Time series trend analysis and prediction of water quality in a managed canal system, Florida (USA)

L Wan¹,³ and Y C Li²
¹Xuzhou Institute of Technology, No. 2 Lishui Road, New city district, Xuzhou 22111, China
²Soil and Water Science Department at Tropical Research and Education Center, IFAS, University of Florida, 18905 SW 280th Street, Homestead, FL 33031, USA

E-mail: hjwanl@163.com

Abstract. Water quality in the Everglades canal system discharging into the St. Lucie Estuary in south Florida is greatly influenced by anthropogenic activities. We analyzed the water quality parameters trend from 1979 to 2014 using seasonal trend decomposition and LOESS of four canals (C23, C24, C25, and C44). We also predicted the value of total nitrogen (TN) and total Phosphorus (TP) and turbidity from 2016 to 2020 using exponential smoothing and additive Holt-winters method. The main goals of this paper were to reveal long-term changes trends in water quality and identify the emphasis of management in the near future. We found that the change range of water quality parameters was different, but change rules were similar. The mean of DO at each station was in the range of 5.0~6.3 mg l⁻¹, and the colour was in the range of 5 ~ 400PCU. The highest specific conductivity was 2250 µS cm⁻¹ in C24 and the lowest was 125 µS cm⁻¹ in C44, and the minimum turbidity was only 0.4 NTU at C25 and the maximum value was 90.3 NTU at C44. The mean value of TN at each station was C24>C44 >C23>C25, and TP was C24>C23>C44>C25. Nutrient export exhibited an upward trend for TN from all the basins and downward trend was present for TP. Water quality was influenced by rainstorm and non-point source pollution. Our predictions showed that TN and TP concentrations in C23 canal were higher than other three canals every month and turbidity of C44 canal was higher than other three canals. Therefore, improved best management practices should be implemented on land that have high TN and TP concentrations in runoff and leachate flowing into the C23 canal.

1. Introduction
Since 1970s last century, researchers carried out a lot of monitoring, modeling, planning, control and restoration on the aquatic environment and ecosystems. And they got abundant time series data from these long-term monitoring programs. How to evaluate water quality trends, identify indicators of water quality characteristics has been one of wide concern. Water quality trend analyses are useful to evaluate the effectiveness of corrective actions, and determine whether watershed programs have been successful in meeting standards set by state and federal governments [1]. Recently, many statistical methods was used to evaluate trends in water quality over time; these include the Seasonal Kendall tests, Chi-square test, the Mann-Kendall and so on [2-6]. However, these methods constrained water quality parameters data to linear or monotonic trends and cannot detect intermediate reversals in direction [7]. Due to human disturbance and other natural factors, changes in water quality cannot be readily identified with linear and monotonic assumptions. To solve this problem, STL (Seasonal Trend...
Decomposition using LOESS method is applied to water quality evaluation, which is a way to deal with the nonlinear and non-parametric statistical method of local trends using locally weighted LOESS fitting [8,9]. The LOESS method is using locally weighted regression and scatterplot smoothing (LOESS). We used the STL methodology to reveal the change trend of water quality in a managed canal system in this study.

It is essential that trend prediction of water quality data is accurate. Such accuracy is key to enabling manager to choose the best management option that meets a large number of identified goals [10]. Recently, water quality modeling using for trend prediction included: wavelet-artificial intelligence hybrid models [11], two-dimensional water quality numerical model MIKE 21 AD [12], multiple regression and multiple linear regressions model [13] and so on. Some researchers have used time series prediction method to predict water quality parameters change trend recently. For example, Parmar et al [14] used statistical analysis and time series prediction to deal with water quality management. Wang [15] used time series prediction method to analyze the variation trend of the eutrophication in Sansha Bay, Ningde City, Fujian Province, China. Time series prediction method is also used in this paper for analysis. The time series prediction method is used in study has two main advantages: (1) it is the most simple method among all of the methods, and this method is based on the historical trends in water quality change trend along the same path, with no structural changes taking place; (2) forecasts water quality directly without considering factors on which water quality change depend [14,16].

The objectives of this study were to (i) using the STL method to identify water quality long-term changes trend in the past 35 years when land use changed (agricultural development culminated in about 2000 and then became stable afterwards with a corresponding decrease in natural lands over the past 35 years [17]) and hydrological events occurred in this basin; and (ii) predict the water quality change trend so as to identify the emphasis of management in the near future.

2. Methods

2.1. Study area

The Southern Indian River Lagoon (IRL) is a biogeographic transition zone from the canal system to Atlantic Ocean, and the St. Lucie Estuary (SLE) is the largest tributary of the IRL. C-44, C-23, C-24, and C-25 canals are managed by the South Florida Water Management District (SFWMD) and the U.S. Army Corps of Engineers (USACE). C-44 canal was completed in 1924 and enlarged size in 1949. C-23, C-24, and C-25 were completed about 1961 [18]. Freshwater discharges from Lake Okeechobee to the primary canals are controlled through seven water control structures S-48, S-49, S-50, S-80, S-97, and S-99 (figure 1). Each station was named by the combination of the structure and the primary canal; for example, the station at S-50 in the C-25 canal is C25S50 [18].

2.2. Data sources

Water quality data (monthly) collected at C25S50, C24S49, C23S48 and C44S80 from January 1979 to April 2014 were downloaded from the SFWMD (http://my.sfwmd.gov/dbhydroplsql/show_dbkey_info.main_menu) database and were used for this analysis. The parameters selected for analysis are: dissolved O$_2$ (DO), specific conductivity, turbidity, color, Total P (TP), Total N (TN= NO$_X$—N + TKN). Analytical methods and detection limits for selected nutrient species are presented in table 1. The total number of water quality samples used for this study was 424 groups. Missing data were supplemented by linear interpolation (missing data=24 groups).
Figure 1. Location of water monitoring stations.

Table 1. Selected water quality parameters analytical methods.

| Water-quality parameter | Detection Limit | Analytical Method# |
|-------------------------|-----------------|--------------------|
| Color                   | 1.0 PCU*0.1 mg l⁻¹ | SM 2120B          |
| Specific conductivity   | 1 mS/cm          | EPA 120.1          |
| NO₃–N                  | 0.004 mg l⁻¹     | SM 4500NO3F        |
| Total P                | 0.002 mg l⁻¹     | SM 4500PF          |
| Total Kjeldahl N       | 0.05 mg l⁻¹      | EPA 351.2          |
| Dissolved O₂           | 0.1 mg l⁻¹       | EPA 360.1          |
| Turbidity              | 0.1 NTU*         | SM 2130B           |

* Platinun–cobalt units.
* Nephelometric turbidity units.
# SM methods are from the American Public Health Association; EPA methods are from the USEPA.

2.3. Seasonal-trend decomposition using loess (STL) method

The times series of indicators were analyzed by Seasonal-Trend Decomposition using LOESS (STL) method. STL occupies much wider range of component patterns than any single parametric methods [19]. LOESS is considered as a smoother and a non-parametric statistical method. In STL, observed time series are decomposed into trend component, seasonal component and residuals component [7].
Trend component with low frequency can be considered as change tendency. Seasonal component with high frequency can be viewed as variations with stable seasonal disturbance. Residuals component with random disturbance can be regarded as irregular variation. STL method is a recursive process and implements three times of LOESS and moving average process in each recursion, respectively [20,21]. Water quality data was analyzed by the STL method using R3.1.0 version (http://www.r-project.org/). In the internal circulation process of STL, \( Y_v, S_v, T_v, R_v \) represented observed value, seasonal component, trend component, and residuals component at \( v \) moment after \( k \) iterations, respectively (figure 2) [21].

![Figure 2. Inner loop procedure of STL.](image)

2.4. *Times series for trend prediction*

Time series models will often make use of natural one-way ordering of time so that values for a given period will be expressed as deriving in some way from past values, rather than from future values [14]. SPSS (19.0) software was used for trend prediction of water quality parameters. We selected exponential smoothing and additive Holt-winters method. Holt-winters method and additive Holt-winters method was suitable for time series with constant seasonal variations. This model was used to forecast the variation of TN, TP and turbidity in the 5 years of 2016~2020.

3. *Results and discussion*

3.1. *Basic statistics of water quality parameters at each station*

The basic statistics of the monthly measured data on water quality were summarized with box plots (figure 3). Dissolve Oxygen (DO) was measured from 8:00 am to 3:30 pm. The mean of DO at each station was in the range of 5.0~6.3 mg l\(^{-1}\). The maximum value (16.6 mg l\(^{-1}\)) and the minimum value (0.3 mg l\(^{-1}\)) of DO were both at C24. The maximum values of DO at four stations all exceeded 12 mg l\(^{-1}\), which showed very high primary productivity in the water. But the minimum values of DO were only 0.3~0.6 mg l\(^{-1}\) which showed the water was sometimes under the condition of low oxygen. The maximum value of color was 400 platinum color units (PCU) at C24, and the minimum value was 5 PCU at C44. The mean value at each station was C24>C23>C25>C44. The highest specific conductivity was 2250 µS cm\(^{-1}\) in C24, and the lowest was 125 µS cm\(^{-1}\) in C44. The turbidity change
in this zone was most likely influenced by rainstorm and non-point source pollution. The minimum value was only 0.4 NTU at C25, and the maximum value was 90.3 NTU at C44. The mean value of TN at each station was C24>C44>C23>C25, and TP was C24>C23>C44>C25. The mean N and P concentration was lower in C25 and higher in C24. Ambient water quality criteria recommendation for the study area (Nutrient Ecoregion XIII, TN 1.27 mg l\(^{-1}\) and TP is 17.5 μg l\(^{-1}\)) (Environmental Protection Agency Office of Water 2000), the concentrations of TN at C23, C24, C25 and C44 was relative lower (median was 1.05 to 1.24 mg l\(^{-1}\)), but TP were relative higher (median was 85 to 218.5 μg l\(^{-1}\)) during the recently years (2011~2014).

Figure 3. Box plots of water quality parameters at each station.
3.2. Trend analysis

During the past 35 years (1979-2014), DO concentration was in trough from 1990 to 1995 (figure 4). After 2010, DO concentration was downward trend. From north to south, DO showed increasing trend. There was no evident relationship between temperature and DO. Color was in trough from 1985 to 1990 and was low in all the 4 stations about 2010. Turbidity was upward trend basically in the past 35 years in the 4 stations, and there were two peak values in early 1980s and late 2000s. When turbidity was high, the total suspended solids concentration was also high. For example, in C25S99 station, the maximum value of turbidity was 170 NTU, and at the same time, total suspended solids got the maximum value 356 mg l⁻¹. Conductivity was downward trend from 1990 to 1995 and upward trend from 2005 to 2010 obviously except C24, but was downward trend at each station recently.
TN showed an overall downward trend in this basin. All the four stations reached the lowest value in 1990. With time passed, the applications of fertilizers decreased and the TN concentrations in water body decreased slowly. Over the past ten years, TN remained decreased in C44 and C24, stable in C25, and increased in C23. There are only two seasons in this area, dry season and wet season. The seasonal change was apparent. The TN concentration value was higher in wet season, and lower in dry season, which was expected due to increased transport of non-point source pollution by hydrologic pathways (i.e., precipitation and leaching) into receiving waters [22].

TP exhibited an overall upward trend in this basin except C24, and had an upward trend after 2010. In C24, although the overall trend was downward, TP increased in the 1980s and reached the maximum value in about 1985. After 1985, TP was downward trend and reached the minimum value in about 2010. In recently years, TP was upward trend in C24. In C23, TP increased until 2003 and then decreased. A similar pattern was observed in C44 though the overall magnitude was smaller than for C23. The same with TN, the TP concentration value was higher in wet season, and lower in dry season. In the last several years when water levels were high, Lake Okeechobee water, which has an elevated nutrient concentration was pumped from these canals and ultimately into the St. Lucie and Indian River estuaries.

The rainfall extremes with very wet years of 1983 (a strong El Nino), 1994-1995 (tropical storms), 2004-2005 (multiple hurricanes made land fall in the study area), and droughts of 1988-1989, 2000-2011, and 2006-2007 were consistent with the peaks and valleys shown in the STL long-term trend [23-25]. So both TN and TP concentration were higher in 2005 or 2006. Between 2007 and 2011, the study area experienced a 5-year prolonged drought. Both TN and TP concentration decreased from 2007 to 2010. Almost all the water quality parameters reached the trough in about 1990, which maybe because the stormwater management best management practices (BMP) was operated in the late of 20th century and the effect was very good in initial stage.

3.3. Trend prediction of TN, TP and turbidity

The minimum stationary $R^2$ was 0.547, which showed the predicted value was satisfied (table 2). The trend predicted value of TN and TP concentration and turbidity monthly from 2016 to 2020 are illustrated in figure 5 (confidence interval=95%). The change trend of predicted value is in accord with
Figure 5. Trend prediction value of TN and TP concentrations and turbidity monthly from 2016–2020.
Table 2. The results of SPSS predictive parameters.

|       | TN stationary | TP stationary | Turbidity stationary |
|-------|---------------|---------------|----------------------|
|       | R²            | BIC           | R²                   | BIC               |
| C25   | 0.642         | 0.001         | 0.633                | 0.000             |
|       | -2.097        | -4.626        | 0.739                | 0.008             |
| C24   | 0.664         | 0.000         | 0.623                | 0.000             |
|       | -1.542        | -4.035        | 0.612                | 0.074             |
| C23   | 0.617         | 0.000         | 0.547                | 0.000             |
|       | -1.688        | -4.422        | 0.672                | 0.230             |
| C44   | 0.759         | 0.013         | 0.636                | 0.000             |
|       | -1.531        | -5.520        | 0.663                | 0.200             |
|       |               |               |                      |                   |

observed value. In the following years, TN is downward trend and TP has a slow upward trend. Turbidity is also upward trend very slowly except C44, and C44 is some significantly as compared. The TN and TP concentration values in C23 are higher than other three canals every month. The average TN concentration predicted values in C23 is 1.35 -1.32 mg l⁻¹, and 1.10-1.08 mg l⁻¹ in C25, 1.17-1.10 mg l⁻¹ in C24, 1.25-1.16 mg l⁻¹ in C44 from 2016 to 2020. The average TP concentration predicted values in C23 is 0.277 -0.295 mg l⁻¹, and 0.156-0.164 mg l⁻¹ in C25, 0.209-0.204 mg l⁻¹ in C24, 0.162-0.168 mg l⁻¹ in C44 from 2016 to 2020. Turbidity of C44 is higher on the whole, but lower in wet season than dry season. In particular, the C44 canal connects Lake Okeechobee, the second largest lake in the contiguous United States, with the SLE, allowing large volumes of lake water released for flood control into the estuary [18]. Maybe flood in Okeechobee lake can decease the turbidity of C44. From the predicted value, the seasonal change of TN and TP concentration is obvious and the values are higher from July to November in every station, so non-point pollution source is still key point for protect water quality. According to the USEPA ambient water quality criteria recommendation for the study area, the TN and TP concentration values are satisfied with the criteria in dry season, but not in wet season. N and P best management practices should become more stringent on farm lands draining into the C23 canal.

4. Conclusions

Time series trend analysis (from 1979 to 2014) and prediction (from 2016 to 2020) of water quality in the Everglades canal system (canals of C25, C24, C23 and C44) were studied. We concluded that (i) The mean of DO at each station was in the range of 5.0~6.3 mg l⁻¹, and there was no evident relationship between temperature and DO. The value was low in about 1995 and high in about 2010. (ii) Turbidity was in the range of 0.4~90.3NTU, and was influenced by rainstorm and non-point source pollution. (iii) Color and conductivity were significant negative correlation. Conductivity was high in the end of 1980 and about 2010, and low in about 1995. (iv) TN concentration was decreasing trend and TP concentration was increasing trend. The mean value of TN at each station was C24>C44 >C23>C25, and TP was C24>C23>C44>C25. (v) According to prediction result of TN, TP and turbidity value, the TN and TP concentration values in C23 are higher every month and turbidity of C44 is higher than other three canals on the whole. (vi) TN and TP concentration values are satisfied with the criteria in dry season, but not in wet season on the whole, so farm managers should institute more nutrient and runoff control on land that have high TN and TP concentrations in runoff and leachate flowing into the C23 canal., especially in the wet season.

References
[1] Kauffman G J and Belden A C 2010 Water Quality Trends (1970 to 2005) Along Delaware Streams in the Delaware and Chesapeake Bay Watersheds, USA Water Air Soil Pollut. 208 345-75
[2] Williams M R, Filoso S, Longstaff B J and Dennison W C 2010 Long-term trends of water
quality and biotic metrics in Chesapeake Bay: 1986 to 2008 *Estuar. Coasts* **33** 1279-99

[3] Chiueh P T, Wu H T and Lo S L 2011 Seasonal water quality trends in the Feitsui Reservoir Watershed, Taiwan *World Acad. Sci. Eng. Technol.* **5** 414-7

[4] Elci S and Selcuk P 2013 Effects of basin activities and land use on water quality trends in Tahtali Basin *Turkey Environ. Earth Sci.* **68** 1591-8

[5] Wilcock R J, Monaghan R M and Quinn J M 2013 Trends in water quality of five dairy farming streams in response to adoption of best practice and benefits of long-term monitoring at the catchment scale *Mar. Freshwater Res.* **64** 401-12

[6] Wu Q and Xia X H 2014 Trends of water quantity and water quality of the Yellow River from 1956 to 2009: implications for the effect of climate change *Environ. Syst. Res.* **3** 1-6

[7] Qian S S, Borsuk M E and Stow C A 2000 Seasonal and long-term nutrient trend decomposition along a spatial gradient in the Neuse River watershed *Environ. Sci. Technol.* **34** 4474-82

[8] Liang Z Y, Liu Y, Sheng H, Yang Y H, Guo H C, Zhao L and He B 2014 Identification of water quality trends in lakes Dianchi using an integrated STL and RSI approach *Acta Scientiae Circumstantiae* **3** 754-62

[9] Obertegger U and Manca M 2011 Response of rotifer functional groups to changing trophic state and crustacean community *J. Limnol.* **70** 231-8

[10] Emamgholizadeh S, Kashi H, Marofoor I and Zalaghi E 2014 Prediction of water quality parameters of Karoon River (Iran) by artificial intelligence-based models *Int. J. Environ. Sci. Tech.* **11** 645-56

[11] Barzegar R, Adamowski J and Moghaddam A A 2016 Application of wavelet- artificial intelligence hybrid models for water quality prediction: A case study in Aji-Chay River, Iran *Stoch. Environ. Res. Risk Assess.* **1** 1-23

[12] Xin X K, Li K F, Finlayson B and Yin W 2015 Evaluation, prediction, and protection of water quality in Danjiangkou Reservoir, China *Water Sci. Eng.* **8** 30-9

[13] Mokarram M 2016 Modeling of multiple regression and multiple linear regressions for prediction of groundwater quality (case study: north of Shiraz) *Model. Earth Syst. Environ.* **3** 1-7

[14] Parmar K S and Bhardwaj R 2014 Water quality management using statistical analysis and time-series prediction model *Appl. Water Sci.* **4** 425-34

[15] Wang Z 2017 Analysis of variation trend of water quality based on time series in Sansha Bay *Environ. Impact Assess.* **39** 76-81,96

[16] Zhai Y, Wang J, Teng Y and Zuo R 2012 Water demand forecasting of Beijing using the time series forecasting method *J. Geogr. Sci.* **22** 919-32

[17] Wan Y, Wan L, Li Y and Doering P 2017 Decadal and seasonal trends of nutrient concentration and export from highly managed coastal catchments *Water Res.* **115** 180-94

[18] Wan Y, Qian Y, Migliaccio K W, Li Y and Conrad C 2014 Linking spatial variations in water quality with water and land management using multivariate techniques *J. Environ. Qual.* **2** 599-610

[19] Hafen R P, Anderson D E, Cleveland W S, Maciejewski R, Ebert D S, Abusalah A, Yakout M, Ouazzani M and Grannis S 2009 Syndromic surveillance: STL for modeling, visualizing, and monitoring disease counts *BMC Med. Inform. Decis. Sci.* **9** 1-11

[20] Lu W X, Zhao Y, Chu H B and Yang L L 2013 The analysis of groundwater levels influenced by dual factors in western Jilin Province by using time series analysis method *Appl. Water Sci.* doi: 10.1007/s13201-013-0111-4

[21] Zhou J, Liang Z, Liu Y, Guo H, Dan H and Zhao L 2015 Six-decade temporal change and seasonal decomposition of climate variables in Lake Dianchi watershed (China): Stable trend or abrupt shift? *Theor. Appl. Climatol.* **119** 181-91

[22] Swanson F J, Scetena F N, Dissmeyer G E, Fenn M E, Verry E S and Lynch J A 2000 Watershed processes-fluxes of water, dissolved constituents, and sediment G E Dissmeyer Ed. *Drinking Water from Forests and Grasslands: A Synthesis of the Scientific Literature* pp
[23] Mallin M A and Corbett C A 2006 How hurricane attributes determine the extent of environmental effects: Multiple hurricanes and different coastal systems *Estuar. Coasts* **6A** 1046-61

[24] Abtew W and Iricanin N 2008 Hurricane effects on south Florida water management system: A case study of hurricane Wilma of October 2005 *J. Spat. Hydrol.* **8** 1-21

[25] Rogers M W and Allen M S 2008 Hurricane impacts to lake Okeechobee: Altered hydrology creates difficult management trade offs *Fisheries* **33** 1,11-17