4^{=}$ concentrations are characteristic for the oligothrophic waters in the Everglades National Park where, indeed, the contamination of fish, fowl and higher trophic level fish-eating organisms is the highest and far above the standards even though the sulfate concentrations are low and at the level of the suspected natural background levels.

The previous 2014 report reported mostly the data, lacked synthesis and it appeared that the new authors were skeptical about the bell-shaped Goldilocks' relationship. Their conclusion was that there were few remedies available to reduce Hg contamnation, the main source of mercury, atmospheric deposition, remained the same, fish contamination by THg did not change either, and nothing could be done because little is known about the relationship of sulfates to MeHg formation. They wrote off the 2013 cutting edge result and claimed that there is no bell shaped relationship between sufate concentration and mercury methylation.

The more comprehensive investigations in the current 2015 report present far more data and provided again a thorough synthesis which in view of the reviewers, has clearly proven the validity of the Goldilocks' relationship although the chapter authors have difficulties admitting that. As in last year's report, authors are looking for clear-cut simple deterministic relations which, obviously, do not exist in a complex system such as Everglades. The phenomenon is clearly probabilistic and multi-parameter in nature. The lead authors in the argumentative section of the chapter starting on page 3B-59 present on Figure 3B-26 a single parameter relationship of MeHg content of fish to $SO_4^{=}$ water concentrations which clearly shows that the band between 1 to 10 mg of $SO_4^{=}$ contains the highest concentrations of the THg contamination of mosquitoes, sunfish and largemouth bass fish tissue. While there may be valid critiques to this study as outlined in the text, the fact that a single parameter model for fish tissue mercury concentrations can corroborate the Goldilocks' window is remarkable. The chapter also noted that 80% of THg is MeHg. Yet, the authors still are very dubious about the MeHg in fish and sulfate content stating (p. 3b-59) that "the ecosystem-levels of Hg accumulation (in fish) is not

predictable from $SO_4^=$ concentrations in marsh surface waters like the Everglades". The reviewers could agree with this statement if a word "alone" was added behind "concentrations".

The relationship reported in the 2013 report for the fish tissue MeHg and $SO_4^=$ is not unimodal, the authors and scientists of the 2013 pointed out that the aromatic hydrocarbons have a strong effect which was not mentioned in the 2015 report. The authors of the 2015 further state that "the data in this and other recent studies document that too many factors influence the biochemistry of $SO_4^=$ and Hg in natural ecosystems to allow any consistent and identifiable areas if high and low fish THg across the spectrum of surface $SO_4^=$ concentration". On Figure 3B-28, however, the authors of this study found an undeniable strong effects of $SO_4^=$ on MeHg formation at low $SO_4^=$ concentration between 1 – 10 mg $SO_4^=$/L, leveling off after 10 mg/L. Additionally some of the comments made in the argumentive section of this report are not corroborated with previously presented date (for example the strong relationship between sulfate additions and methylation but little correlation between methanogen activity and methylation (pages 3B39-40).

At first look, no one would disagree with the above uncertainty statement. The authors themselves on Figure 3b-27 and in text pointed out that alkalinity, pH, dissolved solids (conductance), chlorides, calcium, iron, methane forming bacteria, etc., may have some effect. Strangely, no one mentioned that the atmospheric input of Hg which varies significantly from north to south (but has not changed over the time) may have an impact, although it was thoroughly analyzed and presented in the Chapter. Elemental mercury availability is also likely be important. The authors also performed multi-regression analysis but failed to find a meaningful relationship. This should come as little surprise considering the complexity of the interactions and the simplistic linear relations assumed in liner regression analysis.

Because of the uncertainty, the authors rejected the proposal for establishing a 1 mg/L $SO_4^=$ standard for the Everglades and called more intensive investigations. The statement maybe correct, but the underlying reasons presented in this Chapter are likely the wrong ones.

The reviewers offer the following comments:

- The "Goldilocks" relationship has been confirmed by Figures 3b-26 and 3b-28 and other studies which implies that reducing high $SO_4^=$ concentrations would have an adverse effect on the MeHg concentrations and would enlarge the area with high fish MeHg contamination further to the north from the ENP which has the highest MeHg contamination. This was already pointed out by the reviewers in their assessment of the Chapter 3b in the 2013 report.
- The geographically variable but time invariant atmospheric input (more than 90% of the total Hg input) into the Everglade system should be considered and included in the analysis of the effects.
- The relationship is a multi-parameter nonlinear model for which linear or simplistically linearized multiregression analysis is not appropriate. T he reviewers suggest a two step analytical data mining process that would involve, for example, principle component analysis that would reveal cross-correlation between the determining parameters and more complex (e.g., neural net) building of a model between MeHg and key determinants revealed by the principle component analysis. See attached papers suggesting the

analytical model building process that would be applicable in the Everglades endeavor. This method also allows mapping of the effects. The study described in these papers is very similar to the problem of MeHg in the Everglades but would require a larger (more than 50) data sets.

- We accept that the suggested 1 mg $SO_4^{=}$ /L concentration as a standard is not attainable, unfounded and that trying to achieve it without considering other inputs and processes might bring more harm than benefits.
- Atmospheric Hg deposition is one of the key determinants and international efforts to change it from time invariant to decreasing parameter must be intensified. It is quite likely that switching from coal fired power plants (the largest source of atmospheric mercury) to other sources of electricity, especially green and solar power, may be the most effective solution.
- The multiparameter multistep nonlinear analysis may be necessary and will identify quantitatively the key parameters affecting the MeHg contamination and possibly lead to remedies. Focusing on one parameter only apparently is not leading to an effective solution.
- Ultimately the best solution will be found through development and calibration of process based models and many of the microcosm and mesocosm studies as well as studies focused on microbial activity are a step in the right direction. More studies on how Hg moves through the food chain in the Everglades system are also needed. Reviewers recognize that this approach is slow to achieve results sufficiently definitive for legislative action, but in the end may be the most cost effective solution. However the District must be willing to accept and espouse the results to the appropriate authorities, rather than making attempts to take the easy road and claim that the system is too complex to figure out. A critical reviewer could legitimately claim that the argumentative section is doing just that.

Specific questions and comments by line number as appropriate.

286-287: This sentence led this reviewer to look at the data more closely. Figure 3B-2 shows the temporal variation of Hg within a site, but it is not clear whether the data points in the box plots are of each individual fish or the averages for each year (the caption is not clear). The text in the paragraph in question offers no further guidance. This comes into play when applying statistics to the data. If as suspected the annual *means* are presented and indeed vary by that much, the key might be to compare the variances within a year the variance of the means from year to year. Looking at the data in Figure 3B-3 (in which a direct comparison for a specific site is impossible) it looks as if the within-year variation is similar to the between year. Therefore the simple answer to a lack of clarity might be you are not sampling enough fish at each site and year to reduce the variance relative to the mean to determine the differences you are looking for (or you are using the wrong sampling protocol, size, age class of fish etc). This comment applies to data on fish species Hg concentration data.

458-468: Reviewers of this chapter have questioned previous SFER claims that Hg concentrations in the fish species monitored are continuing to decline and recognize that this year's report backs off this claim. However the point of this paragraph is salient, mercury

concentrations in all monitored fish species, representing three trophic levels, have with few exceptions been high throughout the POR and continue to remain above criteria level.

683-751: The summary provided in these lines is in agreement with the nicely designed mesocosm study described immediately afterward. It does appear that elevated sulfate concentrations in surface water is not a driver of P release from native soils. The line of reasoning is pretty convincing; mesocosm data support this and, more importantly, four identified driving mechanisms for sulfate P release are not relevant to the conditions in most of the Everglades. However some caution must be exercised to not over extend these results; other locations (specifically the STAs) are accumulating sediment relatively high in P content and organic matter, therefore three of the four mechanisms (all but FeS_x production) may be more active in these locations.

Even more caution is warranted when trying to draw conclusions from the MeHg results of the mesocosm study. First, the results are not backed by any identified driving mechanisms and, in fact, run counter to some identified mechanisms as well as results from the microbiology study, which showed a correlation between sulfate concentration and methylation. Additionally, mesocosm results might influenced by relatively low concentrations of elemental Hg available for methylation (data not shown, this is speculation only). More importantly any conclusions drawn from this study include only two samples taken only 1/3 as often as other water quality data. This complex issue will require considerably more research effort.

985-987: Another (more likely?) alternative is that inhibiting methanogens gave SRB a competitive advantage stimulating activity of SRB and therefore *hgcA*.

1023: The word *potentially* is redundant with other words and clauses in this sentence. And makes for a very weak statement.

1126-1130: It does not seem likely that changes in just flow volume would have an influence on sulfate concentration, sulfate loads surely, but not concentration.

1183: The data an Table 3B-6 (and text on lines 1134) indicate that only WCA-2 has non-significant temporal trends so this statement is not true.

1270-1273: While these data do not prove that influent sulfate is biogeochemically active in EPA and ENP they certainly imply high activity for SRB and other sulfur transformations for which there is an overwhelming set of data concluding enhances Hg methylation.

1302-1304 and Table 3B-8: It is a little alarming to note the wide disparity for estimates of sulfur efflux from the EAA in published literature. However it is interesting to note that the James and McCormick study underestimates flow rate compared to two other measurements by a factor of approximately 4.5. It is conceivable that they missed measuring many of the potential efflux points.

1459: For which of the three applications? Which amendment was best? A few details are lacking here.

1498 (and earlier) It is not clear how burning sugar removes sulfur from the soil/water pools. Presumably most returns to the soil via ash or smoke deposition unless there is some proof that it is released in a vapor form.

1530-1543: It would seem to this reviewer that the "shape" of the relationship between fish THg levels and water sulfate concentration is not what we should get hung up on. The Goldilocks' concept implies that Hg methylation is maximized within the range 1-10 mg/L. Presumably it would be less as the outer ranges are approached (creating a bell shaped curve) but I do not believe the relationship is quantified well enough to predict that for methylation let alone how that MeHg might influence fish tissue concentrations, with so many more factors in between methylation and fish tissue concentrations. Considering the many factors involved, the fact that fish tissue concentrations appear to higher in the appropriate range (as is certainly the case in looking at the data in Figure 3B-26) is perhaps the more remarkable conclusion. That some of the data was excluded in the analysis (I assume that the presented data is ***between*** the 25 and 75 percentiles, but that is not what is stated) does raise some concerns but I assume the original publication provides some justification for that. As an aside, were the tails of fish tissue or sulfate data truncated?

1622-1627: WCA-2 is the region with the highest sulfate concentrations, so perhaps that is a confounding covariate in the alkalinity relation.

1713-1720, especially 1716-1718: These statements are simply not supported by the data presented in this report or in other sources. Clearly the point of this section is to suggest that microbial consortia other than SRB could be responsible for methylation and that is justifiable, but to state that methanogens are the dominant consortia even if limited to just periphyton communities is not justified. In fact it is contradicted the study highlighted earlier in this report which showed that methylation was not inhibited when methanogens were but was inhibited when SRB were inhibited.

1741-1742: The phrase "While these results were hypothesized…" seems incongruent with the rest of the paragraph. Results are from this study not other studies.

Figure and table comments:

Figure 3B-9: Why is natural logarithm used to transform the data? Unless there is a theoretical reason to do otherwise standard practice would be to use $\log_{10}$ values to perform statistical analysis. Also the figure has odd "ghost" data points that are hard to see.

Table 3B-2: The row labeled Freshwater ENP is not mentioned so it is not clear what these data represent. The text (lines 536-541) seems to indicate that data were collected from the Tamiami colony was sampled in 2011-2013 but these data are not shown.

Table 3B-4: How was atmospheric load to the Refuge calculated (which station)? Another way of interpreting this data is that depositional Hg is not easily transported by surface flow.

Figure 3B-21: Font in legend too small to read at even 150% magnification

Editorial page and line comments:

239: There is an extra period in the parenthesis

630: Define PWM the first time it is used.

693-694: This run on sentence confused the point of this study.

1016-1243: Numerous minor edits required in this section including lines 1094, 1131, 1132, 1136, 1148, 1191, 1216

1451-1458, esp. 1454 & 1455 Please reword for clarity; the contrast makes no sense.

1471-1507: Several editorial changes needed here including lines 1478, 1488, 1489, 1497

WATER RESEARCH 41 (2007) 4041–4050

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Extracting knowledge on the links between the water body stressors and biotic integrity

Elias Manolakos[a,*], *Hardik Virani*[b,1], *Vladimir Novotny*[c]

[a]*Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA*
[b]*The MathWorks Inc., Natick, MA 01760, USA*
[c]*Center for Urban Environmental Studies Northeastern University, Boston, MA 02115, USA*

ARTICLE INFO

Article history:
Received 10 March 2006
Received in revised form
27 April 2007
Accepted 6 May 2007
Available online 22 May 2007

Keywords:
Index of biotic integrity (IBI)
Self-organizing map (SOM)
Watershed management
Stressors
Fish metrics
Canonical correspondence analysis (CCA)

ABSTRACT

Multi-metric indices of biological integrity (IBIs) are most frequently created by examining single biological metrics along gradients of environmental degradation, and then combining multiple metrics using "best professional judgment" to characterize and calibrate stressor–response relationships. We aim to provide an efficient data analysis and visualization tool to assess the simultaneous effects of anthropogenic stressors on the fish population through the fish metrics and the associated Index of Biotic Integrity (IBI). Kohonen's self-organizing feature maps (SOM), unsupervised neural networks, are employed to pattern the sampling sites in the state of Ohio based on similar metrics characteristics. Canonical correspondence analysis (CCA) allows us then to draw conclusions about the role of the environmental variables in maintaining the perfect abode for fishes. Different visualizations superimposed with SOM clustering are realized to explore the complex interrelationships in the aquatic system and aid watershed managers to comprehend the effects of the environment on the fish.

© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

The Clean Water Act (CWA) passed by the US Congress in 1972 established, as a national goal, achieving and maintaining the chemical, physical/habitat and biological integrity of the Nation's waters. CWA also defined "pollution" as "the man-made or man-induced alteration of the chemical, physical, biological, and radiological integrity of water". Beside pollutant discharge, pollution includes alteration of habitat, hydrologic impacts of changed flow rate and its variability, effect on increasing of thermal fluctuation and magnitude by removing riparian vegetation, drainage of riparian wetlands, impounding stream, stream habitat fragmentation by impassable culverts and bottom change elevations and effects of impoundments (Novotny et al., 2005).

The last biannual State and USEPA reports to Congress, required by Section 305 of CWA, revealed that almost 50% of the Nation's waters do not meet the integrity goals of the CWA and the majority of causes are not chemical pollutants from point sources. Furthermore, several stressors that may

*Corresponding author. Tel.: +1 617 3733021; fax: +1 617 3738970.

Abbreviations: SPSCORE, total species metric score; DADSRNSCORE, darter/sculpins metric score; SUNSCORE, sunfish metric score; SUMINSCORE, sucker/minnow metric score; INTSCORE, intolerant/sensitive species metric score; TOLSCORE, % tolerant metric score; OMNISCORE, % omnivores metric score; INSSCORE, % insectivores metric score; TPIOSCORE, % top carnivore/pioneering metric score; NUMSCORE, fish abundance metric score; SPWNSCORE, %/number simple lithophils metric score; DELSCORE, DELT anomalies metric score.

E-mail addresses: elias@ece.neu.edu (E. Manolakos), hvirani@mathworks.com (H. Virani), novonty@coe.neu.edu (V. Novotny).

[1] This work was performed while the author was a graduate student at ECE Department, Northeastern University, Boston MA 02115, USA.

0043-1354/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2007.05.002

not cause individually harm or, more likely, for which standards are not available, may act in a synergistic fashion and stress biota.

"Integrity" has been defined by aquatic ecologists as the ability of a water body ecological system to support and maintain a balanced integrated adaptive community of organisms having a species composition, diversity and functional organism composition comparable to that of a natural biota of the region (Karr et al., 1986). Following this definition, the measures of integrity are three-dimensional, i.e., physical (habitat conditions), chemical (water composition and sediments) and biological (aquatic organisms) (Novotny et al., 2005). Key measurement endpoints of biotic integrity are composition and numbers of fish and benthic macroinvertebrates (Karr et al., 1986; Barbour et al., 1999) expressed as Indices of Biotic Integrity (IBIs). Human health effects could also be added because humans eat fish, (also water fowl), drink water and swim in it.

The hierarchical layered effect of progression of allochthonous and autochthonous stresses to risks to the integrity, and to the effects of internal habitat and chemical water and sediment risks on the integrity endpoints was advanced in Novotny et al. (2005). Risk progression begins with the landscape and pollution discharge allochthonous stresses divided into four categories: (1) landscape (e.g., imperviousness); (2) land use (e.g., agricultural, population density); (3) hydrologic/hydraulic (e.g., navigation, impoundments, change of hydrology by urbanization); and (4) pollutant loads. These allochthonous stresses at the bottom of the risk propagation pyramid, however, do not directly impact the biota, the main expression of integrity. Aquatic biota is impacted by in-stream (water body) stresses, such as habitat impairment (e.g., embeddedness, lack stream bank refuge, pool and riffle structure), and pollutant concentrations (risks) in water and sediments. In Novotny et al. (2005) we cautioned against using simplistic relationships and models relating indices of biotic integrity to a single stressor or even to multiple landscape stressors, of which the most popular one is percent imperviousness.

We are presenting here an efficient data mining and visualization methodology aiming at assessing the simultaneous effects of multiple anthropogenic stressors on the fish population through the fish metrics. The methodology first uses Kohonen's self-organizing feature maps (SOM) (Kohonen, 2001) and the *k*-means clustering algorithm (Duda et al., 2001) to partition sampling sites for the state of Ohio into groups based on similarity of their fish metrics characteristics. Canonical correspondence analysis (CCA) (Ter Braak, 1986), a community ordination method popular in ecology, is then applied to assess how environmental variables are associated with the formed patterns. Different visualizations superimposed on the SOM are realized to explore the complex interrelationships in the aquatic system.

2. Materials and methods

In this section, after introducing SOM, *k*-means, and CCA methods and the structure of the Ohio EPA data set that was analyzed, we proceed to describe a systematic methodology which combines them in a manner that can provide to watershed managers summarized, easy to visualize information about the biological integrity of an area of interest at different levels of detail.

2.1. *Self-organizing feature maps*

The SOM (Kohonen, 1990) is a popular neural network structure used for data dimensionality reduction and clustering. In essence it performs a structure preserving, nonlinear projection of high-dimensional input data vectors onto the low-dimensional (usually 2D) space of neurons (see Fig. 1). The data input vectors are presented multiple times one by one to the SOM network (multiple epochs). At presentation t, the input vector $\mathbf{x}(t)$ is compared with all the SOM neuron weights using some appropriate distance metric (e.g. the squared Euclidean distance, see Fig. 1). The neuron with the shortest distance to the input vector is declared as the winning neuron, also called the best matching unit (BMU). The weights of the BMU and its neighboring neurons are then updated to further reduce the distance between them and the presented input vector. This operation has the effect of increasing the similarity of the presented data vector and the weights of the neighboring neurons. The same steps are repeated till convergence or for a fixed number of epochs. Using competitive learning the SOM network encodes in its weights a low-dimensional representation of the unknown input data distribution. Weights adaptation is achieved in an unsupervised manner, meaning that no "teacher output" is required. Several practical SOM applications are listed in Kohonen et al. (1996). However, SOMs have been sparsely used in ecology, although some successful cases have been reported recently (Gevrey et al., 2004; Park et al., 2004).

2.2. *Canonical correspondence analysis*

CCA (Ter Braak, 1986) is an ordination technique widely used in ecological modeling (Ter Braak, 1994) to characterize the relationships between species abundance (e.g. fish), environmental variables affecting the species, and sampling sites. CCA is a direct gradient analysis method combining correspondence analysis (CA) with multiple regression, whereby species composition is related to measured environmental variables. A detailed explanation on the interpretation of a typical CCA plot is provided with the results later in Section 3.3.

2.3. *Ohio data set and IBI fish metrics*

Ohio pioneered the integration of biosurvey data, physical habitat data, and bioassays with water chemistry data to measure the overall integrity of water resources. The Ohio data set we have analyzed includes the chemical, habitat and biological data collected by the Ohio EPA between 1995 and 2000 (July–September). It consists of 1848 stations distributed over the entire state. Care was taken to ensure that most of the variables selected for the analysis were adequately represented in the data set with minimum missing values. Also, the time window for synchronising the dates for the

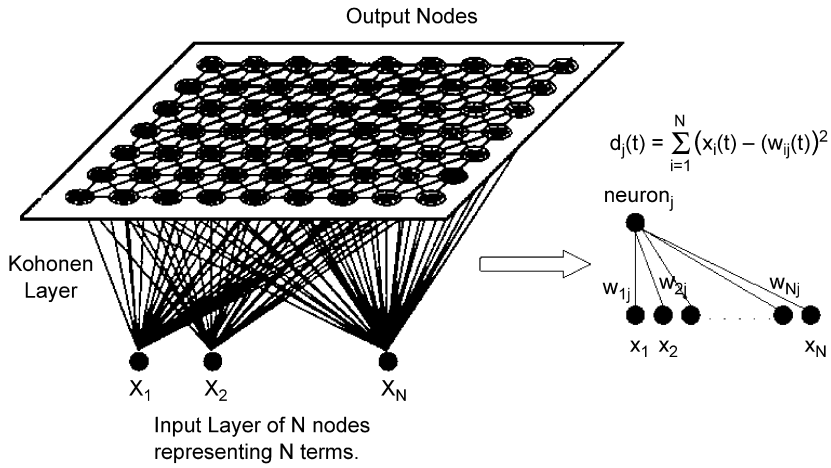


Fig. 1 – Input and output layers of a self-organizing map. The right half shows the weighted connections of a neuron to the N input nodes.

chemical and the biological samples at a particular station was selected to be a week before or after, to capture the effects of the chemicals on the biota. We hypothesised that, in absence of toxic spills generating fish kills (none reported) the water quality changes were relatively slow and up to one week difference between the two types of information (biological and chemical) would still provide representative correlation between chemistry and fish composition. Other parameters such as habitat or land use change very slowly. Since the physical habitat was sampled once per year for a sampling station, the habitat data were duplicated to be accommodated in the data set. The complete information regarding this data set has been included in Virani et al. (2005) and in the Ohio EPA (1999) report. In the report, various exploratory tools such as box-and-whisker plots, scatter plots and multivariate techniques, such as Principal Component Analysis, were used to visualize regional patterns in nutrient concentration and relationships with biological performance parameters.

The Ohio EPA has formulated a list of 12 fish metrics (see list of abbreviations) modified from the ones proposed by Karr (1981), and based on the type of sites: Headwaters, Wading and Boat sites. Each type has a set of metrics to calculate the fish IBI (Ohio EPA, 1987). Karr (1991) eloquently described the reasoning behind the selection of the metrics and selection of the type and fish genera were. The IBI was conceived to characterize the integrity of the streams which was defined at the beginning of this article. The IBI was not intended to be a quantitative measure of any specific pollution but of the overall impact of stressors and the "human disruption" and "degradation" were defined in generic terms. However, Karr's (1991) article lists cases where other authors found correlations of IBIs to some specific disturbances, such as land use, and found them worth to be pursued in further research. Karr himself documented the effect of residual chlorine in the streams after effluent disinfection as one of stressors significantly affecting IBI.

The original IBI defined by Karr et al. (1986) and Karr (1991) uses three groups of metrics. The first group of six metrics in the list of abbreviations evaluate the species *richness and composition*. The suckers, darters, and sunfish species feed on intervertebrates and are in the higher food web groups. Their numbers, normalized by the stream order, show the presence or absence of food (benthic or drifting organisms) which would reflect the degree of disturbance. The number of tolerant and intolerant species is expected to represent the degree of disruption with tolerant ones increasing and intolerant decreasing with the degree of pollution.

The next three metrics in the list evaluate the *trophic composition* of the fish community and is used for assessment of the energy base and the trophic dynamics of the biota. The proportion of omnivores increases as the insectivorous and top carnivorous fish decrease in the degraded systems.

The last three metrics in the list represent the *fish abundance and condition*. The number of fish is expected to decrease with the disturbance. The integrity can be also disrupted by invasive and hybrid species that are not indigenous to the area, e.g., the serious problem with exotic Asian carp proliferation in the Mississippi and Illinois River watersheds widely reported by the information media (Janega, 2006). The DELT (deformities, eroded fins, lesions, tumors) fish anomalies reflect the highest degree of disturbance typically related to severe disruption of integrity by high turbidity, temperature, and chronic effects of priority pollutants.

The IBIs are also related and compared to the reference streams or stream reaches which are the water bodies of the same character as the test site but least impacted by human stresses and disruption.

In the Ohio modification of the IBI the metric related to darters (DADSRNSCORE) combined a number (number of darter species at headwater and wading sites) and a percentage (percent round-bodied suckers at boat sites). The same is also true for SPWNSCORE, where the metric was composed of a number of simple lithophils species for headwaters and percentage of simple lithophils for wading and boat sites. Since it would seem illogical to combine a number and a percentage to represent the same metric, the metrics scores (ranked as 1, 3, and 5 from low to high) were used instead as the input to the SOM.

2.4. SOM+CCA for ecological modeling

Our data mining and knowledge discovery methodology first uses the SOM as an unsupervised clustering mechanism in order to group the sampling sites in the data set according to their fish metrics similarity. We decided to use the fish metrics as the input to the SOM (and not the raw fish assemblages) since the metrics are weighted averages of the raw data and hence are less susceptible to outliers. CCA is then carried out to extract information on the role of the environment variables in explaining the observed variability at different levels of resolution (sampling site, group of sampling sites, state overall). The steps of our methodology, summarized in Fig. 2, are detailed below. All data processing software was developed in MATLAB Version 6.5 (Matlab, 2007). We have also utilized a public domain Matlab SOM toolbox developed at Helsinki University of Technology (Vesanto et al., 2000), which provides functions for data preprocessing, SOM training, and results visualization.

The fish metrics were normalized using log transformation and linearly scaled in the range $[0, 1]$. The 12-dimensional normalized fish metric vectors (one per sampling site) were used as input data vectors to train an SOM. To find the optimal map size we considered a compromise between the topographic and the quantization errors (Kohonen, 2001). The quantization error is defined as the mean of the Euclidean distance of each data vector to its BMUs weight vector and measures map resolution. The topographic error (Kiviluoto, 1996) is calculated as the proportion of all data vectors for which first and second BMUs are not adjacent units in the grid of neurons. Since a very large map size is undesirable (given the size of the data set), we decided to use $9 \times 5 = 45$ neurons, which minimizes the topographic error (0.02) while also resulting in a very small quantization error (0.85). Based on the 1848 sites analyzed this also implies that each neuron contained on the average information for about 40 similar sites. The SOM rough training phase lasted for 20 epochs, followed by the fine-tuning phase for another 100 epochs. Sites representing similar conditions, as judged by their similar fish metrics information, are mapped to the same SOM neuron after the training (site patterning).

The well-known k-means algorithm (Duda et al., 2001) was then used to partition the SOM neurons into k groups (clusters) and assign a cluster label to each neuron. The algorithm initially places the k class centroids in randomly selected positions. Each point (neuron weight) gets associated with the closest centroid. Then the centroid moves to the mean position of the points it represents. Each time the class centroids, move all data points are reclassified, and the same procedure is repeated until convergence. The Davies–Bouldin index (Davies and Bouldin, 1979), which is proportional to the ratio of the sum of within-cluster scatter to the in-between-cluster separation, was used to find the best value for k. The index assigns low values to solutions with compact spherical clusters indicating good clustering performance. The k-means algorithm with $k = 3$ (optimal value) was performed on the SOM weights codebook for 100 iterations. The 2-level hierarchical clustering procedure, SOM for grouping sites with similar metrics followed by k-means grouping of the resulting SOM neuron weight, corresponds to the horizontal path of the workflow shown in Fig. 2, where horizontal solid lines are used to connect the processing blocks.

After site patterning, the environmental variables (related to water chemistry, physical habitat and land use) that are

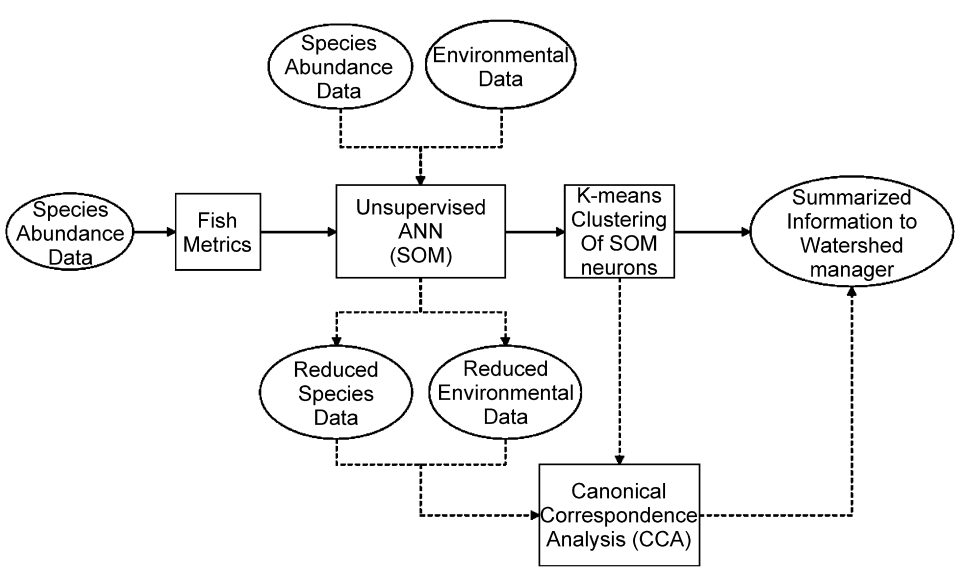


Fig. 2 – Data mining and knowledge discovery methodology combining SOM and k-means hierarchical clustering with canonical correspondence analysis for decision support in watershed management. The horizontal workflow path (hierarchical clustering) is performed first (square blocks connected by solid arrowed lines). It is followed by the vertical workflow path (dashed arrowed lines) in which species and environmental variables data are first patterned on the trained SOM before being analyzed using CCA.

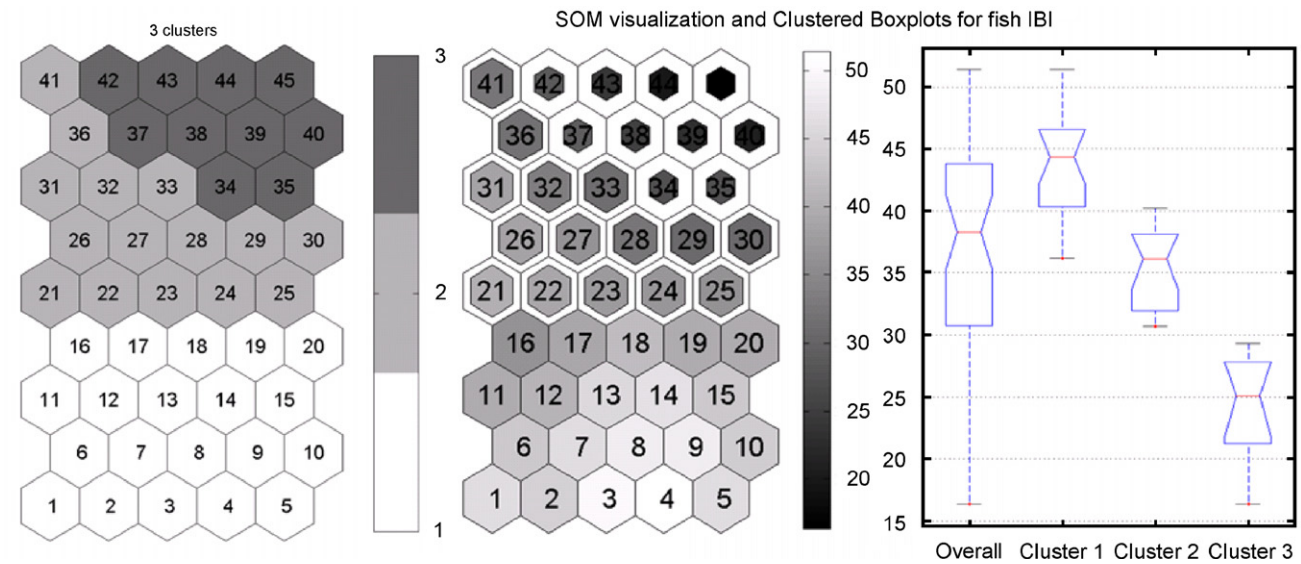


Fig. 3 – (Left) A visual representation of the trained SOM having $9 \times 5 = 45$ neurons (depicted as hexagonal shapes). Each neuron corresponds to a set of sampling sites with similar metrics characteristics. The SOM neurons are partitioned into three clusters shown with different levels of gray. (Middle) The distribution of fish IBI on the SOM neurons. Different hexagonal shapes are used to distinguish visually the neurons of every cluster. Neurons in cluster 3 exhibit low IBI values. (Right) Boxplots summarizing the distribution of IBI over the entire SOM and for each individual cluster. The IBI ranges of each cluster are clearly separated.

part of the data set *but were not used to train the SOM*, have been mapped on the SOM neurons. For each SOM neuron we calculated the average value (over all sampling sites clustered in that neuron) of each environmental variable. Furthermore, we performed the same mapping for fish assemblages, but after log normalizing them to account for their skewed distribution. Upon calculating their mean (over all sites falling in a neuron) fish assemblages are de-normalized and rounded to the nearest integer to obtain the reduced species data set (Fig. 2). At the end of this mapping stage, we have for each neuron the average fish abundances and the mean values of each environmental variable. The original Ohio EPA data set included measurements for 151 fish species collected across the 1848 sampling sites in Ohio. After the species abundance mapping on the trained SOM, 81 species were removed from the analysis due to zero counts in all the SOM neurons. Also, environmental variables CALCIUM and MAGNESIUM had high correlation ($\rho > 0.95$) with HARDNESS and were removed from the analysis. Therefore, the reduced data set consisted of 45 SOM neurons, 70 fish species and 31 environment variables.

3. Results and discussion

CCA was applied on the reduced data set to study the association of environmental variables to sites and species. Since we are interested in extracting and visualizing significant associations we used as inputs to the CCA the SOM neurons (representing groups of sampling sites with similar metric characteristics) and the reduced (averaged) species and variables. Since each neuron carries a cluster label, assigned to it by the second-level k-means clustering, the CCA results can be interpreted at two different levels of resolution, individual SOM neuron and cluster of neurons, as it will be discussed in Section 3.

3.1. *Visualization of the IBI distribution*

Fig. 3 (left panel) provides a visualization of the SOM with $5 \times 9 = 45$ neurons (represented by hexagonal shapes) used for patterning the sampling sites. The neurons are colored based on their cluster label assigned by the k-means algorithm. A detailed visualization of how the fish IBI is distributed over the SOM is provided in the middle panel. Boxplots (Tuckey, 1977) summarizing the IBI distributions (for the state overall and per-cluster) are provided in the right panel. Each cluster is associated with IBI values within a particular range and the three cluster ranges are distinct. This is made clear by noting the IBI cluster means and standard deviations: Cluster 1 (44.11 ± 4.03), Cluster 2 (35.11 ± 3.27) and Cluster 3 (24.45 ± 4.19). These clusters fit the IBI ranking from "poor" to "fair" or "modified" (Cluster 3) to "very good" or "exceptional" (Cluster 1), with Cluster 2 being an intermediate cluster (Ohio EPA, 1987). It is worth noting that these clusters approximately coincide with the Ohio biotic ranking of streams (Ohio EPA, 1987).

Fig. 4 shows the spatial distribution of the sampling sites patterned in each cluster. Regions with poor fish IBI (in Cluster 3) are concentrated in the western part of Ohio, particularly around Toledo and along the Wabash River at the western state border. Almost all small streams in the Toledo area have been channel modified to some degree (Yoder et al., 2000). The Wabash river watershed was designated as Ohio's

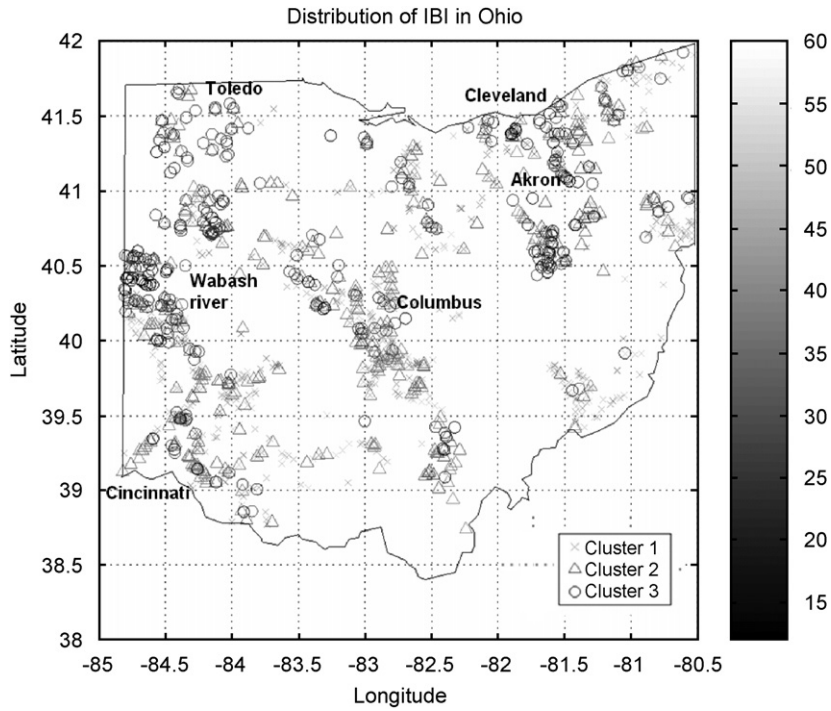


Fig. 4 – Distribution of the fish IBI in Ohio. The neuron's average fish IBI value was used for all sampling sites patterned in the same SOM neuron. The three clusters are distributed across the State. However, Cluster 3 sites (lower IBI) are mostly concentrated in the western part of the state and are dominant in the Toledo area.

worst watershed by the Ohio EPA in 1999. Lack of buffer zones, excessive nutrient and high bacteria levels were attributed as some of the reasons for the poor conditions (Ohio EPA, 1999). The basins around Lake Erie, especially around Cleveland in the Cuyahoga county are also degraded.

3.2. *Visualization of fish metrics distribution*

Each input data vector element has a weighted connection to each and every one of the SOM neurons (Fig. 1). The value of this weight models the influence of an input element (fish metric) to the sites represented by an SOM neuron. The distribution of each fish metric on the SOM (Fig. 5) can be visualized by the 2D map (called a component plane) of the corresponding weights. Visualization of the SOM component planes provides us information about the correlations between individual components, division of data in the input space and relative distributions of the components.

In Fig. 5 we see that almost all the metric scores have a low to high gradient distribution as we move from the top to the bottom half of the SOM. As stated before, the lower the rank of a metric the higher the degree of degradation in the biota in running streams. Darter species are known to be insectivorous, habitat specialists and sensitive to physical and chemical environmental disturbances (Ohio EPA, 1987). These factors make the darters reliable indicators of good water quality and habitat conditions (see how DADSRNSCORE is distributed on the SOM). The metric related to sunfish (SUNSCORE) is primarily a measure of the degradation of pool habitats. Suckers (SUMINSCORE) represent a major component of the Ohio fish community with their relatively long life spans. DELSCORE deals with the incidence of DELT (Deformities, Eroded fins, Lesions and Tumors) anomalies in fish communities, an indication of stress and environment degradation (Barbour et al., 1999; Karr and Chu, 1999). Fig. 5 indicates a highly exacerbated condition for the fish biota in the upper rows of SOM neurons, as also confirmed by the distribution of the fish IBI (Fig. 3).

3.3. *Exploring the relationships between the environment variables and fish metrics*

The CCA results are summarized in a plot of sample scores, species scores, and environmental variable arrows (Ter Braak, 1994). Sites and species scores are represented as points. Environmental variables are represented by lines with arrows pointing in the direction of maximal variation. Variables with lines close to each other and headed in the same (opposite) direction are highly positively (negatively) correlated. Two lines at a 90° angle indicate that the corresponding variables are uncorrelated.

CCA was applied to the reduced Ohio data set (45 SOM neurons, 70 fish species and 31 environment variables) and the results are summarized by Fig. 6. Variable correlations are easily revealed by this figure. For example, CONDUCTIVITY, which measures the amount of ions (electrically charged particles) present in the water, is highly correlated to HARDNESS in this data set. High sulfate values (SO_4) and extremely high conductivity indicate effects of mining disturbance in cluster 3. On the other hand, good riparian

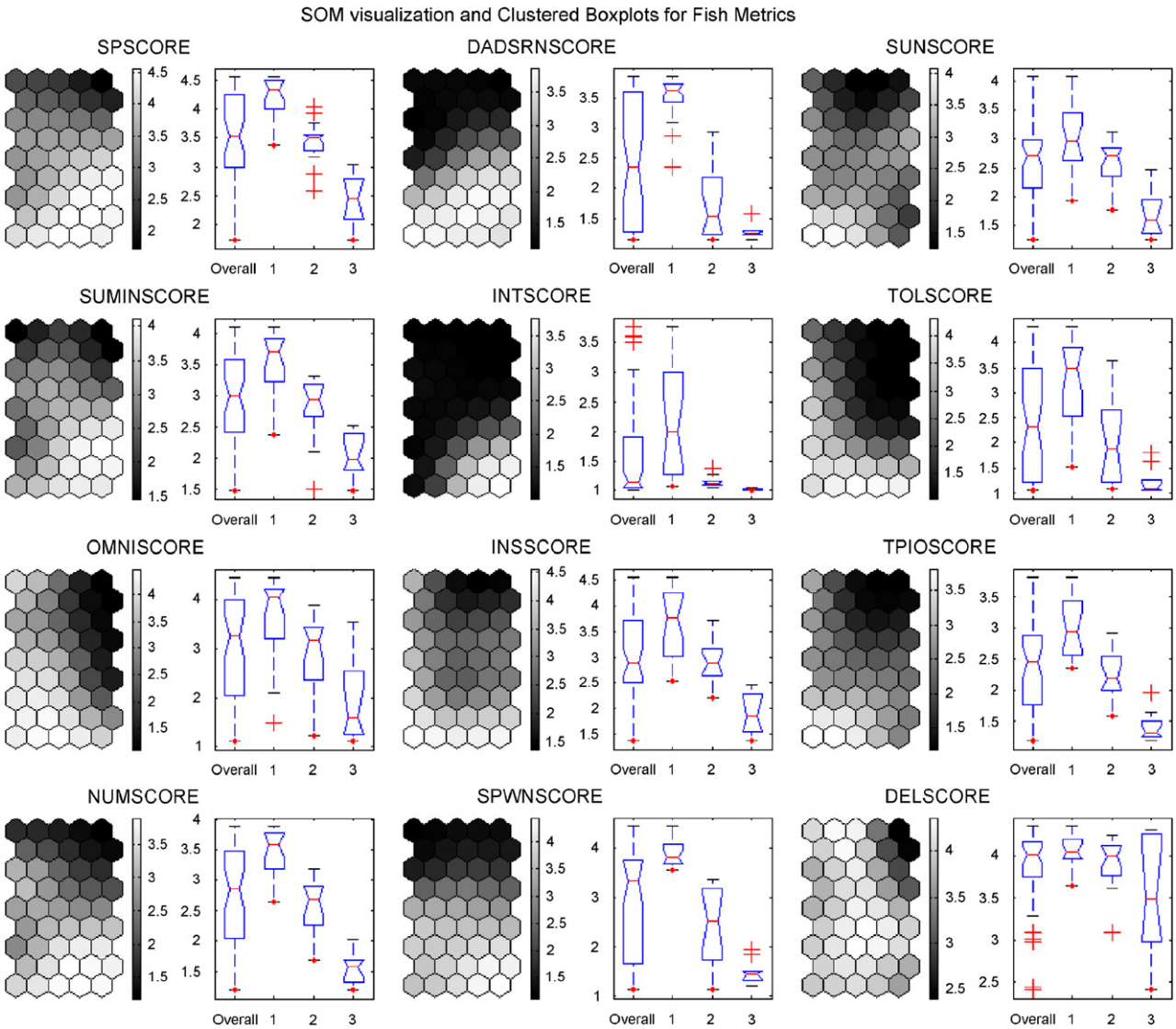


Fig. 5 – Fish metric score component planes visualized on the SOM and the corresponding metric distribution boxplots for the entire SOM and for each cluster separately. The ranges of metrics scores are shown in the corresponding side bar. For most metrics their values increase as we move from top (cluster 3, lower IBI) to bottom (cluster 1, higher IBI) SOM neurons.

zones (RIPARIAN) and presence of tree canopy (COVER) impede the flow of sediment and phosphorus (P) into the streams, with maximum phosphorus retention achieved by wooden riparian buffers (Ohio EPA, 1999). It appears that RIPARIAN and CHANEL as well as GRADIENT and RIFFLE variables are positively correlated, while GRADIENT and EMBEDDEDNESS are negatively correlated.

If the projection of a site's score (neuron score after SOM training) to an environmental variable's line lies on the same (different) side of the origin as the environmental variable's arrow, the value for that variable is larger (smaller) for that site relatively to its overall mean. The magnitude of a projection (measured as the distance from the origin where all variable lines merge) measures the magnitude of the deviation from the expected value. Upon projecting all neuron scores to a variable's line we calculated the median projection. Moreover, since the SOM neurons are partitioned by k-means into $k = 3$ clusters, we also derived the median projection per cluster. This median projection provides an estimate of the importance of an environmental variable in explaining the observed metrics variability within a particular cluster of neurons. After computing the three cluster median projections for each variable, we have assigned to that variable the label of the cluster exhibiting the largest median projection to its line. Through that process most variables got associated with either cluster 1 (higher IBI) or cluster 3 (lower IBI) as indicated by the oval size boxes enclosing variable arrows in Fig. 6. Using this method, each environmental variable's influence on the fish metrics has been assessed at three different levels of resolution (Virani et al., 2005): individual SOM neuron (using the neuron's score projection), cluster of SOM neurons (using the cluster median projection) and overall SOM map (using the overall median projection).

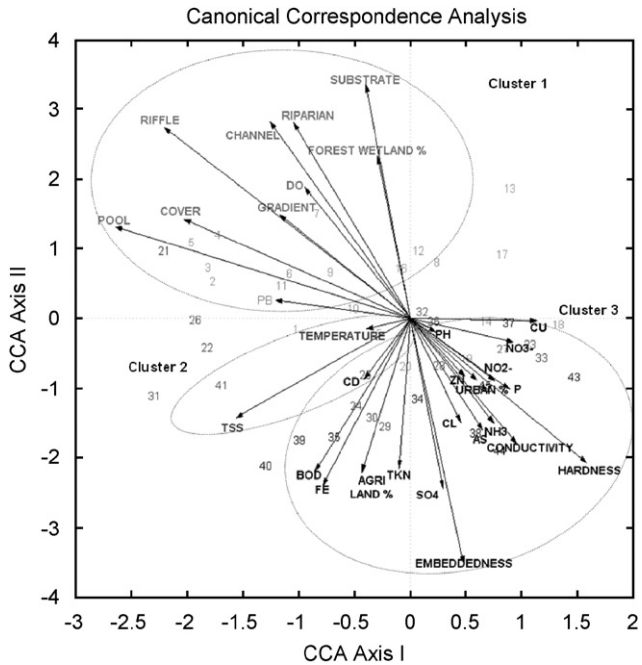


Fig. 6 – Canonical correspondence analysis bi-plot showing the environment variable arrows (scaled by a factor of 4) and the score "tick" of each SOM neuron. Neurons with numbers in the sets {1–20}, {21–33, 36, 41}, {34, 35, 37–40, 42–45} correspond to clusters 1, 2, 3, respectively (see also Fig. 3). Each environmental variable has been associated with the cluster whose neurons, when projected to its arrowed line, exhibit maximal median projection value. Most variables are associated either with cluster 1 (higher IBI) or cluster 3 (lower IBI) as indicated by the corresponding oval boxes enclosing variable arrows. The first two CCA axes account for 28% and 19% of the variance in the data set, respectively.

As shown in Fig. 6, the habitat gradients are in general diagonally across the nutrients and metals, which indicates a negative correlation between the two domains of variables. The analysis confirms the general knowledge that land use has impact on IBIs. The effect can be positive such as that by forests (see e.g., Steedman, 1988), or negative by agriculture (Wang et al., 1997; Yuan and Norton, 2004). Dissolved oxygen (DO) concentration effects are also widely known (see, e.g., Paterson et al., 2003). However, all of these analyses were based on either simple plotting or regression analysis of few independent variables. In the work by Yuan and Norton (2004), the large number of input parameters was preprocessed with principal component analysis to eliminate cross-correlation. The advantage and contribution of our method is that by clustering we can provide a comprehensive quantitative picture as to which parameters are most important under what situation. We also found that fish has a negative response to metal concentrations in water, which is evident from the presence of zinc (ZN), iron (FE), copper (CU), sulphate, etc. in Cluster 3 (dominated by pollutants) which is a common knowledge incorporated into water quality standards manuals (e.g., US EPA, 1986, 1994). The studies from which the standards are derived, related mortality or chronic responses of individual test fish species to a single metal or single pollutant (e.g., DO). Our analysis methodology provides a quantitative impact of the cluster dominating pollutants in a synergistic interaction with other pollutants.

It is also interesting to observe that, based on the length of the arrow in Cluster 3, embeddedness (EMBEDDED) has a strong negative influence on the fish IBI. Embeddedness is strongly negatively correlated to substrate quality and gradient habitat parameters which are dominant in Cluster 1. Yuan and Norton (2004) noted a dominant effect of substrate on IBI in their study. An obvious conclusion is that presence of excessive pollution by pollutants and bottom siltation drive the sites into Cluster 3 while good habitat (less embeddedness) and less pollution by pollutants are the characteristics of Cluster 1. Since clusters are identified based on IBI metrics we can then identify which fish genera are characteristic of the clusters.

The correlation matrix of all the environmental variables to the fish metrics and total indices of biological integrity; i.e. index of biotic integrity (IBI) for fish (Karr et al., 1986), Ohio's invertebrate community index (ICI) for macroinvertebrate community, and Ohio's qualitative habitat evaluation index (QHEI) (Ohio EPA, 1987) was calculated. A visualization of the matrix is provided in Fig. 7 where the lighter the matrix cell the higher the corresponding correlation coefficient value. All metric scores, except the one related to DELT anomalies (DELSCORE), are strongly correlated to the habitat parameters. Forests cover (PER_FORWET) has a positive effect on the indices; agricultural (PER_AG) and urban lands (PER_URBDEV) have a negative correlation to the indices, although the correlation is not high. It is interesting to observe that the three different indices (IBI, ICI and QHEI) exhibit similar correlation coefficients to the environmental variables, although they have not been used to train the SOM neural network.

Apart from providing a substantial knowledge base in terms of the interrelationships between the various variables, our findings are closely related with regression analyses by the Ohio EPA (1999) and others (Steedman, 1988; Rankin, 1995; Wang et al., 1997, 2001; Yuan and Norton, 2003) which provide a validation for this kind of methodology. Furthermore the proposed method identifies the relationships at the level of individual IBI metrics. Each metric provides information on the effects on the most representative fish species under different stresses, and since metrics are clustered based on similarities, the relationship and competition between fish species is revealed. Furthermore, by overlaying environmental and habitat variables on the SOM, the combined impact of these variables is revealed and visualized.

4. Conclusions

- A data mining methodology was introduced that can be used to identify clusters of sampling sites with similar metric characteristics (using SOMs and *k*-means) and then discover the environmental variables that best explain the metrics variability within a cluster (using CCA). This approach does not require calibrated, often subjective metrics scoring based upon professional judgment. New metrics can be added and those found to be statistically

irrelevant, or highly correlated, can be removed and the analysis be repeated. Furthermore the proposed method identifies the relationships at the level of individual IBI metrics. Each metric provides information on the effects on the most representative fish species under different stresses, and since metrics are clustered based on similarities, the relationship and competition between fish species categories is also revealed.

- The methodology has been applied for an extended Ohio EPA database. SOM analysis groups together sites across the state with similar metric characteristics, ultimately reflecting in the fish IBI. The *k*-means neurons clustering analysis that follows, allows us to divide the examination of the entire state of Ohio into three clusters of sites, which helps us to compare cluster statistics and define the key ecological characteristics. The mapping of the environmental variables on the SOM neurons then enabled the identification of parameters (landscape, physical/habitat and chemical) that exhibited similar distribution of impact.
- Results from this research help us identify locations with similar biological traits with an associated overview of the environment. The SOM-based methodology is beneficial as it allows the manager to pick up representative sites within specific basins for monitoring purposes and remediation. The potential benefits are far reaching, not the least of which include being able to direct limited financial resources more efficiently to monitoring and remediation activities. This methodology is currently extended to other states.
- Input–output predictive local models are now being developed, cluster-by-cluster and, possibly, metric-by-metric, using non-linear supervised training methods, such as multilayer perceptrons, etc. Further model development will ensure the characterization of relationships between instream biological community performance and stressors, tuned to account for regional imbalances and resulting in well-developed bioassesment programs, as envisioned in the Clean Water Act.
- The developed methodology is quite generic. If a large database is available, it can be applied to different spatial and temporal scales and provide intuitive visual feedback for exploratory data analyses, e.g., via GIS. It is also being now expanded to other biotic indicators and problems such as macroinvertebrate metric-based indices (e.g. Ohio's ICI). This is particularly important when analyzing patterns in rapidly changing landscapes due to development, or examining large-scale impacts of human activity or climate change on biological integrity of streams.

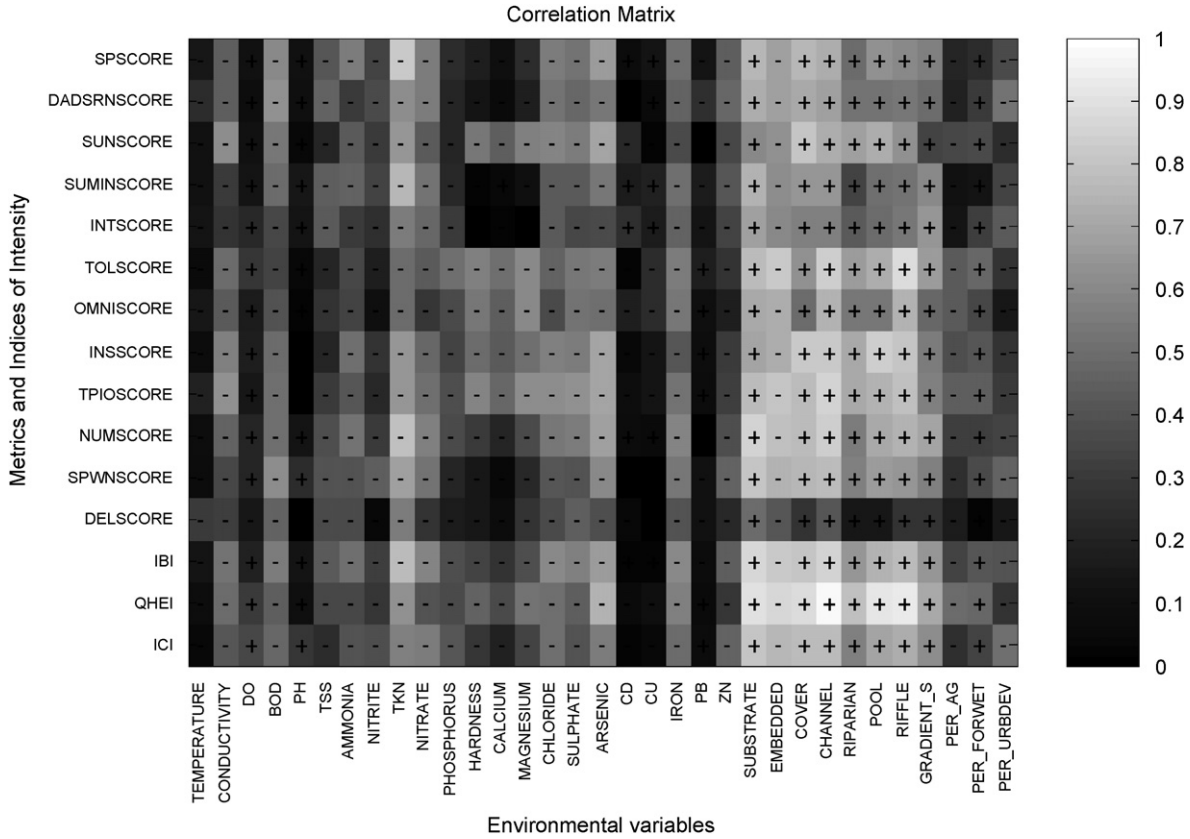


Fig. 7 – Visualization of the matrix of correlations of environmental variables to fish metrics and indices of integrity. The gray level of a matrix cell indicates the absolute value of the corresponding correlation coefficient; the correlation coefficient's sign is also noted inside the cell. Most fish metrics are highly correlated to the habitat variables (located in the right half of the matrix) which are prevailing in cluster 1 neurons (higher IBI).

Acknowledgements

This research has been sponsored by the US Environmental Protection Agency STAR Watershed Program by a Grant No. R83-0885-010 to Northeastern University. The authors greatly appreciate this support. The authors also greatly appreciate contribution of Mr. Ed Rankin of the Midwest Biodiversity Institute (Columbus, OH) who provided the data base. The findings and conclusions contained in this report are those of the authors and not of the funding agencies, nor the STAR program.

REFERENCES

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates, and Fish. US EPA, Washington, DC EPS-841-B-99/002.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Recognition Mach. Intell. 1 (2), 224–227.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification. Wiley, New York.

Gevrey, M., Rimet, F., Park, Y.S., Giraudel, J.L., Ector, L., Lek, S., 2004. Water quality assessment using diatom assemblages and advanced modelling techniques. Freshwater Biol. 49, 208–220.

Janega, J., 2006. Empty net is good mews in search for Asian carp. Chicago Tribune, June 15.

Karr, J.R., 1981. Assessment of biotic integrity using fish communities. Fisheries 6, 21–27.

Karr, J.R., 1991. Biological integrity: a long neglected aspect of water resources management. Ecol. Appl. 1 (1), 66–84.

Karr, J.R., Chu, E.W., 1999. Restoring Life in Running Waters. Island Press, Washington, DC.

Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., Schlosser, I.J., 1986. Assessing Biological Integrity in Running Waters: a Method and its Rationale, Special Publication 5. Illinois Natural History Survey, Champaigne, IL.

Kiviluoto, K., 1996. Topology preservation in self-organizing maps. Proc. IEEE Int. Conf. Neural Networks 1, 294–299.

Kohonen, T., 1990. The self-organizing map. Proc. IEEE 78, 1464–1480.

Kohonen, T., 2001. Self-Organizing Maps. Springer, Berlin.

Kohonen, T., Oja, E., Simula, O., Visa, A., Kangas, J., 1996. Engineering application of the self-organizing map. Proc. IEEE 84 (10), 1358–1384.

Matlab, 2007. Web page. The MathWorks, Inc. 〈http://www.mathworks.com〉.

Novotny, V., Bartosova, A., O'Reilly, N., Ehlinger, T., 2005. Unlocking the relationship of biotic integrity of impaired waters to anthropogenic stresses. Water Res. 39, 184–198.

Ohio EPA, 1987. Biological Criteria for the Protection of Aquatic Life: Volumes I–III. Ohio EPA, Columbus, OH.

Ohio EPA, 1999. Association between nutrients, habitat, and the aquatic biota in Ohio rivers and streams. Ohio EPA Technical Bulletin MAS/1999-1-1.

Park, Y.S., Chon, T.S., Kwak, I.S., Lek, S., 2004. Hierarchical community classification and assessment of aquatic ecosystems using artificial neural networks. Sci. Total Environ. 327, 105–122.

Paterson, M.J., Cada, G.F., Sale, M.J., Eddelman G.K., 2003. Regulatory approaches DOE addressing dissolved oxygen concerns at hydropower facilities. DOE/ID 11071, Oak Ridge National Laboratory, Oak Ridge, TN.

Rankin, E.T., 1995. Habitat indices in water resource quality assessment. In: Davis, W.D., Simon, T.P. (Eds.), Biological Assessment and Criteria. Lewis, Boca Raton, FL, pp. 181–208.

Steedman, R.J., 1988. Modification and assessment of an index of biotic integrity to quantify stream quality in southern Ontario. Can. J. Fish. Aquatic Sci. 45, 492–501.

Ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector method for multivariate direct gradient analysis. Ecology 67, 1167–1179.

Ter Braak, C.J.F., 1994. Canonical community ordination part I: basic theory and linear methods. Écoscience 1, 127–140.

Tuckey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.

Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., 2000. SOM Toolbox for Matlab 5. Technical Report A57, Helsinki University of Technology.

US Environmental Protection Agency, 1986. Quality criteria for water. 1986 EPA-440/5-86-001, Office of Water, Washington, DC.

US Environmental Protection Agency, 1994. Water Quality Standards Handbook, second ed., EPA-823-b-94-005A, Office of Water, Washington, DC.

Virani, H., Manolakos, E., Novotny, V., 2005. Self organizing feature maps combined with ecological ordination techniques for effective watershed management. Technical Report #4, Center for Urban Environmental Studies, Boston, MA.

Wang, L., Lyons, J., Kanehl, P., Gatti, R., 1997. Influence of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries 22 (6), 6–12.

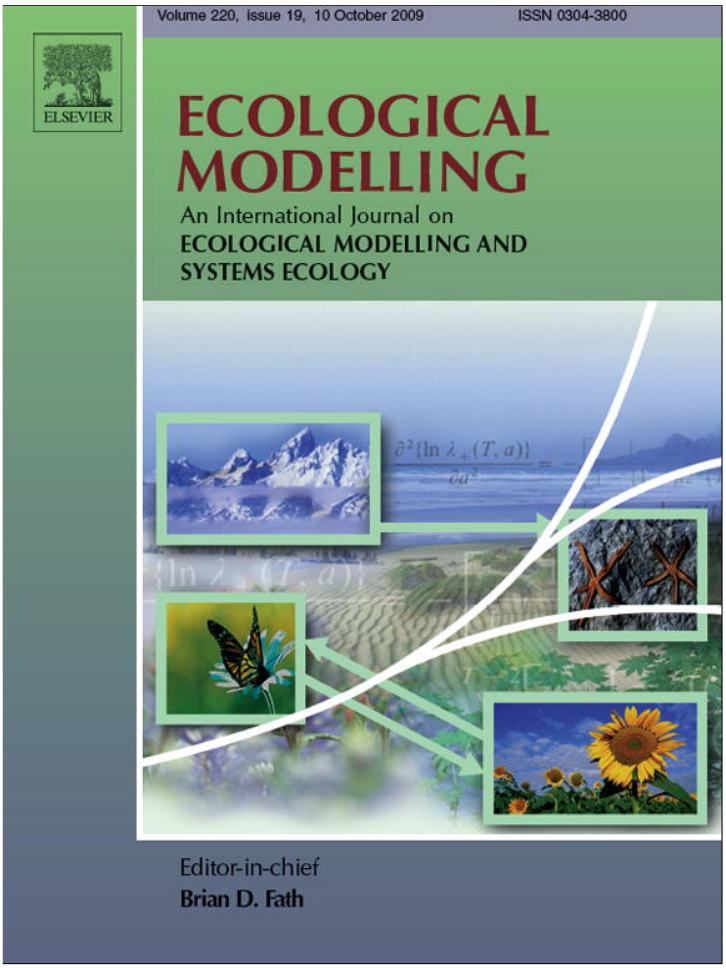
Wang, L., Lyons, J., Kanehl, P., Bannerman, R., 2001. Impact of urbanization stream habitat and fish across multiple spatial scales. Environ. Manage. 28 (2), 255–266.

Yoder, C.O., Miltner, R.J., White, D., 2000. Using biological criteria to assess and classify urban streams and develop improved landscape indicators. In: National Conference on Tools for Urban Water Resources Management and Protection. US EPA, Cincinnati, OH.

Yuan, L.I., Norton, S.B., 2003. Comparing responses of macroinvertebrate metrics to increasing stress. Environ. Monit. Assessment 22, 308–322.

Yuan, L.I., Norton, S.B., 2004. Assessing the relative severity of stressors at watershed scale. Environ. Monit. Assessment 98, 323–349.

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Ecological Modelling 220 (2009) 2393–2406

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Instream and offstream environmental conditions and stream biotic integrity Importance of scale and site similarities for learning and prediction

David Bedoya[a,*,1], Vladimir Novotny[a,2], Elias S. Manolakos[b,3]

[a] Civil & Environmental Engineering Department, Northeastern University, 400 Snell Engineering Center, 360 Huntington Avenue, Boston, MA 02115, USA
[b] Department of Informatics & Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis, Ilissia, Athens 15784, Greece

ARTICLE INFO

Article history:
Received 11 June 2008
Received in revised form 3 May 2009
Accepted 3 June 2009
Available online 17 July 2009

Keywords:
Index of Biotic Integrity (IBI)
k-Nearest neighbors
Self-organizing maps (SOM)
Environmental stressors
Multiple regression models

ABSTRACT

Two different methods to predict biotic integrity were tested and compared in the present paper. The first one tries to predict the fish indices of biotic integrity (IBI) at the state or regional scale based on the most similar observations to a specific target site of interest using the simple to implement *k-nearest neighbors* (or kNN) method. Two different distance functions were considered to find the *k*-nearest neighbors: the Euclidean and the Mahalanobis. The second method was applied on the same datasets and consisted of a step-wise multiple regression. The two modeling approaches yielded similar results but kNN proved to be more time-efficient and very fast computationally for the given dataset sizes, which allowed applying a leave-one-out cross validation.

In an attempt to reveal the importance of scale in the prediction of IBI, regression models were constructed at the state (or regional) scale and at the more refined cluster of sampling sites scale. Clusters of sites were extracted using Kohonen's self-organizing maps (SOM) followed by *k*-means clustering of the SOM neurons. Cluster-level regression models, constructed after site patterning, performed better in IBI prediction than global regression models constructed without any previous site patterning. The importance of identifying groups of sites with similar environmental characteristics affecting the IBI was revealed. The combined use of site patterning and regression modeling for IBI prediction also helped identifying important variables acting at the local scale which remain latent at the global scale.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The main objective of the U.S. Clean Water Act of 1972, is *to restore, and maintain the chemical, physical, and biological integrity of the Nation's waters.* Integrity is defined as the ability to support and maintain a balanced integrated, adaptive community of organisms comparable to that of a natural biota of the region (Karr and Dudley, 1981; Karr et al., 1986). This objective highlights the multidimensionality of the aquatic system; physical, chemical and biological. For many years, most of the emphasis was given to the chemical quality of receiving waters. Great progress in the control of point sources was achieved between 1980s and the end of the last century by using the *National Pollution Discharge Elimination System* (NPDES) permits. Even though the NPDES applies only to some diffuse sources, its application is far more challenging because these are driven by difficult-to-measure parameters such as land use, atmospheric conditions, and sub terrain sources and are affected by meteorological events. Urban or feedlot runoff events are just two examples (Novotny and Olem, 1994).

Even though the dynamism and effects of the highly intertwined and mutually dependant stressors on the biological community (and therefore, on stream health) have been widely studied and proved, the regulatory system has failed to keep pace with this reality and continues to neglect it in fresh water systems. Oversimplification, lack of multidisciplinary work, and patchiness of the U.S. regulatory system have been major impediments in developing a new framework that incorporates this whole spectrum of the problem (Karr, 1991). This becomes more evident when the sources of impairment originate from diffuse/non-point sources. It is very challenging to identify a steady state between land use changes, non-point source pollution, and biotic integrity. Land use changes can result in drastic changes in the water quality (i.e. fertilizers and pesticides transported by runoff), hydrologic regime (i.e. increase of peak flows in paved areas), or increase in siltation

* Corresponding author. Tel.: +1 617 2187147; fax: +1 617 3734419.
E-mail address: dbedoya@coe.neu.edu (D. Bedoya).
[1] Present address: Office 431, Civil & Environmental Engineering Department, Northeastern University, 400 Snell Engineering Center, 360 Huntington Avenue, Boston, MA 02115, USA.
[2] Present address: Office 429, Civil & Environmental Engineering Department, Northeastern University, 400 Snell Engineering Center, 360 Huntington Avenue, Boston, MA 02115, USA.
[3] Present address: Office B23, Department of Informatics & Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis, Ilissia, Athens 15784, Greece.

0304-3800/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolmodel.2009.06.017

Author's personal copy

due to denudation processes that subsequently impair stream habitat. It is highly unlikely that a sole stressor (e.g. a point source) would be responsible for the whole degradation of a water body. Therefore, many other sources of pollution remain unregulated.

Chemical water quality plays a key role in the overall integrity of a stream. However, this only represents one out of its five main integrity components in fresh water systems. Habitat structure and quality, flow regime, energy sources, and biotic interactions are the other four (Karr et al., 1986). System alterations which produce a deviation from the reference state in any of these constituents will have an impact in the biological community due to the propagation of stresses through the stressor hierarchy until the system's endpoint (the biological community) is reached (Karr et al., 1986; Novotny, 2003). The state of a system is usually assessed with environmental indicators, which can be categorized as stressors, exposure and response indicators (Yoder and Rankin, 1998; Yoder et al., 2000). Stressors may include point and non-point loading, land use changes, stream modifications, and other large scale influences that generally result from anthropogenic activities. Exposure indicators include chemical concentrations, whole effluent toxicity, tissue residue, sediment contamination, habitat degradation and other parameters that result in a risk for the biota. A biotic assessment endpoint or response indicator is a direct measure of ecological integrity of a water body because the biota is the point of convergence for all the stresses throughout the ecosystem (Yoder and Rankin, 1998; Yoder et al., 2000; Novotny et al., 2005).

Biological indices are nowadays regarded as valid overall stream health response indicators. However, in the past, most of the biological indices were based on the macroinvertebrate or microscopic community and designed to identify impacts from specific pollutants, especially from organic matter. Because stream degradation may originate from many different sources of pollution, these type of indices were hardly criticized as valid indicators of overall stream health (Doudoroff and Warren, 1957). Valid environmental and biodiversity indicators should be sensitive enough to track changes from reference conditions, applicable in large geographic areas, capable of providing a continuous assessment over a wide range of stress, and differentiate between natural cycles or trends and anthropogenic stress (Ott, 1978). These indices need to be calibrated within homogeneous environmental and geographic units (e.g. ecoregions) because reference species composition, distribution, and richness (i.e. biological integrity) differ between regions. Therefore, there is a need to define what truly ecological health is within each ecological context in order to have truly reliable indicators of streams's overall health (Karr, 1991).

In work by Doudoroff and Warren (1957) fish populations were already targeted as potential indicators of overall ecological conditions in fresh water bodies. Nowadays, one of the most widely used and accepted indices of stream's integrity in the United States is the fish Index of Biological Integrity (IBI) developed by Karr et al. (1986). Karr's IBI is currently accepted as a sensitive index to human impacts and applied successfully to aquatic communities (Richards et al., 1996; Roth et al., 1996; Dyer et al., 1998; Lammert and Allan, 1999; Dyer et al., 2000; Wang et al., 2001; Karr and Yoder, 2004; Yuan and Norton, 2004; Yoder et al., 2005; Manolakos et al., 2007). Many public agencies have adopted it as a framework for their own calibrated version at the state or regional scales (Ohio EPA, 1987; Bode, 1988; Roth et al., 1998; Niemela and Feist, 2000; Lyons et al., 2001; Lyons, 2006). One of the main advantages of biological approaches is that they have environmental memory, i.e. they can reflect impacts that occurred in the past or might reflect impacts which would go unnoticed in standard, periodic water

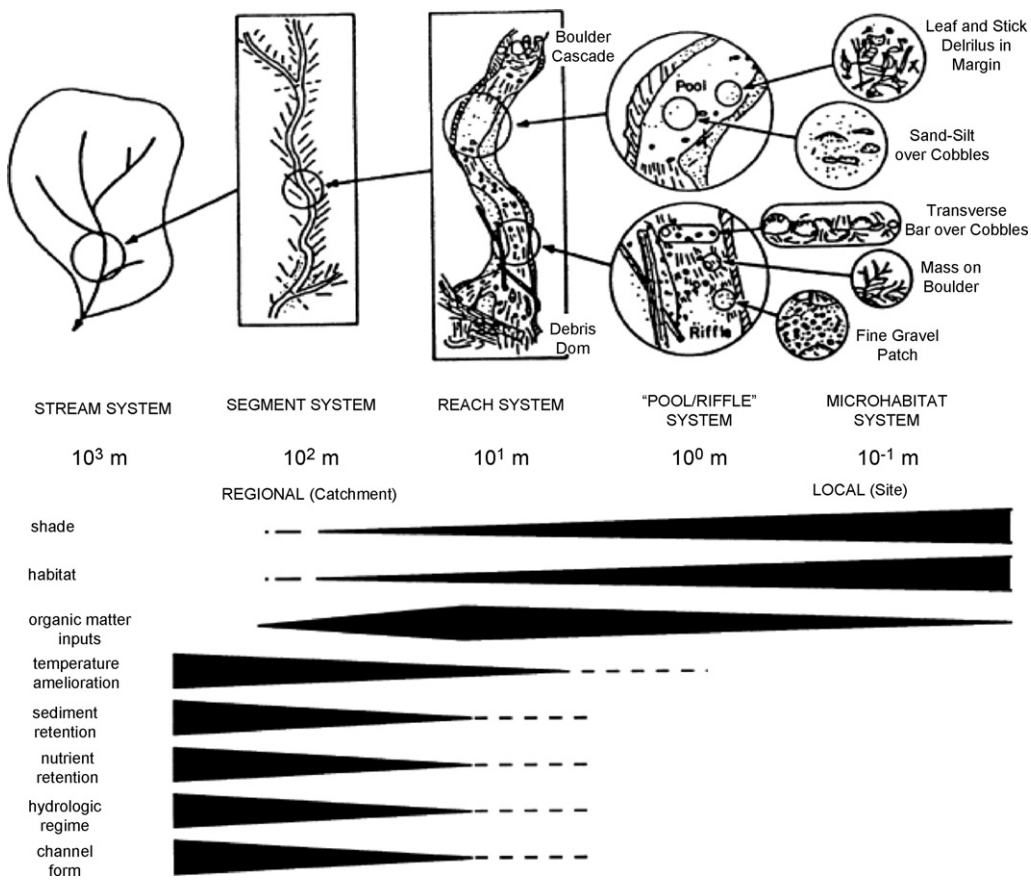


Fig. 1. Landscape influences structure and function across spatial scale. The thickness of the black stripes is proportional to the importance at each scale (adapted from Allan et al. (1997)).

Author's personal copy

column physical and chemical monitoring programs. Traditional chemical criteria do not document biological effects of pollution and do not include many other pollutants (e.g. pharmaceuticals). Moreover, a non exceedance for one chemical is not necessarily an indicator of healthy biota (Karr and Yoder, 2004). Even though the index developed by Karr et al. (1986) is based on fish, numerous IBI based on the macroinvertebrate community also exist and are used to evaluate the effect of specific pollutants (Hilsenhoff, 1987) or the overall health of a stream (Wright et al., 1988; Stribling et al., 1998; Barbour et al., 1999; Southerland et al., 2005).

Fresh water systems are organized in a hierarchical manner (Frissell et al., 1986). This hierarchy is spatially nested. Therefore, a system at one level sets the conditions of subsystems at lower levels. All habitats reside within the watershed environment. At each level, systems can develop and persist predominantly at a specified spatiotemporal scale and have a physical and biological functional role (Frissell et al., 1986). Environmental features in the watershed system are therefore linked to a level of hierarchical organization depending on their geographic scale (Fig. 1). Anthropogenic disturbances affecting high hierarchical levels will, invariably, propagate until the system endpoint is reached (i.e. the biological community). Thus, offtrseam and direct instream human disturbances (e.g. point sources) will ultimately propagate and modify the instream habitat conditions which will end up having an impact on aquatic fauna if enough exposure occurs (Yoder and Rankin, 1998; Yoder et al., 2000; Novotny et al., 2005). Anthropogenic stressors can affect the hierarchical structure of suitable habitat at different levels of this hierarchy and act as *environmental filters*. Only species able to able to *"pass"* all the natural and human-induced filters at the different levels of organization will be observed at a sampling site (Poff, 1997).

Therefore, the hierarchical organization of suitable habitat and biological community generates two main challenges in order to predict the biological integrity of a water body: (1) new efficient, versatile, multivariate mathematical methods need to be developed in order to deal effectively with the large number of inter-twined, cross-correlated, mutually dependant variables affecting the biological community in the hierarchical fresh water system. Mechanistic models or methodologies that pre-assume a certain relationship between the biological community and the explanatory variables cannot capture the high complexity of these systems (Poff, 1997; Borsuk et al., 2003) and; (2) the geographic scale at which the problem is approached will reveal different combinations of variables as most relevant to IBI prediction due to the biological-stressor hierarchical scale correspondence. The relevance of scale and/or geographic distribution of environmental observations in the prediction of stream's biological integrity has been revealed in previous research efforts. For example, Roth et al. (1996), Manolakos et al. (2007) and Dyer et al. (2000) based their research in multiple basins/watersheds and generally identified regional land use or land use-related parameters (e.g. substrate quality) as environmental gradients that constituted an axis of continuous change in their respective study regions. On the other hand, other studies narrowed the research area to one watershed or a portion of one and found more geographically localized environmental variables as the most relevant to IBI (Allan et al., 1997; Lammert and Allan, 1999; Dyer et al., 2000).

In order to address the first challenge, two different methodologies to predict biotic integrity were formulated and tested. For the analyses, three large state- or region-scale databases of indices of biotic integrity and their metrics, as well as accompanying land use, habitat, and chemical parameters were obtained. The first IBI prediction methodology consisted of using the k nearest neighbor (kNN) method with the entire databases. This simple non-parametric method was used first because it is usually considered as a benchmark for subsequent, more elaborate techniques to be compared against. The kNN technique is based on assessing proximity among observations by measuring their dissimilarity. The Euclidean and Mahalanobis distance functions were used for this purpose. A detailed description of the kNN methodology can be found in Jain et al. (1999) and Jain and Dubes (1988). It was used as a first step because it is a very fast, computationally efficient technique that easily allows good model validation by using a leave-one-out approach without drastically increasing the computation time. Since it was performed using the entire databases, it was expected to reveal the main environmental parameters with a significant impact on biotic integrity at larger scales. Once the kNN predictions were performed and an expected prediction performance baseline was established, another methodology was tested. It consisted of a step-wise multiple regression using the best fitting function (linear or non-linear) at each step. The results from both methods were then compared at the same geographic scale.

In order to prove the relevance of scale, the state-wide regression results were compared to the results of regressions obtained using clusters of similar observations. The clusters of sites were extracted using self-organizing maps (SOM) (Kohonen, 2001; Manolakos et al., 2007) followed by SOM-neuron grouping with the *k*-means method (Duda et al., 2000). The first goal of the research was to compare the performance of more traditional approaches (regressions) in identifying critical environmental variables to a simpler and time-efficient technique based on site similarities (kNN) at the state or regional scales. The second goal was to demonstrate the importance of geographic scale in the prediction of biological integrity and develop a methodology able to identify relevant variables at different scales. This was done by running the regression model first with the entire database (state or region level) and then on a cluster (of sampling sites) basis.

2. Data patterning techniques

2.1. Self-organizing feature maps

SOM are considered a type of unsupervised Artificial Neural Network (ANN). The SOM consist of a topologically ordered mapping of the multidimentional input space (in our case vectors of environmental variables) onto a two-dimensional space according to a meaningful order (Kohonen, 2001). SOM are composed of multiple units called cells or "neurons", which represent a homogeneous unit in the SOM environment. Neurons can be further grouped into clusters using similarity functions among the neuron centroids.

A SOM is usually composed of a two-dimensional lattice that represents the SOM cells. In an initialization process, each neuron in the SOM is associated with a random weight vector ($m_i = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{in}]$), which has the same dimension (n) as the input environmental vectors ($x_b = [x_{b1}, x_{bs}, \ldots; x_{bn}]$). Using a dissimilarity function (e.g. the Euclidean distance), each environmental vector (corresponding to a sampling site's observation vector) is associated with the most similar SOM neuron, called the *best matching unit* (BMU). Thus, an initial environmental vector SOM-layout is obtained. Subsequently, the initial neuron-allocated weights (m_i) are updated using a neighborhood function. This function minimizes the overall distance between the neuron itself and its neighbors. The new updated neuron weight is called the *generalized median* (ε). This process is iterated several times (epochs) until convergence or until a certain criterion is met (usually $m_i \cong \varepsilon_i$). After convergence, similar SOM neurons can be further grouped according to their similarity. Grouped SOM neurons constitute the clusters.

2.2. The k-nearest neighbor method

The kNN is a simple search algorithm in which one observation point (which is composed of multiple physical and chemical envi-

Author's personal copy

Table 1
Description of the environmental variables, scores and indices available for each state and their units.

OH (429 sites)	MN (125 sites)	MD Piedmont sites (246 sites)
Water Chemistry	*Water Chemistry*	*Water Chemistry*
Conductivity (Cond) (μmho/cm)	Conductivity (Cond) (μmho/cm)	Conductivity (Cond) (μmho/cm)
Dissolved oxygen (DO) (mg/L)	Dissolved oxygen (DO) (mg/L)	Dissolved oxygen (DO) (mg/L)
pH (standard units)	pH (standard units)	pH (standard units)
Total suspended solid (TSS) (mg/L)	Total suspended solid (TSS) (mg/L)	Nitrate as N (NO_3) (mg/L)
Total phosphorus (P) (mg/L)	Total phosphorus (P) (mg/L)	Temperature (Temp) (°C)
Ammonia as N (NH_4) (mg/L)	Ammonia as N (NH_4) (mg/L)	Sulfate (SO_4) (mg/L)
Nitrite as N (NO_2) (mg/L)	Total nitrogen (TN) (mg/L)	Alkalinity (ANC) (μEq/L)
Nitrogen Kjeldahl (TKN) (mg/L)	Temperature (Temp) (°C)	Dissolved organic carbon (DOC) (mg/L)
Nitrate as N (NO_3) (mg/L)	Turbidiy (Turb) (NTU)	*Habitat and morphology*
Hardness as $CaCO_3$ (Hard) (mg/L)	*Habitat and morphology*	Remoteness score (Remote) (0–20)
Biological oxygen demand (BOD) (mg/L)	Substrate, channel, and cover scores, habitat index (QHEI) (0–100)	Instream habitat (Instrhab) (0–20)
Total calcium (Ca) (mg/L)	Buffer width (MBufWid) (m)	Epifaunal substrate (EpiSub) (0–20)
Total magnesium (Mg) (mg/L)	Mean bank erosion (MBankEros) (m)	Velocity–depth variability (Vel–dpth) (0–20)
Chloride (Cl) (mg/L)	%Undercut (PctUndercut)	Pool quality (Pool) (0–20)
Sulfate (SO_4) (mg/L)	%Woody (PctWoody)	Riffle quality (Riffle) (0–20)
Total arsenic (As) (μg/L)	%Over vegetation (PctOverVeg)	Channel alteration (Chan) (0–20)
Total cadmium (Cd) (μg/L)	%Emerging macrophyytes (Pct Emermac)	Bank stability (BankStab) (0–20)
Total copper (Cu) (μg/L)	%Submerged macrophytes (PctSubMac)	%Embeddedness (PctEmbed)
Total iron (Fe) (μg/L)	%Other cover (PctOtherCov)	%Channel with flow (Ch_flow)
Total lead (Pb) (μg/L)	%Vegetal cover (PctCov)	%Shading in channel (Shading)
Total zinc (Zn) (μg/L)	%Pool (PctPool)	Buffer width (MBufWid) (m)
Habitat and morphology	%Run (PctRun)	Aesthetic quality (Aesthet) (0–20)
Substrate score (Subs) (0–20)	%Riffle (PctRiffle)	Habitat index (PHI) (0–100)
Embeddedness score (Embed) (0–4)	%Pool + run (PctPoolRun)	Thalweg depth (MThalDep) (cm)
Riparian score (Rip) (0–10)	Mean width (MWidth) (m)	Mean width (MWidth) (m)
Instream cover score (Cov) (0–20)	Thalweg depth (MThalDep) (cm)	Maximum depth (MaxDepth) (cm)
Riffle score (Riffle) (0–8)	Mean depth (MDepth) (cm)	Slope (Sl) (%)
Pool score (Pool) (0–12)	Width–depth ratio (WDRatio)	Average flow velocity (m/s)
Channel score (Chan) (0–20)	Sinuosity ratio (Sin)	Woody debris count
Gradient score (Grad) (0–10)	Slope (Sl) (m/km)	Root count
Habitat index (QHEI) (0–100)	%Boulder (PctBould)	*Land use (in drainage area)*
Land use (beyond 100 m buffer area)	%Rock (PctRock)	%Urban land uses (Urban)
%Agriculture (Agri) (25% increments)	%Fines (PctFine)	%Agriculture + barren (Agribarr)
%Forest–wetland (Forwet) (25% inc.)	%Embeddedness (PctEmbed)	%Forest + wetland + water (Forwetwat)
%Urban (Urban) (25% increments)	Mean fines' depth (MFineDep) (cm)	*Biological indices*
Biological indices	*Land use (in riparian area)*	Fish IBI (1–5)
Fish IBI (12–60)	Land use (0–5), riparian (0–15) scores	Benthic IBI (1–5)
ICI (0–60)	%Disturbed LU in 100 m buffer (PctDistLU)	Hilsenhoff index (0–10)
	%Undisturbed LU in 100 m buffer (PctUnDistLU)	
	%Dist. LU in 30-m buffer (PctDistLU30)	
	%Undisturbed LU in 30-m buffer (PctUnDistLU30)	
	Biological indices	
	Fish IBI (0–100)	

ronmental variables measured at a specific site) is compared to a set of observations with the exact same attributes. The objective is to find a specified number of most similar observations (k) to the one being tested. In order to measure the degree of dissimilarity, there exist numerous distance metrics. Some common metrics are the Minkowski distance, the Euclidean distance (which is a particular case of the Minkowski distance), the cosine distance, or the Mahalanobis distance. The latter is particularly interesting because it applies a whitening transformation to the data that avoids or reduces linear correlation distortion among features. Detailed information on these functions can be found in Jain and Dubes (1988) and Jain et al. (1999). The Euclidean (Eq. (1)) and Mahalanobis (Eq. (2)) distances were used in the research with a customized application in *MATLAB*®.

$$ED(X_i, X_j) = \left(\sum_{k=1}^{n} (X_{i,k} - X_{j,k})^2\right)^{1/2} \quad (1)$$

$$MhD(X_i, X_j) = (X_i - X_j)^T \times \sum^{-1} \times (X_i - X_j)) \quad (2)$$

In the above equations, n is the dimension of the data vectors (number of environmental variables) in the database, X_i and X_j are the pair of vectors being compared. Matrix Σ in Eq. (2) is the covariance matrix of the observed data vectors using the selected features.

3. Data and study area

Biological, chemical, and physical data were obtained from the public agencies responsible for their respective biological monitoring programs in Minnesota (Minnesota Pollution Control Agency (MPCA)), Ohio (Ohio EPA), and Maryland (Maryland Biological Stream Survey (MBSS)). In Minnesota, a total of 125 sites sampled between 2003 and 2006 were available. These were distributed in two main regions: the Southern (Des Monies and Minnesota River basins) and the Northern regions (Upper Mississippi and St. Croix River basins). In Ohio, a total of 429 observations collected between 1996 and 2000 were used. These were distributed in five basins: the Western Lake Erie, the Muskingum River, the Wabash River, the Scioto River, and the Little Miami/Middle Ohio River basins. In Maryland, a total of 246 observations collected between years 1995 and 1997 in the Piedmont region were used. The analysis in Maryland focused only in one stratum (i.e. piedmont) because the reference criteria and the metrics used to calculate the Physical Habitat Index (PHI) differ significantly from the coastal and highland strata (Paul et al., 2002). For this reason, we felt a state-wide analysis was not possible in Maryland. The type and format of data available, especially for the physical variables, were quite different from state to state. A summary of the environmental variables recorded for each observation is provided in Table 1. The number of sites in each state is the total number of observations with no miss-

Author's personal copy

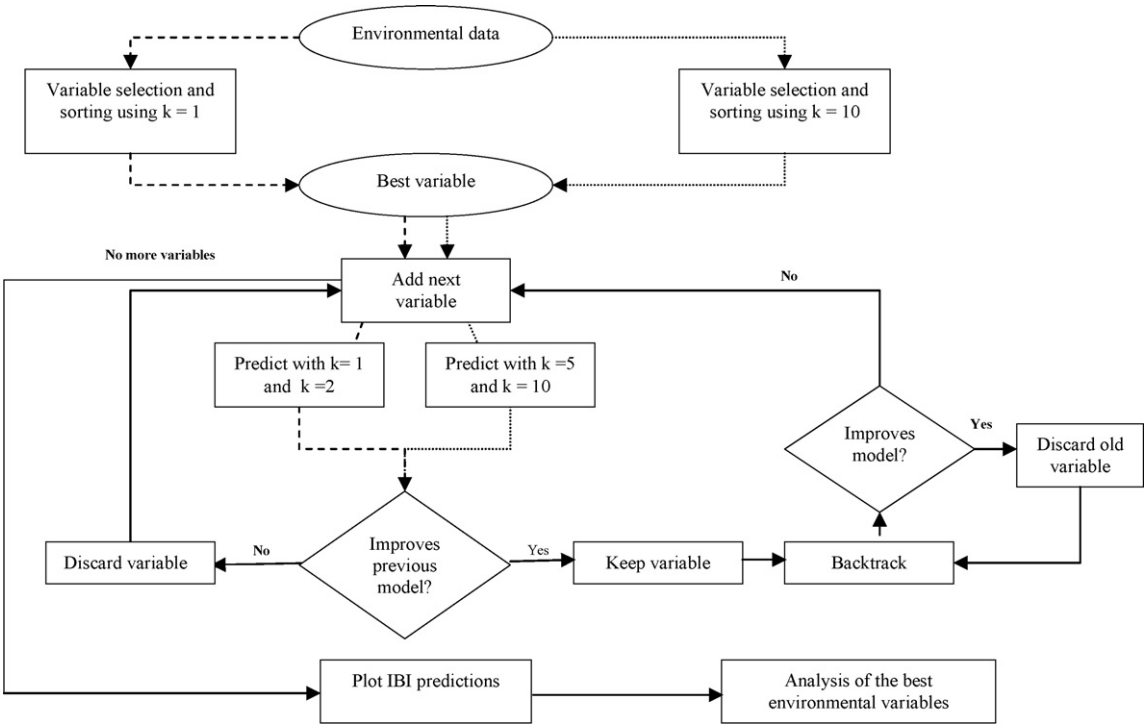


Fig. 2. Flow-chart of the step-wise kNN prediction method. Dashed arrow lines represent the steps followed when the environmental variables are sorted with $k = 1$. Dotted arrow lines represent the steps followed when the variable sorting is performed with $k = 10$. Solid arrow lines depict common steps for both cases.

ing data in any of the fields and it is the number of observations used in our analysis.

The environmental databases used in our research are compiled in periodic sampling events and used by public agencies to assess the state of the streams. These environmental observations reflect the environmental conditions at the sampling time only partially. In these periodic grab samples, healthy biota cannot be assumed, even when chemical criteria are not exceeded (Karr and Yoder, 2004). Other valuable ecological information concerning past polluting events and legacy pollution (e.g. from abandoned mines), non-sampled chemicals, or stressors not present at the sampling time may be missing. This potentially incomplete physical, chemical, and temporal characterization of the ecological system could lead to misleading conclusions because biological information incorporates such information since it is considered a continuous measure. A potential solution to such issue could be the use of legacy pollution from sediment data because they are better long-term indicators than the reported water column data. However, since similar research we found to compare against was based on the same or very similar databases to the ones available (MBSS, 1999; Ohio EPA, 1999a; Dyer et al., 2000; Volstad et al., 2003; Niemela et al., 2004; Manolakos et al., 2007) we decided not to include new information. Moreover, the scope of our research was to compare the accuracy of two predicting methodologies and check the effects of scale on the prediction of IBI; not so much getting into a discussion of how accurately stream's health is captured with the current sampled data in the respective state's monitoring programs.

4. Methodology

4.1. Simple IBI prediction using kNN

Due to the small computation time required, a leave-one-out cross validation procedure was used. Thus, each individual observation was taken out of the database and compared against the rest of the remaining observations one at a time. Once the first observation was compared to the rest of the database, it was reintroduced into the database and the next observation was taken out to repeat the process until all the observations were tested. With this method, there was no need to separate a validation set because each point was validated against the remaining sites in the database. Two different similarity functions were used; the Euclidean and the Mahalanobis distances. Prior to the analysis with the Euclidean distance, the data were log transformed and scaled in the range [0 1]. The steps followed are described below (also see Fig. 2).

(1) *Best metric selection* (*using $k = 1$ and $k = 10$ closest neighbors*): It evaluated prediction capability of each environmental variable alone by comparing the IBI value of the site being tested (one-out) with the average IBI of the identified closest site/s (1 and 10). The variables were then sorted in decreasing order (ranked) for both cases separately (for $k = 1$ and $k = 10$). The r^2 of the linear regression curve between the calculated (average over neighbors) and observed IBI scores determined the variable ranking. One ($k = 1$) closest neighbor was used because by using the closest observation, the extreme values would be predicted more accurately, since few observations in the very low and upper IBI ranges existed. With $k = 10$, observations in the mid IBI range (with larger number of observations) would be predicted more reliably, but not the extremes. Therefore, two lists of sorted variables were obtained (one with $k = 1$ and one with $k = 10$).

(2) *Step-wise predictions using variables from the sorted list for $k = 1$*: Following the variable sorting obtained in step 1 (with $k = 1$) a new variable at a time was included. The similarity function was calculated with the selected variables. The IBI prediction was performed by finding the IBI value (with $k = 1$) or the mean IBI value (with $k = 2$) at each step. If the new IBI prediction with the new added variable (with either $k = 1$ or $k = 2$) improved the previous one, the new variable was kept, otherwise it was not. When a new variable was added, backtracking was performed. Therefore, previously included variables were excluded, one at a time, to see how the predictions were affected. If the exclusion

Author's personal copy

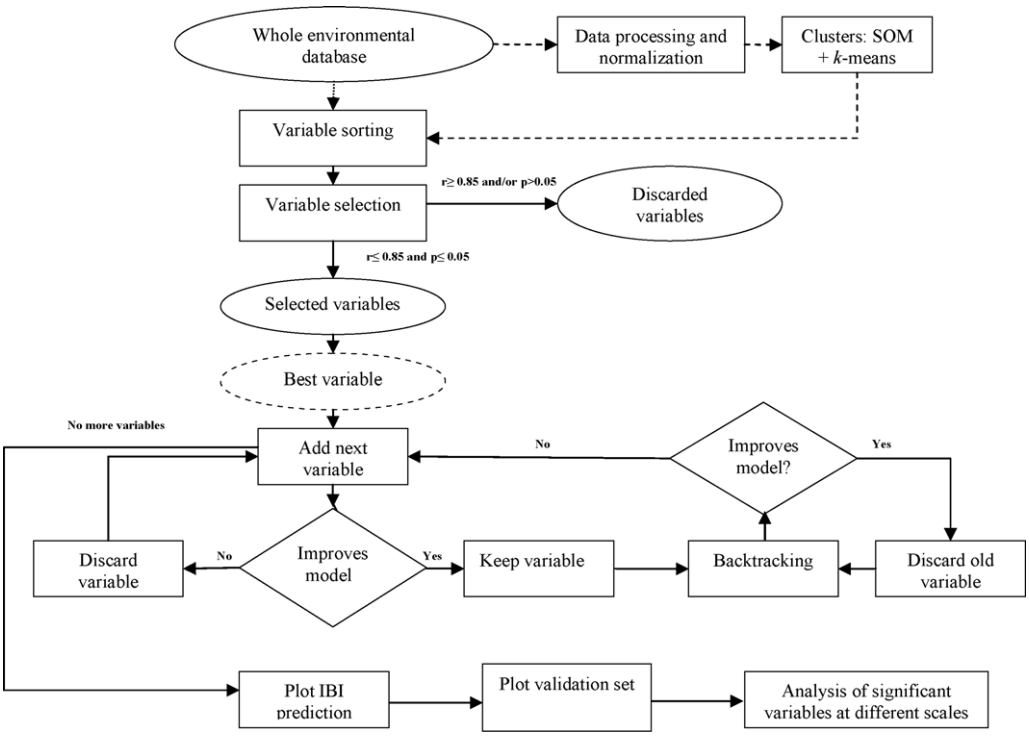


Fig. 3. Flow-chart of the step-wise multiple regression method. Dashed lines indicate steps for the cluster-based model only. Dotted lines indicate steps for the whole database model only. Solid lines are common steps for both methods.

of an old variable improved the prediction, then this variable was eliminated from the model. The reason for backtracking was to minimize the effect of the order in which the variables were introduced in the model, as suggested by Jain and Dubes (1988).

(3) *Step-wise predictions using the variables from the sorted list for k = 10*: It was implemented as step 2 above, except that in this case the average IBI value from the 5 or 10 closest neighbors ($k = 5$ or $k = 10$) was used for the prediction.

4.2. IBI prediction using regression and SOM + regression

Prediction of IBI using multiple regression was performed at the state (or region in Maryland) and at the cluster (of sites) scales. The regression equations were obtained following a step-wise methodology and using 75% of all the available observations in each database for model development. The remaining 25% was kept for model validation. The observations split into subsets was random. A diagram summarizing the different steps is presented in Fig. 3 and described as follows:

(1) *Database clustering using the SOM* (*only in cluster-based predictions*): Each of the databases was clustered using all the available chemical and physical environmental variables shown in Table 1. Biological information (i.e. IBI) was not used in the clustering process. In Ohio, land use data was not used for clustering purposes because this variable was measured in a very crude scale (25% increments). Land use data in Ohio was kept out of the clustering algorithm as a cautionary measure because it could negatively alter the SOM procedure. The environmental data were converted to their natural logarithms and ranged [0–1] before training the SOM. Logging the data is common in ecological modeling to smooth the data and reduce the effects of outliers. After logging the linear scaling from 0 to 1 allows each sensor to have the same dynamic range and same prior "importance" before conducting SOM-based clustering and again controls the influence of outliers. The number of SOM neurons was determined based on the topographic and quantization errors. The quantization error is the average distance (Euclidean) between each data vector and its BMU, while the topographic error is the proportion of data vectors for which the first and second closest SOM cells are not adjacent in the grid of neurons (Kiviluoto, 1996). A compromise between the two types of error had to be made because the quantization error usually tends to decrease as the number of SOM neurons increase, and a very large map size was undesirable given the available data. Hence, the maximum number of SOM neurons was limited to 100. In Ohio, a SOM with 60 (6×10) neurons was used. For Minnesota and Maryland, SOM with 63 (7×9)) and 54 (6×9) neurons were used, respectively.

The next step consisted of finding the optimum number of neuron clusters. The *k*-means algorithm was used for this purpose (Manolakos et al., 2007). The optimal number of clusters found using the Davies–Bouldin index (Davies and Bouldin, 1979) was 3 in Ohio and Minnesota, and 5 in Maryland.

(2) *Selection of a validation set*: 25% of randomly selected observations in each cluster were kept aside for validation. The remaining 75% was used to develop the regression models. The validation sets used for the cluster-based and the state-based regressions were the same in all cases.

(3) *Best metric selection* (*at state and cluster level*): In the regression development datasets, each one of the environmental variables was regressed linearly against the fish IBI score. The environmental variables were then sorted in decreasing order based on the coefficient of multiple determination (r^2). An *F*-test was performed in each case to check the statistical significance of the regressions. Only variables that showed statistic significance ($p \leq 0.05$) were included in the list.

(4) *Linear correlation checking*: The correlation coefficient (r) was calculated for each pair of significant variables selected and sorted in the previous step. In cases in which the variable-

Author's personal copy

Table 2
Summary of the best IBI predictions using the kNN methodology. The different functions (Mh = Mahalanobis; Eu = Euclidean) and selected number of closest neighbors (k) are specified. Final selected variables in each case are also listed.

Location	Similarity function	k	RMSE[a]	r^2	Variables used
OH	Mh	1	7.28	0.51	Subs, Riffle, Cov, Embed, Forwet, Urban
		5	7.76	0.47	Riffle, Pool, Embed, Cd, Cu
	Eu	1	7.77	0.47	Pool, Riffle, Cov, Subs, Rip, Cu, Urban
		5	6.76	0.53	Riffle, Pool, Subs, Rip, Embed, SO_4, TKN, pH, Urban
MN	Mh	1	21.74	0.41	TN, PctWoody
		5	20.52	0.48	LU, Rip, Cond, MFineDep, PctCov
	Eu	1	19.43	0.53	LU, TN, Channel, PctBoulder, PctRun, PctRiffle
		5	19.42	0.54	TN, Cond, TSS, PctUnderCut, MDepth
MD	Mh	2	0.66	0.54	Agribarr, Urban, PHI, ANC, Aesthet, SO_4, DOC
		10	0.63	0.59	Agribarr, Vel-dpth
	Eu	2	0.67	0.54	Agribarr, ANC, Urban, PHI, Sl, DO
		10	0.63	0.59	Agribarr, Urban, ANC, PHI, Vel-dpth, DOC

[a] RMSE for different IBI predictions. IBI scales: 12–60 in Ohio; 0–100 in Minnesota; 1–5 in Maryland.

variable $|r| \geq 0.85$, the least discriminant variable (with smaller r^2) was removed because it was considered not to bring any new relevant information to the system.

(5) *Step-wise regression and backtracking*: This was done by starting the regressions with the best variable from step 3 and adding the next best one at each step. If the new added variable increased the previous r^2 it was kept, otherwise it was discarded and the next variable was tested. When a variable was tested, linear and non-linear regression combinations were evaluated. The function that yielded the highest r^2 was used. Quadratic, logarithmic, exponential, inverse, S-curve, and power functions were the non-linear functions tested. Backtracking was also performed in this case. Steps 2 through 5 were performed using the statistical software *SPSS Version 15*® for Windows.

(6) *Model validation*: The equations obtained in step 5 were tested with the validation sets and the IBI predictions plotted.

4.3. Metal chronic and acute toxicity effects

After the IBI predictions were obtained, a further fine-tuning was attempted by adding a penalty to sites with high metal concentrations (only available in Ohio). The penalty was calculated using an exponential curve (Eqs. (3) and (4)) between the criterion continuous concentration (CCC) and the criterion maximum concentration (CMC). This curve was used because the penalty would be small when the CCC was slightly surpassed but increased steeply as the CMC threshold was approached. The CCC was selected as the lower limit because it represents the threshold below which no negative effects on biota should be observed (Stephan et al., 1985). Since none of the sites in our database exceeded the reported CMC, the final acute value (FAV) was not included in the penalty equation. A description of the CCC, CMC, and FAV can be found in Stephan et al. (1985) and were calculated for each metal using the criteria for the protection of aquatic life by Ohio EPA (Stephan et al., 1985; Ohio EPA, 2008).

$$P = \sum_{i=1}^{n} (e^{\alpha_i \times (\mathrm{CONC}_i - \mathrm{CCC}_i)} - 1) \tag{3}$$

$$\alpha_i = \frac{\mathrm{Ln}(P_{\mathrm{CMC}_i} + 1)}{\mathrm{CMC}_i - \mathrm{CCC}_i} \tag{4}$$

In Eqs. (3) and (4), P is the final, cumulative penalty, n is the number of available metal concentration measurements, CONC is the observed concentration for that metal, P_{CMC} is the maximum penalty when the CMC concentration is reached, α is a coefficient calculated given the boundary conditions of the equation ($P_{\mathrm{CONC} \leq \mathrm{CCC}} = 0$ and $P_{\mathrm{CONC} = \mathrm{CMC}} = P_{\mathrm{CMC}}$). The penalty when CMC is reached (i.e. P_{CMC}) was set at will in each case. The P_{CMC} value that maximized the final IBI prediction was selected.

5. Results

5.1. Predictions using the kNN approach

5.1.1. One or two closest neighbors ($k = 1$ or $k = 2$)

5.1.1.1. Minnesota. In this case, the Euclidean distance outperformed the Mahalanobis ($r^2 = 0.53$ and 0.42, respectively with $k = 1$). In both cases, total nitrogen (TN) and percent disturbed land use in the riparian buffer (LU) were among the most significant variables, but both were not necessarily included in the final prediction model. With the Mahalanobis function, land use quality was discarded in the backtracking process and required only two variables to yield its best possible prediction (see Table 2). The variables used in the best model (Euclidean distance) were related to nutrient loads, land use patterns, stream variability, substrate quality, and channel morphology, which strongly agreed with the variables obtained in the regression model for the whole state (Table 3).

5.1.1.2. Maryland. Both proximity functions performed very similarly, achieving equal final results ($r^2 = 0.54$ with $k = 2$ in both cases) and identified very similar significant environmental variables (see Table 2). Land use in the drainage area and alkalinity were relevant parameters in both cases (like in the whole dataset regression model), and so was the PHI (unlike the regression model). The rest of the selected variables were different. The Mahalanobis function identified aesthetic quality (*Aesthet*) as an important parameter, similarly to the regression model. Even though this is a qualitative parameter, it seems to have high predicting capabilities in Piedmont regions probably because it is a good indicator of urban degradation.

5.1.1.3. Ohio. The Mahalanobis function outperformed the Euclidean ($r^2 = 0.51$ and 0.47, respectively with $k = 1$ in both cases). Habitat parameters (Substrate, Riffle, and Cover) were able to explain a very large portion of the total biologic variability in both cases. Land use in the riparian corridor was also relevant in both cases (i.e. *Urban* and *Rip* in Table 2). The significant habitat variables found with the distance functions in Ohio agree again with those found relevant with the regression approach. No chemical parameters were selected in this case, with the exception of copper using the Euclidean distance function which introduced little overall information beyond the one already included by habitat parameters, which accounted for 86% of the group's variability.

5.1.2. Five or 10 closest neighbors ($k = 5$ or $k = 10$)

5.1.2.1. Minnesota. Again, in this state the Euclidean function outperformed the Mahalanobis function ($r^2 = 0.54$ and 0.48, respec-

Author's personal copy

Table 3
Summary of the step-wise regressions for IBI prediction for the development and validation sets. The variables used in each case are listed together with their coefficients and curve type (in parentheses). Variables in italics in the whole database regressions indicate variables also used in some of the kNN predictions. Results in Ohio after including metal toxicity penalties.

Scale	State	Variables	State	Variables	State	Variables
Whole database	MN	Cond (-3.1×10^{-5}, Q) TN (.066, −2.58, Q) TSS (−.235, L) MWidth (−27.6, I) MThalDep (−496.62, I) PctRun (−.007, .77, Q) LU + 1 (9.5, Lg) Constant (56.81) r^2 = .68 r^2 validation = .52 RMSE* = 27.16	MD	Agribarr/10 (−.035, .488, Q) ANC (−.00048, L) Aesthet (−.006, .15,Q) Vel-dpth (.049, L) MWidth (.02, L) Constant (.597) r^2 = .66 r^2 validation = .58 RMSE* = 0.71	OH	Embed (2.397, −17.843, Q) Subs (.273, L) Pool (.563, L) SO_4 (−3.957, Lg) Hard/100 (−.264, 3.031, Q) TKN (.846, I) Cu (−.087, .749, Q) Constant (66.260) r^2 = .47 r^2 validation = .41 RMSE* = 7.23
Cluster 1	MN	PctDistLU (0.0015, Q) TSS (−11.5, Lg) PctOverVeg + 1 (50.97, I) Rip (0.273, Q) Constant (10.66)	MD[b]	ANC (−0.069, Q) Vel-Dpth (.182, L) Agribarr (−.081, Q) Urban (−.046, .1, Q) Shading (.121, Q) Constant (3.84)	OH	Embed (−4.541, L) SO_4 (−.031, L) QHEI (−967.42, I) pH (−13.80, 219.31, Q) TKN (.686, −5.342, Q) Hard (.018, L) Chan (−.098, 2.742, Q) NO_3 (.01, −.610, Q) NO_2 (9.887, L) Constant (−824.408)
Cluster 2[a]	MN	TN (−0.398, Q) PctEmbed (−0.367, L) MThalDep (−314.2, I) MDepth (1075.9, I) PctPool + 1 (−145.7, I) Constant (113.81)	MD[b]	Agribarr (−.178, .159, Q) ANC (−.223, L) PHI (.293, L) CH_Flow (.188, L) Shading (−.181, L) Constant (3.51)	OH	Riffle (1.703, Q) Pool (.649, L) Zn (−.85, .019, E) Subs (.497, L) NO_3 (.101, −2.074, Q) NO_2 (−293.3, 60.71, Q) Cd (4.097, L) SO_4 (−1.897, Lg) Constant (29.256)
Cluster 3[c]	MN	Temp (−822.4, S) SI (−1.48, L) MThalDep (−623.04, I) TN (−6.69, E) Constant (957.72) r^2 = .81 r^2 validation .59 RMSE* = 25.81	MD	Aesthet (.075, 1.05, P) Urban (−.67, Lg) Cond (-7.08×10^{-7}, Q) EpiSub (.007, .23, E) Constant (4.075) r^2 = .71 r^2 validation .62 RMSE* = 0.66	OH	BOD (23.295, I) Cond (−.001, L) Zn (.001, −.124, Q) TKN (117.23, .021, P) SO_4 (291.707, I) Constant (−98.829) r^2 = .62 r^2 validation = .44 RMSE* = 6.78

L: linear; E: exponential; Lg: natural log; I: inverse; Q: quadratic; P: power; S: S-curve.
[a] Clusters 2–4 in Maryland.
[b] Regressions performed with standardized values.
[c] Cluster 5 in Maryland.
* RMSE for different IBI predictions. IBI scales: 12–60 in Ohio, 0–100 in Minnesota, 1–5 in Maryland.

tively with $k = 5$). Land use degradation in the riparian corridor explained a big percentage of the total biotic variability (around 40% in both cases). However, land use-related variables were removed from the prediction with the Euclidean function after backtracking and including other variables such as TN or conductivity (Table 2).

5.1.2.2. Maryland. Like in the previous case, both functions performed almost identically in terms of predicting capability ($r^2 = 0.59$ in both cases with $k = 10$). However, the Mahalanobis function only needed two variables (Agricutural/barren land and velocity–depth variability) to achieve such result. Conversely, the Euclidean needed six variables. Both variables used with the Mahalanobis function were also used in the regression model and in the Euclidean distance function. Other variables used in the regression model (i.e. ANC) were also included with the Euclidean function.

5.1.2.3. Ohio. Unlike the predictions with $k = 1$, the Euclidean distance obtained better overall results ($r^2 = 0.53$ versus 0.47 with $k = 5$). The Mahalanobis distance needed only five variables opposed to nine needed with the Euclidean function. The latter, matched the variables selected in the regression model very well. The Mahalanobis function only needed three physical habitat parameters (embeddedness, riffle and pool quality) to explain a large part of the total IBI variability (44%). The remaining IBI variability (up to the final 47%) was explained with the inclusion of cadmium and copper concentrations (Table 2).

5.2. Regression models

5.2.1. Predictions at the state-/regional-scales

5.2.1.1. Minnesota. The environmental variables included in the regression model were conductivity, total nitrogen, TSS, land use score, mean width, and mean thalweg depth. They explained 68% of the total variability in the development set and 52% in the validation set (Table 3). In the latter subset, the main source of error was found in those sites with very poor IBI scores (IBI = 0) mainly located in the Des Moines River basin (see most southern observations in cluster 1 in Fig. 4).

Author's personal copy

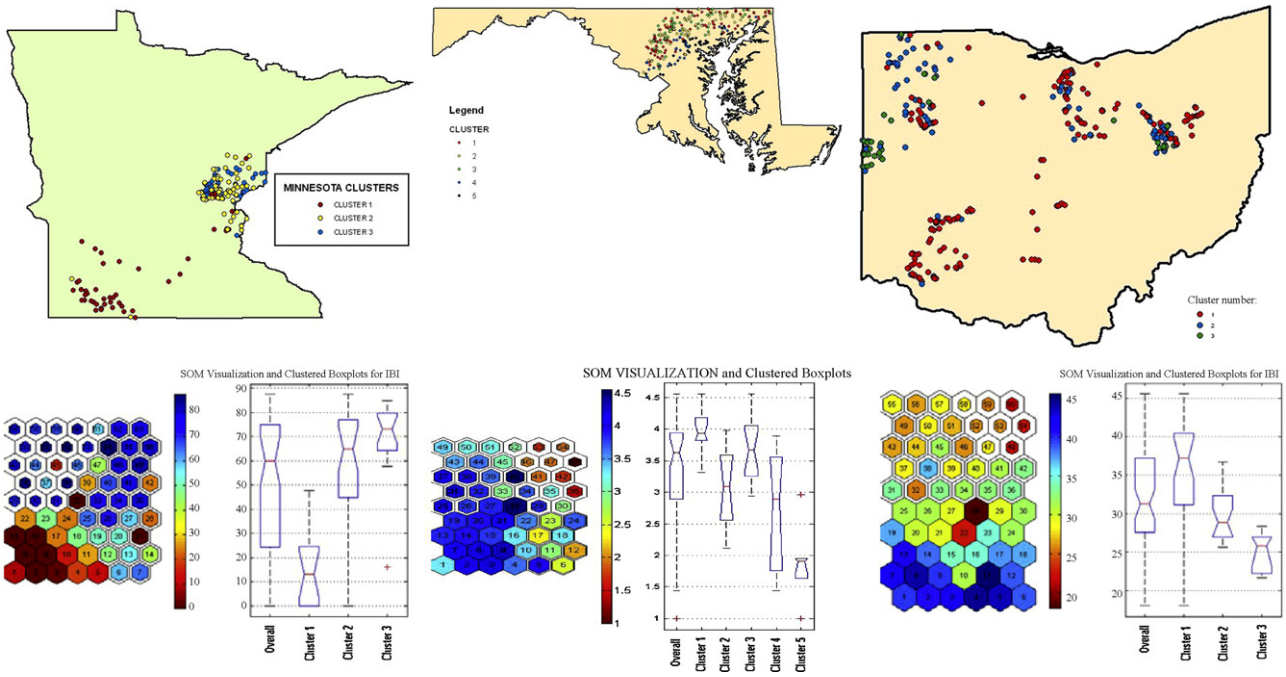


Fig. 4. Site cluster distribution in Minnesota (left), Maryland (Piedmont sites) (center), and Ohio (right). In Minnesota, cluster 1 is concentrated in Southern watersheds. In Maryland clusters 4 and 5 are concentrated in a specific region and, in Ohio, sites located in the same watershed usually belong to the same cluster.

5.2.1.2. Maryland. In this case, a significant portion of the biological variability was explained with five variables; percentage of agriculture/barren lands in the drainage area, alkalinity, aesthetic quality, velocity–depth variability, and mean width. The best model explained 66% of IBI variability in the regression development dataset and 59% in the validation set. Nutrient- and ionic strength-related variables (i.e. NO3-N and conductivity) ranked among the top variables sorted according to their IBI prediction power. However, similarly to what occurred in the kNN based predictions, these variables were not selected in the step-wise process perhaps due to strong cross-correlations with alkalinity and agricultural land uses (ANC-Cond $r = 0.76$; Agribarr-NO_3 $r = 0.69$).

5.2.1.3. Ohio. At the state-scale, 47% of the total variability was explained with the regression development set and 41% with the validation set. The variables included in the final equation were mostly related to substrate quality (i.e. embeddedness and substrate quality), habitat variability (i.e. pool quality), and nutrient-related chemicals (i.e. SO4, hardness, and TKN). Copper concentration, which was the only metal in this model, introduced little improvement in the final result (Table 3).

5.2.2. Predictions at the scale of "cluster of similar sites"

5.2.2.1. Minnesota. The SOM yielded three clusters that were very clearly separated geographically (see Fig. 4). The observations located in the Southern areas (cluster 1, Des Moines and Minnesota River basins) showed significantly lower IBI and habitat quality. They had higher accumulations of fine sediment in the substrate and larger percentages of disturbed land use in the riparian strip. Those areas were also associated with higher total nitrogen and TSS concentrations as well as higher conductivity. Clusters 2 and 3 (located mostly in the Upper Mississippi and St. Croix River Basins) had similar land use and chemical qualities. The main difference between them was due to physical habitat quality, especially substrate and channel characteristics, which translated into significantly higher biological scores in cluster 3, with better habitat quality.

The IBI cluster-based prediction improved significantly the results obtained using the entire state database. The variables selected (different for each cluster) were able to explain 80% of the total variability in the regression development set and 59% in the validation one (Table 3).

5.2.2.2. Maryland. Five clusters of SOM neurons were identified. As shown in Fig. 4, clusters 4 and 5 (with the poorest IBI scores) were concentrated in the lower parts of the piedmont basins discharging in the Chesapeake Bay or the Potomac River (Fig. 4) which indicated a progressive degradation with stream size. IBI scores in cluster 1 were significantly higher than the rest, while cluster 5 scores were significantly lower. Clusters 2–4 showed similar median IBI values, with wide ranges and overlapping among them. For this reason, and because a minimum number of observations was required for each cluster to develop its own equation, clusters 2–4 were merged. In this case, the regression dataset explained up to 71% of the total variability, and a 62% in the validation dataset. The variables used in the final model are presented in Table 3.

5.2.2.3. Ohio. Three very prominent clusters of SOM neurons were found. The IBI distribution was significantly different in all three, having cluster 1 the highest IBI scores and cluster 3 the lowest (see Fig. 4). The best predicting variables differed significantly from cluster to cluster and therefore, the main cluster-specific drivers of biological integrity were identified. The cluster-based predictions outperformed significantly the non-clustered model in the model development set ($r^2 = 0.62$). However, the validation results were very similar to the non-clustered model ($r^2 = 0.44$).

6. Discussion

Two major outcomes derived from our results: (1) the kNN methodology outperformed the more traditional regression approach in both, prediction results and versatility and computational speed; and (2) the relationship between anthropogenic stressors and biological integrity is a function of scale as seen in the regression model.

Author's personal copy

The kNN methodology obtained better overall IBI prediction results in all three states (see Tables 2 and 3). This is especially relevant because all predictions obtained using kNN are validated against the remaining observations in the database, while the regressions are validated using 25% of the sites only. Therefore, the kNN methodology does not require an a priori partition of the database with the associated inherent risk of non-representative sampling. Moreover, the kNN methodology is especially well suited to deal with large number of observations composed of numerous attributes of different nature (e.g. discrete versus continuous data) because it is based on comparison of pairs of observations with the same variables. Moreover, kNN is able to deal with noisy observations (Little and Rubin, 1987; Troyanskaya et al., 2001). More and more database systems such as molecular biology, medical imaging, or genetics are using the efficiency and versatility of similarity search (Seidl and Ktiegel, 1998; Troyanskaya et al., 2001). Techniques such as kNN are necessary for ecological applications because they are able to deal with large databases with highly inter-twined, non-linear variable relationships. Unlike many other exploratory techniques, kNN does not make any pre-conceived assumptions about underlying data relationships, as opposed to the Canonical Correspondence Analysis (CCA), an exploratory technique widely used in ecology, which assumes linearity between explanatory and response variables (Ter Braak, 1986). Moreover, kNN is robust with different types of data and can deal with them all at once (Gustavo et al., 2003) unlike regression approaches, which are very sensitive to data quality and cumbersome to work with when many explanatory variables exist.

Our results confirmed the advantages of the kNN approach in all three databases. The regression results (at the state-scale) versus the kNN results showed how both methods consistently selected very similar top predicting variables (e.g. in Minnesota, TN, conductivity, TSS, percentage of undercut banks, and mean stream depth were the variables selected in the best kNN model, while conductivity, TN, TSS, mean stream width, and mean thalweg depth were the top five metrics in the regression approach [see Tables 2 and 3]). The kNN method proved its robustness and broad applicability in ecological fresh water systems by consistently improving the performance of traditional prediction approaches in all three states. The databases used are independent from each other because each state follows its own monitoring strategy (Ohio EPA, 1999b; Maryland and Maryland, 2000; MPCA, 2004) and has its own habitat and biological index calibration and validation methodologies (Ohio EPA, 1987; Roth et al., 1998; Niemela and Feist, 2000). In all three states, results from the kNN-based approach agreed strongly with results from available literature which attempted to predict IBI using the same or similar datasets.

In Ohio, habitat degradation has been identified as the main source of biological degradation, mainly as a consequence of non-point source pollution from agricultural land uses (Dyer et al., 2000; Manolakos et al., 2007). All the kNN predictions in our research pointed towards this direction by selecting habitat variables as top predictors in all cases in Ohio. In the piedmont regions of Maryland, combinations of agricultural and urban land uses (which were the dominant land uses with a strong, negative cross-correlation [$r = -0.83$]), physical habitat quality (i.e. PHI), and alkalinity set most of the biological quality in this stratum. Similar results were also reported by the MBSS between 1995 and 1997 (Klauda et al., 1998; MBSS, 1999) and are based on the same database used in the current paper. Biological integrity in piedmont basins between 1995 and 1997 was mostly affected by urban impacts (Patapsco and Potomac Washington Metro basins), agriculture (Gunpowder basin), and habitat degradation (Susquehanna, Bush, Gunpowder, Upper Patuxent, and Potomac Washington basins) (MBSS, 1999). In the review from MBSS (1999), unlike our predictions, physical variables not used in the calculation of the stratum's PHI were not included in the analysis (for a description of the Maryland's PHI see Paul et al. (2002)). We included these variables (see Table 1) and our model was still capable of identifying those variables with the greatest impact on IBI discarding variables that introduced noise. As our predictions indicated, urbanization is one of the most important sources of biotic degradation in Maryland's non-tidal streams (Volstad et al., 2003). In piedmont regions, urban land use had a strong negative correlation to IBI ($r = -0.68$). The opposite happened with agriculture ($r = 0.65$).

In Minnesota, the regression-based and the best kNN-based prediction methods (i.e. with Euclidean distance and $k = 5$, see Table 2) agreed strongly and selected very similar predicting variables. Again the kNN approach outperformed regressions and needed just five variables instead of seven needed in the latter. In Northern basins, stream health was mainly driven by combinations of forested and agricultural/range lands. Therefore, total phosphorus concentrations (highly correlated to total nitrogen in our dataset, $r = 0.84$), conductivity, and TSS were the main chemical quality issues in these areas. A clear negative correlation between IBI and substrate quality (e.g. percent of fine sediment) and IBI and stream size (e.g. mean width or depth) have also been reported in this region (Niemela et al., 2004). Conversely, Southern regions suffer from long-term, intense agricultural impacts. Around 80% of the land in the Minnesota and the Des Moines River basins is used for agriculture. In the Minnesota river, substrate degradation, channelization, and organic enrichment were the major impacts on IBI in the early 1990s (Bailey et al., 1992) and early 21st century (Feist and Niemela, 2002). IBI is positively correlated to watershed size and the deleterious effects of habitat degradation and nutrient enrichment seem to have a greater impact on headwater streams where very low IBI scores were observed (Bailey et al., 1992). The Des Moines River basin is also a highly agricultural area and most of the available IBI observations in the basin were equal to zero (essentially no fish present in these sites). These could be explained by a long-term extreme nutrient loading which has maintained the average nitrate levels well above 5 mg/L since 1945. Nitrate concentrations greater than 10 milligrams per liter (the statutory maximum for drinking water supplies (Hallberg, 1987)) have been reported in this river (Turner and Rabalais, 2003). High concentrations of nitrate can be associated to the presence of wastewater treatment plants and other urban discharges or intensive agriculture tile drainage. Negative effects on biota may be observed with mean nitrate concentrations greater than 3–4 mg/L (Ohio EPA, 1999a). Both prediction methodologies successfully identified main sources of reported biological stress derived from agricultural practices, mainly derived from increased nutrient loading and substrate degradation.

In all three states, kNN predictions using the Euclidean distance outperformed or equaled their homologues using the Mahalanobis distance with the exception of Ohio with $k = 1$. The Mahalanobis function usually needed lower number of variables to obtain the best possible prediction. This was especially true when the number of closest neighbors (k) increased (Table 2). Oversimplification due to elimination of correlated variables with the Mahalanobis function is a potential cause that could explain the differences in function performance. As a cautionary measure, we would recommend using it when the goal is the identification of large-scale stressors only (e.g. landscape patterns) or when there is lack of data for certain variables. In the complexity of the natural system, even cross-correlated variables can introduce new, valuable information in the predictions. The Mahalanobis function seems, in some cases, to fail in the identification of underlying natural processes captured by the Euclidean distance function, which is especially well suited for providing estimations in small tight expression clusters (Troyanskaya et al., 2001).

Author's personal copy

The performance of kNN was comparable to other more traditional methods used for prediction purposes in fresh water systems. One of the main advantages of kNN is that no pre-conceived assumptions about the stressor-relationship needed to be made. Such a priori assumptions could lead to misleading results in statistical or regression models (Austin, 2002). However, it is important to keep in mind some of the limitations of methods based on evaluating similarity: (1) these methods are data-driven, therefore a complete dataset with observations covering the whole range of the different variables is needed in order to predict accurately in the whole IBI spectrum. Since extreme events are rare in nature, predictions in the upper/lower boundaries might be less accurate and/or reliable; (2) if a truly similar observation to our query site does not exist, kNN would still give a prediction. In future work, the reliability of a prediction could be assessed based on the distance between the query and the identified most similar site. When more than one neighbor is used, an IBI prediction based on weighted distance to the target site could be introduced; (3) the results from kNN (like other methods) reveal the main stressors at the given study scale and results will vary if this is modified. The best IBI predictors are a function of the scale at which the problem is approached; (4) a step-wise variable selection is necessary because large number of non-relevant variables will lead to inconclusive results because each variable has the same weight. If many non-relevant variables are introduced, the model performance could decrease drastically. Therefore, it is very important to find the underlying meaningful dimensions in our data (Beyer et al., 1999).

In most cases, SOM are only used for organization of multidimensional biological community databases (Kruk et al., 2007; Park et al., 2001, 2007; Manolakos et al., 2007; Chon et al., 1996; Céréghino et al., 2001). In the present manuscript, the advantages of advanced patterning techniques, such as SOM, in high dimensional systems, such as fresh water ecosystems, were exploited to improve the performance of predictive modeling techniques. The importance of scale has been revealed with the segregation of sites into clusters of similar environmental observations (site patterning) using SOM. The relevance of site organization lies in the possibility to frame the effect of local variables because large scale-variables (i.e. state- or regional-scale variables in our research) have little intra-cluster variability. In other words, large-scale variables set the cluster background environmental conditions and therefore the intra-cluster IBI oscillations due to cluster-specific stressors can be revealed. The effect of local variables on IBI must be seen within an environmental context which is determined by larger-scale variables or environmental filters at higher levels of the fresh water system ecological hierarchy (Frissell et al., 1986; Poff, 1997). In our research, background biological quality is mostly set by instream widespread effects derived from anthropogenic landscape changes (e.g. decrease in substrate quality and stream's habitat variability, nutrient load increase, or increase in ionic strength-related parameters). These variables and associated IBI had high state-wide variability (e.g. IBI and percent run in Minnesota state-wide: $\mu_{IBI} = 47.8$, $\sigma_{IBI} = 34.8$; $\mu_{PctRun} = 64.5$, $\sigma_{PctRun} = 22.1$). After clustering, large-scale variables and associated biological responses are grouped into well defined stressor-response units (e.g. three clusters in Minnesota (see Fig. 4): $\mu_{IBI\text{-}1} = 11.23$, $\sigma_{IBI\text{-}1} = 22.83$; $\mu_{PctRun\text{-}1} = 79.85$, $\sigma_{PctRun\text{-}1} = 6.40$; $\mu_{IBI\text{-}2} = 58.56$, $\sigma_{IBI\text{-}2} = 6.93$; $\mu_{PctRun\text{-}2} = 64.13$, $\sigma_{PctRun\text{-}2} = 19.64$; $\mu_{IBI\text{-}3} = 71.42$, $\sigma_{IBI\text{-}3} = 17.98$; $\mu_{PctRun\text{-}3} = 52.70$, $\sigma_{PctRun\text{-}3} = 19.64$. Inter-cluster differences in mean IBI and PctRun were statistically significant according to an ANOVA test [$p < 0.05$]).

As a consequence of the segregation of observations into clusters of similar sites, intra-cluster IBI oscillations due to more localized stressors could be revealed. In Ohio, clusters were graded by the percentage of agriculture and subsequent effects on fish habitat and nutrient loading (mean percentage of agriculture equal to 31%, 50%, and 87% in clusters 1,2, and 3, respectively). Clusters 1 and 2 were mainly driven by habitat and nutrient input. However, cluster 2 also had some impact from metals (zinc and cadmium, see Table 3). Cluster 3 was severely impaired and biological quality was mostly driven by water chemical quality despite having very poor habitat quality as well (mean QHEI equal to 29). The effect of the highly degraded water quality seemed to have a deep effect on the biological community within this cluster beyond degradation of physical habitat, which could indicate acute effects may exist when certain concentration thresholds are reached. Eutrophication (i.e. BOD, TKN) and zinc concentrations were the main chemical sources of degradation. In Ohio, the site cluster segregation improved the overall prediction. This was especially true in the regression development set and not so much with the validation set (Table 3).

The addition of the metal toxicity penalty marginally improved the predictions because very few sites reached the CCC threshold. Metal pollution is no longer considered one of the top threats to attainment of aquatic life uses in Ohio as a consequence of successful regulatory efforts. Moreover, there exists ambient and laboratory evidence that the CCC threshold may be overly stringent (Ohio EPA, 1997). In Ohio, metal concentrations beyond the USEPA reported CCC threshold have been often observed in sites where the Warm Water Habitat (WWH) biological criterion (i.e. IBI ≥ 40) was achieved (19% and 16% of the times for copper and zinc, respectively) (Ohio EPA, 1997). Despite this, some metals were still present in some of the whole state predictions (total copper and cadmium) and cluster-based predictions (i.e. total zinc and cadmium in cluster 2 and total zinc in cluster 3) (see Tables 2 and 3).

In Maryland, three groups were segregated based on increasing agriculture/urban ratio. IBI was positively correlated to agriculture in this stratum and negatively correlated to urban extent, which indicated the more acute impacts from the latter. The lower parts of the basins (clusters 4 and 5) had the highest levels of urbanization and therefore, the poorest mean IBI score (2.7 and 2.3, respectively). Thus, group 1 (with the best integrity, mean IBI equal to 3.81) had the highest percentage of agricultural land use (65%) and the smallest percentage of urbanization (3.95%). Conversely, group 3 had the poorest biological integrity and the largest extent of urban land use (mean agricultural and urban land uses equal to 24.6% and 48.4%, respectively). Within these more homogeneous landscape units, the effect of local variables on group's IBI oscillations such as stream shading in groups 1 and 2, or substrate in group 3 were revealed and helped improve the previous model for the overall stratum (Table 3). The influence of stream shading on IBI has been reported to be more significant as the geographic scale is progressively reduced as indicated in Fig. 1 (Allan et al., 1997).

In the case of Minnesota, the SOM separated the observations in three clusters which were almost perfectly segregated into Southern (cluster 1) and Northern watersheds (clusters 2 and 3) (see Fig. 4). Cluster 1 was mainly composed of sites from the Des Moines and the Minnesota River basins which are strongly dominated by agricultural land uses. Only a few sites from the St. Croix and the Upper Mississippi River basins, whose land use is dominated by a combination of agriculture and forest, belonged to cluster 1. This cluster had very poor biological integrity (mean cluster IBI equal to 11) and was composed of sites with larger drainage areas (average DA equal to 1701 mi^2) which, like in the case of Maryland, could indicate progressive degradation downstream. This cluster was severely affected by both, chemical and physical stressors and had significantly higher levels of total nitrogen and phosphorus, turbidity, suspended solids, conductivity, pH, and percentage of over vegetation which could indicate highly eutrophic conditions. The habitat quality was significantly lower in cluster 1 (QHEI, percent of woody debris) with high levels of disturbed lands in the drainage area and riparian buffers (62% and 47%, respectively). The riparian quality and riparian width were also significantly lower than in

Author's personal copy

clusters 2 and 3 (mean width equal to 6.5 m) with severe fine sediment deposition and sections of slow moving waters (percent fines equal to 74% and percent of pool-run equal to 91%). Therefore, in this cluster the deleterious effects of long-term agricultural basins (Bailey et al., 1992; Feist and Niemela, 2002; Turner and Rabalais, 2003) was revealed with the reduction of scale.

Significant differences in group mean IBI (ANOVA test, $p < 0.05$) were also identified between clusters 2 and 3 in Minnesota (mean IBI equal to 58.5 and 71.4, respectively). Sites from both clusters had similar landscape characteristics (disturbed lands in the drainage area accounted for 8.7% and 9.3% and disturbed land in the 30-m buffer accounted for 1.74 and 7.74% in clusters 2 and 3, respectively), similar riparian quality, and similar water quality (cluster 2 had slightly higher mean conductivity and nutrient concentration but differences were not statistically significant). The main differences between cluster 2 and 3 were found in habitat quality (mean QHEI equal to 61.2 and 71.4, respectively). The habitat parameters with significant inter-cluster differences were related to fine sediment deposition (higher percentage of fine sediments were observed in cluster 2) and stream variability (cluster 2 had many more segments of slow moving waters than cluster 3 (95% was either pool or run in cluster 2 versus 79% in cluster 3). Different reasons could explain these differences. Headwater streams within the St. Croix River basin often originate from peat lands, resulting in dark, tannic acid stained water. These streams are usually low gradient streams that lack riffles and have a glide/pool type of stream morphology. In addition they are typically sinuous, with fine substrates and have a riparian zone comprised of wetland vegetation. On the other hand, lower reaches in the St. Croix River Basin have a riffle/run/pool stream morphology with a variety of substrate types and a wooded riparian zone (Niemela and Feist, 2000). This would explain the greater deposition of fine sediment in cluster 2 compared to cluster 3 under similar landscape and riparian characteristics. Furthermore, cluster 2 also showed significantly higher concentrations of macrophytes (both submerged and emerging). An overabundance of macrophytes can result from high nutrient levels (Crowder and Painter, 1991). However, since the differences in nutrient concentration between clusters 2 and 3 were not significant, the larger number of macrophytes is most likely due to the increased presence of stagnant waters in sites from cluster 2. The macrophyte community is usually targeted to monitor the biological quality of slow moving waters such as lakes, wetlands, or reservoirs (USEPA, 2006).

The St. Croix River basin has historically undergone two rapid modifications: from forest to agriculture in the 19th century, and rapid urbanization since mid 20th century, especially in the Twin Cities area in the southern region of the basin (Andersen et al., 1996). However, forest and agriculture/pasture still remain as main land uses (44% and 34%, respectively) (Fago and Hatch, 1993). From our analysis, it remains unclear if the habitat degradation in cluster 2 is due to the basin's headwater characteristics explained above or due to instream modifications for agricultural/urban purposes. More detailed land use information is necessary to confirm this point.

In summary, the SOM partitioning process successfully separated sites based on their environmental conditions which had in turn, different biological responses as shown in Fig. 4. In Minnesota, biological integrity in cluster 1 was best predicted with landscape related parameters (land use quality and riparian scores), and presence of over vegetation, which indicated the highly eutrophic conditions in the Southern watersheds. Cluster 2 was best predicted with total nitrogen, substrate, stream variability, and stream size/morphologic parameters because of its characteristics explained above. Gradient, stream size, and total nitrogen explained also much of the variability in cluster 3. The presence of nitrogen in the cluster 2 and 3 regression equations could indicate a gradient of biological degradation with a progressive shift from natural to modified land uses. The reduction of scale in these three well defined clusters improved the previous models significantly by revealing local or cluster-based ecological stressors (Table 3). The SOM clustering clearly separated sites based on their anthropogenic impacts: severe landscape modifications coupled with very poor instream chemical and physical qualities in cluster 1; good land use qualities (over 90% of undisturbed lands), with progressively increasing levels of nutrients, and impacted/unimpacted physical habitat in clusters 2 and 3, respectively.

As a concluding remark, it is important to keep in mind that biological monitoring strategies differ from state to state and IBI prediction results may be affected by these differences. Ohio EPA combines a five-year basin approach in which the U.S. Geological Survey 11-digit Hydrologic Unit Code (HUC) watersheds are used as the basic unit of assessment. Sampling locations within this basic unit are obtained by progressively halving the drainage area. Therefore, small watersheds are sampled more frequently than larger ones. Moreover, other fixed monitoring stations are selected based on the location of potential sources of degradation such as point sources or tributaries (Ohio EPA, 1999b). Minnesota's MPCA follows a 10-year basin rotation cycle. Like in Ohio, sampling sites are selected using a geometric "upward" halving approach coupled with a probabilistic selection of 50 random sites per basin (MPCA, 2004). Maryland's MBSS followed a three-to-five-year rotating sampling strategy targeting 17 out of 20 6-digit HUC basins in the 1995–1997 period. Only wadeable streams (first to third order) were sampled. A probability-based lattice was used to select the sampling sites. In this state, the number of sampling sites of each stream order was proportional to the number of miles of the stream order in the state. Thus, at the end, the number of miles per sample in first-, second-, and third-order streams was approximately the same. The MBSS also monitored around 300 fixed sites state-wide targeting local degradation issues (Maryland and Maryland, 2000). Therefore, given the monitoring strategies, the state-wide results could be biased towards stressors affecting smaller-order streams because these are sampled more frequently, especially with the approaches used in Ohio and Minnesota. Stream size related parameters were selected in the state-wide prediction models in Maryland and Minnesota (see Tables 2 and 3). A correlation between IBI and drainage area has also been reported in Ohio (Dyer et al., 2000). The sampling distribution in all three states is well suited for the identification of large-scale stressors (e.g. regional land use) because multiple watersheds with diverse regional characteristics were used. Our results confirmed this. On the other hand, geometric upward aggregation approaches might be better suited to detect local pollution gradients when reduced geographic scales are used (e.g. at the subwatershed level) because the sampling intensity allows capturing the IBI degradation "continuum" by local stressors in background homogeneous conditions (e.g. very similar regional land uses). Probabilistic approaches might fail to detect local gradients due to lack of observations needed to track changes produced by more geographically localized stressors.

7. Conclusions

- A novel IBI prediction methodology based on assessing environmental similarity was presented and implemented. It is based on using the non-parametric *k*-nearest neighbors classifier. Its performance was compared to the traditional regression approach using the same data and geographic scale in three states of the US. The two methods selected similar predicting variables. However, the kNN method outperformed the regression method obtaining more accurate predictions (higher r^2 and lower RMSE) using the whole set of available data and not only a fraction (25% in the

Author's personal copy

regressions). kNN was much faster computationally and allowed applying leave-one-out cross-validation unlike the regression approach. With kNN, more than 50% of the IBI variability was explained at the state-scale in all three states.

- The kNN approach is a technique that can be easily used in highly dimensional systems with inter-twined, cross-correlated variables, such as fresh water systems. An advantage of this non-parametric simple to implement technique is that it makes no a priori assumptions about the relationship between explanatory and response variables.
- Variable selection techniques are advised in order to avoid predictions with highly dimensional vectors that could lead to inconclusive results. Like the regression approach, the kNN-based methodology is able to make predictions at one specific geographic scale only and the predictions are limited by the information carried by the explanatory variables. Results cannot be extrapolated to other settings.
- Both methods identified landscape patterns (i.e. land use) and land use related instream conditions (i.e. substrate quality, stream variability, or nutrient loading) as the best predictors of biological integrity at the state or regional scales. The significant anthropogenic stressors found in this research are the variables with the most relevant impact on IBI given the available databases from the corresponding biological monitoring programs. IBI is a continuous stream health measure able to reflect past polluting events not captured in periodic sampling programs, which reflect only a portion of the whole environmental spectrum. Hence, valuable environmental information for prediction of biological integrity might be missing.
- The importance of scale was clearly revealed after using SOM-based patterning of sampling sites. Observations were clustered into groups of homogeneous land use and land use-related anthropogenic impacts (mainly nutrients, substrate quality, and stream's habitat variability). These groups usually corresponded to well defined geographic areas (i.e. watersheds or basins). The reduced intra-cluster variability of large-scale variables allowed the expression of stressors with an impact at a portion of the sites only (i.e. small-scale stressors). Therefore, the inter-dependence between biological community, anthropogenic stressors, and geographic scale was revealed. The cluster-based IBI predictions improved significantly relatively to using the whole datasets.
- Some of the future research our team will perform is the development of a methodology to predict and/or characterize biological integrity at multiple scales, which would allow water resources managers see the whole "stream's integrity picture". Biological quality in one area will be presented as the consequence of the local stressors but within a background context set by large-scale variables. This would allow determination of effective managing practices by giving a holistic view of biological integrity. Furthermore, we are working on the development of a methodology similar to kNN which would select only "truly similar" neighbors using an agglomerative hierarchical clustering tree structure (Bedoya, 2008). This approach would mitigate the deleterious effects of selecting neighbors with a very different environmental behavior in the IBI prediction process.

References

Allan, J.D., Erickson, D.L., Fay, J., 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37, 149–161.

Andersen, O., Crow, T.R., Lietz, S.M., Stearns, F., 1996. Transformation of a landscape in the upper mid-west, USA: the history of the lower St Croix River valley, 1830 to present. Landscape and Urban Planning 35, 247–267.

Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157 (2–3), 101–118.

Bailey, P.A., Enblom, J.W., Hanson, S.R., Renard, P.A., Schmidt, K., 1992. A fish community analysis of the Minnesota River Basin, Minnesota Pollution Control Agency. Minnesota River Assessment Project 1990–1992, St. Paul, MN.

Barbour, M.T., Gerritsen, J., Snyder, B.D., Stribling, J.B., 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic, macroinvertebrates, and fish, second. ed. EPS-841-B-99/002, US Environmental Protection Agency, Washington, DC.

Bedoya, D., 2008. Biological response to environmental stress. Environmental similarity and hierarchical, scale-dependant segregation of biotic signatures for prediction purposes. PhD Dissertation, Northeastern University, Boston, MA.

Beyer, K., Goldstein, J., Ramakrishnan, N., Shaft, U., 1999. When is "nearest neighbor" meaningful? In: 7th International Conference in Database Theory. Springer Berlin/Heidelberg, Jerusalem, Israel.

Bode, R.W., 1988. Methods for Rapid Biological Assessment of Streams. New York State Department of Environmental Conservation, Albany, NY.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2003. Integrated approach to total maximum daily load development for Neuse River Estuary using Bayesian probability network model (Neu-BERN). Journal of Water Resources Planning and Management-ASCE 129, 271–282.

Céréghino, R., Giraudel, J.L., Compin, A., 2001. Spatial analysis of stream invertebrates distribution in the Adour-Garonne drainage basin (France), using Kohonen self organizing maps. Ecological Modelling 146, 167–180.

Chon, T.S., Park, Y.S., Moon, K.H., Cha, E., Pa, Y., 1996. Patternizing communities by using an Artificial Neural Network. Ecological Modelling 90, 69–78.

Crowder, A., Painter, D.S., 1991. Submerged macrophytes in Lake Ontario: current knowledge, importance, threats to stability, and needed studies. Canadian Journal of Fisheries and Aquatic Sciences 48, 1539–1545.

Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Transactions on Pattern Analysis and Machinery Intelligence 1, 224–227.

Doudoroff, P., Warren, C.E., 1957. Biological Indices of Water Pollution with Special Reference to Fish Populations. Biological Problems in Water Pollution. U.S. Public Health Service, Cincinnati, OH, pp. 144–153.

Duda, R.O., Hart, P.E., Stork, D.G., 2000. Pattern Classification, second edition. Wiley, New York, NY.

Dyer, S.D., White-Hull, C., Carr, G.J., Smith, E.P., Wang, X.H., 2000. Bottom-up and top-down approaches to assess multiple stressors over large geographic areas. Environmental Toxicology and Chemistry 19, 1066–1075.

Dyer, S.D., White-Hull, C., Wang, X., Johnson, T.D., Carr, G.J., 1998. Determining the influence of habitat and chemical factors on instream biotic integrity for a Southern Ohio watershed. Journal of Aquatic Ecosystem Stress and Recovery 6, 91–110.

Fago, D., Hatch, J., 1993. Aquatic resources of the St Croix River Basin. Biological Conservation 76 (211), 213–1213.

Feist, M., Niemela, S., 2002. Evaluating Progress of Biological Condition in Streams of the Minnesota River Basin, Minnesota Pollution Control Agency. Biological Monitoring Program, St. Paul, MN (Available at: http://www.pca.state.mn.us/water/biomonitoring/sf-biostreams-mnriver.pdf).

Frissell, C.A., Liss, W.J., Warren, C.E., Hurley, M.D., 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10, 199–214.

Gustavo, E.A., Batista, P.A., Monard, M.C., 2003. An analysis of four missing data treatment methods for supervised learning. Applied Artificial Intelligence 17, 519–533.

Hallberg, G.R., 1987. Nitrates in groundwaters in Iowa. In: D'Itri, R.M., Wolson, R.G. (Eds.), Rural Groundwater Contamination. Lewis Publishing, Chelsea, MI.

Hilsenhoff, W.L., 1987. An improved biotic index of organic stream pollution. Great Lakes Entomologist 20, 31–39.

Jain, A.K., Dubes, R.C., 1988. Algorithms for Clustering Data. Prentice Hall Inc., Saddle River, NJ.

Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM Computer Surveys 31, 264–323.

Karr, J.R., 1991. Biological integrity: a long-neglected aspect of water resource management. Ecological Applications 1, 66–84.

Karr, J.R., Dudley, D.R., 1981. Ecological perspective on water quality goals. Environmental Management 5, 55–68.

Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., Schlosser, I.J., 1986. Assessing Biological Integrity of Running Waters: a Method and its Rationale. Illinois Natural History Survey, Champaign, IL.

Karr, J.R., Yoder, C.O., 2004. Biological assessment and criteria improve total maximum daily load decision making. Journal of Environmental Engineering-ASCE 130, 594–604.

Kiviluoto, K., 1996. Topology preservation in self-organizing maps. In: IEEE International Conference in Neural Networks, pp. 294–299.

Klauda, R., Kazyak, P., Stranko, S., Southerland, M., Roth, N., Chaillou, J., 1998. Maryland biological stream survey: a state agency program to assess the impact of anthropogenic stresses on stream habitat quality and biota. Environmental Monitoring and Assessment 51, 299–316.

Kohonen, T., 2001. Self-Organizing Maps. Springer-Verlag, Berlin.

Kruk, A., Lek, S., Park, Y.-S., Penczak, T., 2007. Fish assemblages in the large lowland Narew River system (Poland): application of the self-organizing map algorithm. Ecological Modelling 203 (1–2), 45–61.

Lammert, M., Allan, J.D., 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23, 257–270.

Little, R.J., Rubin, D.M., 1987. Statistical Analysis with Missing Data. John Wiley and Sons, New York.

Author's personal copy

Lyons, J., 2006. A fish-based Index of Biotic Integrity to assess intermittent headwater streams in Wisconsin, USA. Environmental Monitoring and Assessment 122, 239–258.

Lyons, J., Piette, R.R., Niermeyer, K.W., 2001. Development, validation, and application of a fish-based Index of Biotic Integrity for Wisconsin's large warmwater rivers. Transactions of the American Fisheries Society 130, 1077–1094.

Manolakos, E., Virani, H., Novotny, V., 2007. Extracting knowledge on the links between the water body stressors and biotic integrity. Water Research 41, 4041–4050.

Maryland DNR and Maryland DE, 2000. Maryland Water Monitoring Strategy (Available at: http://www.dnr.state.md.us/bay/education/pubs.html).

Maryland Biological Stream Survey, 1999. Relative contribution of stressors and their cumulative impact. The state of the streams. 1995–1997 report. EA-99-6. Maryland Department of Natural Resources, Annapolis, MD. (Available at: http://www.dnr.state.md.us/streams/mbss/ea99-6_toc.html).

Minnesota Pollution Control Agency, 2004. Minnesota's water quality monitoring strategy (Available at: http://proteus.pca.state.mn.us/publications/reports/p-gen1-10.pdf).

Niemela, S., Christopherson, D., Genet, J., Chirart, J., Feist, M., 2004. Condition of rivers and streams in the St. Croix River Basin in Minnesota Minnesota Pollution Control Agency. Biological Monitoring Program, St. Paul, MN (Available at: http://www.pca.state.mn.us/publications/wq-bsm2-01.pdf).

Niemela, S., Feist, M., 2000. Index of Biological Integrity guidance for coolwater rivers and streams in the St. Croix River basin, Minnesota Pollution Control Agency, Biological Monitoring Program, St. Paul, MN (Available at: http://www.pca.state.mn.us/water/biomonitoring/sf-ibi-stcroix.pdf.).

Novotny, V., 2003. Water Quality. Diffuse Pollution and Watershed Management. John Wiley & Sons, New York.

Novotny, V., Bartosova, A., O'Reilly, N., Ehlinger, T., 2005. Unlocking the relationship of biotic waters to anthropogenic integrity of impaired stresses. Water Research 39, 184–198.

Novotny, V., Olem, H., 1994. Water Quality: Prevention, Identification, and Management of Diffuse Pollution. Van Nostrand Reinhold, New York, NY.

Ohio EPA, 1987. Biological Criteria for the Protection of Aquatic Life: Volumes I–III. Standardized Field and laboratory methods for assessing fish and macroinvertebrate communities. Division of Water Quality Monitoring and Assessment, Surface Water Section Columbus, OH. Available at: http://www.epa.state.oh.us/dsw/bioassess/BioCriteriaProtAqLife.html.

Ohio EPA, 1997. Using biological criteria to validate applications of water quality criteria: dissolved and total recoverable metals, Technical Bulletin 1997-12-4. Division of Surface Water, Monitoring, and Assessment, Columbus, OH. (Available at: http://www.epa.state.oh.us/dsw/documents/gli_bio.pdf).

Ohio EPA, 1999a. Association between nutrients, habitat, and the aquatic biota in Ohio rivers and streams, Technical Bulletin MAS/1999-1-1,Division of Surface Water, Monitoring and Assessment Section, Columbus, OH (Available at: http://www.epa.state.oh.us/dsw/documents/assoc_load.pdf).

Ohio EPA, 1999b. Ohio EPA five-year surface water monitoring strategy: 2000–2004. Division of Surface Water, Technical Bulletin MAS/1999-7-2, Columbus, OH (Available at: http://www.epa.gov/nheerl/arm/documents/ohiofiveyearmonitstratdraft.pdf).

Ohio EPA, 2008. Water Quality Standards. Chapter 3745-1-07: water use designations and statewide criteria (Available at: http://www.epa.state.oh.us/dsw/rules/01-07.pdf.).

Ott, W.R., 1978. Environmental Indices: Theory and Practice. Ann Arbor Science, Ann Arbor, MI.

Paul, M.J., Stribling, J.B., Klauda, R.J., Kazyak, P.F., Southerland, M.T., Roth, N.E., 2002. A Physical Habitat Index for freshwater wadeable streams in Maryland. Final Report. Maryland Department of Natural Resources. Chesapeake Bay and Watershed Programs. Monitoring and Non-Tidal Assessment, Annapolis, MD (Available at: http://www.dnr.state.md.us/streams/pubs/ea03-4phi.pdf).

Park, Y.-S., Kwak, I.-S., Chon, T.-S., Kim, J.-K., Jørgensen, S.E., 2001. Implementation of Artificial Neural Networks in patterning and prediction of exergy in response to temporal dynamics of benthic macroinvertebrate communities in streams. Ecological Modelling 146, 143–157.

Park, Y.-S., Song, M., Park, Y.-C., Oh, K.-H., Cho, E., Chon, T.-S., 2007. Community patterns of benthic macroinvertebrates collected on the national scale in Korea. Ecological Modelling 203 (1–2), 26–33.

Poff, N.L., 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16, 391–409.

Richards, C., Johnson, L.B., Host, G.E., 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53, 295–311.

Roth, N., Southerland, M., Chaillou, J., Klauda, R., Kazyak, P., Stranko, S., Weisberg, S., Hall, L., Morgan, R., 1998. Maryland biological stream survey: development of a fish Index of Biotic Integrity. Environmental Monitoring and Assessment 51, 89–106.

Roth, N.E., Allan, J.D., Erickson, D.L., 1996. Landscape influences on stream biotic integrity assessed at multiple spatial scales. Landscape Ecology 11, 141–156.

Seidl, T., Ktiegel, H., 1998. Optimal multi-step *k*-nearest neighbor search. ACM SIGMOD Record 27, 154–165.

Southerland, M.T., Rogers, G.M., Kline, M.J., Morgan, R.P., Boward, D.M., Kazyak, P.F., Klauda, R.J., Stranko, S.A., 2005. New biological indicators to better assess the condition of Maryland streams, Maryland Department of Natural Resources. Monitoring and Non-tidal Assessment Division. DNR-12-0305-0100, Annapolis, MD (Available at: http://www.dnr.state.md.us/streams/pubs/ea-05-13_new_ibi.pdf).

Stephan, C.E., Donald, I.M., Hansen, D.J., Gentile, J.H., Chapman, G.A., Brungs, W.A., 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses, U.S. Environmental Protection Agency. Office of Research and Development. Environmental Research Laboratories. Duluth, MN; Narragansett, RI; Corvallis, OR (Available at: http://yosemite.epa.gov/water/owrcCatalog.nsf).

Stribling, J.B., Kessup, K.J., White, J.S., 1998. Development of a benthic Index of Biotic Integrity for Maryland streams, Maryland Department of Natural resources. Monitoring and Non-Tidal Assessment Division. CBWP-EA-98-3, Annapolis, MD (Available at: http://www.dnr.state.md.us/streams/pubs/1998_benthic_ibi.pdf).

Ter Braak, C.J.F., 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67, 1167–1179.

Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., Altman, R.B., 2001. Missing value estimation methods for DNA microarrays. Bioinformatics 17, 520–525.

Turner, R.E., Rabalais, N.N., 2003. Linking landscape and water quality in the Mississippi river basin for 200 years. Bioscience 53, 563–572.

US Environmental Protection Agency, 2006. Lake and Reservoir Bioassessment and Biocriteria. Technical Guidance Document, Office of Wetlands, Oceans, and Watersheds, Office of Science and Technology, Office of Water, Washington, DC (Available at: http://www.epa.gov/owow/monitoring/tech/lakes.html).

Volstad, J.H., Roth, N.E., Mercurio, G., Southerland, M.T., Strebel, D.E., 2003. Using environmental stressor information to predict the ecological status of Maryland non-tidal streams as measured by biological indicators. Environmental Monitoring and Assessment 84, 219–242.

Wang, L., Lyons, J., Kanehl, P., Bannerman, R., 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28, 255–266.

Wright, J.F., Armitage, P.D., Furse, M.T., Moss, D., 1988. A new approach to the biological surveillance of river quality using macroinvertebrates. Verhangen International Verein Limnology 23, 1548–1552.

Yoder, C.O., Miltner, R.J., White, D., 2000. Using biological criteria to assess and classify urban streams and develop improved landscape indicators. In: Proceedings of the National Conference on Tools for Urban Water Resource Management and Protection, U.S. Environmental Protection Agency, Office Res. Dev., OH, EPA/625/R-00/001, Cincinnati, OH.

Yoder, C.O., Rankin, E.T., 1998. Using fish assemblages in a state biological assessment and criteria program: essential concepts and consideration. In: Smith, T.P. (Ed.), Assessing the Sustainability and Ecological Integrity of Fish Assemblages. Lewis Publishers, Boca Raton, FL, pp. 17–56.

Yoder, C.O., Rankin, E.T., Smith, M.A., Alsdorf, B.C., Altfater, D.J., Boucher, C.E., Miltner, R.J., Mishne, D.E., Sanders, R.E., Thoma, R.F., 2005. Changes in fish assemblage status in Ohio's nonwadeable rivers and streams over two decades. American Fisheries Society Symposium 45, 399–425.

Yuan, L.L., Norton, S.B., 2004. Assessing the relative severity of stressors at a watershed scale. Environmental Monitoring and Assessment 98, 323–349.

WATER RESEARCH 45 (2011) 2359–2374

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Prediction of biological integrity based on environmental similarity - Revealing the scale-dependant link between study area and top environmental predictors

David Bedoya [a,*], *Elias S. Manolakos* [b], *Vladimir Novotny* [a]

[a] *Civil & Environmental Engineering Department, Northeastern University, 400 Snell Engineering Center, 360 Huntington Avenue, Boston, MA 02115, USA*

[b] *Department of Informatics & Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis, Ilissia, Athens 15784, Greece*

ARTICLE INFO

Article history:
Received 20 July 2010
Received in revised form
10 January 2011
Accepted 11 January 2011
Available online 28 January 2011

Keywords:
Environmental stressors
Biological integrity
Geographic scale
Environmental similarity
Stressor and biological hierarchy

ABSTRACT

Indices of Biological integrity (IBI) are considered valid indicators of the overall health of a water body because the biological community is an endpoint within natural systems. However, prediction of biological integrity using information from multi-parameter environmental observations is a challenging problem due to the hierarchical organization of the natural environment, the existence of nonlinear inter-dependencies among variables as well as natural stochasticity and measurement noise. We present a method for predicting the Fish Index of Biological Integrity (IBI) using multiple environmental observations at the state-scale in Ohio. Instream (chemical and physical quality) and offstream parameters (regional and local upstream land uses, stream fragmentation, and point source density and intensity) are used for this purpose. The IBI predictions are obtained using the environmental site-similarity concept and following a simple to implement leave-one-out cross validation approach. An IBI prediction for a sampling site is calculated by averaging the observed IBI scores of observations clustered in the most similar branch of a dendrogram –a hierarchical clustering tree of environmental observations- built using the rest of the observations. The standardized Euclidean distance is used to assess dissimilarity between observations.

The constructed predictive model was able to explain 61% of the IBI variability statewide. Stream fragmentation and regional land use explained 60% of the variability; the remaining 1% was explained by instream habitat quality. Metrics related to local land use, water quality, and point source density and intensity did not improve the predictive model at the state-scale. The impact of local environmental conditions was evaluated by comparing local characteristics between well- and mispredicted sites. Significant differences in local land use patterns and upstream fragmentation density explained some of the model's over-predictions. Local land use conditions explained some of the model's IBI under-predictions at the state-scale since none of the variables within this group were included in the best final predictive model. Under-predicted sites also had higher levels of downstream fragmentation.

* *Corresponding author*. Tel.: +1617 314 7116; fax: +1 617 3147115.
E-mail addresses: dbedoya@coe.neu.edu (D. Bedoya), eliasm@di.uoa.gr (E.S. Manolakos), novotny@coe.neu.edu (V. Novotny).
0043-1354/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2011.01.007

The proposed variables ranking and predictive modeling methodology is very well suited for the analysis of hierarchical environments, such as natural fresh water systems, with many cross-correlated environmental variables. It is computationally efficient, can be fully automated, does not make any pre-conceived assumptions on the variables interdependency structure (such as linearity), and it is able to rank variables in a database and generate IBI predictions using only non-parametric easy to implement hierarchical clustering.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Integrity has been defined as the ability of a water body to maintain "a balanced, integrated, adaptive community of organisms having a species composition, diversity and functional organisms comparable to that of a natural biota of the region" (Karr et al., 1986). Biological integrity of streams is usually measured with some version of a calibrated index. One of the most widely used indices in the United States is the *Fish Index of Biological Integrity* developed by Karr et al. (1986). Many public agencies have adopted it as a framework for deriving their own calibrated version at the state or regional scale (Ohio EPA, 1987; Bode, 1988; Roth et al., 1998; Lyons et al., 2001; Lyons, 2006). Karr's IBI and subsequent versions and calibrations are based on a comparison of observed fish abundances and community composition against expected values in reference sites with similar environmental characteristics.

The importance of IBI lies in its sensitivity to disturbances of different nature because the biological community is an endpoint in the ecological river system (Karr et al., 1986; Novotny, 2003). However, the identification of major sources of biological degradation is challenging because the natural habitat is organized as a nested hierarchy of environmental filters with different geographic scales, to which the biological community has adapted (Pickett et al., 1989). Consequently, the geographic scale at which biological integrity is evaluated is of great importance because the stressors identified as most significant to the biological community are those at the highest level in the hierarchy of environmental filters at that particular scale (Poff, 1997). Therefore, measures to improve biological integrity need to be approached in a holistic scale-adaptive manner in order to be effective. The impact of stressors should be viewed within the context of disturbances occurring at larger scales than the study region (i.e. background quality).

The ecological hierarchy in the natural river system is composed of numerous instream and offstream environmental variables which are highly inter-twined and cross-correlated (Novotny et al., 2005). Therefore, changes in one of them will most likely have a cascade effect that may translate into changes in the instream conditions affecting the biological community. For example, land use changes in the watershed will affect, among other variables, sediment and nutrient input which will, in turn, affect physical and chemical instream water quality. If enough exposure of living organisms occurs, these will be negatively affected because they are the system's endpoint (Novotny et al., 2005).

The river system is organized a hierarchy of environmental characteristics and habitat conditions across multiple spatial scales (Frissell et al., 1986). An aquatic habitat is suitable for specific fauna when the different natural environmental filters at different spatial scales are overcome (Poff, 1997). Man-made modifications of any of these natural filters at any scale-level are stressors that will modify the pristine biological integrity of the site (Poff, 1997; Karr et al., 1986.) *Large-scale variables* -or *environmental gradients*- are those which produce a change in biological integrity of the system through their whole range of values within the study area (i.e. spatial scale of the study). These environmental parameters are usually the best integrity predictors at the selected scale. (Bedoya et al., in press; Lannert and Allan, 1999). On the other hand, small-scale variables have also an effect on particular sections of the area of study, but not on its entirety (Bedoya et al., in press; Lanmert and Allan, 1999). Therefore, identification of variables acting as gradients in a study area should be targeted as top priority for remediation purposes. Large-scale variables actually set the background biological integrity of a region and therefore, overall improvement of its biological integrity is always conditioned by them.

Because of the numerous cross-correlated variables potentially affecting IBI and the non-linear variable-to-IBI relationships, development of effective predictive modeling methodologies able to exploit a large number of multi-dimensional environmental observations is paramount. Moreover, new methods to predict biological integrity should not be constrained by any pre-imposed conditions. The methodology we present in this research meets these two key requirements. IBI prediction is performed with a two-phase approach. The first phase consists of ranking variables based on their overall impact on the biological community at the scale of our study area. The second phase consists of a step-wise IBI prediction using environmental variables from different categories (e.g. instream habitat variables). The best predicting variables from each group of variables are then progressively combined in order to obtain an overall improved IBI predictive model.

2. Methodology

2.1. Data and study area

The research reported here was based on 429 observations within the state of Ohio. This dataset was extracted from a larger database compiled by the Ohio Environmental Protection Agency (EPA) during its *Statewide Biological and*

Water Quality Monitoring and Assessment Program (Ohio EPA 2008). The original database made available to our research team consisted of 1848 observations out of which only 429 had information for all the environmental parameters (i.e. sites with no missing data). This research is based on these 429 complete observations. One observation corresponded to one or two visits to the corresponding site with a small time difference between visits (usually one to two weeks). During these site visits, grab samples for chemical water quality analyses were collected and an evaluation of habitat and biological qualities was performed. The data were collected between 1996 and 2000 by Ohio EPA. Most of the sites had at least two observations during this period, although some of them were evaluated just once. Most of the samples were collected in summer months (July through September) with very few (less than 20) in early October. By sampling in the same time period, potential IBI annual fluctuations were avoided to the maximum extent practicable. Sampling activities focused in summertime low flow periods when stress to aquatic biological communities is believed to be greatest (Ohio EPA, 2005). The distribution of the sampling sites across the state of Ohio is presented in Fig. 1.The state of Ohio follows a very systematic sampling strategy. Site selection within the watershed is driven by a stratification of the watershed based on a sequential, systematic halving of the drainage area, such that a census of all streams within the watershed down to a prescribed drainage area size are selected for sampling (Ohio EPA, 2005).

Biological -fish IBI scores- as well as instream environmental parameters -chemical and physical quality- were complemented with offstream parameters obtained with a Geographic Information System (GIS). To our knowledge, all data were collected in base-flow conditions and extreme events (e.g. a spill) were not reported at any sampling site.

For each observation site, biological integrity was measured using the fish IBI. In Ohio, this is a discrete score ranging from 12 (essentially no fish) to 60 (reference conditions). The IBI is calculated as the sum of 12 different metrics (each one an integer score ranging from 1 to 5) that describe the species richness and composition, the trophic composition, and the fish abundance and condition of the fish community (Karr et al., 1986; Ohio EPA, 1987). Instream variables consisted of water quality and habitat quality metrics (Table 2). Habitat parameters consisted of metrics from the Qualitative Habitat Evaluation Index (QHEI) (Rankin, 1989). The QHEI and its metrics are discrete scores with different ranges (see Table 2). The percentage of fine sediment in the river bed (embeddedness) was also available (this variable is not used as a QHEI metric itself, but as a penalizing factor for the QHEI's substrate and channel quality metrics).

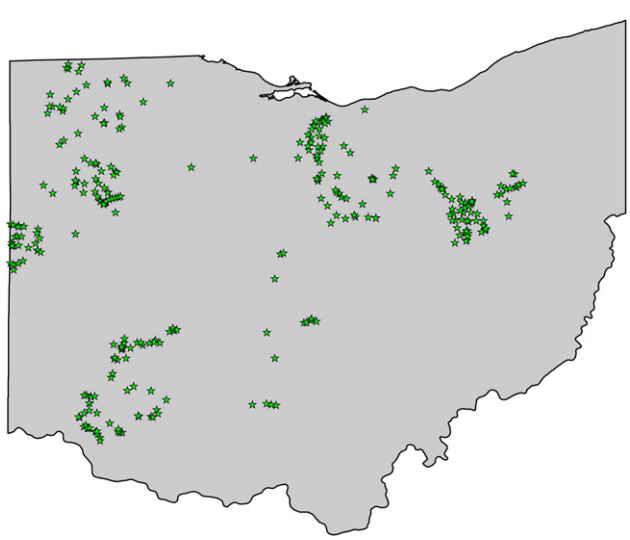


Fig. 1 – Distribution of sampling sites across the state of Ohio.

The offstream environmental variables were grouped into three main categories: upstream land use, stream fragmentation, and point source density and intensity. In order to calculate the upstream land uses, each site's watershed was delineated using a 30-m resolution Digital Elevation Map (DEM) with *ArcGIS Spatial Analyst*®. Subsequently, the percentages of different upstream land use was calculated at two different scales: the regional scale, which included the whole upstream contributing catchment, and the local scale, which included only 2 miles upstream from the sampling site. Land use percentages were calculated for the whole upstream area as well as the 100- and 30-m buffers at both scales. These two buffer widths were selected based on literature values. A buffer width of 30 m is considered the minimum necessary to provide some benefit to the receiving water body such as temperature amelioration (Castelle et al., 1994). Moreover, 30 m was the maximum resolution of the DEM. A buffer with of 100 m was selected because this distance is considered sufficient to perform basic functions such as sediment removal, nutrient removal, and preservation of species diversity (Castelle et al., 1994). Land use percentages were calculated using the *Thematic Raster Summary* function within *Hawth's Analysis Tools for ArcGIS* (Beyer, 2004). Land cover categories as defined in the 2001 National Land Cover Dataset (NLCD) were used (USGS, 2008b) and listed in Table 1. The Open Water (OW) land use category was only calculated for the regional- and local-scale whole catchment areas, not for the buffers. Drainage areas (DA) for each site were also calculated. The fragmentation and point source metrics (Table 2) were calculated using information from the National Hydrography Datasets (NHD) (USGS, 2008a). The *ArcGIS Utility Network Analyst* was used to trace upstream or downstream a specific site. Major dams (i.e. with DA $\geq$ 2.59 Km^2) and point sources (major and minor waste water treatment plants and major industrial dischargers) were obtained from the National Inventory of Dams (USACE, 2005) and the Permit Compliance System database (USEPA, 2008) respectively. In the fragmentation metrics, downstream metrics such as downstream dam frequency in the main channel (DW_MainDf, see Table 2) or upstream metrics such as upstream dam frequency (U_Df, see Table 2) were considered indicators of site accessibility or habitat continuity from downstream or upstream points respectively. Metrics that combined upstream and downstream segments such as average dam frequency (Avg_Df, see Table 2) were considered indicators of habitat size.

Sites were not segregated a priori based on ecoregions or stream size because the model should be able to separate clusters of sites with significant environmental differences. In other words, we wanted to "let the data speak".

Table 1 – Description, percentage quartiles, and individual IBI predicting power for the different NLCD land use categories in the Ohio database. Hay = hay/pasture; ForD = deciduous forest; ForM = mixed forest; ForE = evergreen forest; Shr = shrub/scrub; WetH = herbaceous wetlands; WetW = woody wetlands; Herb = herbaceous; Crop = crops; Bar = barren; DevH = high intensity urban; DevM = medium intensity urban; DevL = low intensity urban; DevO = open urban space; OW = open water; Oth = other land uses.

Regional Drainage Area (RDA)			Regional 100-Meter Buffer (R100)			Regional 30-Meter Buffer (R30)		
Name	Quartiles	R^2	Name	Quartiles	R^2	Name	Quartiles	R^2
Regional Land Use in Contributing Area			Regional Land Use in 100 m Buffer			Regional Land Use in 30 m Buffer		
RDA_Hay	3.16–7.67–14.60	0.385[a]	R100_Hay	2.97–8.20–13.60	0.322[a]	R30_ForD	9.82–20.90–38.01	0.368[a]
RDA_ForM	0.00–0.00–0.03	0.294[b]	R100_DevH	0.00–0.13–0.38	0.320[a]	R30_Hay	3.13–7.35–12.57	0.312[a]
RDA_DevL	1.22–2.39–6.06	0.293[a]	R100_DevO	5.26–6.84–10.33	0.300[a]	R30_Herb	0.23–1.00 -2.08	0.312[b]
RDA_OW	0.10–0.25–0.60	0.292[a]	R100_Herb	0.29–1.00–1.81	0.296[a]	R30_Crop	30.69–50.23–64.82	0.304[a]
RDA_ForD	5.29–9.45–18.60	0.281[a]	R100_ForD	8.22–16.06–30.56	0.292[a]	R30_DevM	0.07–0.26 -0.62	0.259[a]
RDA_WetH	0.00–0.02–0.09	0.261[b]	R100_ForE	0.00–0.09–0.24	0.285[a]	R30_DevL	0.79–1.51–3.48	0.245[a]
RDA_Crop	40.49–60.84–75.58	0.257[a]	R100_DevM	0.10 -0.40 -0.87	0.275[a]	R30_WetW	0.00–0.26–0.95	0.242[a]
RDA_DevM	0.24–0.64–1.64	0.253[a]	R100_DevL	0.94–1.85–3.79	0.274[a]	R30_DevO	4.44–6.26–10.01	0.234[a]
RDA_ForE	0.01–0.08–0.25	0.243[a]	R100_WetW	0.00–0.13–0.58	0.270[a]	R30_ForM	0.00–0.00–0.04	0.232[a]
RDA_Bar	0.00–0.01–0.05	0.239[a]	R100_ForM	0.00–0.00–0.05	0.261[a]	R30_ForE	0.00 -0.05 -0.20	0.225[a]
RDA_Herb	0.35–0.89–1.38	0.225[a]	R100_WetH	0.00–0.03–0.20	0.236[b]	R30_Shr	0.00–0.00–0.02	0.208[a]
RDA_DevH	0.07–0.26–0.75	0.223[a]	R100_Crop	34.65–54.32–70.27	0.231[a]	R30_DevH	0.00–0.07–0.20	0.198[a]
RDA_WetW	0.00–0.04–0.20	0.223[a]	R100_Shr	0.00–0.00–0.05	0.223[a]	R30_Bar	0.00–0.00–0.01	0.182[a]
RDA_Shr	0.00–0.00–0.03	0.221[a]	R100_Bar	0.00–0.00–0.02	0.184[b]	R30_WetH	0.00–0.01–0.36	0.172[b]
RDA_DevO	5.23–6.25–9.76	0.212[a]	R100_Oth	0.00–0.00–0.00	0.012[a]	R30_Oth	0.00–0.00–0.00	0.012[a]
RDA_Oth	0.00–0.00–0.00	0.012[a]						
Local Drainage Area (LDA)			Local 100-Meter Buffer (L100)			Local 30-Meter Buffer (L30)		
Local Land Use in Contributing Area			Local Land Use in 100 m Buffer			Local Land Use in 30 m Buffer		
LDA_DevO	4.97–7.22–13.73	0.289[a]	L100_ForD	7.39–24.43–46.14	0.335[a]	L30_ForD	7.33–29.34–53.65	0.334[a]
LDA_ForD	4.47–13.59–28.73	0.285[b]	L100_DevL	0.30–2.06–6.99	0.272[b]	L30_Crop	6.33–25.24–53.97	0.202[a]
LDA_Hay	0.00 -5.51 - 12.92	0.214[a]	L100_DevO	4.17–7.80–14.68	0.208[a]	L30_DevL	0.00–1.20–6.59	0.186[a]
LDA_OW	0.00–0.19–0.96	0.200[b]	L100_Hay	0.00–3.97–10.85	0.196[a]	L30_Hay	0.00–1.67–9.45	0.161[a]
LDA_DevL	0.42–2.42–11.19	0.183[a]	L100_Crop	8.55–30.06–59.61	0.190[b]	L30_DevO	3.00-6.42-15.23	0.158[a]
LDA_DevM	0.00–0.27–2.01	0.159[a]	L100_Herb	0.00–0.34–1.84	0.117[b]	L30_Herb	0.00–0.00–1.61	0.154[b]
LDA_Crop	17.20 -44.71 -69.59	0.152[a]	L100_WetW	0.00–0.00–1.28	0.113[a]	L30_DevM	0.00–0.00–1.12	0.100[a]
LDA_DevH	0.00–0.00–0.80	0.140[b]	L100_DevM	0.00–0.00–1.78	0.107[b]	L30_WetW	0.00–0.00–2.24	0.092[a]
LDA_Herb	0.00–0.65–1.68	0.130[a]	L100_Shr	0.00–0.00–0.00	0.069[c]	L30_WetH	0.00–0.00–2.24	0.087[a]
LDA_WetW	0.00–0.00–0.47	0.128[a]	L100_ForE	0.00–0.00–0.15	0.064[a]	L30_DevH	0.00–0.00–0.00	0.071[c]
LDA_WetH	0.00–0.00–0.19	0.124[b]	L100_WetH	0.00–0.00–0.63	0.051[a]	L30_ForE	0.00–0.00–0.00	0.066[a]
LDA_Shr	0.00–0.00–0.00	0.117[a]	L100_DevH	0.00–0.00–0.16	0.048[b]	L30_ForM	0.00–0.00–0.00	0.036[c]
LDA_ForE	0.00–0.00–0.25	0.098[a]	L100_ForM	0.00–0.00–0.00	0.042[a]	L30_Shr	0.00–0.00–0.00	0.032[c]
LDA_ForM	0.00–0.00–0.00	0.077[a]	L100_Bar	0.00–0.00–0.00	0.005[b]	L30_Bar	0.00–0.00–0.00	0.004[b]
LDA_Bar	0.00–0.00–0.00	0.020[a]	L100_Oth	0.00–0.00–0.00	0.000	L30_Oth	0.00–0.00–0.00	0.000
LDA_Oth	0.00–0.00–0.00	0.000						

a a = best prediction at 423 branches.
b b = best prediction at 328 branches.
c c = best prediction at 233 branches.

2.2. *Environmental variables ranking*

Environmental variables are divided into two categories: offstream and instream. The offstream category is composed of four groups: local and regional land use -in the whole upstream area and the 30- and 100-m buffers-, stream fragmentation, and point source density and intensity. The instream category is composed of two groups: water and habitat qualities.

The individual predictive power of each environmental variable is initially estimated by obtaining the coefficient of determination [r^2] of the observed (measured) IBI versus a calculated IBI prediction values, generated using a leave-one-out cross-validation approach detailed below.

Let us assume that we are given a database of measurements represented as a matrix X of *n* observations (rows) by *m* environmental variables (columns). One element of that matrix, i.e. one observation of variable v, (let's call it Xi, v without loss of generality) is isolated and will be called the *test* or *query site*. The remaining *n-1* observations of the same variable, namely the measurement ($[X1, v, \ldots Xn-1, v]$) in the same column of the data matrix, are organized in a "dendrogram" tree structure having $[X1, v, \ldots Xn-1, v]$ as leaves. This is done by applying agglomerative Hierarchical Clustering (HC) using the average linkage method and the standardized Euclidean as the distance metric (Jain et al., 1999). When the resulting dendrogram is "cut" at a certain distance from the root (more details on how the cut level is decided are given in the last paragraph of this section) several tree branches are emanating from the cut. Among them we identify the *Most Similar Branch* to the test site *i* when using variable v (to be called $MSB_{i,v}$) as the branch (overall branches B_k defined by the cut) for which the standardized Euclidean distance between

Table 2 – Description and individual IBI predicting power for the water quality, habitat, point source, and stream fragmentation metrics.

Name	Units	Description	R^2	Name	Units	Description	R^2
Water Quality Parameters				Habitat Parameters			
BOD	mg/L	Biological Oxygen. Demand	0.12[b]	Embed	0–4	Embeddedness	0.28[a]
TKN	mg/L	Total Kjeldahl Nitrogen	0.09[c]	Riffle	0–8	Riffle and run quality	0.24[a]
As	μg/L	Total Arsenic	0.08[c]	Subs	0–20	Substrate quality	0.23[a]
NH_4	mg/L	Ammonia as Nitrogen	0.07[b]	Pool	0–12	Pool and glide quality	0.23[a]
NO_2	mg/L	Nitrite as Nitrogen	0.06[b]	Chan	0–20	Channel morphology score	0.21[a]
Mg	mg/L	Total Magnesium	0.04[b]	DA	Km^2	Drainage area	0.19[c]
SO_4	mg/L	Total Sulfate	0.03[a]	Rip	0–10	Riparian and bank qualities	0.17[b]
Cond	μmho/cm	Conductivity	0.02[a]	Cover	0–20	Instream vegetal cover	0.13[a]
DO	mg/L	Dissolved Oxygen	0.02[c]	Grad	0–10	Gradient score	0.09[c]
Hard	mg/L	Hardness (as $CaCO_3$)	0.02[a]	*Fragmentation Parameters*			
Cl	mg/L	Total Chloride	0.02[a]	SITE_Con*	Fraction	Total connected length from observation site/ Total basin network length	0.47[a]
pH	S.U.	pH	0.01[b]	Dfl_MainLen*	m^2/km	Downstream flooded area/Downstream main channel length	0.46[a]
TSS	mg/L	Total Suspended Solids	0.01[b]	Dsto_MLen*	m^3/km	Downstream dam storage/Main channel length	0.44[b]
NO_3	mg/L	Nitrate as Nitrogen	0.01[c]	DW_MainDf*	Km	Main channel downstream length/Number of downstream dams	038[a]
Ca	mg/L	Total Calcium	0.01[b]	Avg_Df*	Km	Mean value between DW_MainDf and U_Df	0.26[a]
Cd	μg/L	Total Cadmium	0.01[a]	U_Df	Km	Upstream network length/Number of upstream dams	0.25[a]
Cu	μg/L	Total Copper	0.01[c]	UPS_Con	Fraction	Upstream connected length/Total upstream length	0.19[a]
Fe	μg/L	Total Iron	0.01[b]	Uflood_len	m^2/km	Upstream flooded area/Upstream network length	0.17[a]
Zn	μg/L	Total Zinc	0.01[b]	UPS_Flooded	Fraction	Upstream flooded area/Drainage area	0.17[a]
TP	mg/L	Total Phosphorus	0.00[a]	UPS_stor_len	m^3/km	Upstream dam storage/Upstream network length	0.16[a]
Pb	μg/L	Total Lead	0.00[a]	UPS_stor_DA	m^3/Km^2	Upstream dam storage/Drainage area	0.15[a]
				Point Source Parameters			
				PS_LTOT	No./km	Number of upstream point sources/Upstream network length	0.26[b]
				PSDisch_LPS	m^3/d/Km	Upstream point source discharge flow/Distance from site to all upstream point sources	0.21[b]
				PS_LPS	No./km	Number of upstream point sources/Distance to all upstream point sources	0.21[a]
				PSDisch_DA	$m^3/d/Km^2$	Upstream point source discharge flow/Drainage area	0.21[b]
				PSDisch_LT	m^3/d/Km	Upstream point source discharge flow/Upstream network length	0.20[b]
				Flow_PS	%	% of upstream network carrying waste water	0.20[b]
				LPS-DA	Km/Km^2	Distance to all upstream point sources/Drainage area	0.18[b]

*Downstream parameters calculated up to the basin outlet All distances were calculated following stream network channels.
a a = best prediction at 423 branches.
b b = best prediction at 328 branches.
c c = best prediction at 233 branches.

the test site's value (Xi, v) and the mean value of variable v over the branch leaves is minimized. See Eqs. (1) and (2) below for the formal definition,

$$MSB_{i,v} = \arg_k \left\{ \min \left(D^k_{i,v} \right) \right\} \tag{1}$$

$$D^k_{i,v} = \begin{matrix} dist[Xi, v, \quad average(Xj, v)] \\ j \in B_k \end{matrix}, \tag{2}$$

where$D^k_{i,v}$ is the standardized Euclidean distance between the test-site's value $X_{i,v}$ and the average $\{X_{j,v}\}$ over the sites residing at the leaves of branch B_k.

Note that depending on where the dendrogram is cut, a resulting branch may contain one or more observations. As "calculated IBI" prediction for site i based on variable v information we will use the mean IBI value of the observations clustered in its corresponding Most Similar Branch (see Eq. (3))

$$IBI^C_{i,v} = \begin{matrix} average \left(IBI^O_j \right) \\ j \in MSB(i, v) \end{matrix}, \tag{3}$$

where "observed IBI" (IBI^O_j) is the measured IBI value for each site j of branch $MSB_{i,v}$ recorded in the database.

The same procedure is repeated for every site i (keeping v fixed) leading to an $n \times 1$ column vector of predicted IBIs for all sites based on information for variable v alone, called IBI^C_v. Then the same procedure is repeated for every variable v giving rise to an $n \times m$ matrix of IBI predictions,IBI^c, having the same structure as the data matrix X.

Finally three different dendrogram tree cuts are applied (leading to 233, 328, and 423 branches respectively) to obtain

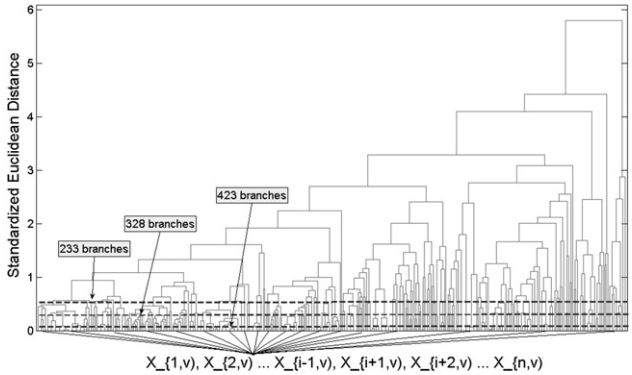


Fig. 2 – Example of a dendrogram built using an array of n-1 observations composed of v environmental variables (v is a one- or possibly multi-dimensional vector of selected environmental variables [v1,v2…,vm]). Test-sites (X1,v,X2,v,…Xi-1,v,Xi + 1,v,…Xn,v) correspond to the leaves of the tree. Three cuts (dashed lines) are determined (see text for details). Each cut generates branches (test site clusters). The Most Similar Branch (MSB) to the test site Xi,v, is determined (see text for details). The average measured IBI of the sites (leaves) belonging to the MSB is used as the predicted IBI for the test site. The same procedure is repeated for each test site and then for each cut. The predicted IBI values are compared to the measured IBI for all sites. The predictive value of variable v is assessed based on the |r2| of the fit of best predictive model (among the three models corresponding to the three different cuts).

three different $n \times 1$ vectors of IBI predictions based on each variable v. Note that the distance from the root for each cut has been selected so that the number of resulting branches is approximately equal to 50%, 75%, or 100% of the total number of available complete observations in the database (see example in Fig. 2). Among the three vectors the one with maximal similarity (in terms of [r^2]) to the observed (measured) IBI_v^O vector is considered as the "best" predictor of IBI using information of variable v alone, and this [r^2] value is assigned as the score of variable v in the rank ordering of the environmental variables.

2.3. *Step-wise IBI prediction*

This step consists of obtaining progressively improved IBI predictions by combining variables from each group separately (see groups in Fig. 3). The "best" variables from each group are combined to find the "best" offstream and instream predictors following the order of variables specified in Fig. 3. Finally, the subset of best offstream and instream predictors are also combined in a similar manner to obtain the overall best set of predictors.

The IBI prediction methodology remains the same as in step 2.2. However, in this case, a step-wise approach is followed. For each group, and following the group's variable ranking obtained in Section 2.2, the best predicting variable is first selected (let's call it v_1 w.l.o.g.). Subsequently, the second best predicting variable in each group (to be called v_2) is also selected to form an array of two-dimensional environmental

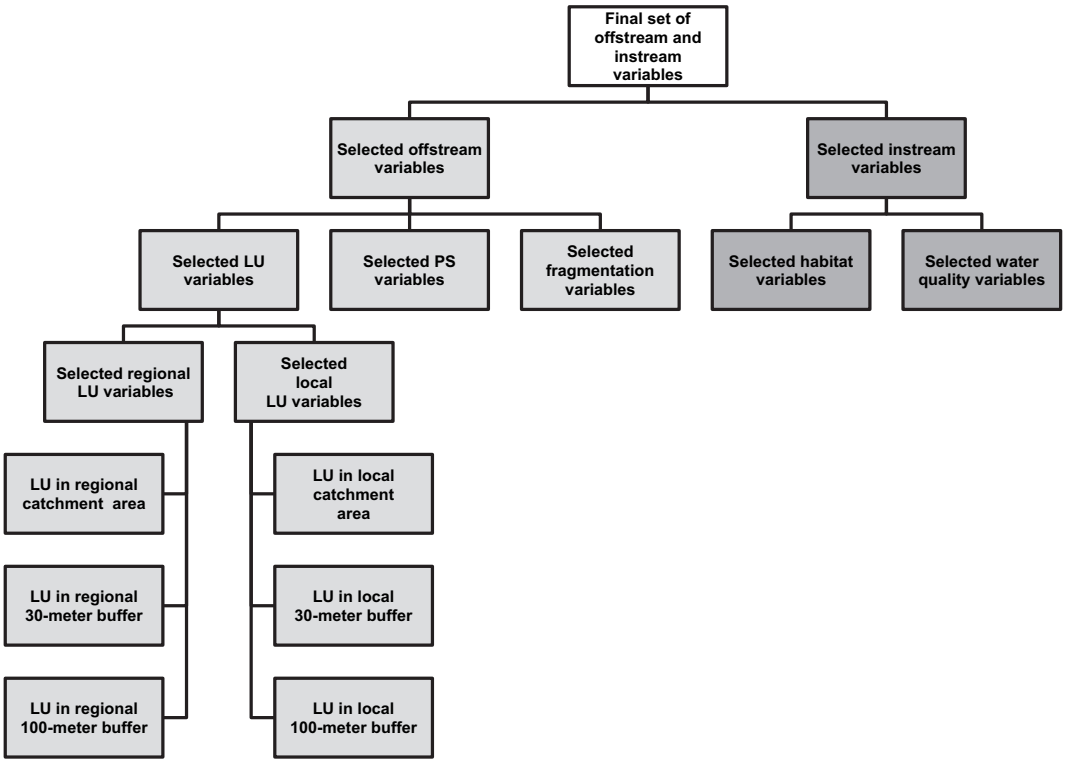


Fig. 3 – Diagram showing the order in which the groups of variables are combined. Dark grey rectangles indicate instream variables. Light grey rectangles indicate offstream variables. White rectangles indicate final model with a mix of both types of variables.

vectors ($\{X_{1,(v1,v2)},X_{2,(v1,v2)},\ldots,X_{n,(v1,v2)}\}$). One at a time, a two-dimensional *test site* is isolated from the rest *n-1* two-dimensional observations (exactly as we did in Section 2.2). The test-site vector is presented again to the branches of the dendrogram built using the remaining *n-1* two-dimensional vectors. The average IBI of the observations clustered in the most similar branch is selected as the test-site's calculated IBI Eq. (3), but now with $v = (v1,v2)$. The selected dendrogram's branch is the one that minimizes the standardized Euclidean distance between the *test-site* vector and the average environmental vector calculated from the observations located in the particular branch (Equations (1) and (2) but now with $v = (v1,v2)$). If the two-variable model improves the previous prediction (increase in R^2) at any one of the three selected cut levels of the dendrogram (i.e. using 233, 328, or 423 branches), the new variable is retained otherwise it is discarded.

Subsequently, the third best predicting variable in each group (v_3) is introduced to form an array of two ($\{X_{1,(v1,v3)},X_{2,(v1,v3)},\ldots,X_{n,(v1,v3)}\}$) or three ($\{X_{1,(v1,v2,v3)},X_{2,(v1,v2,v3)},\ldots,X_{n,(v1,v2,v3)}\}$) dimensional observations (depending on the inclusion or exclusion of v_2 in the previous step). Again, one at a time, a test site is separated from the rest of *n-1* observations and associated to the most similar branch of the dendrogram calculated with the variables used in this step. Improvement in the IBI prediction relatively to the previously tested model results in the inclusion of the last variable in the best set so far, or exclusion otherwise. This procedure is repeated for each group until all variables have been considered for inclusion. At the end of this "greedy" procedure, the "best" combination of predicting variables from a particular group of variables is identified. Although this method does not guarantee to find the globally optimal IBI predictive model it does move step by step toward a model with improved performance as new variables are introduced and it is quite fast to implement.

In this research, strongly cross-correlated variables were not eliminated because the model's performance is not adversely affected when more variables are introduced. Since prediction with the environmental similarity concept is merely based on comparing site environmental vectors with the same vector elements, presence of cross-correlated variables will not affect the performance because the same variables are used for all prediction sites. Therefore, even marginal improvements can be accounted for without jeopardizing model performance. Furthermore, since variables are examined for inclusion in a sequential manner, keeping all variables "in the game" has also the advantage of not retiring prematurely a variable that although is highly correlated to a variable already added to the model may have a possible dependence to a third variable not yet examined for inclusion.

Subsequently, the different groups of predictors are also progressively "merged" using only the "best" variables for each individual group resulted in the previous step. The stepwise IBI prediction methodology used when two groups of variables are combined is identical as before. Fig. 3 shows the order in which the groups of variables are merged.

2.4. *Impact from local stressors*

Mispredicted observations were isolated and tested for statistically significant differences against the group of well-predicted sites. A site was labeled as mispredicted if the calculated IBI fell beyond the $\pm 1.5 \times$ RMSE interval (where RMSE is the Root Mean Square Error of the IBI predictions for all the available observations in the dataset).

Significant differences in water quality (for those sites affected by point sources), point source and fragmentation density and intensity, as well as local and regional land uses were tested using a Student *t-test* at the 95% confidence level.

3. Results

3.1. *IBI predictions with offstream variables*

3.1.1. *Land use*

The top seven dominant land uses in our database were (in decreasing order of median percentage in the watershed [Table 1]): cropland (60.84%), deciduous forest (9.45%), hay/pasture lands (7.67%), urban open space (6.25%), low intensity urban space (2.39%), herbaceous lands (0.89%), and medium intensity urban space (0.64%). All the remaining land uses had a median extent in the watershed smaller than 0.5%.

The local land use sub-model (Local LU model in Fig. 4) was able to account for 49% of the total IBI variability. Results seemed to indicate that proximity to the stream is important because most of the selected variables in the group model were land uses within the buffer zones instead of the whole catchment area.

The regional land use model (Regional LU model in Fig. 4) explained 58% of the total IBI variability. In this case, selected land uses alternated between percentages in the whole catchment and in the buffers. The first selected model variable -percentage of hay/pasture in the drainage area- had also the highest individual IBI prediction power of all local or regional land uses (Table 1). Approximately 95% of the group variability was explained with the top three group variables: hay/pasture in the drainage area and deciduous forest and urban open space in the 30- and 100-m buffers respectively. The other six model variables only accounted for the remaining 5% of the group's variability.

The subsequent merger of the regional and local land use models yielded almost identical results as the regional land use model with the exception of the last two variables (see Fig. 4). The overall land use model –All LU model in Fig. 4- accounted for 60% of the total IBI variability. The two selected local land uses (medium intensity urban lands within the local 30- and 100-m buffers) introduced marginal improvement (1.2% of the group's variability).

3.1.2. *Point source density and intensity*

This sub-model –Point Sources sub-model in Fig. 4- accounted for the smallest IBI variability of all the offstream sub-models. Upstream point source intensity (PS_LTOT) explained 26% of the overall variability and was the first and only metric selected in the step-wise algorithm.

3.1.3. *Stream fragmentation*

Fragmentation density and intensity metrics explained 54% of the overall variability. River network connectivity at the basin scale (SITE_Con) explained 47% of the overall IBI variability (87% of the sub-model's variability). This variable had the

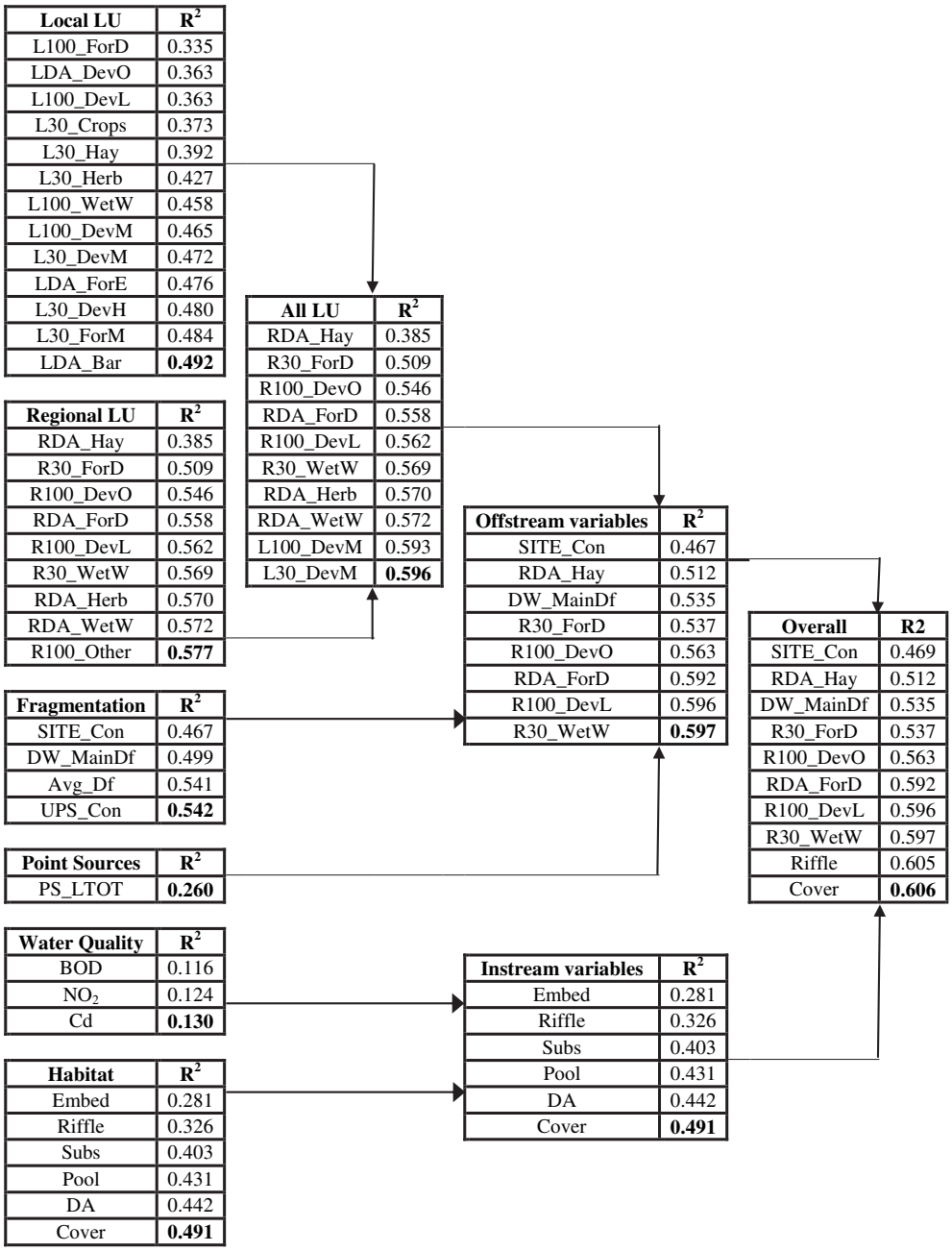


Fig. 4 – Step-wise IBI predictions. R^2 indicates the variability explained after adding a new variable to the model. All results were achieved using a hierarchical tree with 423 branches.

greatest IBI prediction capability of an individual variable overall. Downstream dam frequency, average dam frequency, and percentage of upstream connected network were other selected variables and accounted for the remaining 13% of the group's variability.

3.1.4. *Combination of best offstream variables*

When the best offstream variables were combined (Fig. 3) only fragmentation and regional land use variables were selected (Offstream variables sub-model in Fig. 4). None of the local land uses or point source variables were selected. The best prediction of this sub-model model marginally improved predictions by the overall land use model, accounting for 60% of the total variability. This model used only eight variables instead of ten in the land use model.

3.2. *Predictions with instream variables*

3.2.1. *Instream habitat variables*

The instream habitat sub-model (which included drainage area) explained 49% of the overall IBI variability. Six variables were selected (Habitat sub-model in Fig. 4). The top four predictors, which accounted for 88% of the group's variability, were directly or indirectly related to habitat's substrate

quality (i.e. embeddedness and substrate quality) or habitat variability (riffle and pool qualities). Drainage area (which was positively correlated to IBI) and instream cover explained the remaining IBI variability in the group.

3.2.2. *Water quality variables*

The water quality variables group was clearly sorted in three main clusters. The first one was related to nutrient concentration, especially nitrogen (BOD, TKN, NO_2, and NH_4) which had the group's highest prediction powers. Nitrate (NO_3) and TP concentrations were not ranked among the top chemical IBI predictors, being TP the poorest predictor in the group. The IBI prediction with water quality parameters was the poorest of all. The group's top two variables (BOD, NO_2) were related to nutrient loading and explained 95% of the group's variability. The remaining 5% was explained by cadmium concentrations.

3.2.3. *Combination of best instream variables*

The final *"Instream variables"* model (Fig. 4) yielded the exact same results as the *"Habitat"* model. Therefore, the *"Water Quality"* model did not bring any new valuable information beyond the habitat variables.

3.3. *Final predictions*

The final model –Overall model in Fig. 4- was composed of all the selected offstream variables and only two instream habitat parameters: riffle quality and instream vegetal cover. These two variables only accounted for 1.5% of the group's variability. The final model accounted for 61% of the overall IBI variability, which was a very modest improvement from the *"Offstream variables"* model. IBI prediction plots for the *"Offstream variables"*, *"Instream variables"* and *"Overall"* models are presented in Fig. 5.

3.4. *Local environmental stressors*

In the final *"Overall"* model, a total of 28 sites were above the 1.5 × RMSE threshold, while 27 were below it (see Fig. 5). Among the over-predicted observations, two sites had either extremely high concentrations of copper and zinc or point source density. Since the influence of these two sites in the performance of the *t-tests* was evident, they were removed. The biological quality of these two sites was mostly set by their extremely degraded water quality. After removing sites with outlier local conditions, significant differences among over- and well-predicted observations were identified in the upstream river fragmentation as well as the land use-related metrics (Table 3). Over-predicted sites had more severe upstream fragmentation but also better land use at the regional scale, which was a likely cause of over-prediction. Local land use results were mixed. Over-predicted sites had larger percentages of forested areas in both, the local catchment and buffer areas but they also had larger percentages of hay and pasture lands in the local catchment area. Presence of hay pasture lands in the drainage area was identified by our model as the best IBI predictor and it is negatively correlated to IBI (Table 1).

On the other hand, under-predicted sites (i.e. calculated IBI < observed IBI - 1.5 × RMSE) had consistently significantly lower hardness and hardness-related parameter values in sites with upstream point sources, while sulfate concentration was higher. Under-predicted sites had better land use quality at the local and regional levels (more forested areas,

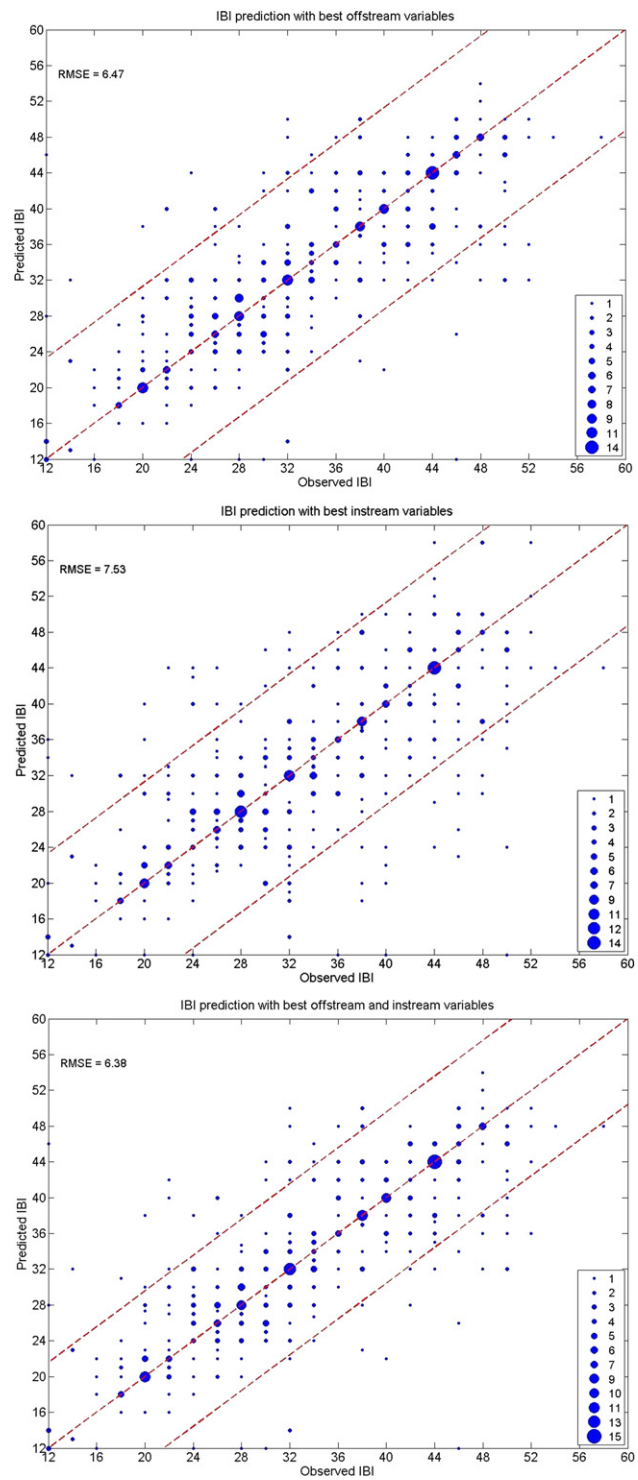


Fig. 5 – IBI predictions with the best offstream variables (top), best instream variables (middle), and best variables overall (bottom). Dashed lines indicate perfect fit line (center) and ±1.5 × RMSE (sides). Dot size is proportional to the number of hits that is indicated in the legend.

Table 3 – List of variables with significant differences between over-predicted sites and sites with an IBI prediction within the ±1.5 × RMSE range.

Variable Name	# of over-predicted sites/# of well-predicted sites	Type of sites	Value in over-predicted sites (95% conf. interval)	Value in well-predicted sites (95% conf. interval)	p
Uflood_len	11/19	NPS + UF	14.2 ± 8.8	2.2 ± 1.3	0.000
UPS_Con	11/19	NPS + UF	40.6 ± 25.7	75.6 ± 10.2	0.003
UPS_stor_len	11/19	NPS + UF	142.5 ± 82.2	17.2 ± 13.4	0.000
Ups_stor_DA	11/19	NPS + UF	0.115 ± 0.078	0.021 ± 0.018	0.003
UPS_Con	28/374	UF	76.6 ± 14.6	89.4 ± 2.5	0.011
L30M_ForD	28/374	ALL	44.2 ± 9.1	30.6 ± 2.72	0.009
L100M_ForD	28/374	ALL	39.8 ± 8.0	26.4 ± 2.4	0.003
L100M_DevM	28/374	ALL	0.24 ± 0.17	2.5 ± 0.58	0.041
LDA_ForD	28/374	ALL	26.5 ± 7.1	17.8 ± 1.8	0.012
LDA_ForE	28/374	ALL	1.1 ± 1.2	0.4 ± 0.1	0.014
LDA_Hay	28/374	ALL	12.6 ± 4.8	8.2 ± 1.1	0.034
R30M_ForD	28/374	ALL	36.5 ± 7.5	23.2 ± 1.8	0.000
R100M_ForD	28/374	ALL	29.0 ± 5.7	19.3 ± 1.5	0.001
RDA_ForD	28/374	ALL	18.1 ± 4.6	12.9 ± 1.1	0.019

NPS = sites without point sources; UF = sites with upstream fragmentation; ALL = all sites.

less urban development and less crop lands). However, the density and severity of impoundments in the downstream section was greater (Table 4). Downstream fragmentation had great impact on IBI in the final model (Fig. 4).

4. Discussion

This methodology proved to be very versatile and time-efficient when large, multi-parameter, environmental vectors are used for prediction of a target variable. The major advantageous difference with respect to more traditional approaches lies on the fact that the presented approach is able to allow easy, unbiased assessment of large, multi-dimensional vectors composed of data of very different nature and measurement ranges such as concentrations of chemical compounds or discrete scores in the case of habitat quality. Because all environmental parameters are standardized prior to perform any IBI predictions, large site environmental vectors composed of parameters of very different nature and measurement range can be compared. Because all vector components are standardized a priori, each of them carries the same weight in the IBI prediction. Another big advantage over some commonly used, traditional prediction techniques such as regression is that model performance and speed is not affected by presence of highly cross-correlated variables since prediction is obtained by a mere comparison of site environmental parameters. Highly correlated variables do not affect

Table 4 – List of variables with significant differences between under-predicted sites and observations with a prediction within the ±1.5 × RMSE range.

Variable Name	# of under/ well-predicted sites	Type of sites	Value in under-predicted sites (95% conf. interval)	Value in well-predicted sites (95% conf. interval)	p
Hard	11/213	PS	247.0 ± 42.1	313.8 ± 13.8	0.033
Mg	11/213	PS	20.9 ± 4.7	28.2 ± 1.6	0.046
SO_4	11/213	PS	135.2 ± 15.5	64.1 ± 25.1	0.042
Dsto_MLen	22/331	DF	1920.8 ± 711.9	1194.8 ± 157.5	0.025
L100_ForD	27/374	ALL	39.5 ± 8.5	26.4 ± 2.4	0.005
L100_ForE	27/374	ALL	1.8 ± 2.0	0.4 ± 0.2	0.002
L30_ForD	27/374	ALL	47.4 ± 9.9	30.6 ± 2.7	0.002
L30_DevL	27/374	ALL	2.0 ± 1.7	6.0 ± 1.0	0.038
L30_ForE	27/374	ALL	3.8 ± 5.8	0.4 ± 0.2	0.000
LDA_ForD	27/374	ALL	29.9 ± 7.8	17.8 ± 1.8	0.000
LDA_ForE	27/374	ALL	1.4 ± 1.3	0.4 ± 0.1	0.000
R30_Crop	27/374	ALL	38.1 ± 1.3	49.1 ± 2.5	0.025
R30_ForD	27/374	ALL	36.3 ± 7.2	23.2 ± 1.8	0.000
R100M_ForD	27/374	ALL	29.4 ± 5.9	19.3 ± 1.5	0.000
RDA_ForD	27/374	ALL	20.0 ± 4.7	12.9 ± 1.1	0.002

DF = sites with downstream fragmentation; PS = sites with point sources; ALL = all site.

the outcome of the distance function because each parameter in the query site environmental vector is compared against the same exact environmental parameter in the rest of observations.

Even though this methodology was not designed to "ordinate" environmental observations, the hierarchical tree used for prediction could be seen as a clustered distribution of similar sites, which is able to depict both explanatory (e.g. land use, water quality) and response variables (i.e. IBI) simultaneously. Traditional techniques such as Polar Ordination (PO) or Non-Metric Multidimensional Scaling (NMDS) do not allow a simultaneous display of both explanatory variables (stressors) and response variables (IBI) on the same two-dimensional grid. Other widely used traditional ordination techniques such as Correspondence Analysis (CoA) allow a simultaneous display of variables as well (Giraudel and Lek, 2001).

The site distribution using the presented methodology is obtained with no a priori relationships between the explanatory and response variables. Some multivariate ordination techniques such as Principal Component Analysis (PCA) or CoA assume linear relationships between the explanatory and the response variables which may not hold true in many cases, leading to well-known problems such as the horseshoe effect (PCA) or the arc effect (CoA) (Giraudel and Lek, 2001).

The model confirmed biological integrity is the result of many inter-twined stressors of different nature acting at different scales. Out of the five main components of biological integrity (energy sources, water quality, habitat structure, flow regime, and biotic interactions) (Karr and Kerans, 1981; Karr et al., 1986; Karr, 1991), the first four were partially or fully represented in our database.

At the study scale, only two groups of stressors were necessary to approximate the best variable combination for IBI prediction: regional land use and stream fragmentation at the basin-level. The relevance of these variables for IBI prediction was consistent with the geographic scale of the study, which had many sites scattered through a wide range of watersheds and within multiple basins. The relevance of the sampling strategy and geographic scale of the study area is paramount (Allan et al., 1997). At a specific scale, relevant variables in the highest possible level of the stressor-response hierarchy reveal as best predictors of biological integrity. It has been proved that when IBI predictions are based on a wide array of observations from different watersheds and stream orders; regional scale variables will emerge as best predictors (Roth et al., 1996). Alternatively, if the study is based on similar types of observations with little regional environmental variability (e.g. same order streams in one watershed), more local variables will emerge as the most significant because the background regional quality for the group of sites is very homogeneous (Lammert and Allan, 1999).

4.1. *Land use*

The model identified regional land use as one of the most important contributors to biological integrity at the state-scale. Generally, the IBI prediction power of the dominant land uses was greater in the buffer zone than in the whole drainage area (Table 1). Hay/pasture and low intensity urban development were the only exceptions among dominant land uses. The effect of these on IBI was more evident in the whole drainage area, especially for hay/pasture.

In Ohio, combinations of hay/pasture and deciduous forest (second and third most dominant land uses) were the most relevant to IBI. Surprisingly, the most abundant land use (i.e. cropland) was not part of the final model or the offstream variables sub-model. This was most likely due to negative, strong cross-correlations between the percentage of crops and deciduous forest. Agriculture and forest have been identified as important contributors to IBI variability (Roth et al., 1996; Wang et al., 1997; Wilson and Xenopoulos, 2008). A positive correlation between quality of fish assemblages and percentage of forested lands -which are negatively correlated to agriculture- has been reported. This correlation held true for both, the drainage area and the regional buffers (Wilson and Xenopoulos, 2008; Stewart et al., 2001).

In most research efforts, different agricultural land uses such as cropland, range and pasture, orchards, or hay are usually merged into one category: agricultural lands (Anderson et al., 1976). In our research, agricultural land uses were not merged. The different sub-categories were kept as originally defined in the NLCD (USGS, 2008b). This revealed hay/pasture lands within the drainage area as a great predictor of biological integrity despite its smaller extent if compared to cropland (average cropland coverage equal to 56.1% versus 9.1% for hay/pasture). Pasture and range lands in the drainage area have been associated with reduced vegetal cover, increased water temperature, nitrate, biomass concentrations, photosynthetic rates, and total suspended solids as well as increased fine sediment loading. A major shift in species composition of the macro-invertebrate community was also observed in areas with pasture lands (Quinn et al., 1997). The presence of rangeland is particularly harmful to aquatic fauna, especially in sites with poor riparian quality (Meador and Goldstein, 2003) and proved the most harmful to the aquatic community in the state of Ohio.

Regional urban land uses played an overall smaller role on the integrity of Ohio streams. The dominant urban land uses (i.e. open space and low intensity development) were mostly relevant at the regional 100-m buffer (Fig. 4). This result agreed strongly with research negatively correlating urbanization along the stream buffers and stream integrity (Stewart et al., 2001; Wang et al., 2001; Morley and Karr, 2002). Urbanization seems to be significant at the local level as well. Medium intensity urbanization in the local buffers was the only local variable present in the final land use sub-model (see *All LU* model in Fig. 4). Medium intensity development was not a dominant land use in local buffers (2.21 and 1.76% in the 100- and 30-m buffers respectively, versus 12.3 and 11.6%, and 5.9 and 5.51% of open space and low intensity urban lands respectively). Even though local open and low intensity developed lands were not selected in any model, this was most likely the consequence of strong correlation with their regional homologues ($r = 0.60$ and 0.57 for open space and $r = 0.57$ and 0.59 for low development in the 30- and 100-m buffers respectively). Nonetheless, new information introduced by the local medium intensity development could indicate that proximity of intense urbanization is an important factor to the site's integrity (Wang et al., 2001; Morley and

Karr, 2002). Around 10–12% of connected imperviousness is considered the threshold beyond which biological quality declines rapidly in watersheds with small or no riparian buffers (Schueler, 1994; Wang et al., 2000, 2001). Selection of medium intensity development in local buffers by our model may indicate that this threshold has been reached.

Another minor contributor in the offstream and final models was the presence of woody wetlands in the 30-m regional corridor. Percentage of this land use in the regional 30-m buffer was present in the final model and its extent in the drainage area was selected in the land use model (Fig. 4). Even though little new variability was explained by this land use, its presence is remarkable because of its little extent (mean percentages equal to a 0.33, 0.70, and 1.07% in the drainage area, 100- and 30-m regional buffers respectively). Woody wetlands seemed to gain importance with proximity to the stream (its individual-based predictive power ranked 12th out of 16 land uses in the drainage area, 9th out of 15 land uses in the 100-m regional buffer, and 7th out of 15 land uses in the regional 30-m buffer). A similar result was reported by Richards et al. (1996), who linked forested wetlands (mean extent equal to 10% in the drainage area) with increased presence of woody debris and other channel characteristics such as bankfull depth. Wetlands regulate surface water flow and site's hydrology (Mitsch and Gosselink, 1986). Their presence is associated with decreased sediment input, nutrients, temperature, ionic strength, and increased resilience to disturbances (Richards et al., 1996; Detenbeck et al., 2000). Of special importance is the presence of wetlands near the receiving water body as the model indicated (30-m buffer was selected over drainage area in the final model). A decrease in wetland-stream distance has been positively correlated to reduced levels of nutrients, ions, and bacteria. Wetland extent has been correlated to decreased lead and high color in downstream lakes. This was found to be especially true in areas with highly fragmented riparian corridors (Johnston et al., 1990; Detenbeck et al., 1993, 2000). In the final model, two regional land use variables, positively-correlated to IBI were selected as final predictors when in very close proximity to the stream (i.e. in the 30-m buffer). On the other hand, negatively-correlated, regional land use variables were usually selected for the whole drainage area or for the 100-m buffer. Even though a definite conclusion may not be inferred from this fact, it may be an indication that preserving watershed-wide natural continuity along a stream's immediate lands may help improve or maintain biological integrity when development occurs beyond these limits.

4.2. *Fragmentation*

The negative effects of stream fragmentation to aquatic species have been widely studied (Reyes-Gavilan et al., 1996; Morita and Yamamoto, 2002; Morita and Yokota, 2002). Stream fragmentation and anthropogenic flow regulation affects a large percentage of streams worldwide, especially in developed countries (Dynesius and Nilsson, 1994; Nilsson et al., 2005). Stream fragmentation by dams has serious consequences for the biological community, preventing fish from reaching upstream habitats and isolating trapped upstream populations. Decreased species richness and risk of extinction of native fauna through demographic, environmental, and genetic stochasticity are some of the consequences fragmented populations face (Morita and Yamamoto, 2002). Moreover, physical barriers are not the only consequence of dams. Usually, hydrologic changes are also associated to impoundments. Alteration of the natural flow regime affects fauna by eliminating or modifying natural habitat conditions, which may generate a shift in species composition and, therefore, biological integrity (Poff and Allan, 1995; Richter et al., 1996; Poff et al., 1997; Fischer and Kummer, 2000; Freeman et al., 2001; Gilvear et al., 2002).

In our research, some fragmentation metrics had the largest individual IBI predicting power overall. This was especially true with metrics that accounted for both, the upstream and downstream habitats or the downstream habitat only. These variables were able to explain around 40% of the total IBI variability by themselves. Upstream fragmentation metrics had far less prediction power and were only relevant in some sites as shown in Table 3.

A potential explanation is that most of the available observations were located well inland and far from the basin outlet (average stream distance to basin outlet = 284.3 Km, minimum distance = 18.35 km, maximum distance = 833 Km). This could have influenced the overall model results since most of the available habitat was located in the downstream section. The fact that most of the available habitat in the available observations was located in the downstream section may have generated strong correlations between overall fragmentation metrics (i.e. metrics including both, upstream and downstream sections) and downstream-only metrics. However, and as mentioned in previous sections, model performance and speed is not negatively impacted by introducing strongly cross-correlated variables. Irrespective of this caveat, the model still selected an overall fragmentation metrics as the most powerful IBI predictor (Table 2), which is a clear indication of the paramount importance of habitat size and continuity on aquatic ecosystems. The impact of a fragmented upstream network was demonstrated when comparing fragmentation levels between well- and mispredicted sites with fragmented upstream networks (Table 3). Statistical differences were identified among these. No statistical differences in the size of these sites' drainage areas were found.

4.3. *Point sources and instream water quality*

Even though most of the nutrient-related parameters were among the best water-quality IBI predictors, nitrate and TP concentrations were not ranked among them, being TP the poorest predictor in the group. A clear relationship between nitrate concentration and IBI has not been found in Ohio. Only concentrations beyond 3–4 mg/L had consistently negative effects on IBI Rankin et al. (1999). The poor prediction power of phosphorus concentrations could be attributed to high concentrations beyond the biomass limiting-nutrient condition (Rankin et al., 1999). The second cluster of variables was composed of ionic strength-related parameters (Mg, Hard, Cl, Cond, SO_4). The third and last group was composed of metal concentrations (Zn, Cd, Fe, Cu, Pb) with the exception of arsenic which had the third highest individual predictive

power of all available chemicals. Other variables such as DO, TSS, or pH had very low predictive power. The first two variables selected in the water Quality Model (BOD and NO_2, Fig. 4) indicate that nutrient input is the main water quality contributor to biological degradation at the study scale. BOD has been identified as a significant source of degradation in Ohio streams (Dyer et al., 2000; Norton et al., 2000, 2002) and is an indication of highly eutrophic conditions. The higher predicting power of BOD and several nutrient-related parameters clearly indicate that eutrophication processes have a significant impact on IBI. The most significant impact of eutrophication on aquatic fauna occurs in the ultimate or "collapse" phase in which oxygen is depleted because it is used to fuel decomposition of massive amounts of decaying algae or phytoplankton. Concentrations of DO prior to the ultimate eutrophication phase fluctuate on a daily basis based on algae photosynthetic or respiration processes (Novotny, 2003). For this reason, for eutrophic systems that have not yet reached system collapse, DO may not always be a good predictor of biological integrity as the model identified. Even though samples were collected during summer months or in early Fall (period in which environmental conditions will be more favorable for algae blooms in Ohio), it is unlikely that all sites with an excess nutrient input were in the ultimate phase of eutrophication.

The third selected variable in the model was cadmium concentration, which provided marginal improvement (see Fig. 4). Metal toxicity is indeed a powerful agent of biological degradation. However, it is only able to explain a significant part of the overall IBI variability at smaller scales such as the upper or lower parts of a watershed (Dyer et al., 2000). This is most likely a consequence of its highly localized nature (i.e. coming from point sources or legacy pollution). None of the chemical variables were included in the subsequent *Instream variables* model. Habitat and sometimes water quality -especially if related to nutrient input- are mostly driven by local and regional land uses. Therefore, in sites with severely impaired habitats (e.g. with a high level of fine sediment due to accelerated denudation processes), the most likely cause of poor water quality is non-point source pollution (i.e. chemicals attached to flushed particles in runoff). This could explain why water quality data did not provide any new information when merged with the habitat model at the study scale.

Point source density and intensity only had a significant impact at the local scale as expected. When outliers were removed, significant differences were not identified between well- and mispredicted observations with reported point sources (therefore, not included in Table 3 or Table 4). Significant differences in water quality were only found in some under-predicted sites with respect to well-predicted sites. Lower hardness levels and higher sulfate concentrations were observed in under-predicted observations (Table 4). From the results, the overall effect of point source pollution on IBI is very small compared to other more ubiquitous stressors directly or indirectly related to land use at the scale of our study. Point sources are a significant factor if they have a significant presence in the area of study. For example, point source pollution has been identified as a significant negative factor in some studies based in only one basin or a portion of it (Dyer et al., 1998a, 2000). Another study based in the whole state of Ohio using only habitat and water quality data, concluded that water quality had a significant impact on IBI in specific clusters of sites only; but confirmed that the most relevant instream parameters at the state-scale were habitat-related (Manolakos et al., 2007).

4.4. *Instream habitat*

Instream habitat and drainage area were able to explain 49% of the overall IBI variability. Substrate-related metrics (i.e. embeddedness and substrate quality), stream variability (i.e. pool and riffle quality), and vegetal cover were the most relevant QHEI metrics. Habitat quality has been identified as the main instream source of IBI variability (Hall et al., 1996; Dyer et al., 1998a; Manolakos et al., 2007). Habitat quality is strongly driven by land use changes in the drainage area and may account for land use-related water quality information such as nutrient input. Our model confirmed this point and the *Habitat* model selected exactly the same variables as the *Instream* model (Fig. 4). Riffle and Cover qualities were selected in the *Overall* model but with very modest contributions to the final outcome. Stream variability, substrate quality, and/or instream cover have been identified as significant contributors to biotic quality in Ohio (Dyer et al., 1998a, 1998b; Yuan and Norton, 2004; Manolakos et al., 2007) and elsewhere (Minshall, 1984; Quinn and Hickey, 1990; Richards et al., 1993; Rabeni and Smale, 1995). Drainage area was positively correlated to IBI, which strongly agreed with the findings by Dyer et al. (1998a) in Ohio.

4.5. *Mispredictions due to local conditions*

Two main sources of IBI overprediction due to local environmental conditions were identified. The first source consisted of higher levels of upstream fragmentation in sites with fragmented upstream networks (Table 3). The second source, which affected all observations, was local land use patterns not included in the final prediction model. Over-predicted observations had significantly higher percentages of forested areas in the drainage area and regional buffer corridor (Table 3). This contributed to high calculated IBI scores. The extent of forested land in the local catchment and buffer zones were also significantly higher in over-predicted sites while medium intensity urbanization in the local buffer was lower. These results were counter-intuitive given the lower observed IBI scores in sites with such good '*land use quality*'. However, these sites had significantly higher percentages of hay/pasture lands in the local catchment area. This land use was identified as the most detrimental to IBI and could explain the over-predictions (i.e. sites with very good regional characteristics, which were included in the final model, but significantly higher levels of pasture lands at the local scale, which was not included in the final model). Significant differences of pasture lands at the regional scale were not present in over-predicted sites.

On the other hand, under-predicted sites had significantly better '*land use quality*' at both scales as well. Therefore, exceptional local land use quality (i.e. significantly higher levels of forested areas combined with smaller percentages of urban and crop lands, see Table 4) is the most likely cause of

IBI under-prediction. The final model didn't include information from local land use variables.

5. Conclusions

- We presented a highly versatile, predictive modeling methodology which has many potential applications in the environmental data analysis field. The new methodology is capable of dealing with a large number of observations with many associated attributes in a very time-efficient manner. Moreover, one of the main contributions of such methodology is that it does not rely on a-priori assumption on the relationship between the environmental variables and the prediction target. Many traditional exploratory techniques in the ecological modeling field make such kind of assumptions (e.g. canonical correspondence analysis is based on linear regressions between the explanatory and response variables). These two main features are of paramount importance because natural systems are composed of many inter-twined, cross-correlated variables with a highly non-linear inter-dependencies. Furthermore, the model performs well when discrete or crudely scaled data is used because it is based on assessing environmental similarities among sites with the same attributes. As a result, it allows using variables with different scoring criteria at once.
- At the state-scale, regional land use and basin-level stream fragmentation are the main predictors of biotic integrity in Ohio. Habitat variables only contributed marginally to model improvement, while instream water quality and point source intensity and density were not able to improve the final model at all. Most of the information from instream water and habitat qualities is introduced into the model by regional land use, which acts as a surrogate variable.
- We revealed the importance of local stressors which were not accounted for in the final model. Over-predictions mainly came from a combination of higher upstream fragmentation, extreme point source density and intensity, and high levels of hay/pasture in the local catchment area. under-predictions mainly came from extraordinary local land use quality which was not accounted for in the model.
- If the 55 mispredicted sites –out of 429 observations- could be disregarded due to unique local conditions, the model would explain 86% of the overall IBI variability. Therefore, in our dataset local stressors accounted for an extra 25% (i.e.86%–61%) of the variability explained by land use and fragmentation metrics. The remaining 14% may be due to sampling errors, data quality issues, or natural randomness (for example, a site with BOD = 24 mg/L; TKN = 3.1 mg/L; TP = 1.29 mg/L; Zn = 180 μg/L; Cu = 39 μg/L; Fe = 19,700 μg/L; or NO_2 = 0.19 mg/L had one of the highest observed IBI scores (50)).
- The results showed how water quality issues from point sources have a small overall impact on the biotic integrity in Ohio. This may indicate a successful control of point sources through the EPA's NPDES Program, which have been top priority for U.S. surface waters since the Clean Water Act of 1972. These results do not indicate that water quality problems from point sources are not relevant anymore in Ohio, but they have shifted from being a widespread issue to a local one at the state-scale.
- Our model identified stream fragmentation and land use change - especially in the regional buffers- as the most important stressors to biological integrity. Habitat degradation and nutrient input are the most direct instream consequences from land use disturbances. Results suggest that in order to achieve the aimed *physical, chemical, and biological integrity of the Nation's waters*, protection and enforcing policies have to refocus towards a more holistic view that goes beyond the traditional point source control. Ecosystem continuum must be kept and watershed- or basin-level- land use planning is necessary to attain such goals, especially in the most immediate lands of a water body. Stressors should be approached in a scale-down manner. This would guarantee that improvements at the local level have successful outcomes because the regional background conditions meet the minimum requirements to attain the targeted integrity.

Acknowledgement

This research has been partially supported by the US EPA/NSF/USDA STAR Watershed Program, Grant No. R83-0885-010 to Northeastern University, Boston, MA. The authors would like to express their gratitude towards Mr. Ed Rankin and Mr. Dennis Mishne from Ohio EPA and Mr. Scott Dyer and Ms. Charlotte White-Hull from The Procter & Gamble Company for their help with the environmental data and valuable advice.

REFERENCES

Allan, J.D., Erickson, D.L., Fay, J., 1997. The influence of catchment land use on stream integrity across multiple spatial scales. Freshwater Biology 37, 149–161.

Anderson, J.R., Harvey, E.H., Roach, J.T., Whitman, R.E., 1976. A Land Use and Land Cover Classification System for Use with Remote Sensor Data. Geological Survey Professional Paper 964. U.S. Government Printing Office, Washington D.C.

Bedoya, D., Manolakos, E.S, Novotny, V., 2010. Characterization of biological responses under different environmental conditions- a hierarchical modeling approach. Ecological Modelling. Ecological Modelling 222 (3), 532–545.

Beyer, H.L., 2004. Hawth's Analysis Tools for ArcGIS. Available at. http://www.spatialecology.com/htools.

Bode, R.W., 1988. Methods for Rapid Biological Assessment of Streams. New York State Department of Environmental Conservation, Albany, NY.

Castelle, A.J., Johnson, A.W., Connolly, C., 1994. Wetland and stream buffer size requirements - a review. Journal of Environmental Quality 23, 878–882.

Detenbeck, N.E., Batterman, S.L., Brady, V.J., Brazner, J.C., Snarski, V.M., Taylor, D.L., Thompson, J.A., Arthur, J.W., 2000. A test of watershed classification systems for ecological risk assessment. Environmental Toxicology and Chemistry 19, 1174–1181.

Detenbeck, N.E., Johnston, C.A., Niemi, G.J., 1993. Wetland effects on lake water-quality in the Minneapolis St-Paul metropolitan-area. Landscape Ecology 8, 39–61.

Dyer, S.D., White-Hull, C., Carr, G.J., Smith, E.P., Wang, X.H., 2000. Bottom-up and top-down approaches to assess multiple stressors over large geographic areas. Environmental Toxicology and Chemistry 19, 1066–1075.
Dyer, S.D., White-Hull, C., Wang, X., Johnson, T.D., Carr, G.J., 1998a. Determining the influence of habitat and chemical factors on instream bioitc integrity for a Southern Ohio watershed. Journal of Aquatic Ecosystem Stress and Recovery 6, 91–110.
Dyer, S.D., White-Hull, C.E., Johnson, T.D., Carr, G.J., Wang, X., 1998b. The importance of space in understanding the risk of multiple stressors on the biological integrity of receiving waters. Journal of Hazardous Materials 61, 37–41.
Dynesius, M., Nilsson, C., 1994. Fragmentation and flow regulation of river systems in the Northern 3rd of the world. Science 266, 753–762.
Fischer, S., Kummer, H., 2000. Effects of residual flow and habitat fragmentation on distribution and movement of bullhead (Cottus gobio L.) in an alpine stream. Hydrobiologia 422, 305–317.
Freeman, M.C., Bowen, Z.H., Bovee, K.D., Irwin, E.R., 2001. Flow and habitat effects on juvenile fish abundance in natural and altered flow regimes. Ecological Applications 11, 179–190.
Frissell, C.A., Liss, W.J., Warren, C.E., Hurley, M.D., 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10, 199–214.
Gilvear, D.J., Heal, K.V., Stephen, A., 2002. Hydrology and the ecological quality of Scottish river ecosystems. Science of the Total Environment 294, 131–159.
Giraudel, J.L., Lek, S., 2001. A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. Ecological Modelling 146, 329–339.
Hall, L.W., Scott, M.C., Killen, W.D., Anderson, R.D., 1996. The effects of land-use characteristics and acid sensitivity on the ecological status of Maryland coastal plain streams. Environmental Toxicology and Chemistry 15, 384–394.
Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACM Computer Surveys 31, 264–323.
Johnston, C.A., Detenbeck, N.E., Niemi, G.J., 1990. The cumulative effect of wetlands on stream water quality and quantity. A landscape approach. Biogeochemistry 10, 105–141.
Karr, J.R., 1991. Biological Integrity: a long-neglected aspect of water resource management. Ecological Applications 1, 66–84.
Karr, J.R., Fausch, K.D., Angermeier, P.L., Yant, P.R., Schlosser, I.J., 1986. Assessing Biological Integrity of Running Waters: a Method and Its Rationale. Illinois Natural History Survey, Champaign, IL.
Karr, J.R., Kerans, B.L., 1981. Components of biological integrity: their definition and use in development of an invertebrate IBI. 1991 MidWest Pollution Control Biologists Meeting. In: Environmental Indicators: Measurement and Assessment Endpoints. U.S. Environmental Protection Agency, Lincolnwood, IL.
Lammert, M., Allan, J.D., 1999. Assessing biotic integrity of streams: effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. Environmental Management 23, 257–270.
Lyons, J., 2006. A fish-based index of biotic integrity to assess intermittent headwater streams in Wisconsin, USA. Environmental Monitoring and Assessment 122, 239–258.
Lyons, J., Piette, R.R., Niermeyer, K.W., 2001. Development, validation, and application of a fish-based index of biotic integrity for Wisconsin's large warmwater rivers. Transactions of the American Fisheries Society 130, 1077–1094.
Manolakos, E., Virani, H., Novotny, V., 2007. Extracting knowledge on the links between the water body stressors and biotic integrity. Water Research 41, 4041–4050.
Meador, M.R., Goldstein, R.M., 2003. Assessing water quality at large geographic scales: Relations among land use, water physicochemistry, riparian condition, and fish community structure. Environmental Management 31, 504–517.
Minshall, G.W., 1984. Aquatic-insect substratum relationships. In: Resh, V.H., Rosenberg, D.M. (Eds.), The Ecology of Aquatic Insects. Praeger Scientific, New York NY, pp. 358–400.
Mitsch, W.J., Gosselink, J.G., 1986. Wetlands. Van Nostrand Reinhold, New York, NY.
Morita, K., Yamamoto, S., 2002. Effects of habitat fragmentation by damming on the persistence of stream-dwelling charr populations. Conservation Biology 16, 1318–1323.
Morita, K., Yokota, A., 2002. Population viability of stream-resident salmonids after habitat fragmentation: a case study with white-spotted charr (Salvelinus leucomaenis) by an individual based model. Ecological Modelling 155, 85–94.
Morley, S.A., Karr, J.R., 2002. Assessing and restoring the health of urban streams in the Puget Sound basin. Conservation Biology 16, 1498–1509.
Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C., 2005. Fragmentation and flow regulation of the world's large river systems. Science 308, 405–408.
Norton, S.B., Cormier, S.M., Smith, M., Jones, R.C., 2000. Can biological assessments discriminate among types of stress? A case study from the Eastern Corn Belt Plains ecoregion. Environmental Toxicology and Chemistry 19, 1113–1119.
Norton, S.B., Cormier, S.M., Smith, M., Jones, R.C., Schubauer-Berigan, M., 2002. Predicting levels of stress from biological assessment data: Empirical models from the Eastern Corn Belt Plains, Ohio. USA. Environmental Toxicology and Chemistry 21, 1168–1175.
Novotny, V., Bartŏsová, A., O'Reilly, N., Ehlinger, T., 2005. Unlocking the relationship of biotic integrity of impaired waters to anthropogenic stresses. Water Research 39, 184–198.
Novotny, V., 2003. Water Quality. Diffuse Pollution and Watershed Management. John Wiley & Sons, New York.
Ohio Environmental Protection Agency, 1987. Biological Criteria for the Protetction of Aquatic Life: Volume I-III. In: Standardized Field and Laboratory Methods for Assessing Fish and Macroinvertebrate Communities. Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, OH.
Ohio Environmental Protection Agency, 2005. Surface and Groundwater Monitoring and Assessment Strategy 2005–2009. Division of Surface Water and Division of Drinking and Ground Waters. Ohio EPA Technical Bulletin EAS/2005-09-1.
Ohio Environmental Protection Agency, 2008. Statewide Biological and Water Quality Monitoring & Assessment. Available at: http://www.epa.state.oh.us/dsw/bioassess/ohstrat.html. Ohio EPA, Division of Surface Water, Columbus, OH.
Pickett, S.T.A., Kolasa, J., Armesto, J.J., Collins, S.L., 1989. The ecological concept of disturbance and its expression at various hierarchical levels. OIKOS 54, 129–136.
Poff, N.L., 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16, 391–409.
Poff, N.L., Allan, J.D., 1995. Functional-Organization of stream fish assemblages in Relation to Hydrological variability. Ecology 76, 606–627.
Poff, N.L., Allan, J.D., Bain, M.B., Karr, J.R., Prestegaard, K.L., Richter, B.D., Sparks, R.E., Stromberg, J.C., 1997. The natural flow regime. Bioscience 47, 769–784.
Quinn, J.M., Cooper, A.B., Davies-Colley, R.J., Rutherford, J.C., Williamson, R.B., 1997. Land use effects on habitat, water quality, periphyton, and benthic invertebrates in Waikato, New Zealand, hill-country streams. New Zealand Journal of Marine and Freshwater Research 31, 579–597.

Quinn, J.M., Hickey, C.W., 1990. Magnitude of effects of substrate particle size, recent flooding, and watershed development on benthic invertebrates in 88 New Zealand rivers. New Zealand Journal of Marine Freshwater Resources 24, 411–428.

Rabeni, C.F., Smale, M.A., 1995. Effects of Siltation on stream Fishes and the potential Mitigating role of the buffering riparian zone. Hydrobiologia 303, 211–219.

Rankin, E.T., 1989. The Qualitative Habitat Evaluation Index (QHEI): Rationale, Methods, and Application. Ecological Assessment Section, Division of Water Quality, Planning, and Assessment. Ohio Environmental Protection Agency, Columbus, OH.

Rankin, E.T., Miltner, B., Yoder, C.O., Mishne, D., 1999. Association between Nutrients, Habitat, and the Aquatic Biota in Ohio Rivers and Streams Ohio EPA Technical Bulletin MAS/1999-1-1, Columbus, OH.

Reyes-Gavilan, F.G., Garrido, R., Nicieza, A.G., Toledo, M.M., Brana, F., 1996. Fish community variation along physical gradients in short streams of northern Spain and the disruptive effect of dams. Hydrobiologia 321, 155–163.

Richards, C., Host, G.E., Arthur, J.W., 1993. Identification of predominant environmental-factors structuring stream macroinvertebrate communities within a large agricultural catchment. Freshwater Biology 29, 285–294.

Richards, C., Johnson, L.B., Host, G.E., 1996. Landscape-scale influences on stream habitats and biota. Canadian Journal of Fisheries and Aquatic Sciences 53, 295–311.

Richter, B.D., Baumgartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. Conservation Biology 10, 1163–1174.

Roth, N., Southerland, M., Chaillou, J., Klauda, R., Kazyak, P., Stranko, S., Weisberg, S., Hall, L., Morgan, R., 1998. Maryland biological stream survey: development of a fish index of biotic integrity. Environmental Monitoring and Assessment 51, 89–106.

Roth, N.E., Allan, J.D., Erickson, D.L., 1996. Landscape influences on stream biotic integrity assessed at multiple spatial scales. Landscape Ecology 11, 141–156.

Schueler, T., 1994. The importance of imperviousness. Watershed Protection Techniques 1, 100–111.

Stewart, J.S., Wang, L.Z., Lyons, J., Horwatich, J.A., Bannerman, R., 2001. Influences of watershed, riparian-corridor, and reach-scale characteristics on aquatic biota in agricultural watersheds. Journal of the American Water Resources Association 37, 1475–1487.

U.S. Army Corps of Engineers, 2005. National Inventory of Dams. Available at. http://crunch.tec.army.mil/nidpublic/webpages/nid.cfm.

U.S.Environmental Protection Agency, 2008. Permit Compliance System Database. Available at. http://epa.gov/enviro/html/pcs/pcs_query_java.html.

U.S. Geological Survey, 2008a. National Hydrography Dataset. Available at. U.S. Department of the Interior. http://nhd.usgs.gov/index.html.

U.S. Geological Survey, 2008b. National Land Cover Database. Available at. Multi-resolution Land Characteristics Consortium U.S. Department of the Interior. http://www.mrlc.gov/index.php.

Wang, L., Lyons, J., Kanehl, P., Bannerman, R., 2001. Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental Management 28, 255–266.

Wang, L.Z., Lyons, J., Kanehl, P., Bannerman, R., Emmons, E., 2000. Watershed urbanization and changes in fish communities in southeastern Wisconsin streams. Journal of the American Water Resources Association 36, 1173–1189.

Wang, L.Z., Lyons, J., Kanehl, P., Gatti, R., 1997. Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. Fisheries 22, 6–12.

Wilson, H.F., Xenopoulos, M.A., 2008. Landscape influences on stream fish assemblages across spatial scales in a northern Great Plains ecoregion. Canadian Journal of Fisheries and Aquatic Sciences 65, 245–257.

Yuan, L.L., Norton, S.B., 2004. Assessing the relative severity of stressors at a watershed scale. Environmental Monitoring and Assessment 98, 323–349.

Closing Comments on 2015 SFER Chapter 3B Regional Mercury and Sulfur Environmental Assessment for the Everglades
Reviewers Stein (AA) and Novotny (A)

In the following closure the reviewers will focus on the issues where there were differences of opinion between reviewers and authors or where reviewers make additional recommendations.

Chapter 3B is attempting to address the highly scientific and complex ecological problem of high concentrations of methyl mercury in fish in the Everglades Protection Area. An interesting puzzle is that the highest concentrations occur in the least polluted areas including Everglades National Park which also receives the smallest mercury inputs and ecologically exhibits highly oligotrophic conditions, factors that would seem to limit methylation potential. A big concern noted by the reviewers in the last several years is the initial postulation in the 2013 SFER and subsequent denial (this year and last) of a unimodal relationship between mercury methylation and sulfate reduction euphemistically called the "Goldilocks' concept". The "cutting edge" research (we believe the research is quite sound) conducted by at least two different teams compiled by the SFWMD but with different philosophy confirmed the fact that was also accepted by the reviewers beginning with the 2013 SFER, that there is no simple unimodal relationship and proposing and/or accepting a low goal/standard such as 1 mg SO_4^{2-} would not solve the methylation problem. In this sense, even the title of the chapter "Mercury and sulfur environmental assessment …" may be misleading because it insinuates that there may be a possibility that by controlling sulfate in the system the problem of high MeHg may be solved.

The present authors seem to be pre-focused on discounting the unimodal relationship that few are advocating and in doing so they claim that there is no relationship. However, this is far from the truth. The Goldilocks' concept discovered by a number of scientists and included in the 2013 SFER and even indirectly confirmed by the new authors of the 2014 and 2015 SFERs, is a convincing model that offers a processed-based explanation of the influence of sulfate concentration on mercury methylation in soils and sediments. However, even in the 2013 SFER report that advocated the Goldilocks' concept, other parameters such as organic carbon, iron etc. were acknowledged to blur the direct influence of the Goldilocks' relationship on the MeHg - SO_4^{2-} couple in sediment and therefore mercury levels in fish as well.

It appears, however, that the reviewers find themselves in the middle of some kind of internal struggle between the previous authors who presented a foundation for a credible model and the new authors who deny that such a model exists. The reviewers note that the authors' response indicates that the external scientists and highly reputable research agencies contributing to the 2013 Chapter 3B elected not to contribute to the 2014 and 2015 reports. For one, the primary author admitted that he is skeptical of the "bell-shaped relationship" and falsely stated that the bell-shaped relationship is not consistently observed in nature. On the contrary, bell shaped relationships are ubiquitous in nature specifically when dealing with living organisms' growth response to stimulation of nourishing effects at low concentrations and toxic effects at high concentrations of many compounds, for example the effects of metals on bacteria and algae. This would also apply to sulfates and other effects on sulfur reducing bacteria.

Strangely, Figure 1A provided by the authors in their response would, to many, be a proof of a multimodal bell shaped effect where sulfate concentrations of less than about 1-2 mg/L would stimulate MeHg production by SRB and would be toxic to SRB at concentrations of more than about 10 mg (probably by generated sulfide toxicity). So this is what determines the 1 to 10 mg/l Goldilocks window and the fact that the highest MeHg fish tissue concentrations are measured within this window is undeniable. Other factors (organic carbon iron etc.) presumably shift the curve to the left or right at specific sites and characterizations of these effects would lead to a better all-encompassing model. The fact that the authors focused only on denying the existence of the Goldilocks' effect and did not investigate the effects, or a priori did not quantitatively look on the effects, of other parameters is puzzling. The methodologies of data mining by, for example, multiple component analyses and neural nets are widely available and they had enough points to investigate the phenomenon using multivariate nonlinear modeling. The authors used only a simple unimodal regression model which obviously failed.

Closure on specific comments: If a comment is not specifically numbered below, the reviewers believe the authors' response satisfactorily addressed the original comment.

Comment #1: The response does not clear our initial confusion and if we have interpreted it correctly seems to confirm the most salient point of our original comment. The response indicates only one composite sample per year per site, yet there are clearly error bars representing one standard deviation on each bar in both figures 3B-2 and 3B-3. How can there be a standard deviation on one value? If the error is associated with repeated technical sampling from the same composited fish tissue sample (our interpretation of the response), then trying to correlate mercury concentrations in fish (at least mosquito fish) to site and temporal differences (or methylation processes) is next to impossible. Using the current sampling techniques the error (apparently instrumentation error which is hard to reduce) is nearly the same as the differences between site and years.

Comment #3: The reviewers appreciate the detailed response provided by the authors. While we have some disagreements on scientific approaches and interpretations provided in this and other comments, it is clear that authors put considerable effort into their responses, are versed in appropriate literature and back their comments with additional information. This is a welcome contrast to the responses in some other chapters.

The first part of the comment, directed toward soil P release due to sulfate in the STA waters and therefore not directly related to this chapter, was merely cautionary. The authors cite two publications and extract a figure nicely supporting their argument that soil P release due to sulfate is not occurring in the STAs. But at some point the sediment P retention capacity may become oversaturated and P may be released. This a common situation with agricultural soil overloaded with fertilizers and also with urban lagoons that after a decade of removing P lose their removal efficiency. The fact that SAV and EAV are growing and adding more organic carbon to the system may explain this continuing high adsorption capacity for (dissolved) phosphates. If SAV and EAV disappear the P removing efficiency will be compromised, but this should be addressed in Chapter 5B.

In regard to differences in observed methylation rates between the reported mesocosm study and the microbial study of Bae et al., as stated above we agree the process is complex and not controlled by one single mechanism, including sulfate reduction. However, the reviewers take the position that A) there clearly is a Hg methylation problem in the Everglades ecosystem B) many studies confirm that sulfate reduction is one process (with supporting definable mechanisms) that results in methylation (and is often the driving factor of methylation in wetland systems) and C) Everglades sulfate concentrations are in the Goldilocks range that these studies suggest it is important; therefore it must be having some influence. That its effect is not as obvious in empirical field studies within the Everglades (as the authors rightfully point out) doesn't discount its importance or, for that matter, refute it as a driving mechanism. Many other covariates are in play in soil methylation, for example, the quantity and type of organic matter providing electrons for SRB, and one (or more of these) might be limiting/masking the sulfate-MeHg couple. More process-based research could shed light on the influence of these co-variates. Our bigger concern is that (in general) the type of empirical studies, looking at fish Hg concentrations for example, is three or more steps down the process, greatly increasing the likelihood that correlations are masked by even more co-variates. That does not discount the value of those field studies, as stated above, but causative relationships are not likely to be developed from these studies.

Comment #4: It is well known from studies on anaerobic digestion of sewage sludge that methanogens and SRB do compete for organic substrate and that certain factors shift this balance (see for example Mizumo and Noike 1094, Mizumo et at. 1998) and to a lesser degree in wetlands (e.g. Capone et al. 1983). Whether that is true in the Everglades environment is unknown, but we believe it does offer another interpretation that should not be dismissed. Therefore we accept authors point 1, but not point 2.

Comment #5: The difference of opinion between authors and reviewers is succinctly stated here. We are much less likely to conclude the relationship between sulfate and MeHg is relatively weak. Certainly not conclusive, but also not weak.

Comment #8: See Comment #5.

Comment #9: As an aside, we agree that not all of our comments require a revision to the text. Hopefully, a comment seeking additional clarification results in modifications to the text (not applicable here). However as stated in our SOW, we are to provide alternative interpretations to the data when we believe they is appropriate and this comment is a (minor) example of that case.

Comment #12: See the latter part of Comment #3. When the reviewers see these data of fish mercury levels they see a general trend that correlates well with the bell shaped Goldilocks' curve for sediment MeHg production. Does that prove there is a direct correlation? No of course not, but in our opinion the chapter's discussion points seem focused on convincing the reader that there is no correlation and that sulfate concentration is a non-factor in the methylation problem in the Everglades ecosystem. The authors never say that directly, but push the discussion toward that conclusion.

References:

Capone, D.G., D.D. Reese, and R.P. Kiene. 1983. Effects of Metals on Methanogenesis, Sulfate Reduction, Carbon Dioxide Evolution, and Microbial Biomass in Anoxic Salt Marsh Sediments. Applied and Environ. Microbiology. 45(5):1586-1591.

Mizuno, O., Y.Y. Li, and T. Noike. 1998. The Behavior of Sulfate-Reducing Bacteria in Acidogenic Phase of Anaerobic Digestion. Water Resources. 32(5):1626-1634.

Mizuno, O., Y.Y. Li, and T. Noike. 1994. Effects of Sulfate Concentration and Sludge Retention Time on the Interaction between Methane Production and Sulfate Reduction for Butyrate. Water Science Technology. 30(8):45-54.

Key Findings and Recommendations (SOW Task 3)

- The best opportunity to reduce mercury levels in higher trophic organisms is to better characterize the factors influencing the methylation of input mercury. Unfortunately, as in last year's report there seems to be focus on debunking a unimodal relationship between sulfate concentration and mercury methylation and, by extension, mercury concentrations in fish. While fish concentrations are the concern, until the mechanisms converting elemental mercury deposition into bio-available methyl-mercury are characterized for Everglades conditions , discovering how to reduce fish mercury concentrations will remain elusive.

- There appears to be considerable scientific debate internal to the District as to the influence of sulfate on mercury methylation. On one hand, District scientists used a credible (Goldilocks') model to provide evidence that intermediate sulfate concentrations enhance mercury methylation. Other scientists, primarily authors of this year's chapter, conclude that the model relationship is weak and not applicable to the Everglades system. The reviewers acknowledge both views as part of appropriate scientific debate, but believe the model, while not all-encompassing, provides a rational framework to better characterize mechanisms of mercury methylation. It would be far better to refine the model than to dismiss it as inappropriate.

  The Goldilocks' model should be further refined for the conditions typical in the EPA and EAA regions of South Florida. Likely fruitful avenues for future research include investigation on the quantity and type of dissolved organic matter and the link between mercury methylation in sediments and its incorporation into biomass. Assessment of mercury concentrations in fish should continue, but attempts to link these concentrations to a simple unimodal relationships or refuting simple relationships should be downplayed.

Chapter 4 (Nutrient Source Controls) Accountability

Dodds 2014 SFER review 29 September 2014

1. This chapter seems a reasonable assessment of accounting for nutrients in the sub basins of the area of interest. Most all of the chapter is reporting with little analysis of why specific BMP's are effective. It is good news that P loads are generally decreasing, but many have not shown recent improvements, and lots of variability exists among years. Could those data be used to make even more improvements at not much costs?
2. The lack of N control in many of the basins is concerning. The natural everglades periphyton mats appear to have substantial nitrogen fixation capacity, at least in some places, indicating that N loading could shift algal communities.
3. In general time series analyses are necessary to predict 1) if any particular year represents much of an improvement, and 2) the actual long term trends against variable processes. For example in the data from table 4.4, there is no apparent improvement after the year 2000. Furthermore the data in 4.4 suggest that the models over predict TP loads by about 50%
4. Figure 4.8 shows why 5 year rolling window might not be the best visualization tool. Here and otherwise a regression with only the baseline or only the post baseline data included may give the best idea of trends, particularly if confidence bands are also plotted.
5. I suggest a very short section on methods including QA/QC which refers to the specific documents where the details of methodology are given in more detail.
6. A little more analyses on concentrations as well as loads could be useful in parsing out effects driven by factors other than discharge-driven differences in loading rates.
7. Also, more detailed analyses of inputs and outputs, as in Chapter 8, may shed more light on mechanisms of removal and avenues for further improvements.
8. Line 60. Reduction mass per year?

rmccafferty

Registered: 09/30
Posts: 1

Posted 10/06 #3

Request for clarification:

Comment #4: "Figure 4.8 shows why 5 year rolling window might not be the best visualization tool. Here and otherwise a regression with only the baseline or only the post baseline data included may give the best idea of trends, particularly if confidence bands are also plotted."

Clarification: Figure 4-8 is not compatible with your comment. Please verify the figure number you are referencing.

Comment #5: "I suggest a very short section on methods including QA/QC which refers to the specific documents where the details of methodology are given in more detail."

Clarification: Please indicate what section of the chapter and specifically what methodology this comment pertains to.

Thank you!

Edit | Delete

wkdodds

Registered: 09/13/12
Posts: 18

Posted 10/07 #4

Sorry..

first query, Comment #4 I mean Figure 4.3. The idea would be use use all the pre data as a way to generate a confidence band (from 1980 to 1988), and then use a linear or non-linear function from that point forward (it looks like maybe a 2nd order polynomial) to fit the remaining data including a confidence band around that regression line. Just running a 5 year mean gives no indication that it is known something changed at the beginning, and no indication of how much certainty there is about improvement over prior conditions and how certain it is the 25% has been exceeded.

second query, Comment #5. I really meant for the nutrient data throughout the chapter. There are many years of analyses here, potentially done in different labs. If a unified QA/QC program was used for all, than very shortly referring to a document with a short description would suffice. If multiple labs and methods were used, then maybe a synoptic table? Since time series are important here, being certain trends are real is necessary.

Edit | Delete

2015_SFER_Ch_4_RTCs_Final Dodds Accountability

Chapter 4 (Nutrient Source Controls) Accountability

Dodds 2014 SFER final review 5 November 2014

Closing comments

- The document is well written and carefully explains the results from the last year.
- The responses adequately explain the points that were not clear. This chapter documents the required metrics well for the most part
- Clarifications might be necessary for readers who are not as intimately acquainted with the regulatory goals, and all the background of the methodology.
- Continued efforts to integrate materials on nutrient source controls across all projects are encouraged.

Review of Draft 2015 SFER Chapter 5A Restoration Strategies - Design and Construction Staus of Water Quality Improvement Projects

Authors: Otto Stein (AA)
Level of Panel Review: Accountability

Accountability Questions:

Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate?

Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report? Are findings linked to management goals and objectives?

The panel may also provide constructive guidance for the District's large-scale programs, particularly as related to water quality assessment and control across the agency. Importantly, this review should specifically focus on report-related scope and generate critiques only on material presented in the SFER and not expand into high-level agency priorities or policies. For example, do the chapters on the Restoration Strategies Projects (5A) and Science Plan Update (5C) adequately provide a description and status of these required programs/initiatives?

Broad Comments:

Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate?

The chapter articulates the status of several construction projects designed to enhance the performance of the STAs so that they meet the WQBEL. The projects are divided ointo three flow paths Eastern, Central and Western. In WY 2014 work progressed in the Eastern and Central paths but apparently no work has been started in the Western path. However 2025 is the anticipated completion date for all projects.

Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report?

The chapter makes liberal use of figures and tables creating an easy-to-follow organization structure.

Are findings linked to management goals and objectives?

The reasons for these new construction projects is documented in the Summart section at the beginning of the Chapter.

Is there any constructive criticism and guidance to offer for the District's large-scale programs?

As was suggested last year it would be logical to have Chapter 5A and 5C be reviewed by the same reviewers. This is probably more critical for review of 5C (which requires knowledge contained in 5A) than for 5A (which could stand alone without 5C). Additionally, since these chapters really focus on the "big picture" of STA performance, it makes sense to assign AA and A reviewers to it.

Editorial page and line comments, suggested text changes in *italics*:

1The interior text of Figure 5A-4 is too small to read properly.

Closing Comments on 2015 SFER Chapter 5A Restoration Strategies - Design and Construction Status of Water Quality Improvement Projects

Reviewer Stein (AA)

The reviewer has no additional comments on the Chapter.

Key Findings and Recommendations (SOW Task 3)

- The chapter articulates the status of several construction projects designed to enhance the performance of the STAs so that they meet the WQBEL. The projects are divided into three flow paths Eastern, Central and Western. In WY 2014 work progressed in the Eastern and Central paths but work on the Western path is not scheduled to start until 2019. This could challenge the 2025 anticipated completion date for all projects. .

- It would be logical to have Chapter 5A and 5C be reviewed by the same reviewers. This is probably more critical for review of 5C (which requires knowledge contained in 5A) than for 5A (which could stand alone without 5C). Additionally, since these chapters really focus on the "big picture" of STA performance, it makes sense to assign AA and A reviewers to it.

Review of Draft 2015 SFER Chapter 5 Performance and Optimization of the Everglades Stormwater Treatment Areas

Authors: Otto Stein (AA) and Peter Dillon (A)
Level of Panel Review: Technical

Technical Questions:

1. Are the findings and conclusions supported by "best available information," or are there gaps or flaws in the information presented in the document?

2. Are there other interpretations of the data and other available information that should be considered by the authors and presented to decision makers? If so, the panel shall identify specific studies that should be addressed or available data to support alternative findings.

The review is divided in to several sections. First are several broad questions and comments regarding the interpretation of reported results. These reflect the general impressions of the reviewers after reading and reflection of the entire document. These are followed by more focused comments and concerns that in many cases back up the first section, but in other cases are merely requests for clarification or better presentation. In general, these were created on a section by section analysis. This is followed by comments on figures and tables, and editorial suggestions. Occasionally a specific comment is qualified in a subsequent one, but left for the authors to respond to, as it is a question that arose as that section was being read.

Broad questions and comments that should be addressed:

The Everglades Stormwater Treatment Areas (STAs) constructed south of Lake Okeechobee were designed and operated to remove phosphorus from runoff entering the Everglades Protection Area. Their large area of 27, 00 hectares (68,000 acres) consist mainly of marshes vegetated by submerged and emerging aquatic vegetation (SAV and EAV). About 83% of the total area provides effective treatment. Following the format adopted last year, the first part of the chapter provides an annual reporting of measured flow and P removal data and measuring the areal extend of SAV and EAV. There are 6 STAs , each divided into several cells, so most of the flow passes through several cells in a plug flow fashion. The overall assessment of performance of the STAs area is given in Table 5B-1. The flow weighted annual removal rates in percent were very good to excellent ranging from 76 to 88 percent. These P removal rates are higher than those expected for designed (constructed) wetlands treating stormwater runoff. It should be noted that the inflow P concentrations were low to average (from 71 to 198 μg/L0 when compared to those, for example, in urban runoff and the STAs were able to reduce the outflow concentrations to approximately 20 μg/L, which is the goal for the ENP. Incidentally, the hydraulic loading rates (HLR) (with exception of STA 5/6) are close to those suggested for free surface constructed wetlands (2.5 – 5 cm/day) in the Water Environment Federation wetland

design manual[1]. The STA 5/6 had much smaller HLR (0.6 cm/day) and P load rate PLR (0.5 g P/ m^2-year) than the other STAs and had the highest P removal (88 %).

The second part of the chapter contains specific studies such as the cause of SAV decline in some portions of STAs, or P removal in very low P concentrations. The last study documented that at very low P concentrations the removal rates are small to negative.

The chapter, although relatively well written with a good organization is greatly suffering from using a mishmash of SI and US units to the point of being absurd. We are living in the twenty first century. Both systems are used in almost all tables. Graphs and tables in the first portion of the chapter use archaic ppb, the second portion is using SI and generally a common approximately equivalent μg/L. Acre-ft, which is an archaic unit of volume, is often used for flow. Because of the use of US units a simple set of equations on page 3b-31 which in the SI system would be dimensionally consistent has to include a set of conversion factors. Conversion to SI units are not provided. They should be provided in parentheses if the writers select to keep the archaic US units. Note: This comment is supplied by the "A" reviewer who is new to this chapter. The "AA" reviewer notes that similar comments has been leveled by virtually every reviewer upon first reading, yet the District still mixes and matches units throughout this chapter year after year. Perhaps it is finally time for District to heed the comment.

The individual reporting for each STA followed a prescribed format. At some STAs vegetation management to control emergence of "undesirable species" such as lettuce and water hyacinth relied on the use of herbicides which apparently was claimed to be safe. Although the water hyacinths are often in Florida a nuisance plant without control, water hyacinths are known as very effective, more than 90%, removal of nutrients and could be let to grow in some cells. They could be removed by cutting and the biomass can be used for production of biofuel. However, this may be very risky on such a large scale.

The chapter also commented on dry-out impact and some difficulties with keeping the wetlands hydrated. In addition to bringing flow by pumping from other parts of the Everglade system and from Lake Okeechobee, a method used in some designed wetlands is to implement recycle of the outflow which may further improve the removal efficiency. On the other hand, the SFWMD study found that periodic shorter duration dry-outs are detrimental to SAV communities but not to EAV wetlands. It could also be pointed out that especially in Europe dry-outs have been also used for preventing and control of highly undesirable cyanobacteria which have been found in many Florida lakes.

A special study starting on p. 3b-53 describes in detail the attenuation of the components that comprise TP, i.e., dissolved ortho phosphate, and soluble reactive phosphate. TDP, TOP and SRP were not defined so the description in this review is just a guess. The logical conclusion of the study was that removal of TOP and SRP is slower than that of TP which contain a large portion of particulate P.

---

[1] Water Pollution Control Federation (1990) *Manual of Practice for Design and construction of urban Stormwater Management Systems* , Alexandria, VA

The organization of the studies on areal extent of SAV in the various STAs is quite cumbersome. In general, these are not separate studies but rather one study with virtually identical methods applied to several field sites. There are some minor differences in the water quality measurements but the dependent variable is always SAV coverage. It seems it would be easier to followif goals, objectives and methods were defined once and the Results and Discussion be organized by STA.

Throughout the SAV studies water quality results are typically presented as outflow and occasionally inflow concentration data. Yet inflow and influent concentration varies temporally and spatially and various cells are of different sizes and shapes making any comparison of simple effluent concentration data dubious and dependent on non-scientific phrases such as "for nearly identical flow rates". Results and analysis should incorporate commonly used normalizing parameters such as influent/effluent ratios, flow weighted means, and hydraulic residence times. Likely this observation is just a symptom of a larger issue with this chapter, "facts" are presented with little to no attempt at synthesis of the results, a comment that has been raised in previous reviews.

On a practical note, the reviewers question the investment the District is making in the use of SAV in so many cells. Clearly this vegetation assemblage is difficult to manage with susceptibility to storms and dry-out, unexplainable die-off, predation, uncontrollable population shifts and generally large areas of open water where plants should be. The District appears to be sold on fact that it is the best community assemblage for P removal, and it clear that very good P removal is possible when everything is working (and growing) as designed. However, the terminated study and much of the data presented in earlier portion in this year's chapter and data from previous years indicate that equal, or nearly equal treatment is possible with EAV or possibly FAV, with far fewer management issues. Until the District can prove that it can grow and maintain stable SAV communities, it should stop or slow additioan conversion of cells to SAV.

Specific questions and comments by line number as appropriate.

172-192 (and in other sections): An explicit description of how oft-mentioned parameters are calculated (as suggested last year) is much appreciated as it removes any ambiguity of what the parameters represent.

286-288: The eastern flow-way of STA-1E was also dry as mentioned in other locations.

365: A better quantitative description than an "enormous" population increase of island applesnails, e.g., # of organisms/m^2, should have been used instead of dramatic but inappropriate term "enormous". This is repeated in other sections of the report. The special study on STA 1-E attempted to explain the emergence of applesnails that decimated SAV in the cell which resulted in a decline in efficiency. While this study reported the facts and consequences it did not attempt to find causes and possible means of prevention.

392 (and similar sections): Note this is a repeat comment from last year. It is unfortunate that bird nesting success is at odds with P removal objectives. While not an expert in avian behavior,

it would seem that these conflicts could be minimized if certain small areas within the STAs were dedicated as bird nesting sites. This could be done by construction of bermed areas or perhaps more effectively by construction of small "islands" within the cells. In fact, areas of higher topography would more closely mimic the original landscape of islands and sloughs and have little influence on the hydraulics of the individual flow-ways or cells if oriented properly. It could be an extremely cost-effective way to minimize habitat and water quality conflicts.

412: In general, the Latin names of all species should be given the first time they are mentioned. While scientific literature protocols would demand this at all locations, the reviewers are fine with use of common names in this Report, provided they can identified with the scientific name when first mentioned.

572-575: Good performance of STA 1 and 2 is attributed to the fact that they have never been farmed which supposedly reduced the P flux from the soil. However, the performance was not different from some other STAs which were farmed.

981: This section describing SAV coverage typically uses mg/L as the unit for P concentration (except Figure 5B-23), while the rest of the document typically uses micrograms/L. Units should be consistent, at least within a section.

984-986: A complete double check of dates mentioned in the document is warranted. Clearly SAV establishment could not have started before the STA became operational as stated here, but in a check of which value is right this reviewer notes that Table 5B-1 stated the STA started in 2004, line 314 states 2005, as does Figure 5B-5. Some confusion is inherent when WY does not match calendar years, but a start date cannot span three years.

986-989: See the comment line 412.

1114-1122: Displaying results only as measured concentrations is probably missing (or masking) many important factors driving performance of the system. Even if flow conditions are similar between the four spatial gradients shown in Figure 5B-27 (flow rates not provided), clearly influent P concentrations are not, as shown in Figure 5B-28. A more telling story might be revealed if relative concentrations (C/C_0) were plotted. This would normalize for concentration variability and to a certain extent, flow variability as well.

1148: The previous two vegetation studies described how the vegetation density measurements were taken, but this one does not. Presumably by the same or very similar methods, but this is not mentioned.

1180-1182: These date suggest that the non-uniform width of this wetland reactor has a negative influence on P removal. Not only should gates on the south side be shut, but perhaps the entire rectangular-shaped extension to the south should be abandoned to the flow path of this cell. It would not be surprising to discover that other non-uniformly shaped cells throughout the STAs have responses similar to this one and, now that expansion of STAs areal distribution is complete, tweaking of the plan form shape within individual flow ways and cells is appropriate.

1199-1207: See the comment line 1148, while a little more information is provided here, it is incomplete. If vegetation sampling is indeed the same thought all these sections, perhaps a detailed description could be given once and subsequent sections reference the method, noting anything that might have been different.

1250-1253: Another alternative is that P removal is achieved by the dominant vegetation types in each of these cells, including the upland/ woody species. Note: Are the data in Figure 5B-43 for the entire flow-ways (including the upstream EAV cells, or for just the SAV cells? If it is for the entire flow-way no meaningful comparisons can be made between vegetation types in the downstream SAV cells since inputs to those cells are not known.

1311-1313: Abundance increased from when, 2011? This is not clear as Figure 5B-45 shows a decrease with time (within the WY 2014). When did the manual inoculations occur? Is the success visible in the data presented?

1311-1350: Why are the April 2014 survey data not shown in Figures 5B-46 and 47? The text indicates the SAV die off was profound and influenced effluent P concentrations yet the graphics suggest a rosy picture with nearly full coverage of SAV. In fact this entire section is very confusing with statements indicting increased SAV coverage (within a limited range) while in general coverage decreased; that die off in cell 2B is what caused its exiting P concentration to increase relative to the adjacent flow-ways when, in fact, all flow way SAV cells had a massive die off. There is no synthesis of the facts relayed in each statement, which when compared to each other, seem contradictory.

1470: What is the P:N ratio? If it is phosphorous to nitrogen, how does its value relate to the rest of the discussion in this paragraph focused on the ratio of live to senesced leaf litter nutrient concentrations?

1476-1477: This is a frighteningly damning statement. It clearly implies the study was terminated for the simple reason it was not providing the result the District would like to see. Virtually all the field data in this and previous reports suggests that when operating properly effluent P concentration are going to approximately 20-25 ppb regardless of retention time, vegetation type, etc. (see for example the data in Figure 5B-27). Considering than influent values in this study are near this background level, these data seem to confirm that even under controlled conditions and regardless of vegetation type, this is the lower limit of possible performance of the STASs. It is interesting to note that the following discussion has some of the best synthesis of the reported data of any study reported in this chapter with potentially important ramifications for future research and operational guidance, yet it was the study that was terminated.

1507: Should the word between above and below be ***and*** or ***to***? The distinction is important to the meaning of the sentence.

1531: Is the litterbag study the same as the decomposition study?

Figure and table comments:

In several figures (typically maps) interior typeset is too small to read at 100% magnification. Either the type should be made bigger relative to the figure or the figure should be made bigger on the page. A good rule of thumb is type size should be as big as the document text at the same magnification. Affected figures include: 5B-20, 5B-25, 5B-26, 5B-30, 5B-33, 5B-34, 5B-45.

Figure 5B-25: What are the test cells identified in the legend?

Figure 5B-36: Clearly the concentration is higher from the hydraulic dead-zone of this cell, but does the shape influence the flow rate of individual culverts? My guess is kinetic head is negligible at all culverts and flow is virtually the same through each culvert (assuming diameter and elevations are equal). However use of flow-weighted concentration values or mass load instead of raw concentration would avoid any ambiguity.

Figure 5B-48: It might be beneficial to shade the temporal region of these data to correspond to the total SAV coverage measured throughout this period of record. Apparently (see comment from lines 1311-1350), part of the story is that the timing of SAV die off varied between flow ways and the effect can be noted the P concentration data. This is not clear from the discussion however.

Editorial comments by line number:

202: …***manufacturers.*** None of these products ***bioaccumulate*** and…

877: were less ***than*** the rates

894: (Cells ***6-2***, 6-3, 6-4, and 6-5)

1039: ***defoliation***…

1216: prior ***to*** being dewatered

1222: southernmost region and ***has a*** lower

1290: geo-references ***stations*** in each cell

1474: seven month ***initial*** period

2015 SFER Chapter 5B Performance and Optimization of the Everglades STAs
Reviewers Stein (AA) and Novotny (A) Note that the A reviewer on our original review was incorrectly listed as Peter Dillon when it was in fact Vladimir Novotny.

In the following closure, the reviewers will focus on the issues where there were differences of opinion between reviewers and authors or where reviewers make additional recommendations. If a comment is not specifically numbered below, the reviewers believe the authors' response satisfactorily addressed the original comment.

Comment #1: The goal of consistently achieving 10 µg/l of total P in the outflow from STA's is unrealistic. As a matter of fact, a study in this chapter on removals at low concentrations documented that at low concentrations the removal is negative. This is especially true for STAs that have several cells in series that imitate plug flow whereby the last low concentration section may not perform as intended.

Comment #3: The statement that English units are standard for the District seems disingenuous. Other units are used in this chapter and most other chapters are rather consistent in the use of metric units. However we acknowledge that English units are still used in practical engineering and hence, while units such as cfs and acre-ft are archaic, prone to making conversion errors and should be replaced, we can, for the time being, accept their usage. We draw the line, however, at the use of ppb and ppm versus mass/volume concentration representations in which metric units would be mg/l or µg/l . The problem with ppb and ppm are that they are mass/mass or volume/volume and do not represent a standard mass/volume concentration definition. Therefore they are not only archaic, they are not explicit and not used anything but layman literature. For this reason the District (or the authors) must abandon ppb and ppm as a unit. The response of authors is not satisfactory.

Comment #4: Harvesting vs. Killing: Killing vegetation by herbicides and leaving it in the wetland will accelerate the natural processes of eutrophication of shallow water bodies to a marsh filled with organic muck and loss of efficiency. Experience with shallow ponds used for treatment of urban stormwater documents that such ponds will be filled with organic sediment in a matter of 10 to 15 years and need to be either dredged or abandoned. Leaving large masses of organic matter in the lagoon is short-sided. The interest in using biomass for biofuel production is rapidly increasing and is a hot topic in this century.

Comment #5: The reviewers were hoping to see a more explicit response to the comment regarding the potential for recycle as a mechanism for preventing dryout and perhaps increasing performance, especially in cases of disturbance causing one cell to behave poorly.

Comment #6: It is lot to ask any reader, but especially a reviewer of an 80+ page document, to remember the definition of an acronym they may not be familiar with defined 10 pages prior, yet we understand the need to use acronyms. The solution is to create a table of acronym definitions that can be referred to as needed. This is a resurrected comment made in consecutive years several years ago. For a year or two following, such a table was provided with this chapter, but now has been abandoned again. This table is as, or more, important than a table of conversion factors.

Comment #8: The response is contradictory with itself! We agree documentation is important, but, in and of itself, serves little purpose. The whole point of the vegetation documentation is, as the last sentence in the response states, to quantify factors affecting performance. Yet if it is related to just concentration at the outflow it is highly unlikely the District will able to assess how or why the vegetation is influencing performance. The normalized variables originally suggested are standard values used to make the assessments the author claim to be shooting for. The response of authors is not satisfactory.

Comment #9: The second part of the authors' response is not clear. What are the treatment functional redundancies? Further, the effort to get very low concentrations in the outflow by adding new cells may not be successful in view of the low concentration study in the chapter. Also because of the buildup of organic matter in the sediment without removal, the efficiency of the cells may be diminishing in the future. The cells have evidently been in operation only a few years.

Comment #13: While the reviewers can see the point of potentially exacerbating bird nesting with the creation of islands in the STAs, there is no data to support that position and it is also possible that it could ease the problem. Perhaps the best to hope is to agree that we disagree on the potential benefit of a limited scale island building campaign.

Comment #14: See our response to Comment 6. Perhaps a table cross referencing scientific and Latin names would help the reader (and reviewer) keep things straight.

Comment 21: The reviewers were referring the completed (nearly completed?) build-out of other STAs. The more important part of the comment was that these data seem to confirm that highly irregular surface area shapes of individual cells lead to hydraulic dead zones (in this case) or possibly short-circuiting in others. We think the need to address this is beyond "potential".

Comment #27: The reviewers did not make this statement lightly though we should not have presumed the District's intentions. However, in our opinion this is probably the most comprehensive and best-analyzed study reported in this chapter of the SFER over the last several years. Yes, as with any experiment, there are unforeseen circumstances that alter the original experimental design (e.g. growth of SAV, use of mesocosms, etc.), but the issues described hardly make the results "uninterpretable". In fact, the interpretations are quite clear to us and not as encouraging as anyone would hope. If the study was indeed terminated for the reasons stated, then a similar study which addresses the operational issues discussed should be reinitiated. Mesocosm studies do have potential scale up issues, but they are an extremely cost effective and proven way to compare alternative factors such as plant species, soil treatments etc. in replicated, controlled conditions and while the magnitude of water quality parameters might not be exactly what would be seen in a full scale system the relative results are almost always what will be seen in the field under ideal conditions. The response of authors is not satisfactory.

Comment 29: The study should be referred to in the same language throughout the document.

Key Findings and Recommendations (SOW Task 3)

- The POR data for P loads and concentrations to and from the STAs indicates that they are, in general, effectively removing large quantities of phosphorous that would otherwise be entering the EPA. Additional experimentation and operational tweaking is still required to assure that these systems will meet the mandated removal targets consistently over the long term.

- From a scientific perspective the design of experiments intended to optimize these systems is sub-standard. More emphasis on the research methodology should be incorporated at the experiment conception level, rather than the "opportunistic monitoring" approach that is too often currently employed. It is not clear why certain well-conceived relatively cost-effective experiments, such as the mesocosm study, have been terminated while other more poorly conceived and costly experiments continue.

- The current operational objectives of the STAs make these locations the ecological "sacrificial lambs" in order to achieve ecological restoration of downstream locations. The District is encouraged to re-examine some operational guidelines to determine if some ecological function of these locations (such as bird nesting) could be better incorporated without compromising other objectives.

- The reviewers acknowledge that the above three findings are identical to the ones stated last year but this represents a certain level of frustration on our part. Year after year we make similar recommendations and suggestions and all too often the same comments and suggestions are still relevant the following year. Sadly, that is the case again this year and we wish to emphasize that fact by this fourth comment.

- There are potential long-term ramifications of killing undesirable plant species in situ by herbicide application and leaving the remaining organic matter to decay. In some STAs the District has spent considerable effort to remove organic muck and native soil to improve performance. Harvesting (especially floating mat vegetation) should be considered as an alternative..

Comments on Chapter 5C: Update for the Restoration Strategies Science Plan
AA reviewer – P. Dillon
Accountability Review

General

a. Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate?
This sub-chapter presents a brief overview of eight studies that are in progress and that are the core of the Science Plan for the Everglades Stormwater Treatment Areas. The Science Plan and the studies all focus on methods to reduce the output of phosphorus from the STAs. Because all of the studies are in an early phase, they understandably have no real findings to present at this time. The study schedules are realistic, the progress to date, although brief, is informative and the future activities consistent with the detailed study plans.

b. Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report?
The material is presented in a very clear and logical manner with the exception of the second project, and is consistent with the detailed Science Plan included on the SFER website. For the second project, I had a difficult time following what was actually going to be done (and why) without reading the detailed study plan on the website. I particularly liked the chart provided with the fourth project on STA 3-4 Performance, Design and Operational Factors. A comparable chart for the other 7 projects would be welcome.

Are findings linked to management goals and objectives?
All of the studies contribute very directly to the management goals and objectives in that they all focus on ways of reducing P output from the STAs.

Is there any constructive criticism and guidance to offer for the District's large-scale programs?
There is considerable information in the scientific literature on nutrient removal in groundwater using e.g. applications of iron-containing mixtures as a barrier to P flux. Some of these methods/substances could be useful in the first project.

2015 SFER
Chapter 5C Update for the Restoration Strategies Science Plan
Reviewer Dillon (AA) – Accountability Review

Closing Comments

The chapter adequately fulfilled its intended purpose of providing an overview of the studies that are in progress as the core of the Science Plan for the Stormwater Treatment Areas. Each project has the potential to contribute to the goal of reducing TP loading to the system, although most are, at this stage, far from complete. The chapter is well-written and clear in its presentation of the current status of these projects. The comments made by the reviewer have been responded to adequately.

Key Findings and Recommendations

1. All of the studies can contribute directly to the management goal of reducing TP loading.

2. Proposed future activities for the coming year are consistent with the long-term plan and the goals of the District, and should proceed.

Review of Draft 2015 SFER Chapter 6
Comments by Walter Dodds "AA" reviewer

Overall this series of reports on ongoing studies is well written and represents an impressive body of work related to conservation of and restoration of the South Florida/ Everglades complex. Most of the comments here are minor points or suggestions for future research.

1. The hydrologic patterns section makes little sense where it is relative to most of the discussion, as most of the text is dedicated to how the patterns influenced the wading bird populations, and some of the discussion on the plant communities. Statements that water levels were "good" or "bad" make little sense without reference to the following section.
2. In the wading bird section, most of the discussion requires the reader to take statements on faith with respect to controlling factors. There appear to be enough data to create a conceptual model of shorebird recruitment (via structural equations modeling or a similar approach), and test it with some leave-some out approaches, creating plots of observed verses predicted nesting rates for each species. The chapter 8 approach to modeling P inputs and retention is an example of such modeling (though the bird modeling will be more variable because more factors may play in). Still, this approach would allow parsing out controlling factors and make for a stronger presentation of the year's results and linkages to hydrologic patterns. It is particularly concerning that tricolor herons are doing so poorly, and good predictive models are really needed to solve this problem.
3. Line 29. Too many significant figures
4. Line 74. What does "widely correlated" mean?
5. Line 335 "Slough as a result of favorable hydrologic conditions" this kind of statement can be found throughout this section. What exactly does favorable mean here, most of it seems to be value judgments based on maintain waterfowl, but explicit explanation and a more scientific approach to the presentation would be better in this entire section.
6. The warmouth movement section is a nice addition with the second year of data. It would be good to combine movement data with gut content data so some of the speculation about moving to shallow areas to feed can be backed up with data.
7. The BOH might also be useful for rapid detection of the extent of harmful algal blooms. It would be nice to see some quantification of the actual spatial resolution of the images and where they begin to blur.
8. The peat bag experiment is interesting and innovative, longer term results will be needed to prove efficacy, I would view this as a preliminary or pilot experiment.
9. Line 858. The discussion has the first mention of herbivory, was this formally measured?
10. Table 6.4 N P and C units, moles per what?
11. Line 980. Do you mean 5 species were responsible for over 70% of the dry mass of leaf fall?
12. Line 1007. Do you mean relatively low soil P concentrations? There seems little pattern in N
13. Line 1037. Interesting observation, is leaf quality changing? Maybe anoxia stimulates growth to a bit then pushes over a threshold and causes death?
14. Line 1044. Soil nutrients do not operate over large scales.
15. Figure 6.28 and others.. Any ideas on why the bimodal distribution of growth rates from head to tail of island? Pretty interesting trend.
16. Line 1090, tense problem

17. Line 1139. It is too bad total N and P are not reported, could these parameters be added in?
18. The Florida Bay story could use a bit more analysis, what is driving the variation in chlorophyll?
19. Figure 6.33 can get rid of y axis 2 decimal points numbers… round to whole number.
20. Figure 6.34 Y axis units?
21. Line 1234. Several quadrats? How many?
22. Line 1285, what does bleed through mean.. Interpolation turns hard boundaries into smoother gradients than they really are?
23. I don't understand figure 6-36 b
24. The entire section on Synoptic Florida Bay Ecosystem Assessment is a little thin on data presentation and analyses.
25. Line 1323. This is a strange methods section, could use a little more information here.
26. The SUBMERGED AQUATIC VEGETATION MODEL ANALYSIS section has nowhere near enough information to be evaluated. The figure has no axes labels and means little.
27. Line 1413. Flow is discharge or water velocity?
28. Line 1538. I don't understand how SF6 can be used as a flow and dispersion tracer without correction for aeration. Does this just mean the flow paths, but no rates?
29. Line 1575. I had not heard of this... it is just a plain cool method!
30. Line 1632. But it should not vary from the fluorescence for that reason…
31. Line 1634. And loss to the atmosphere
32. Line 1664 and elsewhere, use real symbol for microns
33. Fig 6.42 is sort of difficult to read and figure out
34. Line 1707, but not sf6, which could be related to not getting aeration effects out of the picture.
35. Figure 6.44 too small to read effectively.
36. POTENTIAL FUTURE CLIMATE SIMULATION WORKSHOPS: this section is difficult to really evaluate directly because it represents so much work with a modest amount of information, but it is good to see that the team is working on this and a special journal issue will be more generally useful than this report.

Comments by Otto Stein "A" reviewer

Note that Appendix II indicated that the level of review is at the "Technical" level.

The review is divided in to several sections. First are several broad questions and comments regarding the interpretation of reported results followed by questions about specific sections, sentences, phrases, etc. This is followed by comments on figures and tables and lastly, editorial suggestions.

Broad questions and comments that should be addressed:

The chapter provides an overview of the current year's efforts to assess the success of activities designed to restore the ecological health of the greater Everglade region (from the WCAs to

Florida Bay). In general, the chapter provides an appropriate summary of the current year's monitoring activities. The Chapter is well-organized and Table 6-1 provides the reader with a summary of the specific studies reported on more detail throughout the chapter.

The breadth of the district's research program is apparent from the studies presented; from the scientifically dubious, but practically relevant study on floating peat bags to the technology-focused linking of remote sensing and ecological models to predict cattail expansion, to several intensive monitoring programs. Most of the studies are adequately described to give the reader a sense of what was done and what was discovered in an abbreviated format but a couple (mostly those that are more complex) need a little more information. Since the review is technical, we are to provide guidance as to whether we believe the findings and conclusions are supported by the information presented. When the answer to this is no, it is due more to lack of completeness of reporting, than a lack of proper analysis. The more data dense studies (such as the Decomp Physical model) are more likely to be in this category. Suggestions for improvement are given by line and figure number.

Specific questions and comments by line number.

163-167: These two sentences seem contradictory; if the peat drought tolerance criteria is one foot below the ground surface how can the surface be "essentially the same as the threshold"?

333-354: This section seems incongruent with previous discussion and the data in Figures 6-9 and 6-10. The management importance of Taylor Slough is lost to the reader. What are the purpose and/or the target salinities that the South Dade Conveyance System is attempting to remedy? The second paragraph claims salinity has increased above the mean and MFL criteria but this is not supported in the figures. Are the high values collected in WY 2015 (after April 2014?). If so, then the last sentence should reference WY 2016; if not, then something else is wrong.

427-428: What interval is to be returned to?

509-510 (and earlier): What is the difference between tactile and visual foragers?

517: This is a nice study providing data useful to water management decisions.

643: "Unmanned aerial systems" is certainly an obscure term for what I believe you are talking about. Why not "unmanned aircraft" so the reader does not have to guess what you are talking about.

710: What is the "head" of a tree island? The upstream point, high point, or something else?

734: Is Dineen Island a ghost island or a remaining healthy one?

801-805: Need a more clear description of the experimental design.

821-823: This statement begs an obvious question in my mind; Why not get rid of the pot altogether and plant just the root plug into the peat bag?

826-849: This should be reworded to make the experimental design more clear.

891: While the potential utility of this study is clear, analysis of the reported data is typically a comparison of litterfall and growth rates <u>within</u> a healthy tree island or a degraded tree island. However the stated objectives are to <u>compare</u> pristine and degraded islands. More analysis should be dedicated to the stated objective of the study.

964: A measure of diameter cannot be in mm^2. Do you mean basal area calculated from circumference measurements?

1037-1041: The paradox is not well explained, please reword.

1297-1304: The text puts some importance on flow from Trout Creek, but this is not shown on the map in Figure 6-35. Is Trout Creek mislabeled as Joe Bay?
1522-1523: The locations of the backfilled canal sections (and other important features influencing the results of the high flow experiment should be included in a location figure.

1539 (and Figure 6-40: What/where are the C1 and RS1 (and RS2 in later lines)? These are neither identified or located in a figure.

1637: Figure 6:41 does not show results of temporal variation of velocity as indicated in this line.

1654-1658, 1661: Something is missing here. TPP is assumed to mean total particulate P, but this is not stated. Mass of what decreased? How does this contrast to TTP?

1667: Changes of what?

1671: Does construction refer to canal backfilling or something else. A lack of information on what was done at the site hinders the readers ability to grasp the essentials of the discussion in this section.

1719: What canal? This section needs some introductory text to put in context with the rest of this study.

1724: What are DB and UB sites?

Figure and Table comments:

Fig 6-1 through 6-7: Reported data is inconsistent between figures and in some cases not consistent with the corresponding caption. Sometimes upper and/or lower criteria and/or ground surfaces are shown, sometimes not. All graphs should show all important parameters. Note the related comment for lines 163-167.

Table 6-3. The historical range seems to be missing, or is “historical range” and “target” the same thing?

Table 6-4: Basal area units are wrong and the unit tree/ ha (in footnote) is not used.

Figure 6-19: Legends are too small to be legible.

Figure 6-30: The caption is a mouthful; (“the” in the second should be removed) and the definition of the temporal and spatial medians and “interquartile range” are not provided in the text.

Figure 6-32: There appears to be no blue region in the graphic as described in the caption.

Figure 6-33: What does the dashed vertical line represent? Line legends should be included in the graphic, rather than in the caption.

Fig 6-38: The green line is not labeled and the legend (assumed to be species) should be consistent with the text (species rather than genus names). The Y axis (and X axis for that matter) is not labeled.

Figure 6-40: Is SF_6 sulfur hexafluoride, the dye used in the tests or something else? (It is not defined in the figure caption, but is later in the text.)

Figures 6-41, 6-42 and 6-43: All labels and line weights are too small to be easily read.

Figure 6-43: There is too much information in one graph. Panels are not properly identified and caption does not capture the complexity of the data.

Editorial page and line comments:

173: …recession rate of 0.04 foot ***or less*** for a week…?

359: common and ***a*** defining feature

677: image resolution ***at heights*** of 91.4 …

701-743: This introductory section needs editorial help. A couple of sentences are run-ons, and/or contain the wrong verb tense. In some cases (see questions on lines 710, 734) the intent of the authors is lost.

1548: …direction ***were*** measured…

1653: …were ~~the~~ identical…

Final Bullets Dodds and Stein Chapter 6 Everglades Research and Evaluation

- Overall this series of reports on ongoing studies is well written and represents an impressive body of work related to conservation of and restoration of the South Florida/ Everglades complex.

- The chapter provides an overview of the current year's efforts to assess the success of activities designed to restore the ecological health of the greater Everglade region (from the WCAs to Florida Bay). In general, the chapter provides an appropriate summary of the current year's monitoring activities. The Chapter is well-organized and Table 6-1 provides the reader with a summary of the specific studies reported on more detail throughout the chapter.

- The authors have done a good job to responding to suggestions for improvement.

- The warmouth movement section is a nice addition with the second year of data. It would be good to combine movement data with gut content data so some of the speculation about moving to shallow areas to feed can be backed up with data.

- The BOH might also be useful for rapid detection of the extent of harmful algal blooms. In the future it would be nice to see some quantification of the actual spatial resolution of the images and where they begin to blur.

- The peat bag experiment is interesting and innovative, longer term results will be needed to prove efficacy, but the pilot results are intriguing.

- Dispersion is both the process of fast flow paths (which can be traced by the leading edge versus the peak, and slow flow paths (which requires an inter tracer to correct of aeration). In the future better hydrologic data and aeration rates could be obtained by including inert solute tracers with SF6.

- POTENTIAL FUTURE CLIMATE SIMULATION WORKSHOPS: it is good to see that the team is working on this and a special journal issue will be more generally useful than this report.

Comments on Chapter 7: Status of Nonindigenous Species
AA reviewer – P. Dillon
Accountability Review

General

I have used the specific questions asked in the Statement of Work as a framework for this review

a. Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate?
This chapter provides a thorough account of management activities related to eliminating the effects of non-indigenous plant and animal species on the environment of South Florida. Although eradication is the preferred solution, this is often impractical or impossible, and controlling the spread of invasive species may be the most that can be achieved. The report acknowledges that there are many non-indigenous species for which there are no data, but it does seem that attention is given to those with the greatest potential to damage the ecosystem.
It appears that the programmes in place have reduced the extent and impacts of several key invaders, although the region is clearly under pressure from a continuous influx of new species.

b. Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report?
I found this chapter to be well-written, very well-organized and informative. The current status of the most important non-indigenous species is outlined and any control activities that are in place are updated.

Are findings linked to management goals and objectives?
The linkages between the monitoring results and management objectives are very clear. The amount spent ($19 M) on control programmes in one year alone attests to the importance of this issue, as well as to the priority given to controlling these species.

Is there any constructive criticism and guidance to offer for the District's large-scale programs?
Biological controls for melaleuca, water hyacinth and old world climbing fern have been introduced. The scientific literature is full of examples where biological control has had unexpected adverse effects. Chemical controls can certainly backfire as well, although impacts of chemical treatments usually are relatively short-lived. Biological controls that have gone wrong can be permanent. Chemical controls, however, would likely be ineffective for some of the problem species, both plant and animal. It is important that extensive studies be carried out before any biological control agent is introduced.
Secondly, it is obviously easiest to eradicate or control an invading species at he earliest time od discovery. A commitment and plan to eradicate, if possible, any new species upon discovery would seem to make sense.

2015 SFER
Chapter 7 Status of Non-indigenous Species
Reviewer Dillon (AA) – Accountability Review

Closing Comments

This chapter outlines in detail activities related to the management of non-indigenous species in south Florida. Although the task is daunting, some measure of control has been achieved for the potentially most damaging species. However, new introductions occur regularly, making this a never-ending battle.
The chapter is well-written and thorough in its outlining of activities. The work undertaken is directly linked to management goals and objectives. This is essential work if the region is to recover its ecological integrity.
The authors of chapter 7 addressed the reviewer's concerns about the use of biological controls vs. chemical controls very well.

Key Findings and Recommendations

1. Non-indigenous species present a very serious threat to the south Florida environment. Active management approaches involving control of the most damaging species is having some success.

2. The District should continue to use extreme caution when employing biological controls involving the introduction of new species.

3. Rapid response to contain and potentially eradicate newly discovered species should be emphasized across state and federal jurisdictions, and should be built into standard operating procedures.

Dodds secondary reviewer Chapter 8

1. Overall, this is a well done document and the district continues to make progress on problems plaguing the Okeechobee system. I suggest more statistical analyses of temporal trends, more emphasis on algal bloom sampling, and some biotic index work to get at N versus P effects on biotic integrity.
2. Line 87. This is an annual flux… if so it should indicate per year in the units here and elsewhere
3. Line 129. Equipment limitations should be remedied in case of a high water year again.
4. Line 360. Removed P from the basin or sequestered it from the surface water?
5. Table 8.2 and 8.3. Statistical analyses of these data could help indicate if actions are truly having a significant effect and establish confidence bands around the trends. In general it would be good to know if nutrients are being decreased or held steady in the face of increasing loads. With respect to P control, treatment rather than stopping sources is probably not sustainable in the long term
6. Table 8.6 and elsewhere, as usual the mixing of SI and American standard units sets me on edge… but I guess it is a necessary evil
7. Table 8.7 and others, the bars are cool… but the totals row is not really totals, sometimes it is means. If it was always totals then it would always be 100%. As is, this format is somewhat misleading.
8. Line 584. Our experience with autosamplers is that the lines become contaminated, and in low nutrient environments cannot be trusted.
9. Line 634. Statistically lower TP?
10. Table 8.10. Clarify if ratios are molar
11. Figure 8-22, what is the x axis? The bars look weird, flat standard rectangle would be easier to interpret
12. Fig 8.23, Put what all parts of box's whiskers etc in legend.
13. Line 1071. It is hydrologic elsewhere and hydrological here.. pick one and stick with it, preferably the former.
14. Fig 8.24. What are units on y axis?
15. Line 1109. These numbers should be reported as per unit effort or they mean little, so just stick with the data in Figure 8.28
16. Considering the relative potential for toxic blooms, I am surprised more effort is not going into algal bloom sampling.
17. Line 1295. Are these correlations reported or plotted somewhere? I would be interested in more details.
18. The literature cited is rather modest for a 70 page technical document.

Comments on Chapter 8: Lake Okeechobee Watershed Protection Program Annual Update
AA reviewer – P. Dillon
Technical Review

General

As in past years, this chapter includes detailed information about, and analysis of the measures implemented in the past year under the Lake Okeechobee Watershed Protection Plan to reduce phosphorus and nitrogen loading to the lake, about the nutrient status of the lake including external loads and in-lake chemistry, and about the ecology of the lake with respect to submerged and emergent aquatic vegetation, phytoplankton including algal bloom formation, exotic species control, fish, and wading birds.

The chapter is well-written, concise and easy to follow. The conclusions are supported by the data presented. As was the case last year, the writers have produced an excellent summary of the ongoing work on Lake Okeechobee and its watershed, and have provided a very clear picture of the status of the lake. They have demonstrated some progress towards environmental targets, although others remain elusive.

In particular, the greatest concern remains the progress, or lack of it, made towards reducing the TP to the target value of 140 metric tons/year. Although there have been substantial efforts to reduce loads, many of which have had significant success, and many of which have cost large amounts, the total phosphorus load is relatively unchanged in WY2014 compared with recent years and the in-lake TP concentration has not improved. In fact, the calculated load was a little higher in 2014 than in 2013; however, this increase is attributable entirely to the higher flows. I suggest that the authors take an additional approach to interpreting loading data. Loading is the benchmark parameter that we all use but its actual importance is that it is directly related to in-lake concentration of P (or N). Using loading as the key parameter has drawbacks, notably it can be driven more by hydrology than by any actual change in land use, technological fixes, etc. I suggest that an equally useful parameter is annual mean volume-weighted inflow concentration; this "corrects" for year-to-year changes in hydrology and has a more direct relationship with in-lake concentration. This can simply be calculated as the total load divided by the total input (or output) of water (corrected for change in lake storage of water). The load from precipitation (35 mt for TP) is not likely going to be reduced; this input of P (and that of N) as well as of water could be excluded from the calculations to see how the sources that can actually be managed are changing from year to year.

In terms of specific projects, the following comments are provided.

Information on the various construction projects including the Stormwater Treatment Area (STA) projects is very nicely summarized in Table 8-1. Most of these projects are proceeding well and appear to be contributing significantly to reducing the P load. Implementation should lead to significant further reductions in P leading to the lake. The HWTT technology is showing much promise for additional reductions in load and should be continued and possibly expanded.

As mentioned in the review of the 2013 report, the water quality monitoring and the biological monitoring are well-conceived and appear to be well-executed. It seems that there is more

attention paid to the problem of algal blooms than last year, particularly of cyanobacteria, which is good, but more still should be done. The increase in SAV along with the decline of *Chara*, and the generally good diatom:cyanobacteria ratios are positive signs. The fish and wading bird data indicate recovery moving forward, although slowly, from the effects of severe weather events in past years. One very useful addition to the report in following years would be more information on oxygen levels in the lake. Despite its very shallow nature, the lake is so productive that oxygen depletion needs to be considered. More statistical analyses of temporal trends is definitely warranted, as is some biotic index work to better understand N versus P effects on biotic integrity.

Specific comments

l. 88 – units are still somewhat strange. If you use lb/acre for yield (which should in fact include a time dimension, i.e. lb/acre/yr), why mix concentrations in metric? If concentrations are in ug/L or mg/L why change to lb?

l. 129 – this is too important to allow equipment limitations to prevent adequate data collection; the limitations should be remedied in case of a high water year in future

l. 360 - removed P from the basin or sequestered it from the surface water?

l. 411 – TP loads increased but less so than the flow. Therefore the volume-weighted inflow concentration decreased, and this is important because load is only important in that it is directly related to lake concentration. So if inflowing concentration drops then lake concentrations must drop (barring increase in internal sources).

Tables 8-2, 8-3 - statistical analyses of these data could help indicate if actions are truly having a significant effect and establish confidence bands around the trends. In general it would be good to know if nutrients are being decreased or held steady in the face of increasing loads. With respect to P control, treatment rather than stopping sources is probably not sustainable in the long term

Fig 8-5, 8-7 – this is a very nice method of presenting the data graphically

l. 576 – these seem like very important data; why was their collection reduced?

l. 584 – both reviewers have had negative experience with autosamplers; the lines easily become contaminated, and the sample bottles accumulate P on their walls via bacterial and algal formation if left for more than a day. In low nutrient environments they cannot be trusted.

l. 634 – this is an example where statistical analysis would make a much stronger argument – is TP really lower TP?

Table 8-10 - clarify if ratios are molar or mass

l. 790 – it would help the reader who is not familiar with macrophyte ecology if the report identified which macrophytes (and later, same comment for fish) are considered desirable and which are not. Is this based entirely on exotic vs. not exotic?

Fig 8-23 – it would be helpful to put what all parts of Box plots are in the legend.

l 1071 - it is hydrologic elsewhere and hydrological here; pick one and stick with it, preferably the former.

Fig 8-24 - what are units on y axis?

l. 1109 - these numbers should be reported as per unit effort or they mean little, so stick with the data in Figure 8.28

l. 1295 - are these correlations reported or plotted somewhere? More details would be interesting.

l. 1311 – there is now considerable information about bioaccumulation of microcystin in fish – is there any monitoring of the fish? And considering the relative potential for toxic blooms, more effort into algal bloom sampling is warranted.

l. 1353 – this table provides interesting information about various projects but it is not given in enough detail to evaluate

l. 1363 - the literature cited is rather modest for a 70 page technical document.

2015 SFER
Chapter 8 Lake Okeechobee Watershed Protection Program
Reviewers Dillon (AA) and Dodds (A)

Closing Comments

The authors have provided suitable replies to all of our comments on chapter 8. The responses are detailed and reasonable, and have clarified several issues that were raised. There are no points of disagreement between the authors and the reviewers, other than the long-lived but relatively minor issue of mixing metric and non-metric units. As before, we appreciate that some of our suggestions would require additional funding which is not something in the jurisdiction of the authors of the chapter.

Key Findings and Recommendations

1. The Stormwater Treatment Area (STA) projects are making significant contributions to reductions in total phosphorus loading to the lake. Additional work planned for the STAs should lead to more improvements in water quality and should be continued and possibly expanded.

2. The use of HWTT technology has been effective and the technology should be expanded where possible.

3. Despite all of the positive steps taken at significant financial cost, and the projects that have resulted in reduced TP loads, there has been little progress towards reducing the overall TP load to the lake closer to the target value of 140 metric tons. Higher precipitation and runoff in WY2014 resulted in no reductions in P load or in-lake concentration, demonstrating the importance of hydrologic conditions to the TP budget of the lake.

4. The monitoring work has been of high quality, providing valuable long-term data, and it is essential that it be continued.

5. The lake remains in a hypereutrophic state. However, algal blooms were relatively infrequent, with cyanobacteria generally low as determined by the cyanobacteria:diatom ratio. *Chara* has decreased in abundance which, along with increases in SAV, is a positive result.

6. Wading bird foraging and nesting success rates showed improvements from previous years, as did the status of the fishery; this is indicative of further recovery following the major storm events of earlier years.

Chapter 9 **Kissimmee River Restoration and Basin Initiatives**

Dodds AA reviewer technical comments.

This chapter is generally well written and describes an amazing restoration job, one of the biggest ever. It is really commendable the "ecological responses to Phase I of restoration can be found in a recent special edition of the international journal Restoration Ecology dedicated solely to the KRRP". Any time that the research and monitoring of this ecosystem can be published in the peer-reviewed literature, the validity of the project is increased. Such publication should always be encouraged as the peer-reviewed literature is much more likely to survive and serve as a resource to others than reports such as this, which are, while peer reviewed, are classified as grey literature and are not held to the same standards in general.

My largest point here in this review is the lack of attention paid to mechanistic understanding of why low dissolved oxygen might occur. Results are presented without consideration of time of day, and it is well known that the balance between photosynthesis and respiration played off against equilibration via aeration rates plays the dominant role in determining how far from aeration that dissolved oxygen gets, including low oxygen excursions.

Mechanistic models of these processes informed by empirical data are now possible. The research work in the Everglades demonstrated expertise in the region to work with gas tracers (SF6) and inert solutes (fluorescene) that would be required to assess reaeration rates by habitat. The existence of logging optical oxygen probes provides the technology to get diurnal oxygen signals. These data from multiple points in the system could be combined to get estimates of oxygen dynamics and individual points as well as informing a whole-system model of oxygen fluxes to predict the conditions under which low or high dissolved oxygen conditions could occur.

1. line 316. Are there varved sediments anywhere in the area that can be cored to indicate how common anoxic events were historically?

2. As this restoration proceeds and is finalized, is there any provision to plan and modify the project to accommodate future climate change, as is done for the Everglades part of this document (chapter 6)? We are in a no-analog world, and maybe simply restoring to prior conditions will never be entirely possible.

3. Line 487. This is not a totally true statement as ecosystem scientists have found ways to study phenomena at whole-system levels. There are many sub watersheds that could be analyzed in detail for production of BOD. Also small watershed manipulations may shed further light on questions related to production of low oxygen conditions. Considering how much money is put into the actual restoration, the costs of doing fairly extensive and large scale whole-ecosystem experimentation are not that high, and given that such

results are needed to regulate flows in the system to approximate natural conditions, the additional investment would be wise.

4. Line 529. Option b makes sense, in that it that is probably not outside of the range of natural responses to high rain.
5. Line 573. It is somewhat disappointing to see limited post construction monitoring. This is a very important restoration project and one of the general criticisms of wetland restoration has been inadequate post-restoration sampling. Really long term and detailed records are necessary to establish if restorations truly are working.
6. Line 597 Wow! What a fantastic series of pictures!
7. Figure 9.10. This figure is difficult to decipher, any better way to show the data?
8. Table 9.4 Include the number of years total for each, or the number of years met means little. Alternatively the proportion of years the metrics are met could be presented.
9. Line 888. Disappointing that so few metrics are being met, but good that there is some improvement. What is the chance that suggestions for the best ways to meet more of the metrics will actually be adopted and be successful?
10. Table 9.7 and elsewhere. It is a little strange that dissolved oxygen in simply indicated as daytime. The minimum oxygen should be near sunrise and the maximum a bit after noon. Is there any correction for time of day in any of these measurements? Diurnal optical DO sensors are pretty stable and could be deployed to get diurnal swings.
11. Line 1217. Are there any estimates of error in the cover determinations?
12. Figure 9-25 Are the post 2010 data significantly different from the baseline data? Seems pretty hard to get a fix on a 2 year baseline when the data are so variable.
13. Figure 9-32. Source and sink is more common terminology, in my experience, to importer or exporter.
14. Table 9-11. This table does not make sense to me, am I missing something? The land uses that should be the biggest sources are denoted as importers. Does this mean they cause import downstream? If the source-sink terminology is used then the net sources are ones that give off nutrients downstream and those that are sinks intercept nutrients from upstream and let less go downstream. On the other hand, if one is accounting just for inputs and outputs, cropland could be a net importer, because it gets lots of fertilizer and holds back most of it, but the major source of nutrients downstream. I think a conceptual figure of what these numbers mean or some other clarification might help here.

Comments on Chapter 9: Kissimmee River Restoration and Basin Initiatives
A reviewer – P. Dillon
Technical Review

General

This chapter provides a clear summary of restoration activities in the Kissimmee River watershed in HY2014. It is well-written and generally concise enough, although there is some repetitious material in the first 10 or 15 pages. The results are quite encouraging even in cases where specific objectives aren't met; in several cases I don't think that it is realistic to expect them to be met at this stage of the restoration activities. For the most part, the work seems to have been very well done and the data analysis and interpretation appear to be sound.

There are a few areas where the report could be improved. I concur strongly with the AA reviewer (W. Dodds) that error estimates for many of the measured or calculated parameters would add substantially to the strength of the report and the arguments that are made. Most importantly, I had also arrived at the conclusion before reading the AA review that the lack of information concerning the reasons for the oxygen declines was the most significant shortcoming of the chapter.

I have edited my comments below to remove material that is included in the AA reviewer's report.

Specific comments

l. 310 – would it be possible to produce a "natural" hydrograph for comparison, i.e., what expected discharge would be in the absence of management practices?

l. 417 – why would increased discharge lead to lower oxygen rather than higher? Is in-channel aeration going to be necessary or is there any chance that this hypoxia will not occur when the new hydrologic regime is established?

l. 433 – in my limited experience with fish, I've seen bullheads that are extremely tolerant of low oxygen, lasting for extended periods in small ponds with<1 mg/L oxygen, even for days at <0.1 mg/L. Because bullheads were included in the fishkill, it suggests that oxygen alone may not be the sole cause. Any chance that another factor is responsible for the fish kill, perhaps something else that is released from e.g. sediments as a consequence of the low oxygen?

l. 497 – these explanations for the low oxygen seem reasonable; should be easy to check at least some of them either with existing data (#3) or by collecting new data. This is such an important factor in the recovery that it needs to be given high priority.

l. 698 – the expectation about seasonal/monthly patterns is not realistic given natural variability in rainfall patterns. Flow management can offset this variability but can't be expected until the new hydrologic control regime is in place. It would make more sense to compare the 2014 HY rainfall pattern to the long-term rainfall pattern and adjust expectations for each year accordingly, at least for the foreseeable future.

l. 720 – Fig. 9-10 is quite confusing – there must be a clearer way of presenting this.

l. 889 – realistically, this is all that can be expected at this stage

l. 892 – the adaptive management approach is the way to go, at least in the short term where factors such as variability in seasonal rainfall matters

l. 931 – the rationales for decrease in river oxygen levels are not particularly convincing. Until the reason(s) for the anoxic or hypoxic conditions is/are known, there seems to be little chance of improvement. Practices that are useful elsewhere, such as generating passive aeration by e.g. creating riffles are almost certainly impossible in this case. Active aeration is likely unable to handle the high flow rates seen here.

l. 933 – in what way are these other streams reference streams – do they have no urban or agriculture influence or are they simply unchanneled but with the same potential anthropogenic effects?

l. 948 – and night? there is a pronounced diurnal cycle with oxygen, although the late part of the night and early morning are usually the time of lowest oxygen levels, so this may be an optimistic picture

l. 969 – this is critical information and I am glad to see that attention is given to this.

l. 1021 – autosamplers are known to provide fairly poor phosphorus (and ammonium) data, resulting from fouling of intakes, loss of P to the collection bottles' walls, etc., particularly when samples are not recovered daily. Hopefully the autosampler results will be compared to the grab samples taken at the same time to assess this.

l. 1032 – too many significant figures, reduce from 6 to a maximum of 3; same with Table 9-8; this detracts from the reader's confidence in the results

l. 1033 – as always, the mix of metric units with acre-feet appears strange. And later, results are expressed on the basis of square km rather than acres.

l. 1052 – TP loads and concentrations are worse in the latter period everywhere except 65E which fortunately is where it matters most, i.e., the load to the lake downstream hasn't increased if I interpret these data correctly.

l. 1097 – there is a great deal of serial autocorrelation in relationships like this when concentrations at one location in a river are related to those at an upstream or downstream site. For this to be meaningful in any way, randomization tests are needed to establish the effects of the autocorrelation

l. 1428 – excellent!! very happy to see this material in a high quality peer-reviewed journal; it gives strong credibility to the work done

l. 1578 – I don’t understand Table 9-11. The terminology is confusing

RESPONSES TO COMMENTS ON THE DRAFT 2015
SFER –VOLUME I, CHAPTER 9

Final comments

- This chapter is well written and describes an amazing restoration job, one of the biggest ever. It is really commendable the "ecological responses to Phase I of restoration can be found in a recent special edition of the international journal Restoration Ecology dedicated solely to the KRRP". Any time that the research and monitoring of this ecosystem can be published in the peer-reviewed literature, the validity of the project is increased.

- The responses to the reviewer comments are adequate for the most part.

- Much of the rationale behind the comments is to help find ways to strengthen an already strong document, and encourage more funding to continue detailed monitoring and mechanistic studies of responses to restoration.

- Revisions to clarify the DO dynamic measurements, and the simple figure on the terminology used here for the nutrient modeling would be a welcome addition.

- More in depth studies of factors controlling DO would be welcome. Using watersheds under a range of conditions to explore controlling factors would be useful with broader deployment of 24 hour logging sondes. Specific analyses of midday collections linked to the exact time of collection may also help. More robust mechanistic models of DO concentrations should be the ultimate goal.

- Continued funding of monitoring is essential to be certain that the long-term understanding of the effects of the project is possible. While limitations on funding always occur, this is a very important restoration project, and adaptive management demands continued attention to as many relevant response variables as possible. This comment is meant to convey the idea that cuts should not be made to the monitoring budgets in this restoration project.

- Autocorrelation can be spatial or temporal, and this issue should be paid attention to in future research in the area (while acknowledging that it is a difficult issue in stream networks).

2015 South Florida Environmental Report

Chapter 10 – St. Lucie and Caloosahatchee River Watershed Protection Plan Annual and Three-year Updates

Review (AA) by Vladimir Novotny

Accountability Review *(for chapters that are of a more routine nature)*
This level of review targets progress and achievements of expectations in District programs and projects that are generally descriptive or standardized in nature, and may deal with cross-cutting themes or content.

Questions addressed by the review:

– Does the draft document present a defensible account of data and findings for the areas being addressed that is complete and appropriate?
- Is the synthesis of this information presented in a logical manner, consistent with earlier versions of the report?
- Are findings linked to management goals and objectives?

General Comments

The chapter is focusing on two estuaries, the St. Lucie (SLE) and Caloosahatchee River (CRE) Estuaries. This chapter is different in many aspects from the last (WY2013) report that extensively investigated the impact of (intermittent) freshwater releases from Lake Okeechobee where these rivers/estuaries originate and the lake is the largest freshwater contributor. This chapter mainly reports on the River Watershed Protection Plan (RWPP) updates. In accordance with the Northern Everglades and Estuaries Protection Program (NEEPP) the original RWPPs were completed and submitted to the Florida legislature in March 2009 by the three coordinating agencies [South Florida Water Management District (District), Florida Department of Environmental Protection (FDEP), and the Florida Department of Agriculture and Consumers Services (FDACS)]. This chapter fulfills NEEPP requirement to provide the annual and three-year reporting requirements for the RWPPs. The first three-year updates were submitted in 2012 and this chapter is the second three-year update for both watersheds.

The chapter begins with a description of the overall system, including the history of the South Florida estuaries, and the river watersheds to set the stage for the rest of the chapter. The coordinating agencies' efforts to improve the quality, quantity, timing, and delivery of water to the estuaries are highlighted by updates to the Pollutant Source Control Program and the Construction Project. This is followed by a summary of the conditions of the hydrology, water quality, and aquatic habitat based on results of the Research and Water Quality Monitoring Programs, followed by the condition of two ecological indicators—oysters and submerged

aquatic vegetation (SAV)—in both estuaries. The final section outlines the coordinating agencies' future strategies for continued watershed and in estuary restoration activities.

Everglades Past and Present (p. 10-8)

Hydrologically, the rivers are not a part of the Everglades system and the key issues are different from the Everglades where the main problem appears to be excessive, mostly nonpoint loads (surface and atmospheric) of nutrients, mercury and sulfates which was covered in Chapter 3A, 3B and 4. Until a few years ago most projects in the two estuaries had been structural for manipulating the flows, which had very little effect on the Everglades. Both estuaries receive intermittently excess flows from Lake Okeechobee which is also the main feeder of flow to the Everglades Protected Area. However, most flow in the estuaries area originate from the surrounding watersheds.

The chapter in its first portion extensively describes the hydrologic history of the system. Originally, no flow from Lake Okeechobee was entering the estuaries which were 130 years ago much smaller. Hydrological modification of Central Florida connected the Lake by canals with the estuaries which periodically increased fresh water flow laden with nutrients in the estuaries. This resulted in phytoplankton blooms, accumulation of flocculent muck-like sediments, severe losses of seagrass habitats, and decline of oyster beds. The system has become managed but until this century the goal of manipulation and management was flood protection, water supply and land development. As a result of these hydrologic modifications, the natural storage and buffering capacity has been severely reduced.

The section on history (pages 10-8 to 10-11) is very informative but it would fit better into Chapters 1 and 2 than in this last chapter.

Pages 10-12 to 10-18 describe the altered watersheds and subwatersheds of the St. Lucie and Caloosahatchee Rivers and estuaries, their hydrology, land use and present maps. It was pointed out that both coastal watershed runoff and outflows from Lake Okeechobee have profound influence on estuarine physics, water quality, and biotic resources. Table 10-1 shows the land use distribution in both watersheds. Agriculture is the most dominant land use in both watersheds. Considering the fact that until now plans for nonpoint controls have not been implemented the effects of nonpoint sources on the estuaries and their habitat are severe.

Pollutant Source Programs (p.10-19).

In this section the authors describe the pollutant source control programs and their components. As stated herein not much has been happening in the two watersheds until recently. In 2009 the agencies adopted the TMDL (Total Maximum Daily Load –Section 303 of the CWA) plan for both estuaries calling for reduction of nitrogen inputs in the CRE and reductions of both nitrogen and phosphorus in the SLE. Apparently, both estuaries are nitrogen limited. The TMDL adoption triggered several planning efforts included in the Basin Management Action Plan (BAMP) that

outlined the measures to meet the TMDL goals for the estuarine watershed. These plans were adopted by the agencies carrying out the programs in the Caloosahatchee River Watershed in 2012 and the St. Lucie River programs in 2013. The remediation programs include Best Management Practices (BMPs) on agricultural lands, storm and wastewater infrastructure upgrades, and regulatory programs focusing on water quality. Most of these measures are in planning stages; hence, not much water quality improvement has taken place in the water bodies and the first results may be seen in the next three years assessment report in 2017. Some key remediation programs envision full operation even after 2017. A part of the planned actions is improving the water quality model for the systems which will lead, following the adaptive planning concept, to better enumeration of the goals and forecasting of the impact of remediation on water quality and, hopefully, on the biota.

In many states, water quality improvement plans and TMDL action plans for agricultural watersheds rely on voluntary participation of dischargers (farm and cattle operators) that was in most cases supported by incentives and sometimes subsidies. In contrast, the BAMP plans will rely on mandatory actions and the District and the Florida Environmental Protection Department have been granted authority to an initiate Environmental Resource Permitting program. This program and its implementation will have national implications and, if successful as it should be, will serve as an example to other states and help to invigorate the stalled TMDL programs for agricultural watersheds in many states. The recent assessment of the TMDL programs by the US Government Accountability Office concluded that TMDL programs relying only on voluntary participation of agricultural dischargers are ineffective. In Florida, the Department of Agriculture and Consumer Services (FDACS) prepares, with agricultural dischargers, commodity specific water and water quality BMPs in which agricultural and silvicultural (orchards) operators (producers) enroll by submitting a Notice of Intent (NOI) to FDACS, along with an accompanying checklist of the practices applicable to the operation. To date, BMP manuals for citrus, vegetable and row crop, nursery, sod, cow/calf, equine, and specialty fruit and nut operations have been adopted. Currently operators of 40 % agricultural and nonagricultural land area in the Caloosahatchee River watershed and 70% in the St. Lucie watershed received discharge permits.

Another program for control of nonpoint pollution is the Urban Fertilizer Rule which requires that fertilizer application on lawns and other green urban areas must comply with the statewide rule. This rule was apparently adopted in 2007 but no detail about compliance and success were presented in the the 2015 SFER.

The Florida Department of Environmental Protection also manages the federal program of permitting the discharges from Municipal Separate Storm Sewer Systems (MS4) which requires applying for system wide discharge permits, monitoring the discharges and preparing abatement plans with BMPs. All reissued Phase I permits include a new section on TMDL implementation and require enhanced tracking of load reductions achieved through implementing the permit's stormwater management program including nutrient load reductions from street sweeping

activities. The research by the reviewer (Novotny, 2003), thirty years old National Urban Runoff Program (US EPA, 1983) and extensive research in the US and abroad documented that traditional infrequent street sweeping is not effective to substantially reduce the nutrient loadings. Other BMPs such as treatment ponds and biofilters are more effective. Street sweeping provides aesthetical and street cleaning benefits but produces collected solids that may be contaminated and have to be disposed (in landfills?). Vacuum sweepers are more effective than older mechanical rotating broom sweepers.

The 2009 Florida legislature promotes the use of Florida-friendly landscaping to conserve and protect the state's water resources. It is not clear whether or not this Florida program is based on the national more sustainable urban landscape concepts known under the name Low Impact Development (LID) which is a combination of sound hydrologic designs minimizing volumes of storm runoff and quantity of pollution by implementing storage, infiltration, biofilters, roof gardens and many other practices, combined with the aesthetic enhancement of the community (subdivision). Because LID programs have been tremendously successful worldwide, a Manual of Practice specific to Florida should be prepared and made a guide to Florida communities under this program.

The biosolids rule, Chapter 62-640, F.A.C, was revised in 2010 which restricts and limits the use of fertilizers on agricultural lands. This legislation requires permitting and nutrient management plans in the Northern Everglades and forbidding lower quality and harmful fertilizers. Last year all sites covered by this legislation are required to be permitted in accordance with the revised rule. A similar rule has been also mentioned in the report for manure management and disposal but details were not reported in the chapter. It should be pointed out that manure management has become a worldwide problem because soils receiving excessive amounts of solid and liquid manure, especially from large animal operations (e.g., dairy) become oversaturated and overloaded and loose the capacity to retain phosphorus and nitrogen. There is now a trend, mainly in Europe and China, but also in the US to co-digest manure with municipal sludge and other high organic liquids to produce methane and energy (for example, in Madison, WI) (Burns & Moody, 2009). The residual solids can be converted along with other dry or dried organic solids to additional gas (syngas) by pyrolysis.

Watershed efforts (pages 10-23 to 10- 39).

These works and actions were initiated by the TMDL (Total Maximum Daily Load – Section 303d of the Clean Water Act) for both Rivers/Estuaries. TMDLs were adopted for the Caloosahatchee Estuary in 2010 and the St. Lucie Estuary in late 2009. The TMDLs were followed by preparation of Basin Management Action Plans (BMAP) which for the Caloosahatchee Estuary was adopted in 2012 and that for the St. Lucie River and Estuary Basin was adopted in 2013. This implies that most actions have not yet been implemented and the first major effects can be expected in the next 2017 three year review and updates.

One of the cornerstone activities of the planning effort for the Caloosahatchee Estuary will be modification of the water quality model which will lead to redefining the TMDL – the maximum load in the sections of the estuary and subwatersheds. This is important because the model must be capable of backtracking the maximum (nutrient) load to various actions planned in the watershed that will be accepted by the stakeholders.

Construction projects described on pages 10-24 to 10-33 are expected to reduce fresh water and nutrient inputs into the estuaries, improve hydrology, water quality, and aquatic habitats in both the watersheds and estuaries. The projects include local storm water retrofits, Stormwater Treatment Areas (STAs), reservoirs, and habitat restoration. Hence, the construction projects are local (subregional) and regional. An overview of the large regional projects is presented in Table 10-3 and dispersed mostly local projects are listed in Table 10-4. The regional projects are primarily storage reservoirs and storage/treatment wetlands. Of interest is the innovative Floating Aquatic Vegetative Tilling lagoon in which high density floating aquatic vegetation will be allowed to develop during growing season and then the water will be released and the plants will be dried and tilled into the lagoon soil. The table does not identify the FAV species. Water hyacinths are invasive FAV ubiquitous to Florida that under controlled conditions provide high nutrient removal and, subsequent to treatment, biomass that could be used for production of biofuel/energy, e.g., by co-digestion and/or pyrolysis with manure and municipal sludge.

The dispersed water management projects listed in Table 10-4 will provide shallow water storage, retention, and detention to enhance Lake Okeechobee and estuary health by reducing discharge volumes and nutrient loading and groundwater recharge. Table 10-4 does not include information on the type of the facility (e.g., lagoon, wetland, infiltration basin, etc.).

The overall storage needs (page 10-40) to reduce fresh water flow and nutrient loads into the two estuaries were estimated by modeling. The assumptions included in the model included: no additional water can be sent south to EPA, water supply for existing users must be maintained, and additional storage will be needed in each of the estuary watersheds to address local basin runoff. The results were overwhelming; approximately 900,000-1,300,000 ac-ft (1.11 – 1.6 Km^3) of storage was needed in the Lake Okeechobee watershed to help manage lake levels and reduce discharges to estuaries. Additional million acre-feet (1.233 Km^3) of storage is needed in the Everglades system.

Concluding remarks on the Sections dealing with proposed and planned remediation projects (pages 10-19 to 10-42). This section is a long and detailed list of many projects which may have a chance to reduce the fresh water flow (control flooding) and nutrient loads. This appears to be the most ambitious program in the nation that will require cooperation of many agencies and stakeholders under the leadership and coordination of the three leading agencies: South Florida Water Management District, Florida Department of Environmental Protection, and Florida Department of Agriculture and Consumer Services. There are many participating agencies from the US Army Corps of Engineers to Local Governments, to World Wildlife Fund for Nature

(note to the authors: update the name of WWFN). But the reviewer could not find it the text of the chapter who is the actual leader and how the efforts will be coordinated. This effort may need a special coordinating entity with a director, department heads and liaisons to the key player and executors of the tasks otherwise the efforts will be fragments. The same may be true for the efforts being conducted in the EPA. Is the District this coordinating agency? Where is the organization chart? Who is in charge and who are the leaders?

Almost all projects have been initiated after the adoption of the TMDL studies in 2012 and 2013; therefore, no results were available for this report.

The chapter is very detailed and the reader is overwhelmed by hundreds (what it seems) of acronyms and abbreviations, some of them unidentified, some identified dozens of pages before and it takes an effort to identify what all these letters mean. A good remedy would be to periodically spell out the acronyms which will be featured on the page and two or three subsequent pages again.

The chapter uses US units, almost solely limited to acre-ft. This unit was identified at the beginning of the section and it probably may not be needed to put metric equivalent at every time it appears in the text. Note, that for very large flows and volumes in the world hydrological literature, in addition to m^3, also Km^3 (10^6 acre-ft = 1.23 Km^3) is used. Metric prefixes with US units, e.g., kacre-ft for 10^3 acre-ft as it occurred in another chapter, are not appropriate.

Watershed research and water quality monitoring

This is the next section of Chapter 10 that focuses on the quality of water and health of biota in the two estuaries. Oysters, sea grasses are salient features of the estuaries and their density and health are indicators of good quality. By definition, estuaries are characterized by mixing of sea water brought in by tides and fresh water coming from land. Salinity of the water is a good indicator to which health of marine and estuarine biota can be related.

Nutrients brought in by freshwater cause eutrophication which is characterized by excessive growths of algae and algal blooms and, in extreme, by cyanobacteria and red tide microorganisms. Oyster physiology, survival, and growth are optimal when salinity fluctuates from 8 to 25 in many estuaries (Note, the salinity unit should be identified when it first appears in line 1181).

The section of the chapter from page 10-43 to page 10-62 describes the conditions in the St. Lucie River and Estuary and section from page 10-63 to 10-82 deals with the Caloosahatchee River and Estuary. One immediate observation is that these long sections are identical in format (not an impediment), include word-by word repetitive text and identical figures but with different data. This high repetitiveness could be avoided in the final editing. For example sections on rainfall could be unified with data combined in tables with two columns, one for SLE, the other for CRE and figures for both water bodies. The same applies to sections on Methods, Inflow,

and maybe others. For example, the long term flow contributions and percentages of flow from Lake Okeechobee could be combined in one table.

Regarding the effect of fresh water discharges from Lake Okeechobee and the watersheds, the impacts on salinity of the estuaries (Figure 10-15 for SLE and 10-27 for CRE) is profound, i.e., when the inputs of fresh water are large, salinity is dramatically reduced to close to zero (=freshwater quality). Patterns of annual TP and TN loading to the SLE followed total freshwater inflow. Table 10-6 for SLE indicates that salinity in the SPE was in WY 2014 64.4% within the target salinity range 8 to 25.

Figure 10-18 showed that the nutrient inputs (nitrogen and phosphorus) followed the same pattern and were closely related to fresh water inflows. Nutrient loading exhibited an overall increase with freshwater inflow from WY2012–WY2014. The ratio between the average annual N and P loads from the watershed was close to N/P = 7 to 8/1. The report identified the ratio as 4.9/1 but in any case this ratio indicates that the estuary eutrophication could be nitrogen limited but the watershed measures to reduce nutrient inputs should focus on both nutrients.

The monitoring and data on nutrients and water quality confirmed that the SLE is distressed and the status of the estuary may be even worsening. The improvements may not be expected until several years from now when the planned BMPs, storages and other remediation measures begin to fully function.

Water quality in CRE. The first ten pages of this section (10-63 and 10-70) are similar to those for SLE; same wording, similar graphs, just the values are different. As pointed out above the Sections “Method” and “Rainfall” could be unified. Similarly to SLE, freshwater discharges have tremendous impact on salinity as shown on Figure 10-27. When large fresh water discharges from Lake Okeechobee and the watershed occur, salinity is reduced to levels typical for freshwater. It means that tidal mixing is not sufficient to bring salt water into the estuary. It appears that extended freshwater salinity conditions had an effect on sea grasses which in the WY2014 were reduced at the Boys Scout Island site. Fresh water inputs have also an effect on N and P concentrations, as documented on Figure 10-30, which are directly and closely proportional to freshwater inputs.

In addition to N and P concentrations, water quality is also expressed by Chl-*a* concentrations. Because sea-water has a very low to negligible nutrient concentration, the gradient of nutrient concentration decreases from upstream to downstream. This may under certain circumstances create a zone with algal blooms and poor oyster and sea grasses habitats. The TMDL goal of keeping annual average of N concentrations below 0.72 mg/L was exceeded in 2012, 2013 and 2014 Water years by about 10%, more in wet periods (Table 10-7). As expected there is no downward temporal trend yet because the planned actions have not been implemented. Average Chl-a concentrations have risen in the last three years from 12.8 µg/L in WY 2012 to 17.1 µg/L in WY 2014. These values may be misleading because they apparently represent the annual

averages and not the concentrations during the periods preceding and during algal blooms. Indeed, Figure 10-20 shows that Chl-a in the upstream profile HR-1 exceeded periodically 60 µg/L which would indicate algal blooms. The species of algal blooms were not identified in the text.

The effect of fresh water input on the CRE is more pronounced than that for SLE. During the larger fresh water inputs the near zero salinity levels are longer lasting than those in the SLE. Consequently, the effect on sea grasses was significant at Sites 2 and 4 where sea grasses disappeared in WY2013 and were reduced at Site 3 (Figure 10-33).

Total nutrient loading to the CRE came in fresh water inputs from Lake Okeechobee, the East and West Caloosahatchee sub-watersheds, and the Tidal Basin and closely followed the freshwater inflow (Figure 10-30). However, the lake is eutrophic and outflow is rich in nutrients. The authors reported that wet season nutrient concentrations were higher than in the dry season due to a combination of increased inflow and TN loading and increased water temperature and rates of internal cycling.

The long term N/P load ratio detected from Figure 10-30 was 10/1, still borderline as to the limiting nutrient. Annual averages of the Chl-*a* in the 2012-2014 three year period were the highest in the middle CSO4 station (Table 10-9) and during wet seasons were as high as 80 µg/L (Figure 10-32) in the upstream CSO1 which could indicate algal blooms. Algal blooms were reported during wet seasons. Oyster and sea grass populations were better than those in the SLE. Chl-*a* values during WY 2014 were reduced compared to the long-term averages throughout the CRE in the wet seasons of WY2013 and WY2014.

Remarks on Watershed Research and Monitoring. The authors of this section developed a very good reporting protocol and graphics that will also be useful for future reporting and following the progress in improving water quality and biota of the estuaries. The format is also similar to the last year Chapter 10 in the 2014 SFER. However, as stated several times in this review the water quality management and remediation programs in the watersheds are just beginning or even still in planning phases. Initiation of the programs is lagging behind the ongoing programs similar programs in the Lake Okeechobee watershed and in the Everglades Protection Area. It seems that the fresh water nutrient rich contribution of Lake Okeechobee into both estuaries is smaller than that from watersheds and the lake flow is intermittent. Nevertheless, the fresh water inputs have a profound, currently detrimental, effect on water quality and ecological health of both estuaries. Comparison of trends in water quality did not reveal any improvement. It is, therefore, paramount that the TMDL document is rapidly updated by a better model and the planned remedial action included in the BMAPs are immediately initiated in an adaptive management scheme.

The water quality of both estuaries is not good, algal blooms in some sections are common. The same conclusion was reached on the last year monitoring work. As the chapter concluded the

ecological impacts to the St. Lucie and Caloosahatchee Estuaries—including higher nutrient loading, undesirably low salinity, algal blooms, and seagrass and oyster mortality—were significant and public has a notable interest to seek action to resolve the situation.

Several repetitive sections could be combined and data presented in comparative tables.

Moving Forward Section (pages 10-82 – 10-86)

This is the last section of the Chapter. The actions outlined in the Chapter were initiated following the TMDL documents for the estuaries. However, it appears that immediately after the TMDLs were issued, the modeling and, in this context, the reliability of the TMDL goals were questioned which resulted in a new modeling effort and ongoing attempt for modifications. Attacking the model is a common nationwide tactic of those affected by the TMDL restriction that put a greater emphasis on economic impacts of TMDL sometimes ahead of environmental improvement. It is logical that the TMDLs have a great deal of uncertainty which should have been incorporated in the Margin of Safety. It is hoped that the normally expected calls for better modeling have not delayed the planning and implementation efforts. These efforts should be conducted using adaptive management approaches wherein modeling improvement is an integral part of the remediation but the remediation is not delayed because there is an overwhelming evidence of impairment and most measures are common sense.

It is noted that the FDEP along with its stakeholders successfully adopted BMAPs for TN for the Caloosahatchee Estuary (FDEP 2012) and for TP and TN for the St. Lucie River and Estuary. Both BMAPs were reported to be in the first five year phase during which projects will be implemented but planning and other activities (e.g., new modeling) will also be taking place in order to set the stage for the second phase of the BMAP which will begin in 2017. FDEP estimates that the first five-year adaptive iteration of the BMAP in the St. Lucie River and Estuary Basin will reach 30 percent of the required reductions for TP and TN and in the Caloosahatchee estuary BMAP, the total nitrogen load in the tidal watershed will be reduced approximately by 40 percent of the original TMDL required load reduction. This implies that some progress may be noticed in the WY2017 report but the goal will not be met. FDEP itself expects “modest improvements in water quality trends in the watershed tributaries as well as each estuary “ in the first five years.

The last section then describes in detail the projects and actions already described in the Sections *Pollutant Sources Programs* (p. 10-19) and *Watershed Efforts* (p.10-23).There are some repetitions between these sections and the concluding section. The key actions (some already mentioned in the preceding text) included rule changes that would expand permitting processes currently in place for the Lake Okeechobee watershed to the watershed and tidal area of the St. Lucie and Caloosahatchee Rivers and Estuaries. As it was emphasized in this review, remediation efforts relying mostly on voluntary participation are ineffective.

In the FDACS sphere of responsibility non-structural BMP practices under the owner-implemented category generally include nutrient and irrigation management, maintenance of vegetative buffers to protect water features from sediment- and nutrient-laden runoff, and location of livestock feeding/mineral stations away from water features. These actions are cost-shared between the operators and FDACS. The problem may be the level of funding that would provide enough incentive to the operators to participate in the program. FDACS apparently has only $ 3 million annually for cost sharing. Consequently, only about 55 percent and 71 percent of the acreage classified as agricultural land use in the St. Lucie and Caloosahatchee watersheds, respectively, are enrolled in the FDACS BMP program which may be considered as insufficient. Indeed, the level of participation has not changed much since the last year report and the FDACS has requested to increase the level of funding to $10 million/year. Increasing the level of cost – sharing should be considered with a caveat. Implementing BMPs will bring significant economic and social benefits to the farmers. Hence, education of farming operators about the benefits is paramount.

In conclusion, the District and FDEP recommended both short- and longer-term actions to help improve the quantity, timing, and delivery of water throughout the system. Short-term solutions included efforts such as storing water on public and private lands in the watershed and operational approaches to moving water through the system more quickly. The suggested longer-term solutions were focused on completing projects that are already under way or planned in order to assist with storing water, improving water deliveries south, and minimizing damaging discharges to the Northern Estuaries.

There is no doubt that all efforts (1) Pollution prevention and protection of the Everglades Protected Area, (2) Lake Okeechobee pollution prevention and restoration, and finally (3) Northern Everglades program encompassing the St. Lucie and Caloosahatchee watersheds and water bodies are the largest most comprehensive and ambitions pollution control and ecological restoration program today in the world. The past ten years have seen significant improvements in the first two efforts (Lake Okeechobee and EPA) but the efforts are not finished and must continue in perpetuity and include considerations of the ongoing threats of global climatic change. The actions in St. Lucie and Caloosahatchee watersheds are just beginning.

This chapter is closely linked to the management goals and objectives. The synthesis protocol has been adequately developed; it is logical and follows the previous reports. This year report is more comprehensive and more detailed and following the protocols to the point that some sections are repetitive. The detailness is understandable because this is a three year review report.

The data presentations in tables and graphics were excellent but authors should make sure that lettering in some reduced figures is legible.

Literature

Burns, R.T. and L. B. Moody (2009) Current State of Manure Anaerobic Digestion in the U.S. and Beyond. Conference on Energy Production from Anaerobic Digestion of Dairy Manure. Madison, WI. September 28-29, https://science.house.gov/sites/republicans.science.house.gov/files/documents/hearings/102109_Burns.pdf

Novotny, V. (2003) *WATER QUALITY: Diffuse Pollution and Watershed Management*, J. Wiley, Hoboken, NJ

U.S. Environmental Protection Agency (1083) *Results of the National Urban Runoff Program, Vol. 1, Final Report*, U.S. EPA, Washington, DC

REVIEW CLOSURE BY VLADIMIR NOVOTNY (AA)
DRAFT 2015 SFER – VOLUME 1, CHAPTER 10

The chapter is focusing on two estuaries, the St. Lucie (SLE) and Caloosahatchee River (CRE) Estuaries. It describes the overall system, including the history of the South Florida estuaries, and the river watersheds to set the stage for the rest of the chapter. The coordinating agencies' efforts to improve the quality, quantity, timing, and delivery of water to the estuaries are highlighted by updates to the Pollutant Source Control Program and the Construction Project. This was followed by a summary of the conditions of the hydrology, water quality, and aquatic habitat based on results of the Research and Water Quality Monitoring Programs, followed by the condition of two ecological indicators—oysters and submerged aquatic vegetation (SAV)—in both estuaries. The final section outlined the coordinating agencies' future strategies for continued watershed and estuary restoration activities.

Unlike the other chapters in the 2015 SFER which describe progress with ongoing actions on the trends of various pollutants, most of the key actions in the watersheds focusing on improvement water quality are in the planning stages following the relatively recent adoption of the TMDL Reports by the responsible agencies. Previous extensive water works in both watersheds focused mainly on flow manipulation and developing storage for these purposes. The TMDL adoption triggered several planning efforts included in the Basin Management Action Plan (BAMP) that outlined the measures to meet the TMDL goals for the estuarine watershed. These plans were adopted by the agencies carrying out the programs in the Caloosahatchee River Watershed in 2012 and the St. Lucie River in 2013, hence, none of them are in place or fully operating and their full effect cannot be expected before 2020, possibly even later. Hence, the current poor (in some aspects) water quality situation does not meet the TMDL load goals as it was emphasized by the reviewer and extensively pointed out in the comments on the 2015 SFER submitted by Mr. G. Gofort on 10/22/2014. Mr. Gofort, being very familiar with the local situation and the process, pointed out that the current loads of N and P in the St. Lucie watershed are 6 times greater than the TMDL allowed loads, and N loads in the Caloosahatchee watershed are 35% more than the TMDL, respectively. He also pointed out that the current flow manipulation diverting 170,000 acre-ft/year (205×10^6 m^3) carried only 3% of the nutrient load to the Everglades Protection Area. Mr. Gofort has presented a set of thoughtful comments and suggestions that should be taken seriously and addressed by the authors and the participating agencies. Most of the nutrient loads to the estuaries originate from the watersheds and increasing substantially the nutrient load to the EPA by greater diversions is not a good solution. The correct solution is contained in the proposed plans that focus on implementing large and

dispersed storage and BMPs to capture and remove nutrients in runoff, control of fertilizer applications, sustainable land management, including Low Impact Development practices. However, the amount of storage north of Lake Okeechobee that would help manage lake levels and south of the lake that would reduce diversion is immense and overwhelming.

Consequently, the current high loads exceeding the TMDL will require extensive planning that, hopefully, will not be delayed by the effort to revise the TMDL model. The planners should take the adaptive approach to planning and continued developing and implementation of common sense measures concurrently with the fine tuning of the models and the final design. This aspect was addressed in the reviewer's report and highlighted also in authors' response *Comment #24* which stated that the TMDL model for the Caloosahatchee estuary is scheduled to be refined over the next few years; however, DEP requested that the stakeholders provide activities and projects that would begin reducing TN loads ("immediately?", "as soon as possible?"; questions by the reviewer).

The specific comment by this reviewer were extensively discussed and responded by the authors; there were 29 *Comments* included in the authors' responses to the reviewer's assessment of the chapter. Many were clarifications and short answers to the review. If the comments were adequately addressed by the authors they are not included in this closure. In general, the authors satisfactorily responded to most of the statements included in the review.

In *Comment 1* the authors defended the inclusion of an extensive historical background in (last) Chapter 10 which would have been in the reviewer's opinion more appropriate for an introductory Chapter 1of the SFER. The final decision as to where historical background should be included should be done in final overall editing.

Comment 7 explains the connection between the Florida-friendly landscaping program and the worldwide urban landscape programs known in the US under the title Low Impact Drainage (LID) which are currently funded by the Florida DEP. It is not clear if LID Florida Guidelines were prepared by the agency.

The projects that recently began and are currently planned are numerous and the overall scale of the undertaking is very large (*Comment 11*). These projects are carried by three agencies. The authors did not answer reviewer's concern about the lack of a unified coordination body, or "who is in charge". They stated that this is a cooperative effort but the responsibility is with each agency. Apparently, a coordinating oversight and organization body has not been established.

The reviewer's concerns with units used in the chapter, a large number of difficult to follow acronyms and definition of parameters (*Comments 14-16*) as well as repetitiveness of sections (*Comment 17*) were accepted by the authors who will make appropriate adjustments and additions to the text as well as editing lettering of some figures for legibility (*Comment* 29).

The reviewer asked for more detailed synthesis of the water quality data for both estuaries regarding the effect of nutrients on eutrophication processes that past reports reported as serious (*Comments 18 – 21*). The authors explained that this chapter is actually the three year overall assessment of the programs of the River Watershed Protection Plans required by NEEPP. While average values of chlorophyll-*a* were reported, the authors did not provide detailed analyses. Considering the fact that the chapter is already long and overwhelmed by the three year reporting materials, limiting the detailed synthesis and trend analyses in this year's report is understandable. The reviewer had an opportunity to review the previous estuary program chapters and noted that cutting edge research and syntheses of the studies on estuary dynamics, oyster population, eutrophication potential and occurrences and causes of algal blooms were provided therein.

In the concluding Comments 26 and 27 the reviewer expresses a concern with the level of funding for cost-share with stakeholders/dischargers for implementing the Best Management Practices on their lands. It was also noted that the level of participation is slowly increasing. The proposed programs will not only benefit the biota and recreationists of the estuaries but will also bring benefit to the stakeholders such as reduced cost of fertilizers or, potentially, a possibility to produce energy from manure by co-digestion with municipal sludge. Education of farmers and farming enterprises and to a lesser degree suburban homeowners using fertilizers on lawns and introducing them to the environmental sustainability of xeriscape without fertilizers is paramount. While the SFWMD is implementing advisory programs with farmers, education programs in other states with long soil conservation and nonpoint pollution abatement history could be reviewed. These are the programs by the Natural Resources Conservation Service which through its agents and offices in each county provide on-site education and advisory service to farmers in Midwestern US and Great Lakes states (Wisconsin, Indiana, Ohio, etc.), University extension services, etc.

KEY FINDINGS AND RECOMMENDATIONS BY VLADIMIR NOVOTNY (AA) DRAFT 2015 SFER – VOLUME 1, CHAPTER 10A

- It is the reviewer opinion that this chapter fulfills NEEPP requirements to provide the three-year reporting assessment for the RWPPs.
- Most of the programs included in the River and Watershed Protection Plans initiated after adoption of the TMDL reports are in the first phases of implementation and some still in planning phases. Therefore their beneficial effects on water quality have not yet materialized.
- Stakeholders asked for updating and revisions of the TMDL model which may last several years. It is important that this requirement does not slow down the programs. Adaptive management approach starting with rapid implantation of common sense programs is needed. It is an indisputable fact documented in the previous annual reports and other documents that water quality in the estuaries is not good, suffering from the adverse effects of advanced eutrophication exhibited in some years by algal blooms (for example, in 2012), loss of SAV, and oyster beds.
- In overall, the current three components of Everglades protection and restoration programs (1) Pollution prevention and protection of the Everglades Protection Area; (2) Lake Okeechobee pollution prevention and restoration, and (3) the St. Lucie and Caloosahatchee watersheds and estuaries are the largest most comprehensive and ambitious nonpoint and stormwater pollution control and restoration programs today in the world.
- It is important that the level of participation of stakeholders/dischargers, communities and transportation departments (significant source of nitrogen and toxics) must increase from the current level. Voluntary programs without permitting and incentives are not efficient. The authors postulated that the actual participation may be larger than that reported.
- Developing organizational structure to improve coordination among the three agencies involved and other key entities (Universities, NRCS, Communities, stakeholders) in this ambitious project is suggested.

Preliminary Comments on Draft 2015 SFER – G. Goforth 10/22/2014

Ch. 2 – Hydrology

1. The chapter omitted discussion of the most environmentally destructive water management operations conducted jointly by the SFWMD and USACE in more than a decade - the operations that took place between May 2013 and October 2013 associated with the destructive Lake Okeechobee regulatory releases to the estuaries. During WY2014, more than 1,600,000 acre feet of Lake Okeechobee regulatory releases were discharged to the St. Lucie River and Estuary, the Caloosahatchee River and Estuary and the Lake Worth Lagoon, yet the chapter omits discussion of these (Figures 1 and 2 below).
    a. More than 418,000 acre feet of Lake Okeechobee regulatory releases were discharged to the St. Lucie River and Estuary; more than 1,150,000 acre feet of Lake Okeechobee regulatory releases were discharged to the Caloosahatchee River and Estuary. More than 100,000 acre feet of Lake Okeechobee regulatory releases were discharged to the Lake Worth Lagoon. While the estuaries were being devastated by these harmful Lake releases, less than 170,000 acre feet of Lake Okeechobee regulatory releases were discharged to the Everglades – the pre-development destination for excess Lake water.
    b. These releases were the subject of numerous WRAC and other public meetings, generated multiple citizen protests during which thousands of citizens expressed their outrage, resulted in the formation of a State Legislative Committee which held multiple public hearings, and was the focus of a federal Congressional hearing.
    c. Suggest that the chapter summarize the water management operations during WY2014 leading up to the regulatory releases, document the volumes of water discharged to the estuaries, and identify short-term steps the agency is conducting to reduce these devastating impacts in the future.
    d. The District should prepare an After Action report documenting the flows, nutrient loads and sediment loads associated with the Lake regulatory releases. The partial report prepared by the District in October 2013 was incomplete and contained numerous errors and omissions (see my WRAC presentation slides 42-45 for specifics).

2. As contained in the 2012 EFA Watershed permit for the STAs, Ch. 373.4592 directs the SFWMD to send 28 percent more water to the Everglades than occurred during the 1979-1988 period, including an average annual volume of Lake regulatory releases of 250,000 acre feet. As shown in Figure 3 below, the actual Lake flows to the Everglades

have decreased. Suggest that the chapter document assessment of water management operations compared to this Legislative direction, and identify all the operational, policy, legal restrictions to moving additional Lake flow to the Everglades.

3. Table 2-2. Suggest that the chapter explain the inordinately high pumping volumes in WY2014, which does not intuitively synch with prior discussion that WY2014 had about average rainfall.

4. Line 691 – Suggest that the chapter detail the Lake regulatory releases
    a. Suggest that the chapter discuss the influence of LORS2008 on WY2014 operations
        i. Temporary schedule set in place due to concern of HHD safety
        ii. Lowered top end of schedule by 1.25 ft
    b. Suggest that the chapter discuss the reasons why the SFWMD did not follow LORS2008 guidance to send maximum amount practicable to WCAs.
    c. Line 708 – Suggest that the chapter explain why Lake outflow was 170% of average while inflows were only 130% of average.
    d. Suggest that the chapter document the volume of Lake releases sent for water supply to the EAA, L-8 basin, Caloosahatchee River basins, C-51 Basin, STAs, Lower East Coast, etc. and compare this to the volume of Lake water sent to
        i. the Everglades for water supply
        ii. the estuaries through regulatory releases.

5. Line 737 – Missing details of devastating Lake releases to St. Lucie River and Estuary.

6. Line 742 – Suggest that the chapter mention that C-44 Basin runoff often is also routinely sent to Lake Okeechobee when the lake stage is 14.5 or lower.

7. Line 748 – What is meant by: “Rainfall in Martin/St. Lucie rainfall area has improved from WY2013 by 5 inches.”? Should “improved” be “increased”?

8. Line 755 - Missing details of the devastating Lake regulatory releases to the Caloosahatchee River and Estuary.

9. Line 778 – Suggest that the chapter document the approximately 170,000 acre feet of Lake regulatory releases to the STAs and Wildlife Management Areas.

10. Line 797 - Suggest that the chapter document the quantity of Lake regulatory releases to the WCAs. Evaluation of the WCA stages suggests that additional Lake water could have been sent to the WCAs, particularly WCA-1 and WCA-3A prior to the devastating Lake releases to the estuaries (Figure 4 below). Suggest the chapter discuss/show the available capacity that was in the WCAs leading up to, during and after the May-October releases to the estuaries, and discuss why more releases weren't made to the WCAs when LORS2008 guidance suggested that course of action.

Figure 1 – Lake releases discharged to estuaries and Everglades

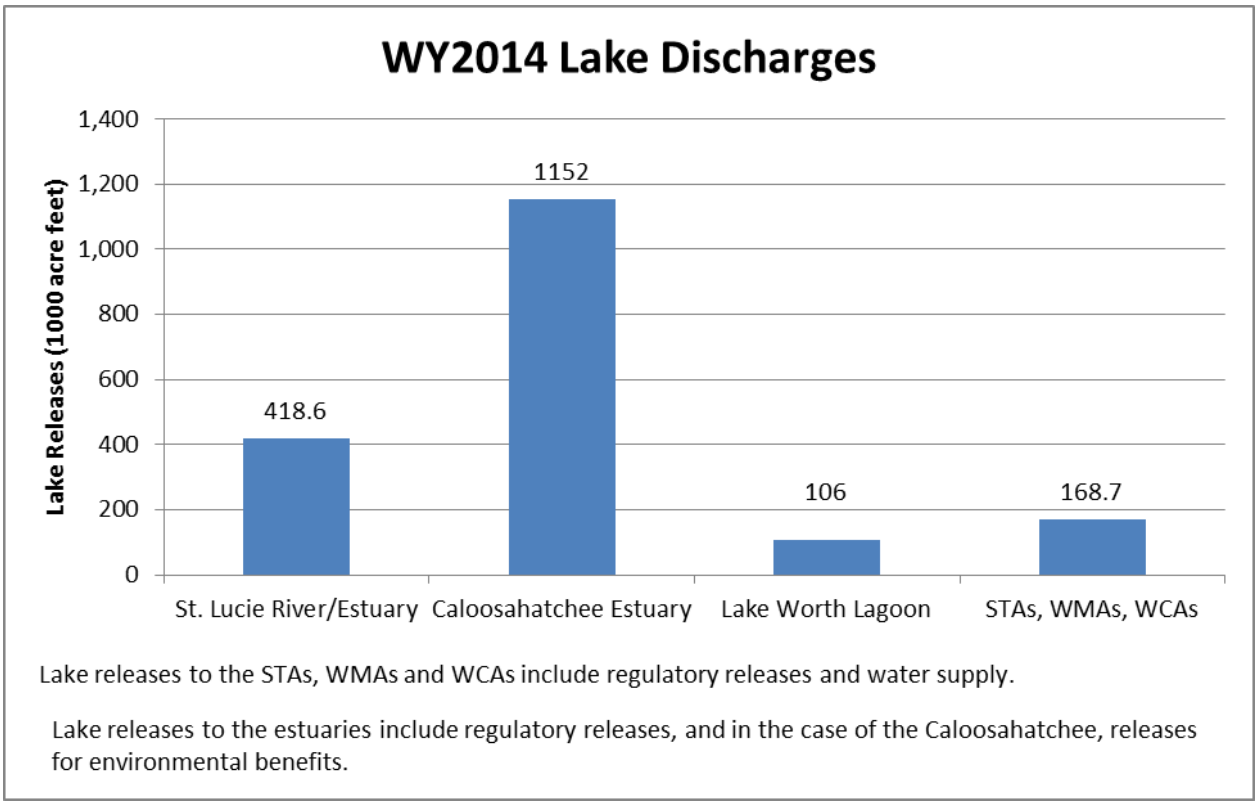


Figure 2. WY2014 flow distribution

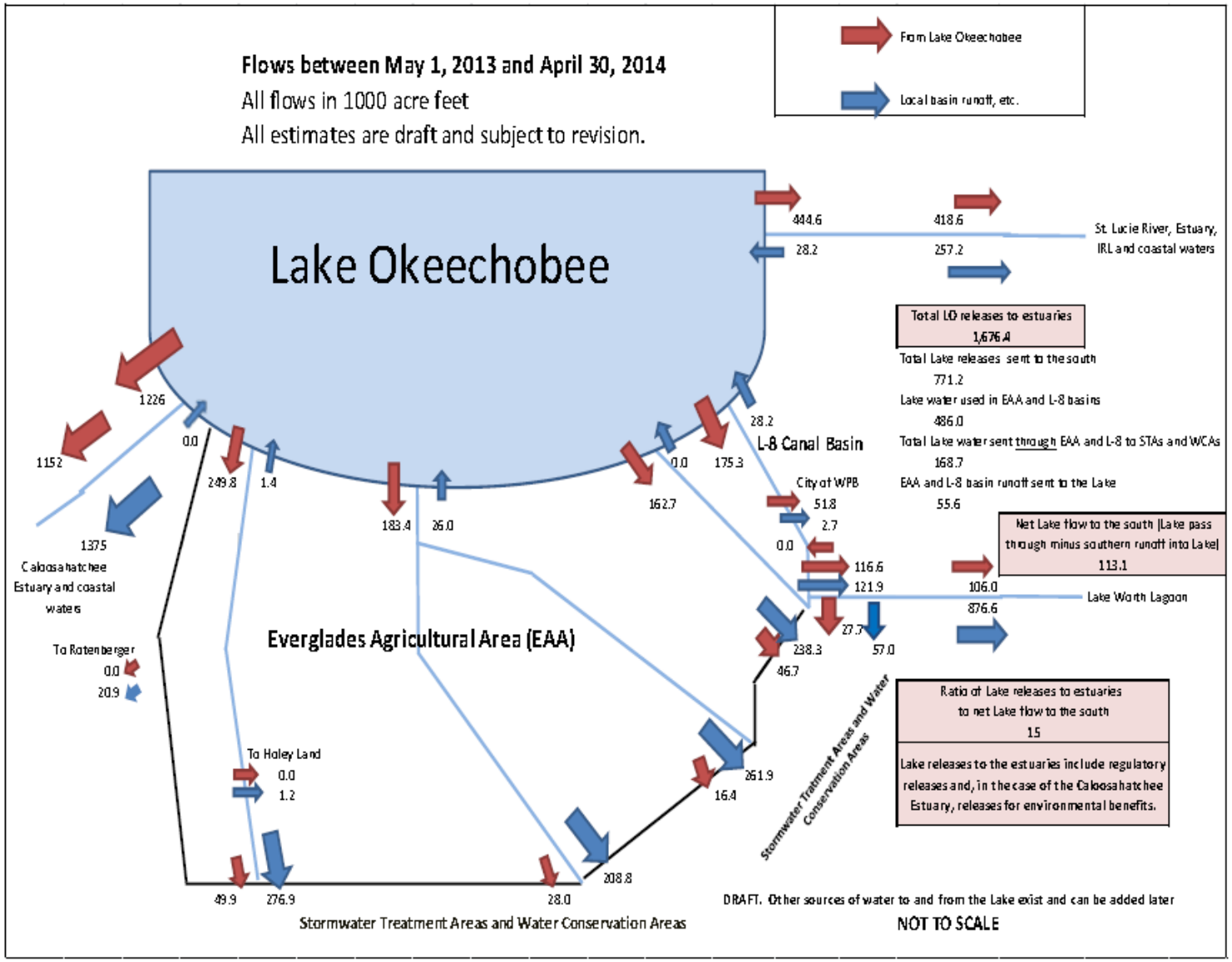
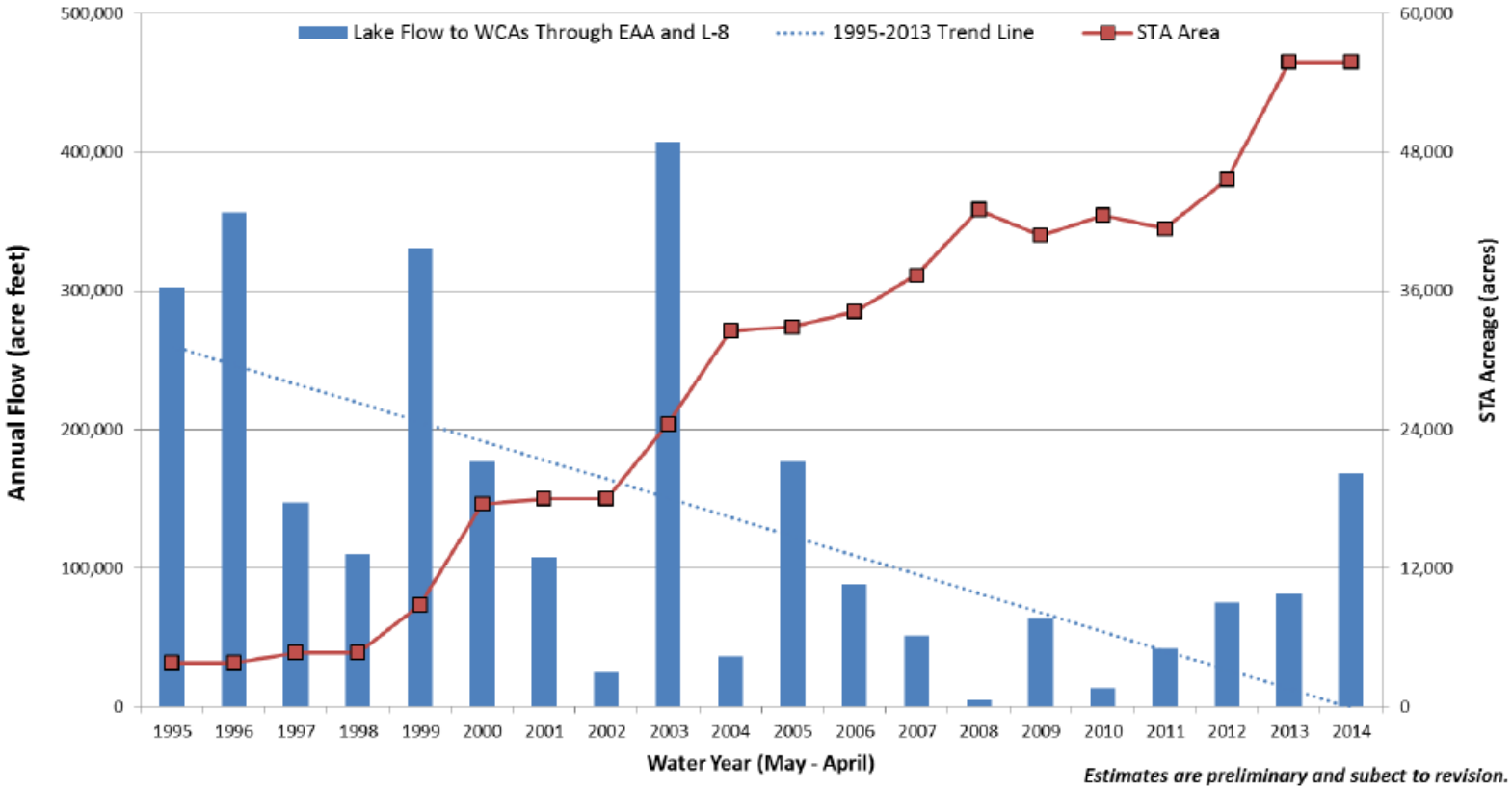


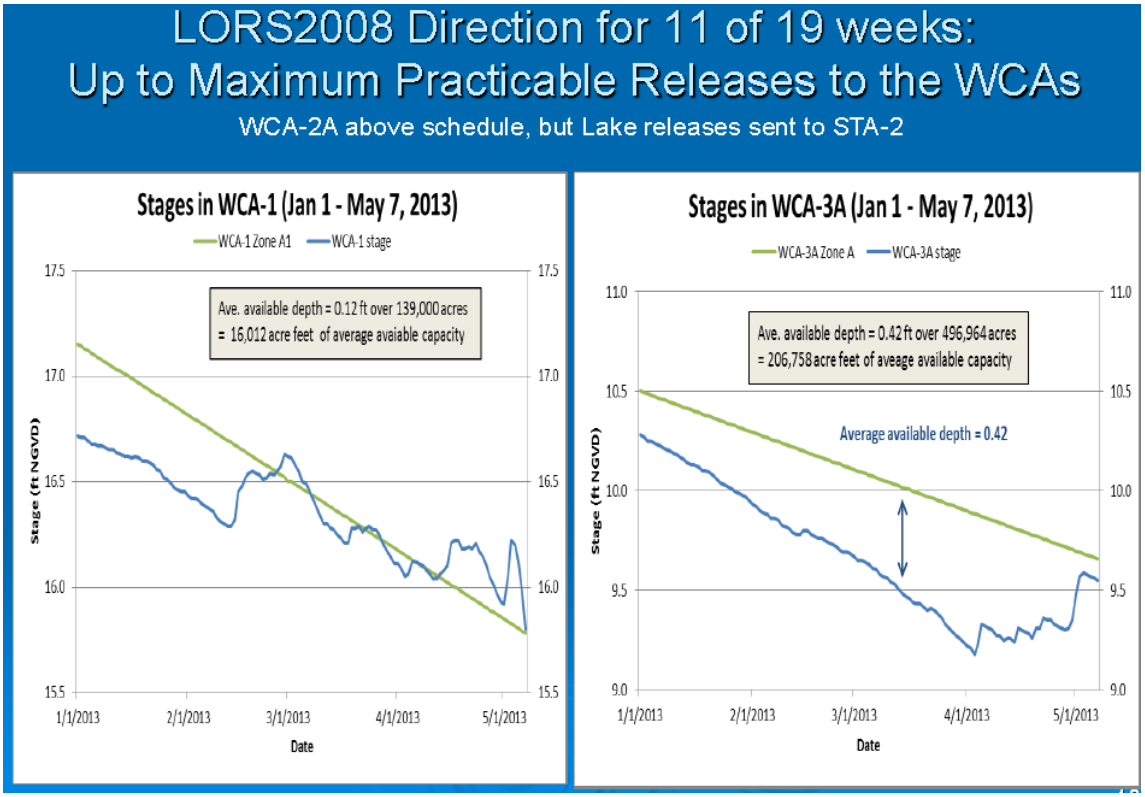
Figure 3.

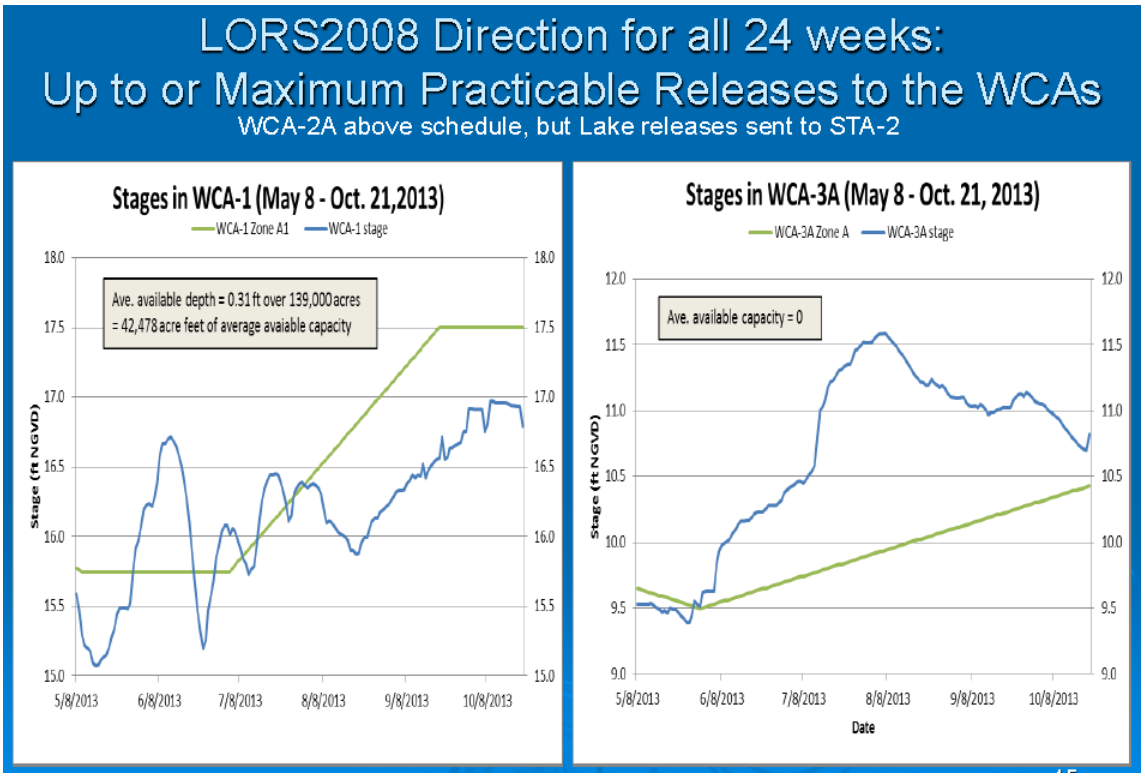


Lake releases south include regulatory releases and water supply that pass through the EAA and L-8 basins via the West Palm Beach, Hillsboro, North New River, Miami and L-8 canals; includes water not treated in STAs.

1979-1988 Base Period Lake flows to WCAs = 100,931 acre feet per year.

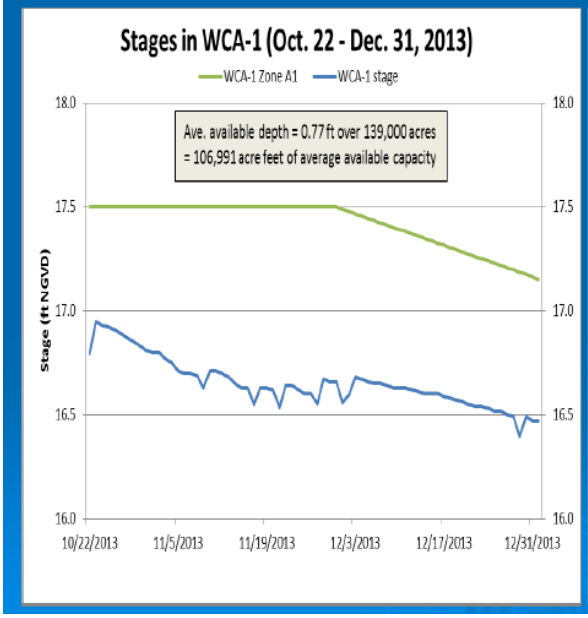
Figure 4 – WCA stages and available capacity

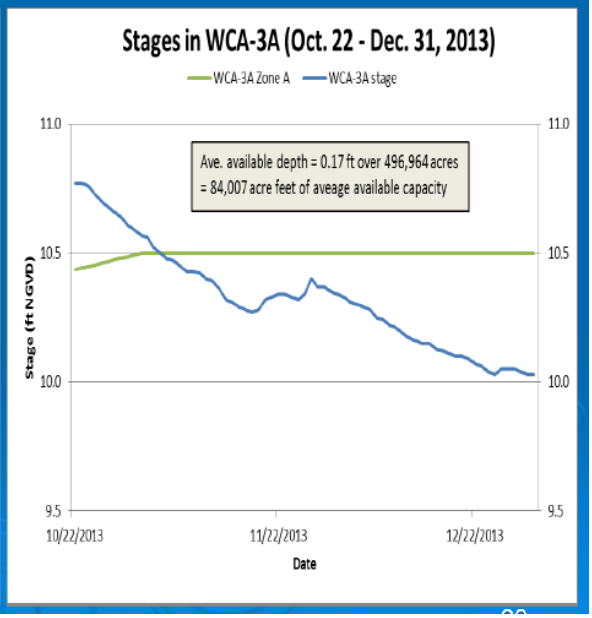




LORS2008 Direction for 9 of 10 weeks: Up to or Maximum Practicable Releases to the WCAs

WCA-2A above schedule, but Lake releases sent to STA-2 at request of Corps.





http://sfwmd.websitetoolbox.com/post/chapt
West Palm ... John Good... Chapter ...
File Edit View Favorites Tools Help
Suggested Sites Web Slice Gallery
Page Safety Tools

Edit | Delete

paul.julian

Registered: 09/06/11
Posts: 4

Posted 10/20 · Edited #3

Authors responses to peer-review panel comments on draft Ch. 3A, posted by P. Julian (10/20/2014).

Attached Files

2015_SFER_Ch3A_RTCs.pdf (415.91 KB, 2 views)

Edit | Delete

ggoforth

Registered: 10/01
Posts: 7

Posted 4 hours ago #4

Suggest that the chapter document and discuss the fact that despite sending approximately 170,000 acre feet of Lake releases to the Everglades carrying approximately 5 metric tons of TP, the impact on water quality in the EPA was insignificant.

Edit | Delete

paul.julian

Registered: 09/06/11
Posts: 4

Posted 3 minutes ago #5

Noted, text will be revised to highlight the volume of water entering the Everglades and the resulting water quality.

Thank You.

Edit | Delete

Previous Topic | Next Topic Selected Posts: -- Select --

Subscribe Reply

Quick Navigation:
Select a Forum

125%

ggoforth

Registered: 10/01
Posts: 7

Posted 22 hours ago #5

Volume I, Appendix 4-2

Included in the calculations summarized in Table 2 was a TP value of **2,500 ppb** reported for the S-351 autosampler on 4/8/2014 (collection dates of 4/1 – 4/8).

- This was the highest TP value at this location reported in the 41-yr period of record (beginning April 9, 1973)
  - This value is 25 times higher than the mean value for the 41-yr record for this station (92 ppb), equal to 25 times the standard deviation; often a value 3 times the standard deviation higher than the mean is classified as an outlier
  - This value is 25 times higher than the median value for the 41-yr record for this station (80 ppb)
  - This value is classified as a statistical outlier using the Maximum Normed Residual test (Snedecor and Cochran 1989), and every other outlier test I applied.
  - The prior autosampler value was 89 ppb (1/28th of the 2,5000 value)
  - The following autosampler value was 185 ppb (1/13th of the 2,5000 value)
  - The grab samples for the same period were 72 ppb and 143 ppb
  - The 7-day autosample value for G-434 (inflow to STA-2 on the NNR Canal downstream of S-351), which received Lake inflow during that 7-day period, was 28 ppb
- The use of this value resulted in a **7-day** estimate of TP load of 19,756 kg
  - which is more than the cumulative load for the entire 358 remaining days of WY2014
  - which is more than the cumulative load for the entire WY2012
  - which is more than the cumulative load for the entire WY2013
- The use of this value resulted in an estimate of the annual flow-weighted mean TP concentration of 170 ppb
  - Which is almost 50 percent higher than the WY2014 annual near-shore TP concentration reported in Table 8-10 of the draft SFER (115 ppb)
  - Which is almost 50 percent higher than the WY2014 annual Lake Okeechobee TP concentration reported on page 8-36 of the draft SFER (118 ppb)
  - Which is more than the WY2014 FWM concentration at S-352 – a phenomenon which has never occurred in the last 20 years
  - Which is almost twice the WY2014 FWM estimated by using the 143 ppb grab sample collected on the same day (87 ppb)

Due to the preponderance of evidence that this value is an outlier, suggest it be excluded from WY2014 calculations.

Edit | Delete

Preliminary Comments on Draft 2015 SFER – G. Goforth 10/22/2014

Ch. 5 – STA Performance

1. The chapter is well-written and contributes another year's performance information regarding the most effective treatment wetlands on the planet! However, the chapter omitted discussion of the benefits and impacts of Lake regulatory releases on the STA performance (see below).

2. The STA management and operations staff should be commended for diligently investigating ways to send additional Lake regulatory releases to the Everglades in order to reduce the devastating releases to the estuaries. For example, through the 2013-2014 dry season, approximately 20 percent more Lake water was sent to the Everglades than during the prior dry season.

3. A critical omission from the chapter is documentation and discussion of the influence of approximately 170,000 acre feet of Lake Okeechobee regulatory releases on the performance of the STAs. See the STA chapters in previous SFERs for precedence (e.g., Goforth et al. 2004, Goforth et al. 2005, and Germain et al. 2012).
    a. An important finding from the WY2014 operations is that although the volumes of Lake water treated were the 2nd highest in 20 years (and the highest since WY2003), the impacts to STA performance were not significant (the chapter documents that WY2014 was one of the best performance years to date). This finding should go a long way to dispelling the fear that treating Lake releases at relatively low hydraulic loading rates of 0.5-1.5 cm/day will cause an adverse impact to STA performance, particularly when spread out over the entire year as called for in LORS2008. Outflow TP concentrations from Lake Okeechobee have decreased significantly since the 2004-2005 hurricanes, when Lake loading on STA performance was last investigated.
    b. Suggest that the chapter document the flows and TP loads associated with Lake releases to each individual STA. In addition to total flows and loads, suggest expressing the Lake releases in terms of hydraulic loading rate (cm/day) and nutrient loading rate (g/m2/yr). For reference, the chapter should compare WY2014 Lake releases to the WY2003 Lake release that overloaded STA-1W (approximately 6 ft per month).
    c. Suggest that the chapter document why certain STA treatment cells were off-line during the devastating May – October 2013 Lake releases to the estuaries. For example, which cells had operations impeded by nesting birds, USACE construction, "resting" of treatment vegetation, vegetation management

activities, etc. Suggest that the chapter provide a rationale why these cells were not receiving the full hydraulic loading they were designed for before and during the emergency operations necessitating Lake releases to the estuaries.

4. As contained in the 2012 EFA Watershed permit for the STAs, Ch. 373.4592 directs the SFWMD to send 28 percent more water to the Everglades through the STAs than occurred during the 1979-1988 period, including an average annual volume of Lake regulatory releases of 250,000 acre feet. Suggest that the chapter document the reasons why this legislative mandate has not been met despite an additional 18,000 acres of treatment area and relatively low outflow concentrations from Lake Okeechobee. Suggest that the discussion address the fact that the 2012 STA permits do not require compliance with the WQBEL for another 10 years.

5. At one time, all of the STA Operational Plans contained guidance to maintain a minimum depth of 6 inches to avoid dryout. Suggest that the chapter document the operations that achieved (or didn't achieve) this operational guidance, and document the number of cells and treatment area that were affected by dryout.

6. Line 11 – Summary: suggest that the section summarize the Lake regulatory releases to the STAs.

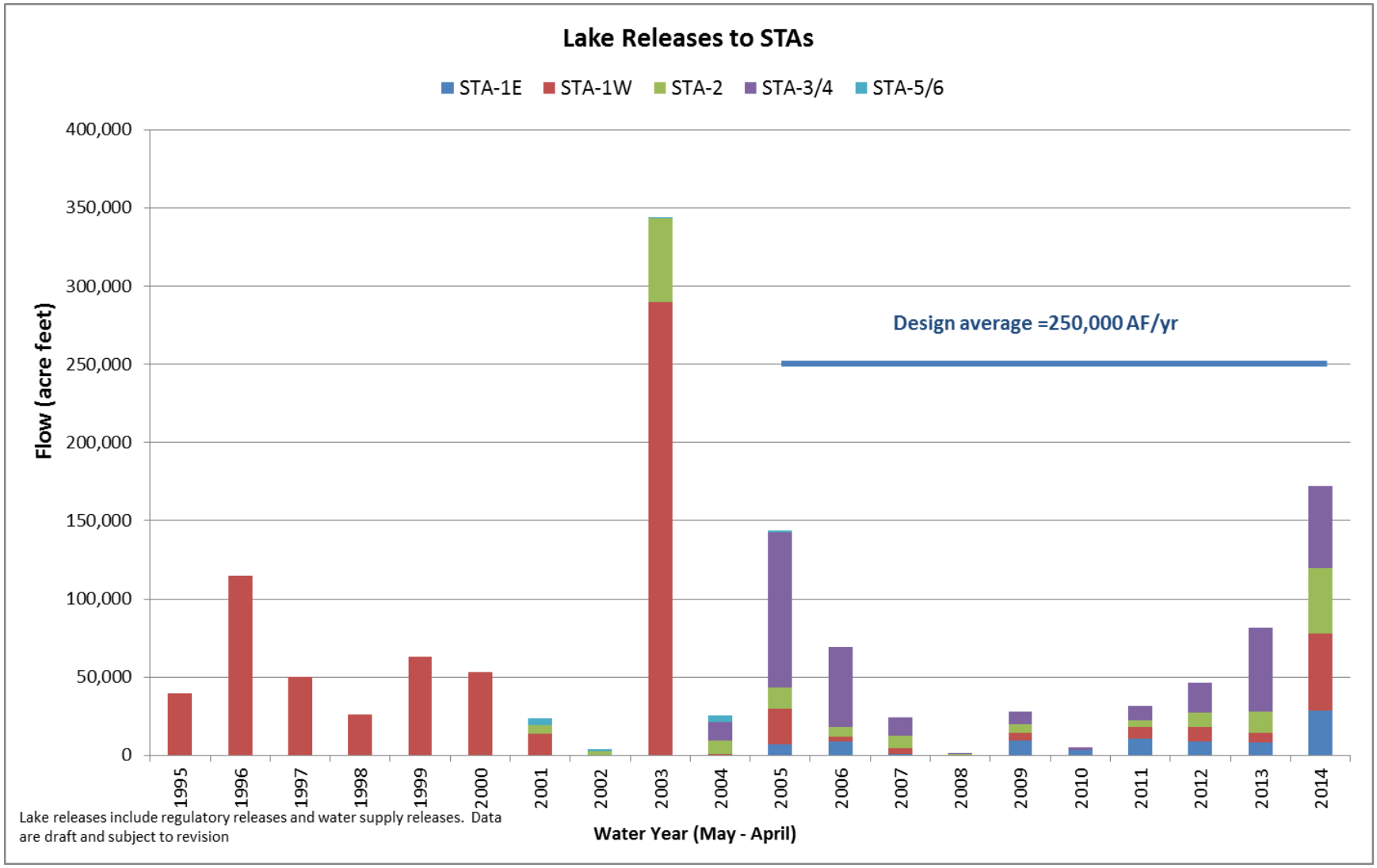
7. Line 290 – The initial water management operation contained in the Avian Protection Plan is to maintain a minimum depth of 6 inches to discourage nesting. Suggest that the chapter document the operations that achieved (or didn't achieve) this operational guidance, and document the number of cells and treatment area that had operations impeded due to nesting birds. When will the District obtain the necessary special purpose permits to avoid operational impedance in the future?

8. Line 355: Lake regulatory releases were also sent to STA-1E in May and October 2013 and January-April 2014. Based on the STA performance over the last two years, it appears that STA performance improved as a result of the Lake releases; suggest that this be documented and discussed.

9. Line 483: Lake regulatory releases were also sent to STA-1W in April 2014. Based on the STA performance over the last two years, it appears that STA performance improved as a result of the Lake releases; suggest that this be documented and discussed.

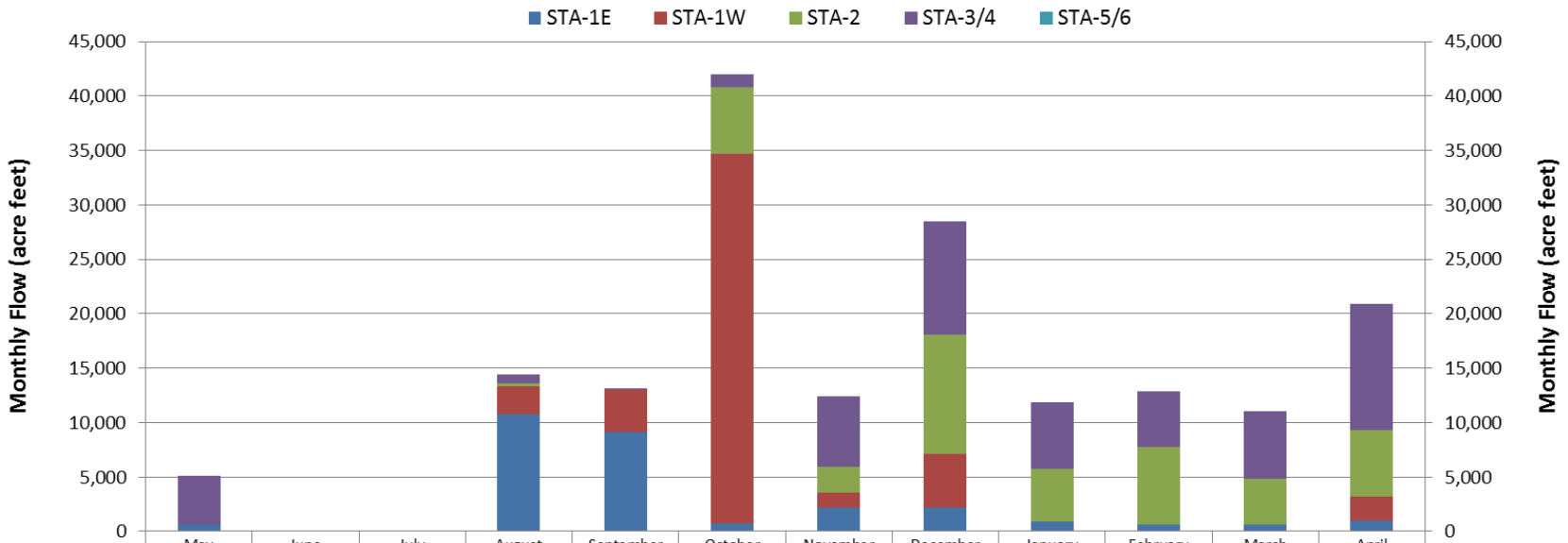
10. Line 600: Lake releases to STA-2 began in October 2013, not November 2013. What, if any, impact on STA performance can be attributed to Lake regulatory releases? Based on the STA performance over the last two years, it appears that STA performance improved as a result of the Lake releases; suggest that this be documented and discussed.

11. Line 722: Lake releases to STA-3/4 began in May, and continued during August and October. Based on the STA performance over the last two years, it appears that STA performance improved as a result of the Lake releases; suggest that this be documented and discussed.


Lake Releases to STAs
STA-1E
STA-1W
STA-2
STA-3/4
STA-5/6
Flow (acre feet)
400,000
350,000
300,000
250,000
200,000
150,000
100,000
50,000
0
Design average =250,000 AF/yr
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Water Year (May - April)
Lake releases include regulatory releases and water supply releases. Data are draft and subject to revision

Estimates are preliminary and subect to revision.

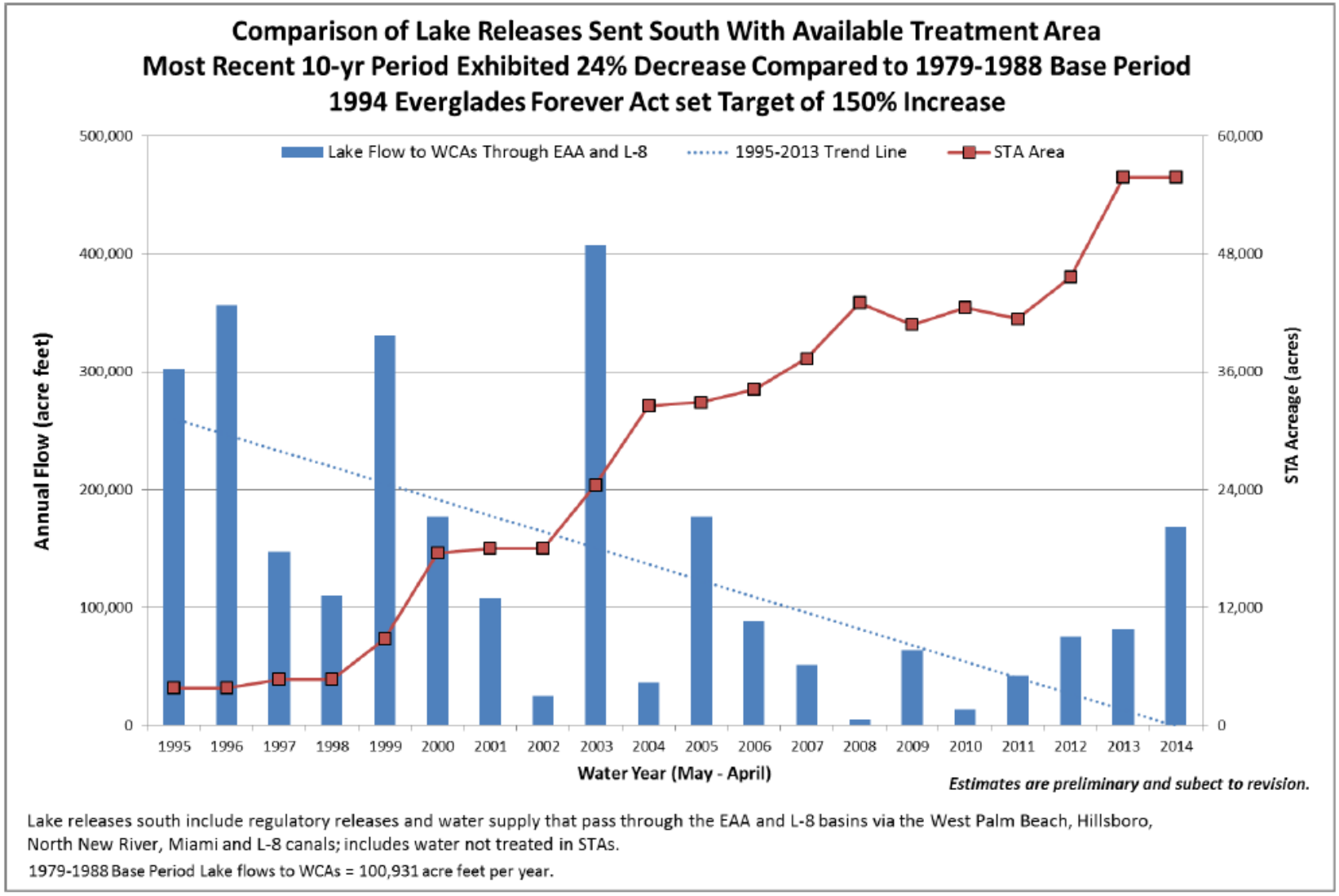
Lake Releases to the STAs



	May	June	July	August	September	October	November	December	January	February	March	April
STA-5/6	0	0	0	0	0	0	0	0	0	0	0	0
STA-3/4	4,474	0	0	821	0	1,210	6,415	10,383	6,110	5,120	6,183	11,601
STA-2	0	0	0	259	0	6,092	2,394	10,997	4,842	7,059	4,227	6,067
STA-1W	0	0	0	2,499	3,913	34,008	1,385	4,853	0	0	0	2,239
STA-1E	593	0	11	10,805	9,133	704	2,175	2,219	890	651	624	960

Lake releases to the STAs include regulatory releases and water supply.

Figure 3.


Comparison of Lake Releases Sent South With Available Treatment Area
Most Recent 10-yr Period Exhibited 24% Decrease Compared to 1979-1988 Base Period
1994 Everglades Forever Act set Target of 150% Increase
Lake Flow to WCAs Through EAA and L-8
1995-2013 Trend Line
STA Area
Annual Flow (acre feet)
STA Acreage (acres)
500,000
400,000
300,000
200,000
100,000
0
60,000
48,000
36,000
24,000
12,000
0
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Water Year (May - April)
Estimates are preliminary and subect to revision.
Lake releases south include regulatory releases and water supply that pass through the EAA and L-8 basins via the West Palm Beach, Hillsboro, North New River, Miami and L-8 canals; includes water not treated in STAs.
1979-1988 Base Period Lake flows to WCAs = 100,931 acre feet per year.

Preliminary Comments on Draft 2015 SFER – G. Goforth 10/22/2014

Ch. 8 – Lake Okeechobee

1. The chapter is well-written and contains useful information. However, the chapter lacks discussions of
    a. the benefits and impacts of the WY2014 Lake regulatory releases, and
    b. the status of the rehabilitation of the Herbert Hoover Dike

2. The District's 2005 SFER anticipated that the Lake's TMDL would be achieved by 2015, as mandated by the State's 2000 Lake Okeechobee Protection Act. What is the current estimate when the Lake's TMDL will be achieved?

3. A critical omission from the chapter is the documentation and discussion of the destructive impacts of Lake regulatory releases on the St. Lucie and Caloosahatchee Rivers and Estuaries. These releases were the subject of numerous WRAC and other public meetings, generated multiple citizen protests during which thousands of citizens expressed their outrage, resulted in the formation of a State Legislative Committee which held multiple public hearings, and was the focus of a federal Congressional hearing. During WY2014, more than 1,600,000 acre feet of Lake Okeechobee regulatory releases were discharged to the St. Lucie River and Estuary, the Caloosahatchee River and Estuary and the Lake Worth Lagoon.
    a. Flows and Nutrient Loads from Lake Releases
        i. St. Lucie River and Estuary:
            1. Flow - more than 418,000 acre feet
            2. Nitrogen – 1.43 million pounds - almost 6 times the TMDL
            3. Phosphorus – 154,000 pounds - almost 6 times the TMDL
            4. Total Suspended Solids – 19 million pounds (rough estimate)
        ii. Caloosahatchee Estuary:
            1. Flow – more than 1,150,000 acre feet
            2. Nitrogen – 4.25 million pounds – 36% more than the TMDL
            3. Phosphorus – 236,000 pounds
            4. Total Suspended Solids – 20 million pounds (rough estimate)
        iii. Lake Worth Lagoon:
            1. Flow - More than 100,000 acre feet
            2. Phosphorus - 37,800 pounds of TP
        iv. Total to estuaries: 426,800 pounds of TP
    b. By contrast, while the estuaries were being devastated by these destructive Lake releases, less than 170,000 acre feet of Lake Okeechobee regulatory

releases were discharged to the Everglades – the pre-development destination for excess Lake water. These Lake flows carried only 5 metric tons of TP to the Everglades due to prior treatment in STAs … less than 3% of the load to estuaries. During this period, TP concentrations in the Refuge (WCA-1) ranged from 6-8 ppb, and inflows to Everglades National Park were approximately 8 ppb.

c. Suggest the chapter document the water management operations during WY2014 leading up to the regulatory releases as well as documenting the volumes of water and the mass of nutrients and sediment discharged to the St. Lucie and Caloosahatchee River and estuaries.
d. Suggest the chapter document the destructive environmental and economic impacts to the coastal regions associated with these releases, and compare those impacts to the benefits of the regulatory releases.

4. Another critical omission is discussion of the status of the rehabilitation of the Herbert Hoover Dike by the US Army Corps of Engineers (USACE). The destructive Lake regulatory releases to the estuaries during WY2014 are a product of the current LORS2008, which is a direct result of state and federal government's efforts to reduce the risk of dike failure.

5. Line 115: Lake regulatory discharges were also made via Lake Culvert 10A into the L-8 Canal and eventually to Lake Worth Lagoon.

Preliminary Comments on Draft 2015 SFER – G. Goforth 10/22/2014

Ch. 10 – River Protection Plans

1. The chapter is extremely well-written, comprehensive and contains very useful information for stakeholders: outstanding work by the authors. However, there are three omissions from the chapter which should be addressed in the final version.

2. The first critical omission from the chapter is the documentation and discussion of the destructive <u>nutrient and sediment loading impacts</u> of Lake regulatory releases on the St. Lucie and Caloosahatchee Rivers and Estuaries. These releases were the subject of numerous WRAC and other public meetings, generated multiple citizen protests during which thousands of citizens expressed their outrage, resulted in the formation of a State Legislative Committee which held multiple public hearings, and was the focus of a federal Congressional hearing. During WY2014, more than 1,600,000 acre feet of Lake Okeechobee regulatory releases were discharged to the St. Lucie River and Estuary, the Caloosahatchee River and Estuary and the Lake Worth Lagoon.
    a. Flows and Nutrient Loads from Lake Releases
        i. St. Lucie River and Estuary:
            1. Flow - more than 418,000 acre feet
            2. Nitrogen – 1.43 million pounds - almost 6 times the TMDL
            3. Phosphorus – 154,000 pounds - almost 6 times the TMDL
            4. Total Suspended Solids – 19 million pounds (rough estimate)
        ii. Caloosahatchee Estuary:
            1. Flow – more than 1,150,000 acre feet
            2. Nitrogen – 4.25 million pounds – 36% more than the TMDL
            3. Phosphorus – 236,000 pounds
            4. Total Suspended Solids – 20 million pounds (rough estimate)
        iii. Lake Worth Lagoon:
            1. Flow - More than 100,000 acre feet
            2. Phosphorus - 37,800 pounds of TP
        iv. Total to estuaries: 426,800 pounds of TP
    b. By contrast, while the estuaries were being devastated by these harmful Lake releases, less than 170,000 acre feet of Lake Okeechobee regulatory releases were discharged to the Everglades – the pre-development destination for excess Lake water. These Lake flows carried only 10,100 pounds of TP to the Everglades due to prior treatment in STAs … less than 3% of the load to estuaries. During this period, TP concentrations in the Refuge (WCA-1) ranged from 6-8 ppb, and inflows to Everglades National Park were approximately 8 ppb.

c. Suggest the chapter document the water management operations during WY2014 leading up to the regulatory releases as well as documenting the volumes of water and the mass of nutrients and sediment discharged to the St. Lucie and Caloosahatchee River and estuaries.
d. Suggest the chapter document the destructive environmental and economic impacts to the coastal regions associated with these releases.
e. Nutrient and sediment loads to the estuaries are not currently presented as part of the monthly staff updates to the Governing Board – could they be included?

3. Line 89: Dispersed Water Management Projects ("DWMPs")
    a. Suggest that references to storage volumes from DWMPs include the estimated reduction in flows and nutrient loads **at the Lake (or estuary)** instead of just presenting the estimated storage at the project location, since the attenuation between the project location and the Lake (or estuary) may render any project benefits moot. Intuitively, the attenuation factor (defined as the measured flow reduction at the waterbody divided by flow reduction at the DWMP site) would be 1.0 for a DWMP located on the waterbody, and decreases to 0 the farther upstream the project site is located due to water reuse and floodplain losses/attenuation.
    b. At the January 2014 Governing Board meeting, the SFWMD Executive Director told the Governing Board he had directed staff to perform an evaluation of the DWMPs to ensure they are truly effective at reducing flow and removing nutrients from the waterbody of interest. When will the results of this evaluation be available?
    c. Table 10-1 estimates annual storage benefits of DWMPs.
        i. Are discharges being monitored for flow and nutrient loads?
        ii. What are the **actual** annual reductions measured at the project location and at the downstream waterbody?
        iii. Are the landowners held accountable for achieving the estimated annual benefits, i.e., if the projected benefits are not achieved, what are the ramifications?
        iv. What is the current annual volume of water being discharged from the lands to be used as the DWMP?
        v. How were the flow reductions estimated?
        vi. Are the estimated reductions projected for Lake Okeechobee (or the River basins), or estimated reductions from the landowner's property? As you know, in general the further the discharge point is from the Lake, the smaller the actual flow reduction would be.

vii. Are these reductions permanent, i.e., the water will be held and used on site, or temporary, i.e., released back to receiving waters after a storm event through surface or groundwater discharges?

viii. How will the actual flow reductions be monitored?

ix. What accountability for achieving the estimated flow reductions does the District place on each DWMP? For example, if the actual flow reductions do not meet the estimated flow reductions, will the landowner return a portion of the public funds paid for the DWMP?

d. What are the nutrient reduction benefits of the DWMPs listed in Table 10-1 **measured at the Lake or estuary**?

i. How cost effective are the projects listed in Table 10-1 relative to regional treatment projects such as an STA? How were the load reductions estimated?

ii. Are these reductions permanent, i.e., the water will be held and used on site, or temporary, e.g., released back to receiving waters after a storm event?

iii. How will the actual nutrient load reductions be monitored?

iv. If the actual nutrient load reductions do not meet the estimated load reductions, will the landowner return any of the funds paid for the DWMP?

v. What accountability for achieving the estimated nutrient load reductions does the District place on each DWMP owner?

vi. What is the current annual phosphorus load and phosphorus concentration in waters being discharged from the lands to be used as the DWMP?

vii. What is the current Works of the District (e.g., 40E-61) permit limit for phosphorus for the same land?

viii. When evaluating DWMPs, does the District consider whether or not the property is meeting the existing Works of the District permit condition?

e. For each DWMP listed in Table 10-1:

i. What is the total cost, separated into initial cost and continuing costs?

ii. If the landowner has a consumptive use permit, or a pending permit application,

1. What is the maximum annual volume of water allocated by the permit, or requested in the permit application?
2. When evaluating DWMPs, does the District consider whether or not the landowner has a consumptive use permit, or a pending application?

f. In their October 2013 Board presentation, District staff reported that the average cost for a DWMP was $163 per acre foot of water per year.
    i. What basis did the District use to establish this value as a fair cost?
    ii. Has that average cost been revised? If so, what is the current average annual cost per acre foot of water per year?

g. The latest 5-year average annual phosphorus loads discharged to Lake Okeechobee exceeded 420 metric tons, more than four times the Total Maximum Daily Load (TMDL) established for surface water inflows to Lake Okeechobee. The District has a Lake Okeechobee Protection Plan, which contains water quality projects that will reduce the nutrients entering the Lake. According to staff's presentation at the January 2014 Governing Board meeting many of these projects are not being completed due to insufficient funds.
    i. Which projects in the District's Plan could be implemented if the funds obligated to the DWMPs were instead expended to finish the projects?
    ii. What is the estimated average cost per pound of phosphorus removal for the water quality projects in the Lake Okeechobee Protection Plan?
    iii. How do these costs compare to the unit costs for the DWMPs?
    iv. Are the DWMPs located in tributaries identified in the Protection Plan as needing additional storage?
    v. The District's 2005 annual report anticipated that the Lake's TMDL would be achieved by 2015, as mandated by the State's 2000 Lake Okeechobee Protection Act. What is the current estimate when the Lake's TMDL will be achieved?

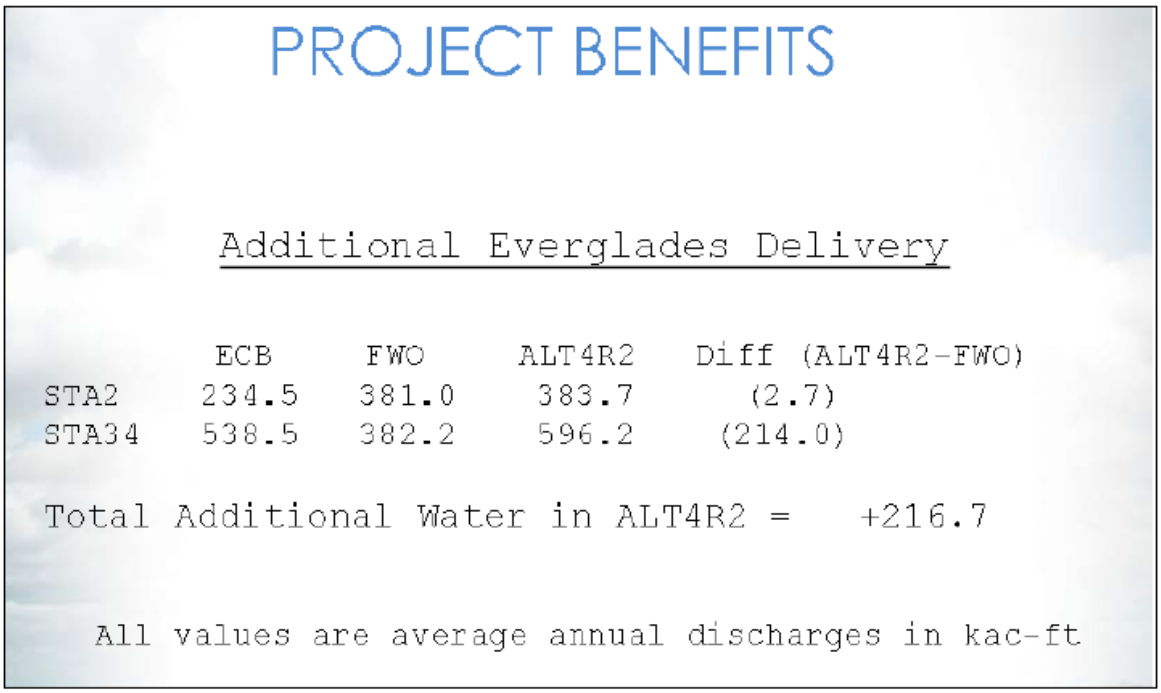
4. Line 2061: A second critical omission in the St. Lucie River Watershed Protection Plan (SLRWPP) is the lack of a project to treat Lake Okeechobee regulatory releases prior to discharge into the St. Lucie River and Estuary. Without a dedicated project to remove nutrients from Lake Okeechobee releases to the St. Lucie River and Estuary, there is no possibility of achieving the TMDL for the C-44 Canal, or achieving the desired nutrient concentrations in the estuary.

5. Central Everglades Planning Project (CEPP). A third critical omission is lack of clarification on the benefits of CEPP to reducing Lake releases to the estuaries.

    a. Line 1064:
        i. Suggest that all mention of additional Lake water to the Everglades through CEPP elements document the <u>absolute</u> value of Lake water sent to the Everglades, since it isn't clear what the "additional flow" is relative

to. For example, in the August 2013 WRAC presentation (see below), what is the estimated Lake releases to the Everglades in the "FWO" scenario, and what is the estimated Lake releases to the Everglades in the "ALTT4R2" scenario? Doesn't the FWO scenario include 250,000 acre feet per year of Lake releases to the Everglades? This section conflicts with information presented in Line 2427, which states the absolute magnitude of Lake releases to the Everglades is only 210,000 acre feet (average annual) – well below the (250,000 + 215,000 = 465,000 AF/yr) estimate.

ii. What are the projected reductions in Lake releases to the estuaries as a result of CEPP? As you know, there is not a 1:1 relationship between the quantity of additional Lake water sent to the Everglades and a reduction in Lake water sent to the estuaries.



Source: SFWMD presentation on CEPP at the August 8, 2013 WRAC meeting

b. Line 1064: Suggest that any mention of additional Lake water to the Everglades through CEPP elements clearly state any water quality constraints, e.g., SFWMD staff Tom Teets stated to the Governing Board that additional Lake water will be sent through CEPP elements only if the WQBEL is met. Since no STA has ever achieved the WQBEL over a 5-yr assessment period, there is reason to doubt any

additional lake water will be sent to the Everglades, and the entire $1.8 billion CEPP project will be used solely for conveying EAA runoff.

c. Line 2427: This information (210,000 AF/yr of Lake releases to the Everglades) conflicts with the information beginning on line 1064 – "additional 215,000 AF/yr"

6. Lines 1353, 1433, and 1448: Suggest that these discussions include Lake Okeechobee regulatory releases since they are a dominant source of flow, nutrient loads and sediment loads to the rivers and estuaries.

7. Line 1638: Similar comments apply to the CRE as was provided for the SLRE above regarding documenting the impacts of the WY2014 Lake regulatory releases.

8. Line 2191: C-44 Project.
    a. Please cite and make available the technical reports documenting the estimated nutrient load reductions of the C-44 project.
    b. Please cite and make available the technical reports documenting the estimated flow reductions of the C-44 project.
    c. Since the C-44 Project is not authorized to capture and treat Lake Okeechobee releases, what is the estimated volume of Lake water that may incidentally be captured and treated by the project?
    d. When will the operations plan for coordinating the operations of S-308, C-44 inflow/outflow pumps and S-80 be developed?

9. Line 2248: Will the activities identified in Table 10-10 be completed before additional public funds are allocated for DWMPs? If not, how can the SFWMD justify expenditure of public funds for projects that may not be effective, cost-effective, or helpful to achieving the nutrient reduction goals of the River Protection Plans?

ggoforth

Registered: 10/01
Posts: 7

Posted 21 hours ago #4

Volume III, Appendix 3-1

1. Line 96: One structure with arguably the lowest TP concentration leaving STA-3/4 is G-388 (the PSTA outflow pump station; see Figure 2), which has exhibited annual flow-weighted mean concentrations ranging from 8 to 13 ppb for the last six years. Why isn't this structure listed in Table 3-1? Why isn't this structure listed in the EFA Watershed permit or the NPDES permit for the STAs? Is the District operating outside the authorization of these permits by discharging through this structure? Why doesn't the District request a permit modification to include this structure in those permits?

Edit | Delete

Previous Topic | Next Topic Selected Posts: -- Select --

Subscribe Reply

Quick Navigation: Select a Forum

Rivers Coalition
Save Our River! Stop the Discharges!
P.O. Box 2627, Stuart, FL 34995 772-225-6849

October 23, 2014

The Rivers Coalition, representing over 70 organizations, endorses Gary Goforth's comments regarding the South Florida Water Management District's Annual Report. The River's Coalition calls for the SFWMD to address the obvious shortcomings and their failure to adequately report the adverse impacts of the Lake Okeechobee releases in their Annual Report.

Sincerely,

Leon Abood, Chairman
Rivers Coalition