WATER RESEARCH 40 (2006) 3800–3810

Available at www.sciencedirect.com

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

Assessment of seasonal variations in surface water quality

Y. Ouyang[a,*], P. Nkedi-Kizza[b], Q.T. Wu[c], D. Shinde[b], C.H. Huang[d]

[a]Department of Water Resources, St. Johns River Water Management District, P.O. Box 1429, Palatka, FL 32178 1429, USA
[b]Soil and Water Science Department, University of Florida, Gainesville, FL 32611 0290, USA
[c]College of Natural Resource and Environment, South China Agricultural University, Guangzhou, China
[d]Agronomy College, South China University of Tropical Agriculture, Danzhou City, Hainan, China

ARTICLE INFO

Article history:
Received 15 February 2006
Received in revised form
1 August 2006
Accepted 26 August 2006
Available online 27 October 2006

Keywords:
Principal component analysis
Seasonal variation
Surface water quality

ABSTRACT

Assessment of seasonal changes in surface water quality is an important aspect for evaluating temporal variations of river pollution due to natural or anthropogenic inputs of point and non-point sources. In this study, surface water quality data for 16 physical and chemical parameters collected from 22 monitoring stations in a river during the years from 1998 to 2001 were analyzed. The principal component analysis technique was employed to evaluate the seasonal correlations of water quality parameters, while the principal factor analysis technique was used to extract the parameters that are most important in assessing seasonal variations of river water quality. Analysis shows that a parameter that is most important in contributing to water quality variation for one season may not be important for another season except for DOC and electrical conductance, which were always the most important parameters in contributing to water quality variations for all four seasons.

1. Introduction

Pollution of surface water with toxic chemicals and eutrophication of rivers and lakes with excess nutrients are of great environmental concern worldwide. Agricultural, industrial, and urban activities are considered as being major sources of chemicals and nutrients to aquatic ecosystems, while atmospheric deposition could be an important source to certain constituents such as mercury and nitrogen. The concentrations of toxic chemicals and biologically available nutrients in excess can lead to diverse problems such as toxic algal blooms, loss of oxygen, fish kills, loss of biodiversity, and loss of aquatic plant beds and coal reefs (Vousta et al., 2001). Nutrient enrichment seriously degrades aquatic ecosystems and impairs the use of water for drinking, industry, agriculture, and recreation and for other purposes.

Pollution of the lower St. Johns River (LSJR) with contaminants such as nutrients, hydrocarbons, pesticides, and heavy metals comes from both point and non-point sources, which are the results of storm water runoff, discharge from ditches and creeks, groundwater seepage, aquatic weed control, naturally occurring organic inputs, and atmospheric deposition. The degradation of water quality due to these contaminants has resulted in altered species composition and decreased overall health of aquatic communities within the river basin (Campbell et al., 1993; Durell et al., 2001). With increased understanding of the importance of drinking water quality to public health and raw water quality to aquatic life, numerous efforts have been devoted to restoring the health of the LSJR and preventing its further pollution during the last several decades. One of such critical efforts was the development of the surface water monitoring network.

Surface water quality monitoring within the LSJR has been conducted by various agencies, and at varying levels of intensity since 1956. The primary objectives are to identify water quality problems, describe seasonal and spatial trends

*Corresponding author. Tel.: +1 386 312 2320.
E-mail address: ouyangy@ufl.edu (Y. Ouyang).
0043-1354/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.watres.2006.08.030

for developing qualitative and quantitative models of riverine ecosystem, and determine permit compliance. Since its inception, the monitoring network has become one of the most critical efforts in the assessment of surface water pollution in the LSJR and has been a significant resource for others working to prevent pollution of the river. However, although such long-term survey and monitoring programs are very critical to a better knowledge of hydrology, geochemistry, and pollution in the LSJR, they produce large sets of data that are often difficult to interpret and are not fully explored.

The problems of data reduction and interpretation, characteristic change in water quality parameters, and indicator parameter identification can be approached through the use of the principal component analysis (PCA) and principal factor analysis (PFA). PCA and PFA are multivariate statistical techniques used to identify important components or factors that explain most of the variances of a system. They are designed to reduce the number of variables to a small number of indices (i.e., principal components or factors) while attempting to preserve the relationships present in the original data. Details for mastering the arts of PCA and PFA are published elsewhere (Davis, 1986; Manly, 1986; Wackernagel, 1995; Tabachnick and Fidell, 2001).

In recent years, the PCA and PFA techniques have been applied to a variety of environmental applications, including evaluation of ground water monitoring wells and hydrographs, examination of spatial and temporal patterns of surface water quality, identification of chemical species related to hydrological conditions, and assessment of environmental quality indicators (Shine et al., 1995; Vega et al., 1998; Yu et al., 1998; Perkins and Underwood, 2000; Tauler et al., 2000; Voutsa et al., 2001; Gangopadhyay et al., 2001; Bengraine and Marhaba, 2003; Ouyang, 2005). These studies provided invaluable insights into the applications of the PCA and PFA techniques to environmental management and protection studies.

Characterization of seasonal changes in surface water quality is an important aspect for evaluating temporal variations of river pollution due to natural or anthropogenic inputs of point and non-point sources. In addition, pollutants entering a river system normally result from many transport pathways including storm water runoff, discharge from ditches and creeks, vadose zone leaching, groundwater seepage, and atmospheric deposition. These pathways are seasonal-dependent. Therefore, seasonal changes in surface water quality must be considered when establishing the pollutant load reduction goals (PLRGs) and developing the total maximum daily loads (TMDLs). The aim of this study is to apply the PCA and PFA techniques to evaluate the seasonal correlations of water quality parameters and to extract those parameters that are most important in assessing seasonal variations of the LSJR water quality.

2. Materials and methods

2.1. Study area

The LSJR basin is located in northeast Florida, USA between 29° and 30° north and between 81.13° and 82.13° west (Fig. 1). It is an area of approximately 7200 km^2. The LSJR is a sixth order, dark-water river estuary, and exhibits characteristics associated with riverine, lacustrine, and estuarine environments. The average gradient of the river is 0.022 m/km with average tidal amplitude of 1.5 m at the ocean inlet. The land uses within the basin largely consist of residential, commercial, industrial, mining, livestock, pasture, row crops, forestry, and water. Series of water quality problems have been identified and addressed since the 1950s. These include point and non-point source pollutants such as nutrients, hydrocarbons, pesticides, and heavy metals (Campbell et al., 1993; Durell et al., 2001).

2.2. Data

In this study, twenty-two (22) monitoring stations located in the main stem of the LSJR (Fig. 1) and sixteen (16) physiochemical parameters (Table 1) obtained from each station were used for analysis. These datasets were collected during March 1998–March 2001 by staff from the SJRWMD. The datasets contain the agency monitoring stations, measured parameters and values, latitude and longitude information, and dates of data collection. The stations are located at places where conditions are most representative and homogeneous, away from transitional areas such as point source mixing zones and near-shore regions (SJRWMD, 1994). The sampling timing is planned in advance without regard to capturing temporary events.

A thorough review of the existing datasets reveals that the data were collected at different times of the day and/or different days of the month. In addition, some stations have been used to collect the physiochemical parameters for about 20 years, while others have just been used to collect the same parameters in recent years. For the purpose of this study, a 3-year time period with a seasonal mean value for each parameter was selected, considering the data availability for PCA and PFA techniques.

2.3. PCA and PFA analysis

The PCA and PFA were performed on SAS (statistical analysis system) software, version 8, developed by SAS Institute Inc. (1999), using the PRINCOMP and FACTOR modules. In mathematical terms, PCA and PFA involve the following five major steps: (1) start by coding the variables $x_1, x_2, \ldots, x_p$ to have zero means and unit variance, i.e., standardization of the measurements to ensure that they all have equal weight in the analysis; (2) calculate the covariance matrix **C**; (3) find the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_p$ and the corresponding eigenvectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_p$; (4) discard any components that only account for a small proportion of the variation in datasets; and (5) develop the factor loading matrix and perform a varimax rotation on the factor loading matrix to infer the principal parameters. In this study, only those components or factors exhibiting an eigenvalue of greater than or close equal to one were retained (Voutsa et al., 2001; Bengraine and Marhaba, 2003).

In order to distinguish the variations of each parameter for a given season, the data was divided into four distinct temporal databases. Winter corresponded from 21 December

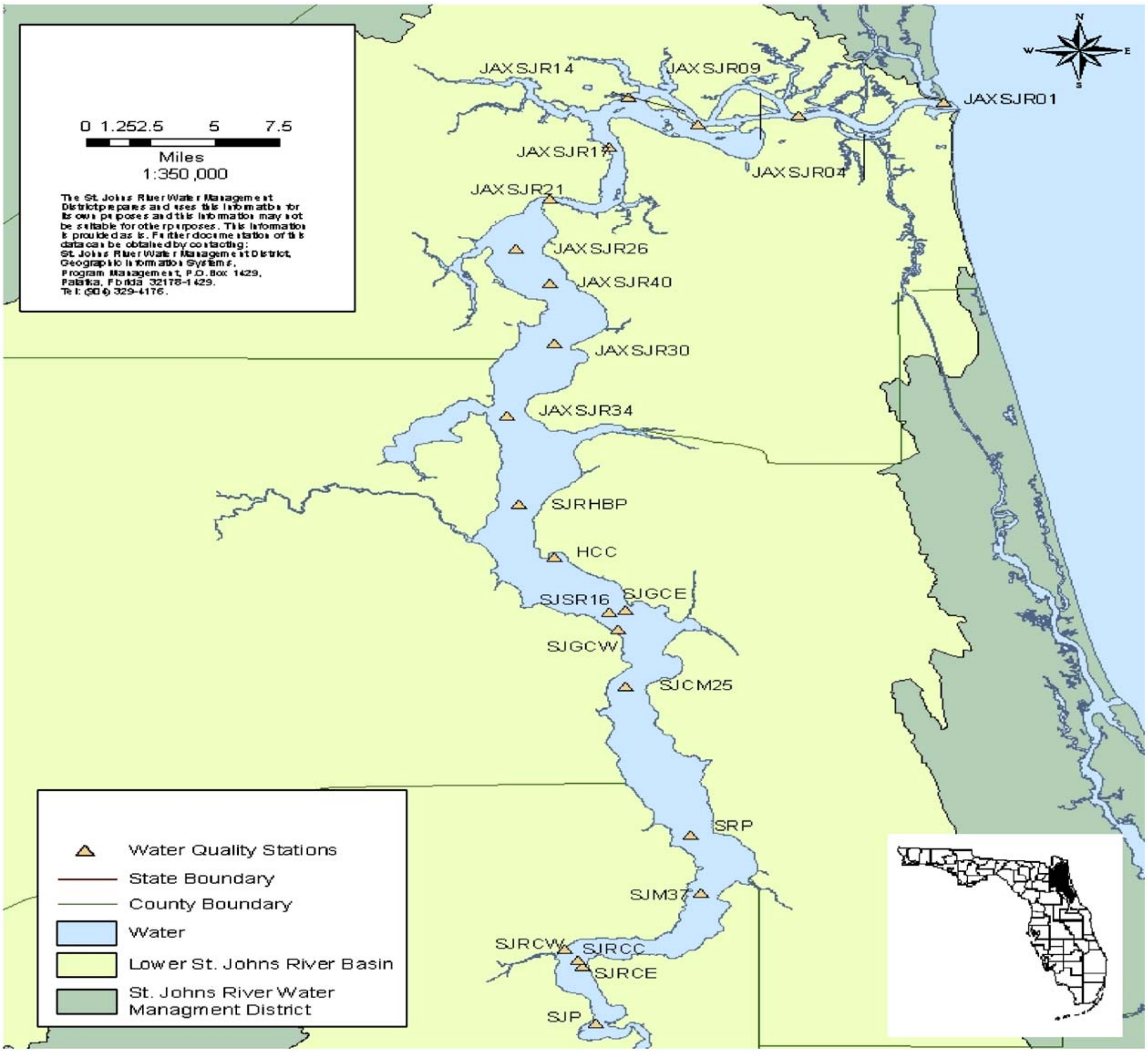

Fig. 1 – Location of the lower St. Johns River, Florida, USA The symbol "△"represents the surface water-monitoring stations.

Table 1 – The water quality parameters associated with their abbreviations and units used in this study

Parameter	Abbreviation	Unit
Water temperature	T	°C
Color	Color	Platinum-cobalt
Electrical conductance	EC	umhos/cm
Dissolved oxygen	DO	mg/l
5 day BOD	BOD	mg/l
pH	pH	pH units
Total alkalinity, (as $CaCO_3$)	Alkal	mg/l
Salinity	Sal	ppt
Total ammonia-nitrogen	TNH_3	mg/l
Total Kjeldahl-nitrogen	TKN	mg/l
Dissolved nitrite+nitrate	DNO_x	mg/l
Total phosphorus	TP	mg/l
Orthophosphate-phosphorus	PO_4	mg/l
Total organic carbon	TOC	mg/l
Dissolved organic carbon	DOC	mg/l
Turbidity	Turb	ntu

to 20 March, spring from 21 March to 20 June, summer from 21 June to 20 September, and fall from 21 September to 20 December. Therefore, four seasonal separation principal components or factors were performed.

3. Results and discussion

3.1. *Seasonal correlation of water quality parameters*

Data in Table 2 provide the seasonal correlation matrix of the water quality parameters obtained from the PCA. In general, the river water temperature had relatively weak to fair correlations, i.e., most of the correlation coefficients are less than 0.7 (absolute value) with other parameters for the entire four seasons. In spring, the correlation coefficients between temperature and other parameters were less than or equal to 0.55 except for turbidity (0.65). Such correlations had slightly changed in summer with a positive increase in correlations with TKN (0.67), TOC (0.65), and DOC (0.62) as well as a

Table 2 – Correlation matrices

	T	Color	EC	DO_2	BOD	pH	Alkal	Sal	TNH_3	TKN	DNO_x	TP	PO_4	TOC	DOC	Turb
Spring																
T	1															
Color	0.1645	1														
EC	−0.2338	−0.9486	1													
DO_2	0.0621	0.9351	−0.9103	1												
BOD	0.1232	0.5521	−0.5856	0.6482	1											
pH	0.2093	0.2296	−0.2165	0.4053	0.802	1										
Alkal	−0.2514	−0.8977	0.8647	−0.8105	−0.212	0.0388	1									
Sal	−0.3027	−0.9601	0.9688	−0.92	−0.5401	−0.2293	0.9199	1								
TNH_3	0.4722	−0.3078	0.1993	−0.5044	−0.3348	−0.3874	0.1575	0.2258	1							
TKN	0.2998	0.9507	−0.9559	0.9105	0.693	0.3691	−0.8244	−0.9576	−0.2001	1						
DNO_x	0.5504	−0.3743	0.3165	−0.5355	−0.7059	−0.5406	0.031	0.2542	0.6518	−0.3762	1					
TP	0.5086	−0.5904	0.5125	−0.7243	−0.6147	−0.4408	0.3369	0.4809	0.7918	−0.5202	0.9118	1				
PO_4	0.5535	−0.3814	0.3219	−0.5715	−0.6635	−0.5348	0.0656	0.2729	0.7937	−0.3613	0.974	0.9417	1			
TOC	0.3259	0.969	−0.9544	0.9087	0.5033	0.2166	−0.9373	−0.9928	−0.2041	0.9581	−0.2179	−0.4522	−0.2345	1		
DOC	0.2446	0.9818	−0.9615	0.9192	0.5242	0.2055	−0.928	−0.9868	−0.2172	0.9635	−0.285	−0.5049	−0.2878	0.9926	1	
Turb	0.6503	−0.2024	0.1263	−0.2464	−0.3417	−0.2424	−0.0558	0.0469	0.419	−0.0909	0.7505	0.7376	0.6923	−0.0233	−0.0995	1
Summer																
T	1															
Color	0.5122	1														
EC	−0.6519	−0.7955	1													
DO_2	0.3394	0.6547	−0.8032	1												
BOD	0.4068	0.6571	−0.829	0.8901	1											
pH	0.0851	−0.1467	−0.0919	0.4486	0.441	1										
Alkal	−0.6117	−0.6152	0.6212	−0.2536	−0.1308	0.4698	1									
Sal	−0.7089	−0.8086	0.9815	−0.7766	−0.8109	−0.055	0.6504	1								
TNH_3	−0.2188	−0.217	0.3177	−0.3869	−0.3146	−0.3155	0.1218	0.243	1							
TKN	0.6741	0.8101	−0.9508	0.7923	0.8569	0.0841	−0.5656	−0.9766	−0.2431	1						
DNO_x	−0.0065	−0.4327	0.5789	−0.8548	−0.7687	−0.4671	0.0063	0.5647	0.1044	−0.5875	1					
TP	−0.2143	−0.5016	0.7102	−0.7963	−0.8538	−0.2602	0.06	0.7239	0.0649	−0.7606	0.8382	1				
PO_4	0.2263	−0.1817	0.3327	−0.6306	−0.6883	−0.4356	−0.424	0.3026	0.02	−0.3619	0.8076	0.8388	1			
TOC	0.6565	0.8525	−0.9088	0.6665	0.6265	−0.2024	−0.8329	−0.938	−0.1864	0.9108	−0.4347	−0.5538	−0.0889	1		
DOC	0.6212	0.8379	−0.9066	0.6819	0.6434	−0.2062	−0.8078	−0.933	−0.1898	0.9154	−0.4434	−0.5767	−0.1225	0.9925	1	
Turb	0.1068	0.2279	−0.1771	0.1817	0.3157	−0.0679	0.1139	−0.2298	0.3269	0.349	−0.2473	−0.4656	−0.4185	0.1843	0.239	1
Fall																
T	1															
Color	0.4796	1														
EC	−0.4114	−0.9664	1													

Table 2 (*continued*)

	T	Color	EC	DO_2	BOD	pH	Alkal	Sal	TNH_3	TKN	DNO_x	TP	PO_4	TOC	DOC	Turb
DO_2	−0.3476	0.3734	−0.3984	1												
BOD	0.6165	0.5574	−0.4511	0.3422	1											
pH	−0.3555	−0.7822	0.8253	0.0135	−0.066	1										
Alkal	−0.0733	−0.6934	0.8258	−0.2435	0.0832	0.8394	1									
Sal	−0.3911	−0.9592	0.9989	−0.408	−0.4433	0.8247	0.8353	1								
TNH_3	−0.2303	−0.5331	0.6	−0.3521	−0.4822	0.3835	0.4914	0.6184	1							
TKN	0.5375	0.9613	−0.9828	0.3409	0.556	−0.7794	−0.7553	−0.9801	−0.6196	1						
DNO_x	−0.4136	−0.3305	0.1343	−0.2663	−0.7432	−0.1555	−0.408	0.1088	−0.1058	−0.1997	1					
TP	−0.6499	−0.4982	0.316	−0.0563	−0.7253	0.1121	−0.2067	0.2855	−0.0432	−0.3842	0.9063	1				
PO_4	−0.5589	−0.3458	0.1535	−0.1135	−0.7443	−0.0918	−0.3942	0.1238	−0.0929	−0.2345	0.9686	0.9623	1			
TOC	0.4014	0.9717	−0.9963	0.3953	0.4599	−0.8207	−0.8209	−0.9957	−0.6094	0.9776	−0.1405	−0.3162	−0.1507	1		
DOC	0.3349	0.9549	−0.9925	0.438	0.3872	−0.8241	−0.857	−0.9924	−0.5947	0.9625	−0.0883	−0.2505	−0.0901	0.9917	1	
Turb	−0.6732	−0.5469	0.4062	0.0873	−0.493	0.2873	0.0587	0.3742	−0.1187	−0.463	0.6413	0.8338	0.7212	−0.413	−0.3627	1
Winter																
T	1															
Color	0.6964	1														
EC	−0.6873	−0.9263	1													
DO_2	0.6388	0.8063	−0.9042	1												
BOD	0.8115	0.5279	−0.4962	0.6125	1											
pH	0.4424	0.0832	−0.1116	0.4157	0.6399	1										
Alkal	−0.2419	−0.7081	0.8153	−0.6639	0.0666	0.2586	1									
Sal	−0.6882	−0.9242	0.9993	−0.895	−0.4851	−0.0975	0.8234	1								
TNH_3	−0.6163	−0.7291	0.8266	−0.7231	−0.3863	−0.0541	0.7398	0.8386	1							
TKN	0.588	0.89	−0.9311	0.8092	0.3502	−0.1228	−0.8298	−0.9321	−0.7716	1						
DNO_x	0.0444	0.1268	−0.1641	0.2072	0.1103	0.3215	−0.1995	−0.172	−0.2591	−0.0818	1					
TP	−0.4233	−0.1573	0.2029	−0.3989	−0.711	−0.5656	−0.285	0.1771	−0.003	−0.0896	0.2124	1				
PO_4	−0.5058	−0.2367	0.2193	−0.3929	−0.7277	−0.5452	−0.2768	0.1927	0.0301	−0.1436	0.2926	0.9481	1			
TOC	0.3999	0.7057	−0.7011	0.5078	0.0758	−0.3572	−0.7101	−0.7046	−0.5439	0.8627	−0.4687	0.0683	−0.0183	1		
DOC	0.6716	0.9416	−0.9731	0.9062	0.5573	0.1092	−0.7578	−0.9703	−0.8068	0.9287	0.1756	−0.2407	−0.2651	0.6589	1	
Turb	−0.5155	−0.0469	0.0412	−0.1294	−0.7217	−0.5848	−0.4589	0.0307	−0.0155	0.1523	−0.0852	0.7469	0.667	0.3247	−0.0579	1

negative increase in correlations with EC (−0.65), alkalinity (−0.62), and salinity (−0.71). The positive increase in correlations with the organic-related parameters such as TKN, TOC, and DOC was a result of a faster decomposition of organic matter with higher temperature in summer. However, a negatively increase in correlations with the mineral-related parameters such as alkalinity and salinity could be attributed to the dilution effect during the wet summer season. The correlation coefficients between temperature and other parameters fell below the absolute value of 0.67 in fall, indicating relatively fair correlations. It is interesting to observe that a good correlation (0.81) between temperature and biochemical oxygen demand (BOD) was found in winter. This indicates that the BOD in the river was temperature-dependent during the winter months. It should be noted that few efforts have been devoted to investigating the correlations among the variables used in this study in an estuarine environment like the LSJR. Bengraine and Marhaba (2003) investigated the annual physical and chemical characteristic changes of the Passaic River, New Jersey, USA using data collected in 1998. These authors found that the water temperature had very weak correlations with other water quality parameters such as pH, color, DO, TOC, and BOD, which were, in general, similar to our findings although no seasonal analysis was performed by these authors.

Water color is an important index of river TOC and DOC contents. Data in Table 2 reveal that water color had very strong correlations (0.83–0.98) with DOC for all of the four seasons. Similar correlations (with a slightly decrease) were found between water color and TOC (0.71) in winter. These findings were confirmed by our previous study (Ouyang, 2005). This occurred because the rate of TOC load from watersheds into the LSJR was low during the dry winter season. Strong correlations between water color and DO were found in spring (0.94) and winter (0.81), whereas poor correlations were observed in summer (0.33) and fall (0.37). Bengraine and Marhaba (2003) reported that very poor correlations was found between water color and DO (−0.06) as well as water color and TOC (0.16) in their annual analysis, these correlations were much poorer than those obtained in our study. The discrepancies could occur partially because of different river environments and partially because of no seasonal analysis by Bengraine and Marhaba (2003).

Very strong correlations between DO and organic-related parameters (i.e., TKN, TOC and DOC) were also found in spring, but the correlations were moderately reduced in summer and profoundly in fall, and finally recovered in winter (Table 2). That is, the correlation coefficients between DO and TKN, DO and TOC, and DO and DOC were, respectively, 0.91, 0.91, and 0.92 in spring, 0.79, 0.67, and 0.68 in summer, 0.34, 0.40, and 0.44 in fall, and 0.81, 0.51, and 0.91 in winter. These data imply that DO was not always highly correlated with TKN, TOC, and DOC in the LSJR. Therefore, seasonal variations should be considered when using DO as an indicator to evaluate surface water quality. It should be noted that several studies have been devoted to investigating the river water quality (Canfield et al., 1984; Vega et al., 1998; Bengraine and Marhaba, 2003; Muslim and Jones, 2003). However, theses studies had either no seasonal analysis (e.g., Canfield et al., 1984; Bengraine and Marhaba, 2003) or not included the correlations between DO and organic-related parameters (i.e., TKN, TOC, and DOC), which made it difficult for comparisons to our findings.

Little research has been devoted to investigating the relationship between BOD concentrations and watercolor changes. Results from the PCA showed that only small variations in correlation coefficients (ranging from 0.52 to 0.65) were observed between BOD and water color for the entire four seasons. In contrast, large seasonal variations in correlation coefficients between BOD and organic-related parameters (i.e., TKN, TOC, and DOC) were found in the LSJR. The correlation coefficients between BOD and TKN, BOD and TOC, and BOD and DOC were, respectively, 0.69, 0.50, and 0.52 in spring, 0.86, 0.63, and 0.64 in summer, 0.56, 0.46, and 0.39 in fall, and 0.35, 0.08, and 0.56 in winter. Similar correlation patterns among BOD, TOC, and DOC were also observed in fresh water lakes in Florida (Canfield et al., 1984; Haven, 2003). It is interesting to note that a good correlation (0.80) occurred between BOD and pH in spring.

Although the water pH had good correlations with EC (0.83), alkalinity (0.84), and salinity (0.83) in fall, such correlations became very poor in spring, summer, and winter. In general, alkalinity always had good correlations with EC and salinity for the entire four seasons. Specifically, such correlations were weaker in summer than in the rest of the three seasons. This could have occurred due to the dilution effects on alkalinity during the wet summer season. Data in Table 2 also shows that DNO_x had very good correlations with TP and PO_4 in spring, summer, and fall, but the correlations became weak in winter. This could be a result of the low nitrification rates during the low temperature in winter.

3.2. *Temporal variations of water quality parameters*

In PCA, eigenvalues are normally used to determine the number of principal components (PCs) that can be retained for further study. A scree plot for the eigenvalues obtained in this study shows a pronounced change of slope after the third eigenvalue (Fig. 2). Cattell and Jaspers (1976) and Vega et al. (1998) suggested using all of the PCs up to and including the first one after the brake. Therefore, the first four PCs will be used for further analysis. These four PCs have eigenvalues greater than or close to unity and explain 96.7%, 91.8%, 96.7%, and 93.2% of the total variances of information contained in the original dataset, for spring, summer, fall, and winter, respectively.

Projections of the original variables on the subspace of the PCs are called component loadings and coincided with the correlation coefficients between PCs and variables. In other words, the component loadings are the linear combinations for each principal component, and express the correlation between the original variables and the newly formed components. The component loadings can be used to determine the relative importance of a variable (or parameter in this study) as compared to other variables in a PC and do not reflect the importance of the component itself.

Component loadings of the first two retained PCs for each season are presented in Figs. 3 and 4. In spring, the principal component 1 (PC1) explained 56.8% of the total variance and was positively and largely contributed by organic-related

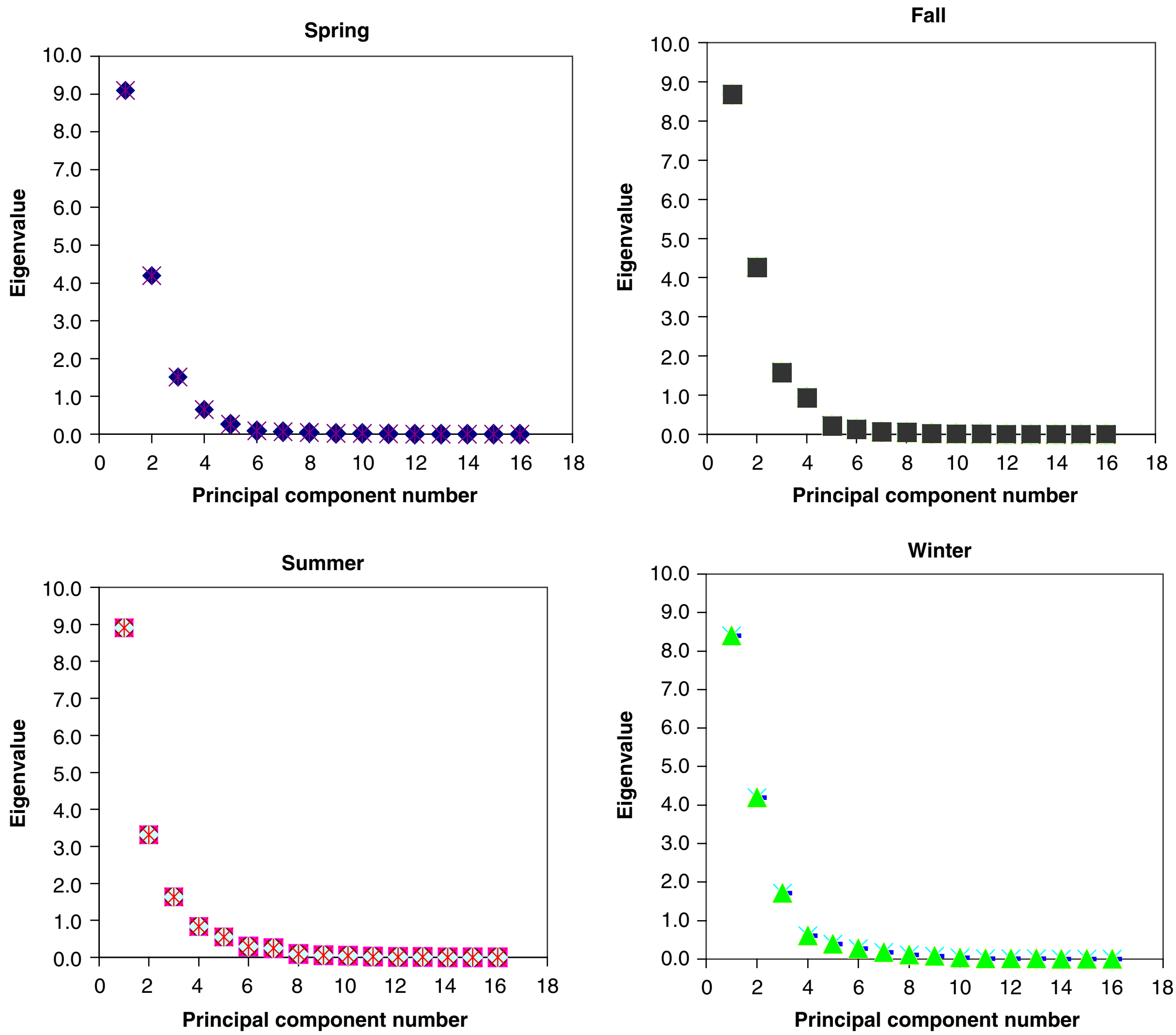

Fig. 2 – Scree plot of the eigenvalues of principal components in spring, summer, fall, and winter.

parameters (i.e., TKN, TOC, and DOC) and physical parameters (i.e., color, DO, and BOD) and was negatively affected by mineral-related parameters (i.e., alkalinity, salinity, and EC) and inorganic nutrients (i.e., TNH_3, DNO_x, TP, and PO_4^{-3}). Therefore, this component seems to measure the preponderance of physical and organic-related water quality parameters over the mineral and inorganic nutrient-related water quality parameters. This component also reveals that the water temperature and turbidity were less important in accounting for river water quality variations in spring since the loading (eigenvector) coefficients were low for these two parameters.

PC2 explained 26.8% of the total variance and was positively and largely contributed to by TNH_3, DNO_x, TP, PO_4^{-3}, temperature, and turbidity and was negatively and largely due to pH, alkalinity, and salinity (Fig. 3b). This component distinguishes the importance of anthropogenic inputs (e.g., TNH_3, DNO_x, TP, and PO_4^{-3}) and physical parameters (e.g., temperature and turbidity) over the natural inputs (e.g., pH, alkalinity, and salinity).

Similar component loading patterns were obtained for PC1 and PC2 in summer (Fig. 3c and d) except for turbidity. That is, PC1 (which explained 55.6% of the total variance) was positively contributed by the physical and organic-related parameters (i.e., color, DO, BOD, TKN, TOC, and DOC) and was negatively affected by the mineral and inorganic nutrient-related parameters (i.e., alkalinity, salinity, EC, TNH_3, DNO_x, TP, and PO_4^{-3}), whereas the PC2 (which explained 20.7% of the total variance) was positively contributed by anthropogenic inputs (i.e., DNO_x, TP, and PO_4^{-3}) and temperature and was negatively impacted by natural inputs (i.e., pH and alkalinity).

Unlike the cases for PC1s in spring and summer, the PC1 in fall, which explained 54.2% of the total variance, was positively contributed by mineral and inorganic nutrient-related parameters (i.e., alkalinity, salinity, TNH_3, DNO_x, TP, PO_4^{-3}, and pH) and was negatively participated by the physical and organic-related parameters (i.e., color, DO, TKN, TOC, and DOC) (Fig. 4a). Such opposite results (as compared to those in spring and summer) further reveal a highly seasonal variation of water quality parameters in this dynamic river system.

In winter, the PC1, which accounted for 52.5% of the total variance, was positively influenced by organic-related parameters (i.e., TKN, TOC, and DOC) and physical parameters

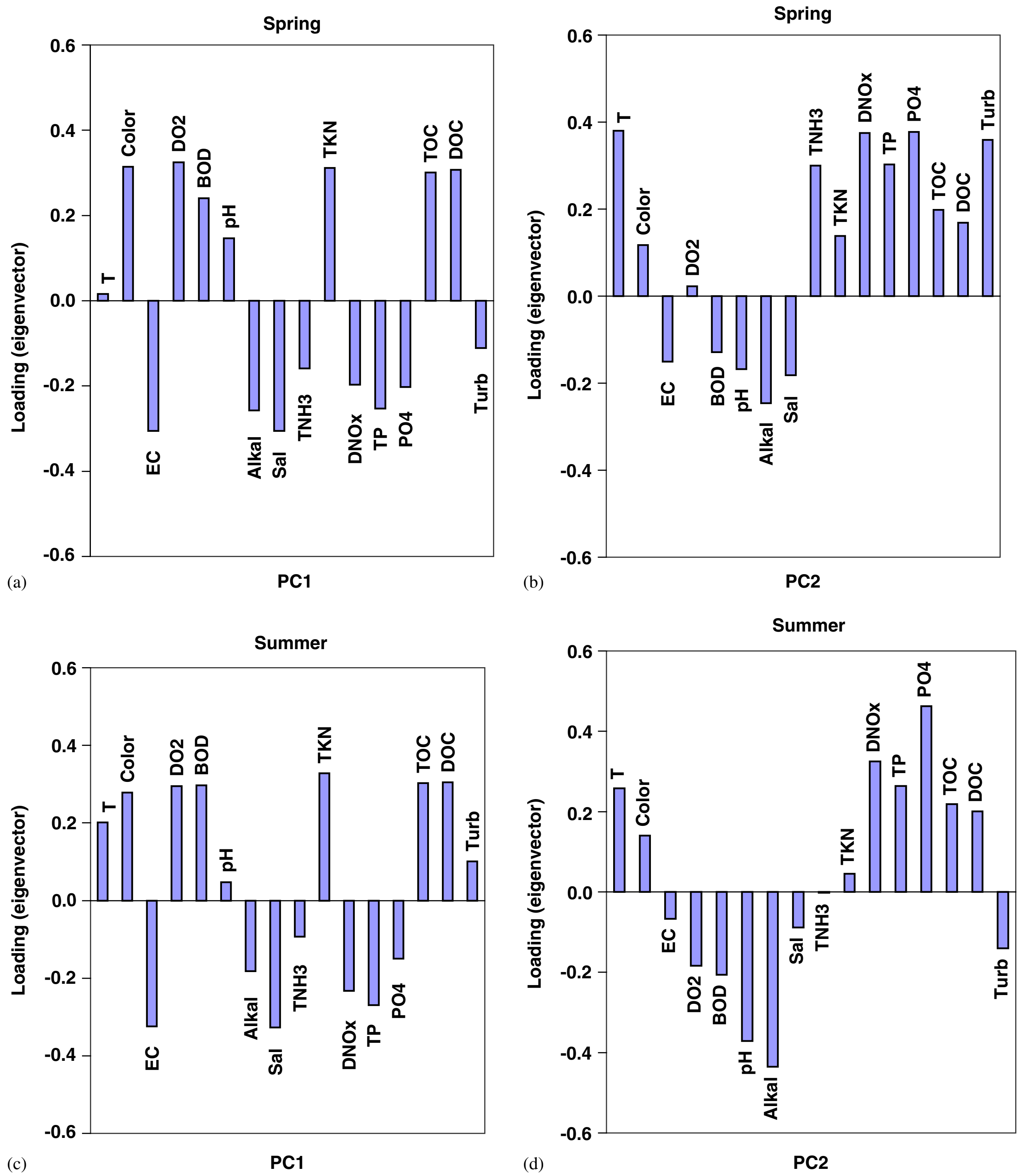

Fig. 3 – Component loadings for the first component (PC1) and the second component (PC2) in spring and summer.

(i.e., temperature, DO, BOD, and color) and was negatively affected by the mineral and nutrient related parameters (i.e., alkalinity, salinity, EC, TNH_3, TP, and PO_4^{-3}) as indicated in Fig. 4c. This component also demonstrates that pH, DNO_x, and turbidity were less important in accounting for river water quality variations in winter since the loading (eigenvector) coefficients were low for these three parameters.

Vega et al. (1998) investigated the seasonal and polluting effects on water quality of the Pisuerga River (Duero basin, Spain) using exploratory data analysis. These authors reported that the overall component loadings (i.e., no seasonal loading provided) for 22 experimental variables used in their study were 46.1% and 19.0%, respectively, for PC1 and PC2. These values were lower than those from our study. In addition, the PC1 in their study was mostly contributed by chloride, bicarbonate, sulfate, conductivity, dissolved solids, hardness, calcium, potassium, magnesium, and sodium, whereas the PC1 in our study was largely contributed by organic-related parameters (i.e., TKN, TOC, and DOC) and physical parameters (i.e., color, DO, and BOD). We attributed

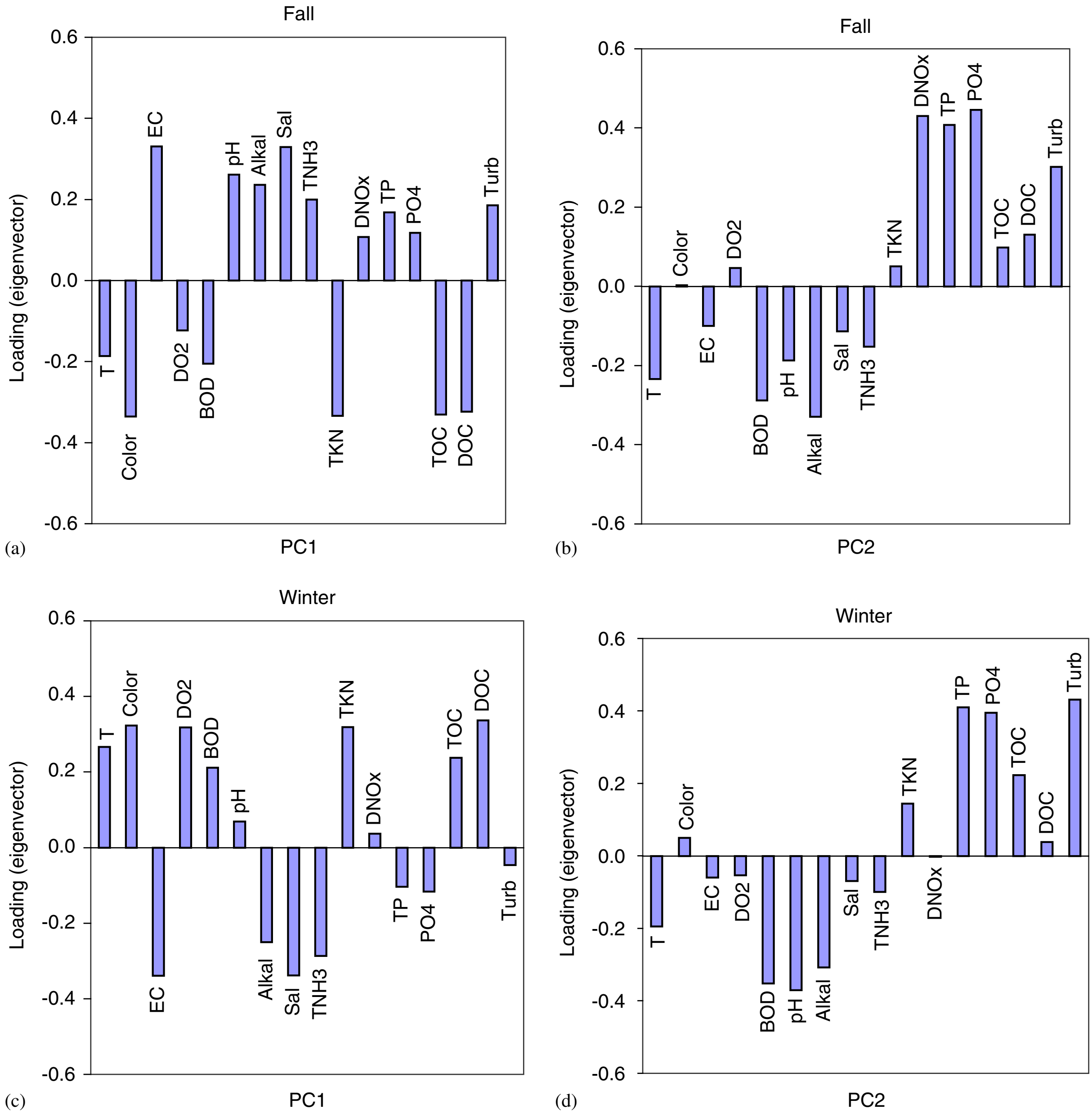

Fig. 4 – Component loadings for the first component (PC1) and the second component (PC2) in fall and winter.

the discrepancies to the different river environments and different water quality parameters as well as to the different time periods (i.e., seasonal vs. overall) used in each study. Results suggested that water quality variables that play important roles in influencing river water quality in one environment may not be important in another environment.

3.3. *Identification of important seasonal water quality parameter*

As can be seen in Figs. 3 and 4, PC1 and PC2 for all of the seasons were highly influenced (negatively or positively) by most of the variables, thus hindering the interpretation regarding which parameters are more important than the others in influencing water quality variations within a given season. Therefore, the PFA is needed to circumvent the ambiguity in the data.

Data in Table 3 show the rotated correlation coefficients for the first four factors in each season. The reason to retain the first four factors for analysis is that these four factors account for 97.2%, 93.8%, 98.1%, and 94.7% of the total variances in spring, summer, fall, and winter, respectively. The rest of the 12 factors accounted for only small percentages of the total variances and had very low and insignificant correlation coefficients. In this study, any water quality parameter with an absolute correlation coefficient value $>95\%$ was considered to be an important parameter contributing to seasonal variations of the LSJR water quality.

The most important water quality parameters that can be used to evaluate seasonal variations of the LSJR water quality are given in Table 4. This table was compiled based on the 95%

Table 3 – Rotated factor correlation coefficients for each season

Variable	Factor1	Factor2	Factor3	Factor4
Spring				
Alkal	−0.96306	−0.07118	0.15123	0.10428
BOD	0.41243	−0.32335	0.8009	0.03543
Color	0.96165	−0.17912	0.11963	−0.07139
DNO_x	−0.16525	0.86451	−0.41863	0.12313
DO_2	0.87743	−0.27799	0.26432	−0.23151
DOC	0.98682	−0.08077	0.11502	−0.01972
EC	−0.95618	0.10905	−0.14923	−0.02065
pH	0.08256	−0.14349	0.94846	−0.17078
PO_4	−0.17395	0.82221	−0.38214	0.3313
Sal	−0.98002	0.02336	−0.14753	0.06255
T	0.27897	0.82888	0.34301	0.15785
TKN	0.93102	−0.0856	0.31629	0.0367
TNH_3	−0.14069	0.52802	−0.16954	0.81372
TOC	0.98727	0.00807	0.13081	−0.05834
TP	−0.41088	0.80962	−0.24623	0.32554
Turb	−0.01785	0.9017	−0.10379	−0.08898
Summer				
Alkal	−0.90236	−0.26929	0.2066	0.16047
BOD	0.5075	−0.69468	0.26554	0.15174
Color	0.76803	−0.23826	−0.10138	0.10582
DNO_x	−0.27951	0.87099	−0.25892	0.00661
DO_2	0.53335	−0.69465	0.26727	0.02897
DOC	0.95135	−0.22508	−0.1197	0.10237
EC	−0.85982	0.41332	−0.07039	−0.0253
pH	−0.21044	−0.37922	0.85296	−0.05751
PO_4	0.11383	0.96174	−0.06013	−0.20117
Sal	−0.89449	0.38968	−0.0757	−0.057
T	0.75278	0.19399	0.31131	0.10719
TKN	0.84453	−0.42028	0.08303	0.19544
TNH_3	−0.16696	0.04926	−0.15841	0.1988
TOC	0.96518	−0.20495	−0.09959	0.03714
TP	−0.38959	0.87563	−0.01754	−0.20691
Turb	0.10711	−0.26404	−0.03787	0.91295
Autumn				
Alkal	0.91455	−0.37382	−0.03261	0.0069
BOD	−0.21279	−0.74356	0.10827	−0.56318
Color	−0.90868	−0.36699	0.06759	−0.15543
DNO_x	−0.04526	0.96876	−0.19462	−0.00025
DO_2	−0.27429	−0.10441	0.91644	−0.17005
DOC	−0.96898	−0.10857	0.14354	−0.15155
EC	0.96213	0.17009	−0.09062	0.17922
pH	0.91494	−0.07054	0.26502	−0.05287
PO_4	−0.02261	0.99586	−0.01262	0.02444
Sal	0.96209	0.14041	−0.10136	0.1973
T	−0.31602	−0.56933	−0.60541	−0.255
TKN	−0.91734	−0.24902	0.01328	−0.255
TNH_3	0.48366	−0.10584	−0.10313	0.80814
TOC	−0.95796	−0.17186	0.08973	−0.19492
TP	0.16755	0.97175	0.09386	−0.01655
Turb	0.34031	0.75605	0.24763	−0.21101
Winter				
Alkal	−0.89442	−0.37145	−0.04053	0.14908
BOD	0.35378	−0.73812	0.12121	0.39811
Color	0.91404	−0.12819	0.01105	0.20053
DNO_x	0.13539	0.12964	0.97071	0.03391
DO_2	0.87323	−0.30619	0.09233	−0.06143
DOC	0.96187	−0.19198	0.07219	0.07802
EC	−0.97975	0.13986	−0.04255	−0.10808
pH	−0.02172	−0.55115	0.30645	0.14179
PO_4	−0.09743	0.93674	0.20278	−0.02366
Sal	−0.98044	0.11558	−0.04715	−0.12566
T	0.56262	−0.43562	−0.02091	0.64911
TKN	0.96054	−0.00675	−0.20951	0.07907
TNH_3	−0.81778	−0.02577	−0.13552	−0.22518
TOC	0.7607	0.17736	−0.60402	0.08261
TP	−0.07632	0.97423	0.10913	0.09298
Turb	0.13004	0.82075	−0.19258	−0.35069

Table 4 – Most important water quality parameters for each season

Season	Positively correlated parameter	Negatively correlated parameter
Spring	Color, DOC, TOC	Alkalinity, EC, salinity
Summer	DOC, PO_4^{-3}, TOC	Alkalinity
Fall	DNO_x, EC, PO_4^{-3}, TP, salinity	DOC, TOC
Winter	DOC, TKN, TP	EC, salinity

These parameters were selected with factor correlation coefficients greater than 95%.

selection criterion. In spring, the organic-related parameters (i.e., water color, DOC, and TOC) as well as the mineral-related parameters (i.e., alkalinity, EC, and salinity) are the most important parameters in contribution to water quality variations in the LSJR although the organic-related parameters were positively correlated while the mineral-related parameters were negatively correlated in water quality seasonal variation. These organic- and mineral-related parameters may be interpreted as representing influences from natural inputs.

Only three parameters (i.e., DOC, PO_4^{-3}, and TOC) were identified as the most important parameters and positively contributed to water quality variations in summer (Table 4). In fall, the inorganic nutrients (i.e., DNO_x, PO_4^{-3}, and TP), mineral-related parameters (i.e., EC and salinity), and organic-related parameters (i.e., DOC and TOC) were found to be the most important parameters for the river water quality variations. The inorganic nutrients may be interpreted as representing influences from anthropogenic inputs. It should be noted that a distinct difference was obtained between fall and spring. The organic-related parameters were negatively correlated and the mineral -related parameters were positively correlated to water quality variations in fall, whereas the opposites were true in spring. During winter, the nutrient-related parameters (i.e., PO_4^{-3}, TKN, and TP) and DOC were positively correlated and the mineral-related parameters (i.e., EC and salinity) were negatively correlated to water quality variations. Data in Table 4 further reveal that DOC and EC were always the most important variables contributing to water quality variations in the LSJR for all four seasons.

This study demonstrated that a water quality parameter that is important in contribution to water quality variation for one season may not be important for another season. Therefore, when selecting water quality parameters for the establishment of pollutant load reduction goals (PLRGs) and the development of total maximum daily loads (TMDLs), the seasonal water quality parameter variations must be considered.

4. Conclusions

1. In this study, surface water quality data for 16 physical and chemical parameters collected from 22 monitoring stations along the main stem of the lower St. Johns River (LSJR) in Florida, USA from 1998 to 2001 were analyzed, using the PCA and PFA techniques. Results from the PCA show that river water temperature had relatively fair to weak correlations with other water quality parameters for the entire four seasons except for BOD in winter, which had a correlation coefficient of 0.8. This occurred because a decrease in water temperature decreased BOD due to the low biological activities in winter.
2. Strong correlations between the DO and the organic-related parameters such as TKN, TOC and DOC were found in spring (>0.90), but the correlations were reduced moderately in summer (<0.79) and sharply in fall (<0.44), and finally recovered in winter (0.51–0.91). Strong correlations between DO and water color was also found in spring (0.94) and winter (0.81), whereas weak correlations were observed in summer (0.65) and fall (0.37). The data indicate that DO was not always highly correlated to TKN, TOC, DOC and water color. Therefore, seasonal variations should be considered when using DO as an indicator parameter to evaluate surface water quality in the LSJR.
3. Large seasonal variation in correlations between BOD and organic-related parameters were found in the LSJR. The correlation coefficients between BOD and TKN, BOD and TOC, and BOD and DOC were, 0.69, 0.50, and 0.52 in spring, 0.86, 0.63, and 0.64 in summer, 0.56, 0.46, and 0.39 in fall, and 0.35, 0.08, and 0.56 in winter, respectively.
4. In general, alkalinity had good correlations (0.62–0.92) with EC and salinity for the entire four seasons but was somewhat weaker in summer (0.62). This occurred because of the dilution effects on alkalinity due to the wet summer season.
5. Results from the PFA show that a parameter that is important in contribution to river water quality variation for one season may not be important for another season. Therefore, when selecting water quality parameters for the establishment of pollutant load reduction goals (PLRGs) and the development of total maximum daily loads (TMDLs), the seasonal variation of parameters on river water quality must be considered.

REFERENCES

Bengraine, K., Marhaba, T.F., 2003. Using principal component analysis to monitor spatial and temporal changes in water quality. J. Hazard. Mater. B 100, 179–195.

Campbell, D., Bergman, M., Brody, M.R., Keller, A., Livingston-Way, P., Morris, F., Watkins, B., 1993. SWIM Plan for the Lower St. Johns River Basin. St. River Water Management District, Palatka, FL, January.

Canfield Jr., Linda, S.B., Hodgson, L.M., 1984. Relations between color and some limnological characteristics of Florida lakes. Water Resource. Bull. 20, 323–329.

Cattell, R.B., Jaspers, J., 1976. A general plasmode (No. 30-15-5-2) for factor analytic excercises and research. Mult. Behave. Res. Monogr. 67, 1–212.

Davis, J.C., 1986. Statistical and Data Analysis in Geology, second ed. Wiley, New York.

Durell, G.S., Seavey, J.A., Higman, J., 2001. Sediment Quality in the Lower St. Johns River and Cedar-Ortega River Basin: Chemical Contaminant Characteristics. Battelle, Duxbury, MA, March.

Gangopadhyay, S., Gupta, A.D., Nachabe, M.H., 2001. Evaluation of ground water monitoring network by principal component analysis. Ground Water 39, 181–191.

Haven, K.E., 2003. Phophorus–algal bloom relationships in large lakes of south Florida: implications for establishing nutrient criteria. Lake Reservoir Manage. 9, 222–228.

Manly, B.F.J., 1986. Multivariate Statistical Methods: A Primer. Chapman & Hall, London.

Muslim, I., Jones, G., 2003. The seasonal variation of dissolved nutrients, chlorophyll *a* and suspended sediments at Nelly Bay, Magnetic Island. Estuarine Coastal Shelf Sci. 57, 445–455.

Ouyang, Y., 2005. Application of principal component and factor analysis to evaluate surface water quality monitoring network. Water Res. 39, 2621–2635.

Perkins, R.G., Underwood, G.J.C., 2000. Gradients of chlorophyll *a* and water chemistry along an eutrophic reservoir with determination of the limiting nutrient by in situ nutrient addition. Water Res. 34, 713–724.

SAS Institute Inc., 1999. SAS Proprietary Software Release 8.2. SAS Institute, Cary, NC.

Shine, J.P., Ika, R.V., Ford, T.E., 1995. Multivariate statistical examination of spatial and temporal patterns of heavy metal contamination in New Bedford Harbor marine sediments. Environ. Sci. Technol. 29, 1781–1788.

SJRWMD, 1994. The Lower St. Johns River Basin Water Quality Monitoring Program Atlas. Lower St. Johns River Basin Management Section, St. Johns River Water Management District, May.

Tabachnick, B.G., Fidell, L.S., 2001. Using Multivariate Statistics. Allyn and Bacon, Boston, London.

Tauler, R., Barcelo, D., Thurman, E.M., 2000. Multivariate correlation between concentrations of selected herbicides and derivatives in outflows from selected US Midwestern reservoirs. Environ. Sci. Technol. 34, 3307–3314.

Vega, M., Pardo, R., Barrado, E., Deban, L., 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water Res. 32, 3581–3592.

Voutsa, D., Manoli, E., Samara, C., Sofoniou, M., Stratis, I., 2001. A study of surface water quality in Macedonia, Greece: speciation of nitrogen and phosphorus. Water Air Soil Pollut. 129, 13–32.

Wackernagel, H., 1995. Multivariate Geostatistics. An Introduction With Applications. Springer, New York.

Yu, J.C., Quinn, J.T., Dufournaud, C.M., Harrington, J.J., Roger, P.P., Lohani, B.N., 1998. Effective dimensionality of environmental indicators: a principal component analysis with bootstrap confidence intervals. J. Environ. Manage. 53, 101–119.