

Water Quality Monitoring Using Remote Sensing and an Artificial Neural Network

Yirgalem Chebud · Ghinwa M. Naja ·
Rosanna G. Rivero · Assefa M. Melesse

Received: 2 February 2012 / Accepted: 12 June 2012 / Published online: 10 July 2012
© Springer Science+Business Media B.V. 2012

Abstract In remotely located watersheds or large waterbodies, monitoring water quality parameters is often not feasible because of high costs and site inaccessibility. A cost-effective remote sensing-based methodology was developed to predict water quality parameters over a large and logistically difficult area. Landsat spectral data were used as a proxy, and a neural network model was developed to quantify water quality parameters, namely chlorophyll-*a*, turbidity, and phosphorus before and after ecosystem restoration and during the wet and dry seasons. The results demonstrate that the developed neural network model provided an excellent relationship between the observed and simulated water quality parameters. These correlated for a specific region in the greater Florida Everglades at $R^2 > 0.95$ in 1998–1999 and in 2009–2010 (dry and wet seasons). Moreover, the root mean square error values for phosphorus, turbidity, and chlorophyll-*a* were below 0.03 mg L^{-1} , 0.5 NTU , and 0.17 mg m^{-3} , respectively, at the neural network training and validation phases. Using the

developed methodology, the trends for temporal and spatial dynamics of the selected water quality parameters were investigated. In addition, the amounts of phosphorus and chlorophyll-*a* stored in the water column were calculated demonstrating the usefulness of this methodology to predict water quality parameters in complex ecosystems.

Keywords Ecological restoration · Landsat TM · Neural network · Remote sensing

1 Introduction

High phosphorus and nitrogen concentrations mainly originating from fertilizer-rich agricultural runoffs and effluents from wastewater treatment plants are threatening many worldwide ecosystems (Reed-Andersen et al. 2000). Increases of the primary productivity have already been observed in several waterbodies (Reddy et al. 1996). Kissimmee River in Florida is one of these ecosystems characterized by high phosphorus and nitrogen loadings originating mainly from agricultural runoffs. The situation has been aggravated by the channelization of the river in 1960, transforming the 167-km-long meandering river into a wide canal only 90 km long (Koebel et al. 1999). Water quality parameters, generally used as indicators of ecosystem degradation or restoration, have been intensively monitored over the vast wetlands of the Kissimmee basin and the Kissimmee River to assess the river state. The South

Y. Chebud · G. M. Naja (✉) · R. G. Rivero
Science Department, Everglades Foundation,
18001 Old Cutler Road,
Miami, FL 33157, USA
e-mail: mnaja@evergladesfoundation.org

Y. Chebud · A. M. Melesse
Department of Earth and Environment,
Florida International University,
Miami, FL 33199, USA

Florida Water Management District (FWMD) database (South Florida Water Management District 2010a) shows that about 1,834 sampling stations are currently monitoring the water quality status in this basin. This monitoring program is necessary but not sufficient to capture pulses of water quality changes, since some of the sampling stations are set only for biweekly or intermittent measurements. However, frequent and extensive monitoring of waterbodies and wetland areas presents logistical difficulties and challenges in terms of cost and manageability (Dekker et al. 1996).

Remotely sensed data coupled with the application of a Geographic Information System (GIS) was proposed as a complementary technique to conduct dynamic monitoring and prediction of water quality parameters at operational levels (Glasgow et al. 2004). Several authors suggested that water quality data could be retrieved from satellite imagery upon training existing monitored data. Hu et al. (2004) applied a space-borne optical remote sensing technique on the shallow waterbodies of South Floridian estuaries. The same method was also used by Ritchie et al. (2003) to examine the Mississippi River basin. These studies have investigated water quality parameters namely chlorophyll-*a* (Chl-*a*), suspended sediments, and total phosphorus (TP) at multiple spatial and temporal scales. In these studies, Chl-*a*, turbidity, and TP served as indicators of water quality status due to the fact that these parameters can be used as a proxy for phytoplankton biomass growth rate and hence eutrophication (Wool et al. 2007).

The application of remote sensing techniques to landscape analysis and habitat monitoring has advanced in the past three decades (Schowengerdt 2007). The use of this technique in ecology and water quality monitoring is considered as providing an “out of the box” solution to resource managers (Sawaya et al. 2003). The objectives of this paper are (1) to develop a new cost-effective technique coupling remote sensing with an artificial neural network to monitor water quality parameters (Chl-*a*, turbidity, and TP) of large waterbodies and (2) to assess the effectiveness of this new methodology while examining water quality parameters in Kissimmee River (South Florida) during four selected days of the wet and dry seasons pre (1998–1999) and post (2009–2010) restoration. The purpose of this paper was not to evaluate the effect of Kissimmee River restoration on water quality but to present a cost-effective methodology to track

water quality parameters changes. This methodology would optimize the cost of monitoring while significantly reducing the number of grab samples.

2 Technical Background

Several studies reported that Chl-*a* and turbidity observations could be retrieved from combining the visible, mid-infrared, and infrared bands (Akbar et al. 2010). Because of the characteristic color of Chl-*a*, its relationship with electromagnetic radiation at 500 and 600 μm wavelengths was documented using optical sensors at laboratory scale (Ritchie et al. 2003) and at field scale (Sass and Creed 2008). Ritchie et al. (2003) reported the signatures of Chl-*a* and suspended sediments and showed their higher reflectance at the green and thermal bands. The same paper indicated that the reflectance value increased when the concentration of Chl-*a* increased. This was also true for the suspended sediments within the same band (Ritchie et al. 2003).

Different water quality parameters were retrieved using specific space-borne sensors with radiometric resolutions such as Landsat Thematic Mapper (TM) (Zhang et al. 2003), Satellite Aperture Radar (Ritchie et al. 2003; Sass and Creed 2008; Oyama et al. 2009), Moderate Resolution Imaging Spectroradiometer (MODIS) (Hu 2009), and Sea-viewing Wide Field-of-View Sensor (Soto et al. 2009). The possibility of retrieving TP water quality data was documented by Shafique et al. (2006) over the Ohio River using hyperspectral bands of air-borne optical sensors at 554, 564, 710, and 740 nm wavelengths. Han and Jordan (2005), when mapping Chl-*a* concentration in Pensacola Bay (Florida), suggested the use of MODIS and Landsat TM as preferable because they are economical and available for public use. Furthermore, MODIS has higher radiometric and temporal resolution (capturing data on a daily basis), whereas Landsat TM has an advantage in its higher spatial resolution of 30 m (to be compared to 250 m for MODIS) enabling to obtain observations for individual canals and rivers.

Landsat TM has seven bands of which three are in the visible range, three in the near and mid-infrared, and one in the thermal region. According to Akbar et al. (2010), Chl-*a* water quality parameter is potentially separable at a defined ratio using the green (0.50–0.60 μm) and red bands (0.60–0.70 μm) of Landsat

TM. The greenish characteristic of Chl-*a*-driven algal system would have higher reflectance at the green band and higher absorption at the red band. Turbidity is a measure of the light-scattering properties of water and is caused by the presence of suspended solids such as clay, silt, and finely divided organic matter (Nas et al. 2010). Significant relationships have been reported between turbidity and suspended sediments and radiance or reflectance from spectral Landsat TM bands. Using in situ studies, Ritchie et al. (2003) reported that wavelengths between 700 and 800 nm (band 4) were most useful for determining suspended sediments and turbidity in surface waters. Brezonik et al. (2005) determined that suspended sediments have the highest contrasting reflectance at the same band (band 4). Total phosphorus is a combined measure of inorganic, organic, and dissolved forms of phosphorus and can be directly related to the biomass of phytoplankton (i.e., suspended algae and cyanobacteria, typically estimated by chlorophyll-*a* concentration) and indirectly related to water turbidity (Swanson and Zurawell 2006). In Akbar et al. (2010), the TP levels were determined through a combination of the green (band 2) and red bands (band 3) of Landsat TM. In Wu et al. (2010), a combination of Landsat TM bands 1, 2, and 3 and their respective ratios was used to correlate empirically with the in situ TP measurements.

Despite knowing the water quality parameter signatures, the practical retrieval of their concentration from shallow waterbodies using optical sensors is often challenging due to the background effect from the reservoir floor, embankments, rocks, and sediment deposits. Hu (2009) and Zhang et al. (2003) indicated that water quality parameter retrieval methods were found somehow dependent on the waterbody classification (based on its designated use) and water depth requiring multiple bands with different information content. Zhang et al. (2003) indicated the need of comparing multiple approaches namely deterministic, semi-empirical, and data-driven regression analysis methods when retrieving water quality parameters. In Zhang et al. (2003), a framework for Chl-*a* data retrieval was suggested from the blue and green band combinations of the Landsat TM imagery upon calibrating empirical relationships. Ritchie et al. (2003) suggested the applicability of a nonlinear relationship model from the results obtained after examining several waterbodies in Arkansas and Mississippi, USA. Brezonik et al. (2005) used a regression analysis of

multiple band combinations and showed that no single band combination has uniqueness, implying the need to explore all possible band combinations.

Neural network approaches were used by Jensen (2005), demonstrating the superiority of this method over the traditional statistical methods. To capture both linear and nonlinear relationships, Sudheer et al. (2006) also suggested a neural network approach as a more flexible method to retrieve water quality parameters. The advantage of neural networks over nonlinear regression approaches is attributable to the fact that the latter demands a prior knowledge of the parameter relationships while the neural network does not. This is vital in view of the complexity of retrieving water quality parameters affected by atmospheric and other background factors under non-ideal contexts causing uncertainties to be accounted for.

The most widely used neural network model is the McCulloch and Pitt model (Hu and Hwang 2002) constituted of neurons accumulating a weighted sum of multiple inputs called “input function” that will later be subjected to a transfer function called “activation function” to synthesize the outputs as represented in Fig. 1. The net function of the McCulloch and Pitt’s neuron model (Hu and Hwang 2002) is represented as a weighted linear function as indicated in Eq. (1). Other higher order net functions could be used for complex models replacing the single input term by the product of two or more inputs.

$$u = \sum_{j=1}^N w_j x_j + \theta \quad (1)$$

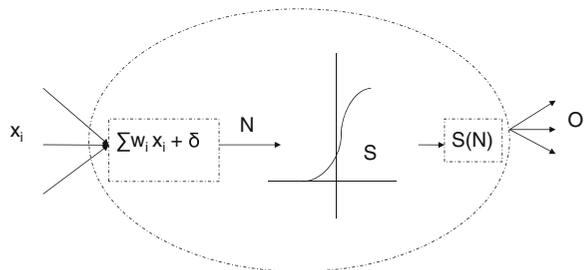


Fig. 1 Sketch of the neural network modeling approach. N is the input function at each neuron that gives the network to model the linear behavior; w the weight for each input x ; δ is the noise which controls the activation level of the sigmoid; S is a sigmoid function as described in Hu and Hwang (Environmental Protection Agency 2000) that serves to model nonlinear relationships (Environmental Protection Agency 2000); and O is the output

where w_j are weight parameters, x_j are network inputs, θ is the bias (threshold), and N is the number of inputs.

The net output of the McCulloch and Pitt's neuron model is a nonlinear transformation of the weighted linear function u (Eq. (1)) by a sigmoid distribution called activation function as shown in Eq. (2).

$$f(u) = \frac{1}{1 + e^{-\frac{u}{T}}} \quad (2)$$

where T is a target parameter used by the sigmoid distribution to fit the output distribution.

Other models that serve the same purpose of transformation could be linear, Gaussian radial basis, inverse tangent, and hyperbolic tangent. Among the different types of activation functions, Hu and Hwang (2002) rated the sigmoid type as the most commonly used.

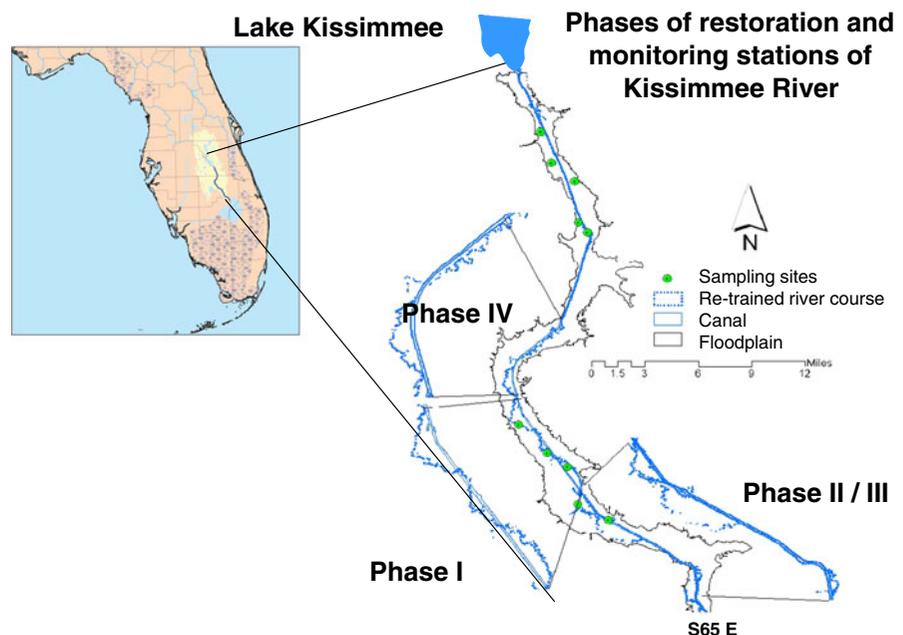
3 Methodology

3.1 Study Area

The selected area for this study was the Kissimmee River basin in South Florida (North of Lake Okeechobee and South of Lake Kissimmee), within the Northern Lake Okeechobee watershed (Fig. 2).

This study was conducted focusing on the lower Kissimmee River basin where the meandering river channel was straightened by a canal (C-38) in 1960 with a resulting flood plain of 16,200 hectares dried out. An important restoration strategy was developed in 1990, implemented by the SFWMD over multiple phases starting from 1999. Three construction phases are now complete, and continuous water flow has been reestablished to 37 km of the meandering Kissimmee River. According to the SFWMD report (South Florida Water Management District 2010b), the first phase (I) backfilled 12.9 km of the C-38 canal, while the second (II) and third (III) phases backfilled 14.5 km of the canal. The fourth (IV) phase reestablished flow to 9.7 km of reconnected river channel. So far, this resulted in 4,960 hectares of wetlands restored. In the present work, the upper, middle, and downstream sections of the river were selected to represent phases IV, I, and II/III of the restoration phases, respectively, as illustrated in Fig. 2 in order to understand the changes that occurred in the river course. Technically, the restoration involves water release in the river original meandering route, backfilling of canals, excavating the river channel, and removing some water control structures and locks. The four phased restoration cross-sections, the pre and post restoration channels, and the floodplain are indicated in Fig. 2.

Fig. 2 Maps of Florida and of the Kissimmee basin showing the different restoration phases of the Kissimmee River. The locations of the monitoring stations are also represented as well as Lake Kissimmee and S65E station



The present work has focused on four selected days during the dry (November 1 to April 30) and wet (May 1 to October 31) seasons before (dry 1998–wet 1999) and after restoration (dry 2009–wet 2010). The selected years are considered as average years as could be noticed when comparing the average temperature and rainfall. The average annual temperature and rainfall of the selected years are around 22 °C and 114 cm, respectively. The temperature and rainfall regional historical average (more than 50-year period of record) are around 22 °C and 113 cm.

3.2 Data Compilation and Processing

The stretch from Lake Kissimmee outlet down to S65E monitoring station, constituting both the old and retrained courses of the river, was delineated from the 2006 land-use GIS database from the SFWMD using ArcGIS 9.3 suite (www.esri.com) (Fig. 2). After delineating the river bank and the flooding zone, the Normalized Difference Vegetation Index (NDVI) values of the extracted image were analyzed. According to Bhagat and Sonawane (2011), waterbodies from small-sized dams to large lakes could be delineated using NDVI. Density slicing of band 5 was also effective for delineation of waterbodies on riverine floodplains (Frazier and Page 2000). Typically, NDVI values range between -1 and 1 , with values above 0 indicating the presence of vegetation and values below 0 indicating the presence of water. In this study, Kissimmee River surface areas with NDVI values below 0 were identified and depicted as waterbodies. For the flood plain, the highest NDVI negative value of -0.2 was selected as a threshold to identify flooded areas and the presence of water (Kiage and Walker 2009; Ma et al. 2007). Since the flooded zone was vegetated, the NDVI showed an excellent contrast of water versus vegetation for delineation. Moreover, at the river shoreline where a sharp waterbody boundary exists, the NDVI presented a good visual contrast very useful for delineation. Upon delineation, water quality data from the monitoring stations were downloaded from the SFWMD web site (DBHYDRO database (South Florida Water Management District 2010a)), overlaid onto the river basin and selected for further data screening and collection. Accordingly, from the 1,834 monitoring stations reported in Section 1, a total of 25 stations were identified along the Kissimmee River. From these stations, six stations have only TP

(lacking Chl-*a* and turbidity) and the other seven stations were inactive since 1997 or earlier. Therefore, only 11 stations were used to capture water quality parameters namely Chl-*a*, turbidity, and TP as indicated in Fig. 2. Water quality data were downloaded from the DBHYDRO database for the dates matching the satellite pass time (dry season of 1998–1999: December 15, 1998; wet season of 1998–1999: July 24, 1999; dry season of 2009–2010: November 27, 2009; wet season of 2009–2010: July 9, 2010) (Table 1). The water quality monitoring for Chl-*a*, turbidity, and TP in South Florida are conducted biweekly at most stations, thus meeting the temporal resolutions of the Landsat TM which passes overhead every 16 days. Most of the selected water sampling dates for the dry and wet seasons of 1998–1999 and 2009–2010 were within 96 h of the satellite pass time and hence considered acceptable (Zhang et al. 2002) (Table 1).

Landsat TM images have been widely used for environmental monitoring, resource management, and land cover/land use classification (Haack et al. 1987; Park and Stenstrom 2006). Those images provide adequate spatial resolution for regional coverage with rather negligible image distortion. The Landsat images collected from the US Geological Survey—Earth Resources Observation and Science (USGS-EROS, <http://eros.usgs.gov/>) were already preprocessed for geometrical correction and hence did not need any further action. The digital hydrometry of the Kissimmee River was compared with maps derived from the satellite imagery for validation—the results showed a precise alignment of the maps with the river hydrometry. Kissimmee River (Florida) lies in a vast wetland system where high humidity (water vapor) affects the visible spectrum. Indeed, the high humidity makes it difficult to obtain cloud-free images (Yang 2009) particularly during the wet season in Florida (Rivero et al. 2009). For water quality assessment, it is critical to use imagery without cloud cover or haze because it can affect the spectral radiometric responses and cause erroneous results (Olmanson et al. 2008). Generally, USGS-EROS preprocesses all US Landsat TM images to at least a 10 % cloud cover level. In rare cases where atmospheric and illumination effects remained on the preprocessed Landsat images, an atmospheric correction was conducted using the Atmospheric Correction for Satellite Imagery (ATCOR 2) methodology (<http://www.geosystems.de/atcor/index.html>) implemented into the ERDAS Imagine 9.3

Table 1 Water quality data obtained from the different sampling stations identified in Fig. 2

Satellite passing date	Dry season			Wet season			Dry season			Wet season		
	December 15, 1998			July 24, 1999			November 27, 2009			July 9, 2010		
Water sampling	December 17, 1998			July 22, 1999			December 1, 2009			July 8, 2010		
Station name	Chl- <i>a</i>	Turb	TP	Chl- <i>a</i>	Turb	TP	Chl- <i>a</i>	Turb	TP	Chl- <i>a</i>	Turb	TP
KREA91	2.5	1.31	0.019	30.1	2.53	0.098	1.1	32	0.048	NA	3.1	0.085
KREA92	3.5	4	0.171	14.4	2.1	0.06	0.9	35	0.034	2	2.4	0.148
KREA93	4.6	1.31	0.118	11.3	2.9	0.038	7	8.5	0.074	4	5	0.151
KREA94	2.4	0.605	0.075	72.7	3.84	0.067	8	0.073	5	5	8.1	0.175
KREA95	7.1	1.04	0.043	17.3	3.8	0.047	5	4.1	0.044	6	1.6	0.048
KREA97				3.8								
KREA98	5.4	0.714	0.034	5.7	0.761	0.025	8	10.1	0.083	4	5.6	0.138
KREA79	7.2	1.99	0.036	26.9	1.81	0.048	19	0.9	0.031	3	5.2	0.079
S65	12.16	23.7	0.151	9.6	5.25	0.096	8	1.9	0.031	14	3	0.111
S65A	NA	2.08	0.085	2.6	1.38	0.089	5	1	0.03	6	3.1	0.126
S65D	1.4	1.3	0.112	2.9	1.46	0.27	2	4	0.064	4	2.8	0.078

The sampling dates are indicated as well as the satellite passing date. Data from KREA91, KREA92, KREA94, KREA98, S65, and S65D stations were used for the training stage

Chl-a chlorophyll-*a* (milligram per cubic meter), *TP* total phosphorus (milligram per liter), *Turb turbidity* (nephelometric turbidity units), *NA* not applicable

interface (www.erdas.com). The atmospherically corrected images were then used to capture a subset of the Kissimmee River before and after restoration.

The river has a width larger than 107 m enabling thus the capture of 3 or more pixels (each with 30 m dimension) of the imagery within the river. After overlaying the water quality monitoring stations onto the reflectance images, each pixel was read from the seven Landsat TM bands. The reflectance values from the different bands were used as inputs to train the neural network while the corresponding readings of the water quality parameters were used as targets as shown schematically in Fig. 3. The training and simulation were conducted using the MATLAB software (www.mathworks.com).

The neural network technique was applied after splitting the data into 60 % for the training and 40 % for the validation purposes. The iteration was made changing the hidden number of neurons from three to five while minimizing the root mean square error (RMSE) at the training stage. The validation was tested separately setting the maximum RMSE TP criterion for less than 0.05 mg L⁻¹ corresponding to the 10 % of the maximum contaminant level (MCL) value

published by the US-Environmental Protection Agency (Environmental Protection Agency 2000). The *Chl-a* and turbidity maximum RMSE values were set at 10 % of the average observed data. Because of the lack of sampling data, testing was not performed in

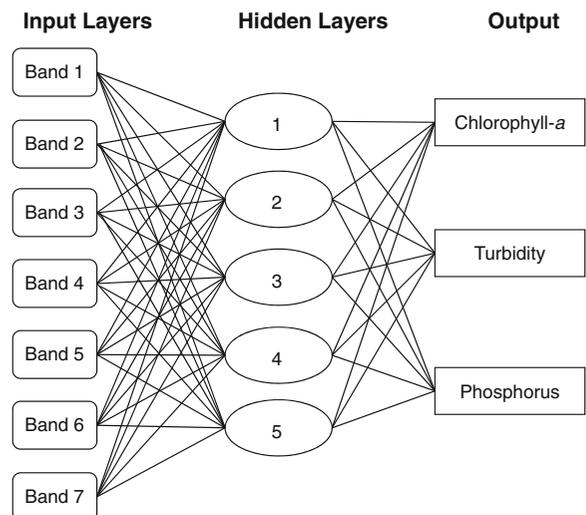


Fig. 3 Schematic diagram of the input and output layers of a neural network model

the present paper. Based on the training and validation results, the model testing with additional DBHYDRO water quality data from subsequent years and seasons would be investigated to determine the performance of the neural network model.

4 Results and Discussion

4.1 Remote Sensing Image Quality and Neural Network Results

The Landsat TM images obtained from the 2009 to 2010 wet and dry seasons showed only little residual haziness with much more improvement after haze removal and atmospheric corrections. The 1998 dry season image had cloud coverage at the northern tip portion of the river where no river restoration was conducted and that portion could thus be removed from the analysis. The image taken in the 1998 wet season contained scattered clouds (not over the river course) and the haze removal and atmospheric correction had decreased the cloud effect in this image from 10 to 1 % considered as an acceptable value (Tucker et al. 1985).

The results from the neural network technique showed that the RMSE values for TP were below 0.03 mg L^{-1} for all seasons at the neural network training and validation phases (Table 2). Similarly, the average RMSE values for Chl-*a* and turbidity were

0.17 mg m^{-3} (ranging from 0.03 to 0.54 mg m^{-3}) and 0.5 NTU (ranging from 0.03 to 1.38 NTU), respectively, at the training stage and at the validation stage (Table 2) below the 10 % level of the observed data (0.91 mg m^{-3} and 0.51 NTU, respectively). The regression coefficients (R^2) obtained after graphically fitting the observed and simulated results were above 0.95. Figure 4 presents some of these graphical fit results for the three selected water quality parameters, indicating the relatively good agreement between the monitored and the simulated values and revealing that the present methodology can be used as a suitable technique to predict water quality data in the Kissimmee River. The residual error between the relative size and the variation in the model predicted values ranged from 5 to 20 %, the highest being for TP during the 1998–1999 dry season as illustrated in Fig. 4.

Several other authors attempted to employ remote sensing image data in water quality mapping. Nas et al. (2010) used the Landsat imagery to investigate spatial water quality patterns in Lake Beysehir. Their best model led to regression coefficient values ranging between 0.61 and 0.71 when comparing the measured and estimated values of water quality parameters. Wu et al. (2010) attempted to develop empirical models to estimate TP concentration in the mainstream of the Qiantang River in China using Landsat data. The optimal regression model produced a coefficient of 0.77. Olmanson et al. (2008) conducted a 20-year

Table 2 Root mean square error (RMSE) and averaged weights in the neural network inputs

		Band 1	Band 2	Band 3	Band 4	Band 5	Band 6	Band 7	RMSE calibration	RMSE validation
1998 Dry	Chl- <i>a</i>	0.57	0.85	0.15	0.44	0.49	0.11	0.41	0.2	0.03
	Turbidity	0.55	0.98	0.48	0.73	0.49	0.48	0.063	0.44	1.38
	TP	0.43	0.88	0.68	0.79	0.88	0.56	0.42	0.03	0.03
1998 Wet	Chl- <i>a</i>	0.65	0.72	0.75	0.93	0.49	0.7	0.32	0.54	0.14
	Turbidity	0.03	0.2	0.0019	0.3	0.97	0.64	0.69	0.03	0.44
	TP	0.073	0.64	0.25	0.35	0.59	0.9	0.056	0.03	0.03
2009 Dry	Chl- <i>a</i>	0.073	0.064	0.51	0.65	0.96	0.39	0.13	0.22	0.14
	Turbidity	0.35	0.76	0.13	0.11	0.73	0.33	0.39	0.83	0.83
	TP	0.91	0.96	0.14	0.22	0.98	0.95	0.17	0.03	0.03
2010 Wet	Chl- <i>a</i>	0.99	0.5	0.36	0.4	0.68	0.23	0.4	0.03	0.04
	Turbidity	0.99	0.076	0.5	0.94	0.99	0.17	0.89	0.03	0.03
	TP	0.26	0.42	0.42	0.62	0.46	0.49	0.68	0.03	0.03

The unit of the RMSE is the unit of the quantity being estimated

Chl-*a* chlorophyll-*a* (milligram per cubic meter), TP total phosphorus (milligram per liter), Turb turbidity (nephelometric turbidity units), NA not applicable

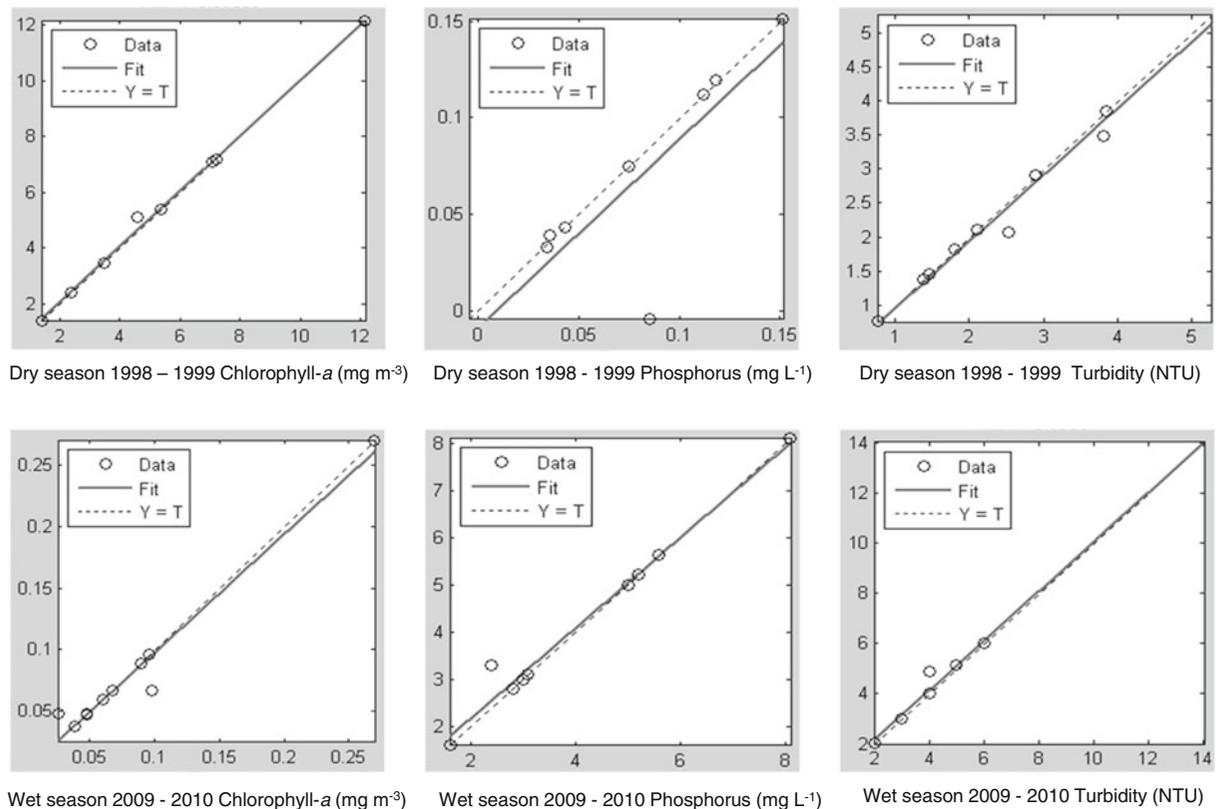


Fig. 4 Graphical fit results of the three water quality parameters (phosphorus, chlorophyll-*a*, and turbidity), 1-1 line of predicted (*Y*) and targeted (*T*) for the different seasons of 1998–1999 and 2009–2010

assessment to evaluate the accuracy of a Landsat method to obtain a comprehensive spatial and temporal coverage of key water quality characteristics that can be used to detect trends at different geographic scales. In the same study, the reliability of the data was evaluated by examining the precision of repeated measurements on individual lakes within short time periods and by comparing water clarity computed from Landsat data to field-collected Secchi depth data. The agreement between satellite data and field measurements of Secchi depth within Landsat paths was strong (average R^2 of 0.83 and range 0.71–0.96).

The nonlinear neural network model enables to produce more accurate estimations than can be obtained with the more conventional methods. Especially so when only sparse, gapped data are available for model training (Abrahart and See 2000; Li et al. 2010; Srivastava et al. 2006). The main advantage of using the neural network is that it allows considering nonlinear multiparameter relationship between reflectance from the different spectral bands and the

water quality parameters without prior knowledge of the parameter relationships (Ceyhun and Yalcin 2010). The weight of each of the bands used in the linear part of the neural model averaged over the multiple neurons in each model is presented in Table 2. Since water quality parameters are expected to change within the 10-year span examined (before and after Kissimmee River restoration), the neural network allocated different weights for the same band and season throughout the two examined seasons and years. The results indicated that band 1 weighted the highest during the wet season in 2009–2010. The results also indicated that the weight of bands 3 and 7 were mostly lower when compared to other band weight contributions and showed closer ties. Similarly, band 2 and band 5 showed dominant weights and closer ties averaging for each parameter type over the years. This could be attributable to the fact that similar information content was captured by the different bands, and unlike other multivariate methods, the neural network would not remove redundancy (Benediktsson et al. 1990). Most

principal component analysis approaches have confirmed this redundancy and concluded that the overlap of spectral characteristics within the different bands is a common phenomenon. Peculiar examples are the overlap of band 2 and band 3 and of band 5 and band 7 with band 6, the latter could be even derived from the former two (Sidjak and Wheate 1999; Zhou and Wang 2008). Indeed, although each individual spectral band reveals different characteristics of the examined parameters, the dataset as a whole contains a degree of redundancy due to the correlation between the spectral bands (Kavzoglu and Mather 2000). Taking into account this redundancy and after comparing the neural network model to a traditional regression model, Wang et al. (2008) estimated the error to below 25 % using the neural network while estimating water quality parameters such as suspended solids, dissolved oxygen, total nitrogen, TP, and Chl-*a*. Table 2 also shows that, except during the 2009 dry season, band 2 had the highest weight for Chl-*a* concentration determination. This is quite normal knowing that Chl-*a* has the greatest reflectance at green wavelengths (band 2). The thermal band (band 6) did not play a significant role in the Chl-*a* determination since this band measures the amount of infrared radiant flux emitted from surface water. Hadjimitsis et al. (2006) determined that thermal bands of Landsat TM sensor are best suited to monitor temperature variations in inland waterbodies. The examined small-sized inland waterbody did not have a significant spatial variation of water temperature to impact the weights of the thermal band.

4.2 Water Quality Parameters Monitoring

The main purpose of the Kissimmee restoration was the reestablishment of hydrologic conditions and fluctuations while restoring the Kissimmee River and its floodplain. However, the impacts of the floodplain restoration on the water quality remain uncertain (South Florida Water Management District 2010b). It is expected that the flooding of the natural wetland zones adjacent to the Kissimmee River would impact the selected water quality parameters (TP, Chl-*a*, and turbidity). Settlement of sediments, attenuation of nutrient levels, and change of vegetation pattern could be expected in the long run. Colangelo (2007) examined the response of the river metabolism to restoration of flows and reported that the post-restoration dissolved oxygen of the Kissimmee River increased from <2 to

4.70 mg L⁻¹ thus meeting the target values derived from free flowing and minimally impacted reference streams. The present work focused on Chl-*a*, TP, and turbidity as water quality parameters.

Chl-*a* is an important parameter considered as an indicator of a waterbody trophic status. Spatial patterns of Chl-*a* tend to be complex, being a result of biological as well as physical and chemical factors. Figure 5a represents the levels of Chl-*a* during the 4 days of the dry and wet seasons of 1998–1999 and 2009–2010. The modeling results showed that during 2009–2010, Chl-*a* decreased from the 5–40-mg-m⁻³ level to the 0–5-mg-m⁻³ level during the dry season and from the 15–40-mg-m⁻³ level to the 0–15-mg-m⁻³ level during the wet season (Fig. 5a). During the two selected days representative of the 1998–1999 dry and wet seasons, the Chl-*a* across the river showed higher values (25–40 mg m⁻³) at the river bank. These values decreased to 5–15 mg m⁻³ in the center of the river. This agrees with the physical process of sedimentation along the banks and algae encroachment (Sabater et al. 2008). During the two selected days representative of the 2009–2010 season, much of Chl-*a* was retained in the meandering section of the river course thus decreasing the Chl-*a* levels in the river downstream (Sabater et al. 2008). While Colangelo (2007) reported a general increase in the concentration of Chl-*a* pre- and post-restoration, the same authors did not compare the amount of Chl-*a* stored in the water column during the pre- and post-restoration periods. Indeed, the results obtained from the maps generated in the present study were further processed to obtain the amount of Chl-*a* stored in the water column as summarized in Table 3 (taking into account the monitored water depth and the pixel area 30 × 30 m). The water flow and stage values were not available for all the stations and thus not considered. When examining the upper section of the river, the amount of Chl-*a* decreased by 67.5 and 88.9 % when comparing the four selected days of the 1998–1999 and 2009–2010 dry and wet seasons, respectively. Except for the middle section of the river and when comparing the 1998–1999 to the 2009–2010 wet and dry seasons, the same decreasing tendency could be observed for the middle and lower sections of the river (Table 3). However, it is worth noticing that these results are based on comparison between 2 days of the wet and dry seasons before and after restoration. A longer time period should be investigated for better establishment of the general

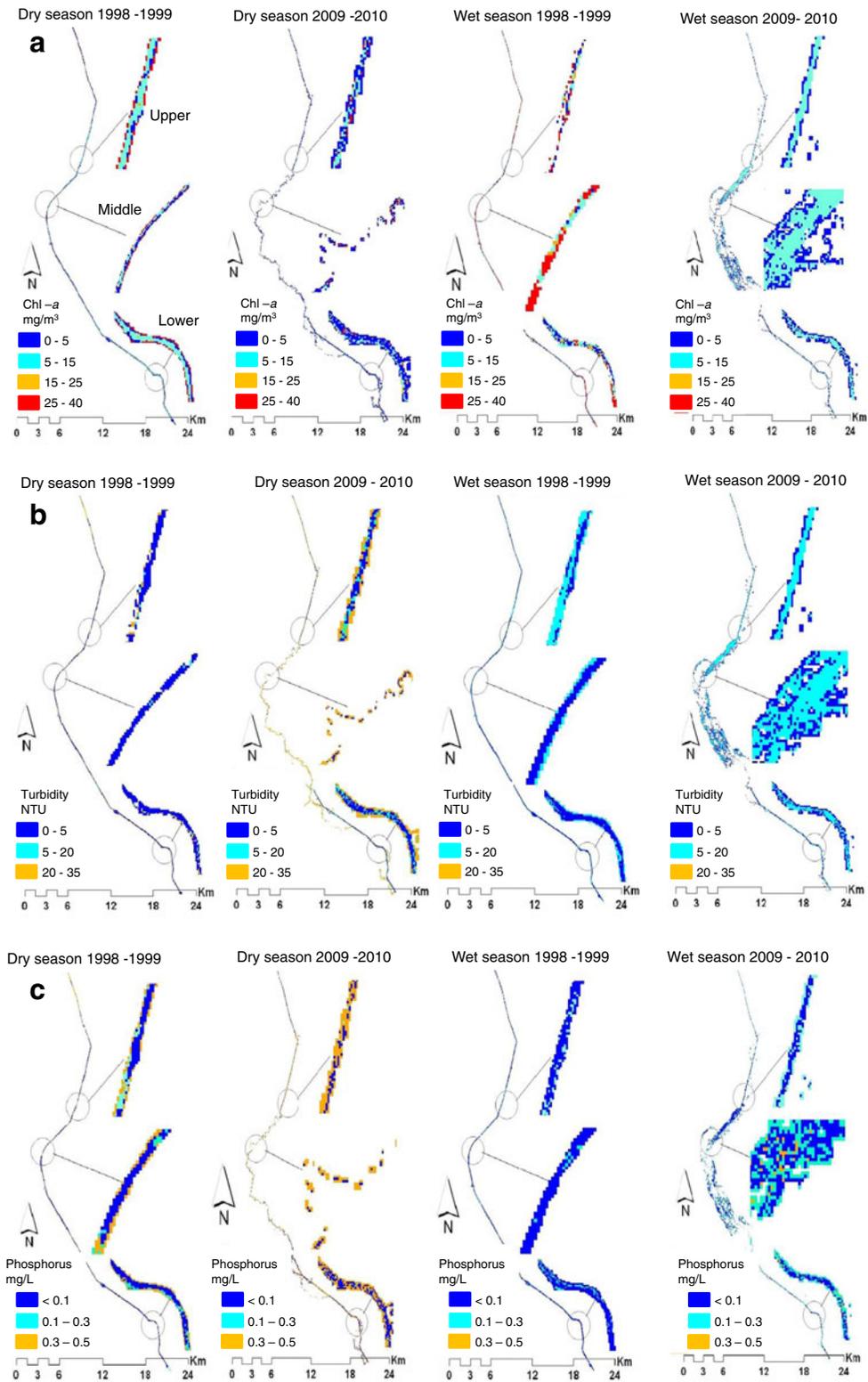


Fig. 5 Modeling results of chlorophyll-*a* (a), turbidity (b), and phosphorus (c) levels during the wet and dry seasons of 1998–1999 and 2009–2010

Table 3 Percentage of increase (+) or decrease (–) in the phosphorus and chlorophyll-*a* amounts stored in the water column in the upper, middle, and lower sections of the river

	Phosphorus (%)	Chlorophyll- <i>a</i> (%)
Upper section of the river		
1998 dry–2010 dry	–78.8	–67.5
1998 wet–2010 wet	+7.4	–88.9
Middle section of the river		
1998 dry–2010 dry	–99.8	–49.8
1998 wet–2010 wet	+2,190	+59.1
Lower section of the river		
1998 dry–2010 dry	–85.2	–76.0
1998 wet–2010 wet	–33.5	–94.8

increase or decrease tendency for Chl-*a* in the water column.

The second modeled parameter was the turbidity generally caused by the presence of suspended and dissolved matter such as clay, finely divided organic matter, plankton, other microscopic organisms, and organic acids. Unlike Chl-*a*, turbidity did not exhibit a cross-sectional concentration variation and showed a decreasing trend from upstream to downstream as observed in Fig. 5b. This could be presumably due to the fact that Kissimmee River meanders in the flat plain which gives enough time for the sediments to settle down with a resulting decrease in turbidity toward downstream locations.

Similar to Chl-*a*, the TP concentration decreased from the bank towards the center, confirming their theoretical relationship in trophic waterbodies (Smith and Shapiro 1981) (Fig. 5c). During the day selected as representative of the 2009–2010 wet season, TP relocated from the river course to the wetlands (higher concentrations in the floodplains generally considered as nutrient attenuation zones). This elevated TP levels could originate from the legacy phosphorus of the dried-up wetlands. Indeed, after transforming the meandering Kissimmee River into a canal, the primary land use of the drained floodplain was as cattle pasture and included also some sod farms (Van der Valk et al. 2009) that would certainly leave behind important legacy phosphorus deposits in the soil. The overall trend in TP stored in the water column presented in Table 3 is not conclusive since an increase and a decrease in the TP levels were observed in different sections of the river. It is worth emphasizing that in the present study, Chl-

a (a proxy for phytoplankton biomass) was used to determine TP levels while relating the measured TP to the reflectance values using the artificial neural network.

5 Conclusions

Optical sensors such as Landsat TM could be used as a cost-effective technique to monitor water quality parameters at a basin scale. The developed neural network model reduced the uncertainty from the exclusion of any of the bands and also captured both the linear and nonlinear complex relationships. The neural network model provided an excellent relationship between the observed and simulated water quality parameters correlated at $R^2 > 0.95$. The results showed that the root mean square error values for phosphorus, turbidity, and chlorophyll-*a* were below 0.03 mg L^{-1} , 0.5 NTU , and 0.17 mg m^{-3} , respectively, for all seasons at the neural network training and validation phases. The trends for changes in the selected water quality parameters were established while noticing that these results were based on comparison between 2 days during the wet and dry seasons before and after restoration. A longer time period should be investigated to test the neural network, determine its performance, and assess the impacts of Kissimmee River restoration on water quality parameters. The general increase or decrease tendency of selected water quality parameters in the water column will then be established, taking into account the difference in water depth and surface area before and after restoration.

Acknowledgments The authors thank the South Florida Water Management District (SFWMD) for the DBHYDRO data used in this study. We also thank the Darden Foundation and the Batchelor Foundation for their financial support.

References

- Abrahart, R. J., & See, L. (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological Processes*, *14*, 2157–2172.
- Akbar, T. A., Hassan, Q. A., & Achari, G. (2010). A Remote sensing based framework for predicting water quality of different source waters. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVIII—Part 1, ISPRS, Calgary: Canada.
- Benediktsson, J., Swain, P., & Esroy, O. (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *28*, 540–552.
- Bhagat, V. S., & Sonawane, K. R. (2011). Use of Landsat ETM + data for delineation of water bodies in hilly zones. *Journal of Hydroinformatics*, *13*, 661–671.
- Brezonik, P., Menken, K., & Bauer, M. (2005). Landsat based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). *Lake and Reservoir Management*, *21*, 373–382.
- Ceyhun, O., & Yalcin, A. (2010). Remote sensing of water depths in shallow waters via artificial neural networks. *Estuarine Coastal and Shelf Science*, *89*, 89–96.
- Colangelo, D. J. (2007). Response of river metabolism to restoration of flow in the Kissimmee River, Florida, USA. *Freshwater Biology*, *52*, 459–470.
- Dekker, A. G., Zamurovic-Nenad, Z., Hoogenboom, H. J., & Peters, S. W. M. (1996). Remote sensing, ecological water quality modeling and in situ measurements: a case study in shallow lakes. *Hydrological Sciences*, *41*, 531–547.
- Environmental Protection Agency, US-EPA. (2000). *Ambient water quality criteria recommendations, information supporting the development of state and tribal nutrient criteria*. Washington DC: EPA 822-B-00-014.
- Frazier, P. S., & Page, K. J. (2000). Water body detection and delineation with Landsat TM data. *Photogrammetric Engineering & Remote Sensing*, *66*, 1461–1467.
- Glasgow, H. B., Burkholder, J. M., Reed, R. E., Lewitus, A. J., & Kleinman, J. E. (2004). Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. *Journal of Experimental Marine Biology and Ecology*, *300*, 409–448.
- Haack, B., Bryant, N., Adams, S., & Yang, X. (1987). An assessment of Landsat MSS and TM data for urban and near-urban land-cover digital classification. *Remote Sensing of Environment*, *21*, 201–213.
- Hadjimitsis, D. G., Hadjimitsis, M. G., Clayton, C., & Clarke, B. A. (2006). Determination of turbidity in Kourris Dam in Cyprus utilizing Landsat TM remotely sensed data. *Water Resources Management*, *20*, 449–465.
- Han, L. H., & Jordan, K. J. (2005). Estimating and mapping chlorophyll-a concentration in Pensacola Bay, Florida using Landsat ETM plus data. *International Journal of Remote Sensing*, *26*, 5245–5254.
- Hu, C. M. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, *113*, 2118–2129.
- Hu, C. M., Chen, Z. Q., Clayton, T. D., Swarzenski, P., Brock, J. C., & Muller-Karger, F. E. (2004). Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: initial results from Tampa Bay, FL. *Remote Sensing of Environment*, *93*, 423–441.
- Hu, Y. H., & Hwang, J. N. (2002). Introduction to neural networks for signal processing. In Y. H. Hu & J. N. Hwang (Eds.), *Handbook of neural networks signal processing* (pp. 1–30). Boca Raton: CRC.
- Jensen, J. R. (2005). *Introductory digital image processing, a remote sensing perspective*. Upper Saddle River: Pearson Prentice Hall.
- Kavzoglu, T., & Mather, P. M. (2000). The use of feature selection techniques in the context of artificial neural networks. In Proceedings of the 26th Annual Conference of the Remote Sensing Society. Leicester, UK.
- Kiage, L. M., & Walker, N. D. (2009). Using NDVI from MODIS to monitor duckweed bloom in Lake Maracaibo, Venezuela. *Water Resources Management*, *23*, 1125–1135.
- Koebel, J. W., Jones, B. L., & Arrington, D. A. (1999). Restoration of the Kissimmee River, Florida: water quality impacts from canal backfilling. *Environmental Monitoring and Assessment*, *57*, 85–107.
- Li, X., Nour, M. H., Smith, D. W., & Prepas, E. E. (2010). Neural networks modeling of nitrogen export: model development and application to unmonitored boreal forest watersheds. *Environmental Technology*, *31*, 495–510.
- Ma, M., Wang, X., Veroustraete, F., & Dong, L. (2007). Change in area of Ebinur Lake during the 1998–2005 period. *International Journal of Remote Sensing*, *28*, 5523–5533.
- Nas, B., Ekercin, S., Karabork, H., Berktaç, A., & Mulla, D. J. (2010). An Application of Landsat-5TM image data for water quality mapping in Lake Beyşehir, Turkey. *Water Air and Soil Pollution*, *212*, 183–197.
- Olmanson, L. G., Bauer, M. E., & Brezonik, P. L. (2008). A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sensing of Environment*, *112*, 4086–4097.
- Oyama, Y., Matsushita, B., Fukushima, T., Matsushige, K., & Imai, A. (2009). Application of spectral decomposition algorithm for mapping water quality in a turbid lake (Lake Kasumigaura, Japan) from Landsat TM data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 73–85.
- Park, M. H., & Stenstrom, M. K. (2006). Using satellite imagery for stormwater pollution management with Bayesian networks. *Water Research*, *40*, 3429–3438.
- Reddy, K. R., Flaig, E. G., & Graetz, D. A. (1996). Phosphorus storage capacity of uplands, wetlands and streams of the Lake Okeechobee Watershed, Florida. *Agriculture, Ecosystems and Environment*, *59*, 203–216.
- Reed-Andersen, T., Carpenter, S. R., & Lathrop, R. C. (2000). Phosphorus flow in a watershed-lake ecosystem. *Ecosystems*, *3*, 561–573.

- Ritchie, J. C., Zimba, P. V., & Everitt, J. H. (2003). Remote sensing techniques to assess water quality. *Photogrammetric Engineering and Remote Sensing*, 69, 695–704.
- Rivero, R. G., Grunwald, S., Binford, M. W., & Osborne, T. Z. (2009). Integrating spectral indices into prediction models of soil phosphorus in a subtropical wetland. *Remote Sensing of Environment*, 113, 2389–2402.
- Sabater, S., Artigas, J., Duran, C., Pardos, M., Romani, A. M., Tornes, E., & Ylla, I. (2008). Longitudinal development of chlorophyll and phytoplankton assemblages in a regulated large river (the Ebro River). *Science of the Total Environment*, 404, 196–206.
- Sass, G. Z., & Creed, I. F. (2008). Characterizing hydrodynamics on boreal landscapes using archived synthetic aperture radar imagery. *Hydrological Processes*, 22, 1687–1699.
- Sawaya, K. E., Olmanson, L. G., Heinert, N. J., Brezonik, P. L., & Bauer, M. E. (2003). Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. *Remote Sensing of Environment*, 88, 144–156.
- Schowengerdt, R. A. (2007). *Remote sensing: models and methods for image processing*. New York: Academic.
- Shafiqe, N. A., Fulk, F., Autrey, B. C., & Flotemersch, J. (2006). *Hyperspectral remote sensing of water quality parameters for large rivers in the Ohio River basin*. In *The Second Interagency Conference on Research in the Watersheds*. Otto, North Carolina: Coweeta Hydrologic Laboratory.
- Sidjak, R. W., & Wheate, R. D. (1999). Glacier mapping of the Illecillewaet icefield, British Columbia, Canada, using Landsat TM and digital elevation data. *International Journal of Remote Sensing*, 20, 273–284.
- Smith, V. H., & Shapiro, J. (1981). Chlorophyll–phosphorus relations in individual lakes. Their importance to lake restoration strategies. *Environmental Science and Technology*, 15, 444–451.
- Soto, I., Andrefouet, S., Hu, C., Muller-Karger, F. E., Wall, C. C., Sheng, J., & Hatcher, B. G. (2009). Physical connectivity in the Mesoamerican Barrier Reef System inferred from 9 years of ocean color observations. *Coral Reefs*, 28, 415–425.
- South Florida Water Management District—SFWMD. (2010a). DBHYDRO Available Online: http://www.sfwmd.gov/dbhydrosql/show_dbkey_info.main_menu. West Palm Beach: SFWMD.
- South Florida Water Management District—SFWMD. (2010b). Environmental Report, Chapter 11: Kissimmee Basin, Available Online: http://www.sfwmd.gov/portal/page/portal/pg_grp_sfwmd_sfer/portlet_sfer/tab2236037/2010%20report/v1/chapters/v1_ch11.pdf. West Palm Beach: SFWMD.
- Srivastava, P., McVair, J. N., & Johnson, T. E. (2006). Comparison of process-based and artificial neural network approaches for streamflow modeling in an agricultural watershed. *Journal of the American Water Resources Association*, 42, 545–563.
- Sudheer, K. P., Chaubey, I., & Garg, V. (2006). Lake water quality assessment from Landsat thematic mapper data using neural network: an approach to optimal band combination selection. *Journal of the American Water Resources Association*, 42, 1683–1695.
- Swanson, H., & Zurawell, R. (2006). *Steele Lake water quality monitoring report*. Edmonton, Alberta: Monitoring and Evaluation Branch, Environmental Assurance Division, Alberta Environment.
- Tucker, C. J., Townshend, J. R. G., & Goff, T. E. (1985). African land-cover classification using satellite data. *Science*, 227, 369–375.
- Van der Valk, A. G., Toth, L. A., Gibney, E. B., Mason, D. H., & Wetzel, P. R. (2009). Potential propagule sources for reestablishing vegetation on the floodplain of the Kissimmee River, Florida, USA. *Wetlands*, 29, 976–987.
- Wang, J. P., Cheng, S. T., & Jia, H. F. (2008). *Application of artificial neural network technology in water color remote sensing inversion of inland waterbody using TM data*. In *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. Istanbul: ISPRS Congress, Proceedings of Commission IV.
- Wool, T., Abrose, R., & Martin, J. (2007). *WASP7 multiple algae model theory and users' guide supplement to water quality analysis simulation program (WAS) user documentation*. Washington, DC: U.S Environmental Protection Agency, Office of Research and Development.
- Wu, C. F., Wu, J. P., Qi, J. G., Zhang, L. S., Huang, H. Q., Lou, L. P., & Chen, Y. X. (2010). Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. *International Journal of Remote Sensing*, 31, 2309–2324.
- Yang, X. (2009). *Remote sensing and geospatial technologies for coastal ecosystem assessment and management*. Berlin: Springer.
- Zhang, Y. Z., Pulliainen, J. T., Koponen, S. S., & Hallikainen, M. T. (2002). Application of an empirical neural network to surface water quality estimation in the Gulf of Finland using combined optical data and microwave data. *Remote Sensing of Environment*, 81, 327–336.
- Zhang, Y. Z., Pulliainen, J. T., Koponen, S. S., & Hallikainen, M. T. (2003). Water quality retrievals from combined Landsat TM data and ERS-2 SAR data, in the Gulf of Finland. *IEEE Transactions on Geoscience and Remote Sensing*, 41, 622–629.
- Zhou, Z., & Wang, X. (2008). Quantitative remote sensing researches on water quality of the Weihe River based on SPOT-5 imagery. In *The Proceedings of Information Technology and Environmental System Sciences: Itess 2008*, 2, 538–542.