STATISTICAL ESTIMATION OF THE SAINT JOHN RIVER SURFACE WATER QUALITY USING LANDSAT8 MULTI-SPECTRAL DATA

Essam Sharaf El Din*
PhD Candidate at Department of Geodesy and Geomatics Engineering
University of New Brunswick (UNB)
Fredericton, Canada
esharaf@unb.ca

Yun Zhang
PhD Professor at Department of Geodesy and Geomatics Engineering
University of New Brunswick (UNB)
Fredericton, Canada
yunzhang@unb.ca

ABSTRACT

Surface water quality assessment has been traditionally performed using laboratory analysis; however, these techniques are costly and resulting in irregular sampling over time and space. In contrast, remote sensing is an appropriate tool for estimating concentrations of surface water quality parameters (SWQPs) and providing a consistent spatial and temporal coverage. It is indispensable to estimate concentrations of both optical and non-optical SWQPs, such as total suspended sediments (TSS) and dissolved oxygen (DO), on a regular basis to offer the proper treatment for water bodies. Remote sensing estimation of DO has not yet been performed because non-optical SWQPs are less likely to affect the reflected radiation. However, concentrations of non-optical variables may be correlated with optical variables which have the potential to affect water colour, the reflected radiation, and consequently can be detected by satellite sensors. Therefore, this research attempts to develop an integrated Landsat8 band rationing and stepwise regression (SWR) approach to estimate concentrations of optical and non-optical SWQPs in the Saint John River. Significant correlation between the Landsat8 surface reflectance and concentrations of both TSS and DO was obtained. Compared to previous studies, it was found that our developed models performed very well in predicting TSS and DO with an $R^2 > 0.89$. It is promising to develop appropriate models for SWQP estimation using our approach, and to show the possibility of achieving this task without being dependent on sampling events.

Keywords: water quality; surface water quality parameters; remote sensing; Landsat8; band rationing; stepwise regression

INTRODUCTION

Surface water quality has been deteriorated due to human, agricultural, and industrial activities. Traditionally, surface water quality assessment relies on point sampling at fixed stations by collecting water samples and analyzing in the laboratory. Disadvantages of point sampling can be overcome by utilizing satellite images which potentially offer wide area coverage and continuous marine measurements (Murdoch, Baron and Miller 2000).

Remote sensing estimation of surface water quality parameters (SWQPs) is carried out by modeling the corresponding relationship between satellite multi-spectral information and concentrations of different SWQPs. The Landsat TM/ETM+ and Moderate Resolution Imaging Spectroradiometer (MODIS) satellite images were utilized to develop multiple regression models for retrieving concentrations of turbidity, chlorophyll, secchi disk depth, and suspended solids over various water bodies (Alparslan, et al. 2007, He, et al. 2008, Mancino, et al. 2009, Liu, et al. 2010, Bresciani, et al. 2011, Krista, et al. 2015, Xiang, et al. 2016).

These sensors were designed mainly for earth observation and consequently their signal to noise ratio for low reflectance water surface was inadequate to obtain reliable information about SWQP concentrations. Moreover, almost all of the available publications attempted to estimate only optical SWQPs, such as turbidity and total suspended sediments. Furthermore, simple linear regression of single bands can provide acceptable correlation between satellite data and concentrations of SWQPs. The advantages of using single bands have been confirmed by other researchers such as (Poets, et al. 2010). However, there is no obvious agreement between the researchers on which bands are the best to predict the concentrations of SWQPs. Additionally, several water quality studies were carried out on highly polluted areas; however, other slightly polluted water bodies, such as the Saint John River (SJR), have not been taken into account.

Based on the previous findings, our focus in this research is to estimate concentrations of both optical, such as total suspended sediments (TSS), and non-optical SWQPs, such as dissolved oxygen (DO), from the Landsat8 (L8) satellite data over the SJR. L8 data have been acquired by a recent satellite sensor and new multi-spectral bands have been added to support water quality studies. Moreover, L8 Band rationing was found to be a good tool for estimating concentrations of SWQPs due to its ability to enhance spectral contrast between different targets, and to remove much of the effect of illumination in the analysis of spectral differences.

Furthermore, the proposed regression-based technique is the stepwise regression (SWR) since it has been found to be efficient in the applications of models' prediction and averaging (Derksen and Keselman 1992).

The identified objectives of this research are to: (1) develop a L8-based-SWR model to estimate concentrations of each optical and non-optical SWQP and (2) to generate a spatial concentration map for each SWQP over the selected study area. A L8-based-SWR approach is developed for the first time to map concentrations of DO, which have not been mapped before with the Landsat data or any other optical instrument.

THE PROPOSED METHOD

The flowchart of estimating optical and non-optical SWQPs using the developed L8-based-SWR approach is shown in Fig. 1.

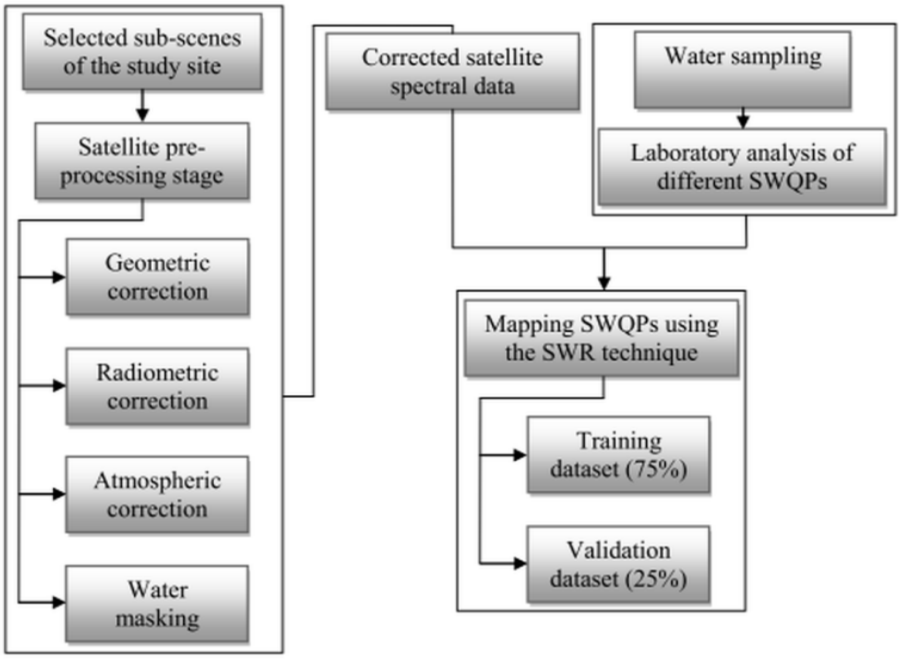


Fig. 1. Flowchart of the proposed methodology

Selected study site

The selected study site is about 70 km long of the SJR as shown in Fig. 2. Peak flows on the SJR occur during the spring season and last several weeks. However, periods of low flow occur during the summer (Arseneault 2008).

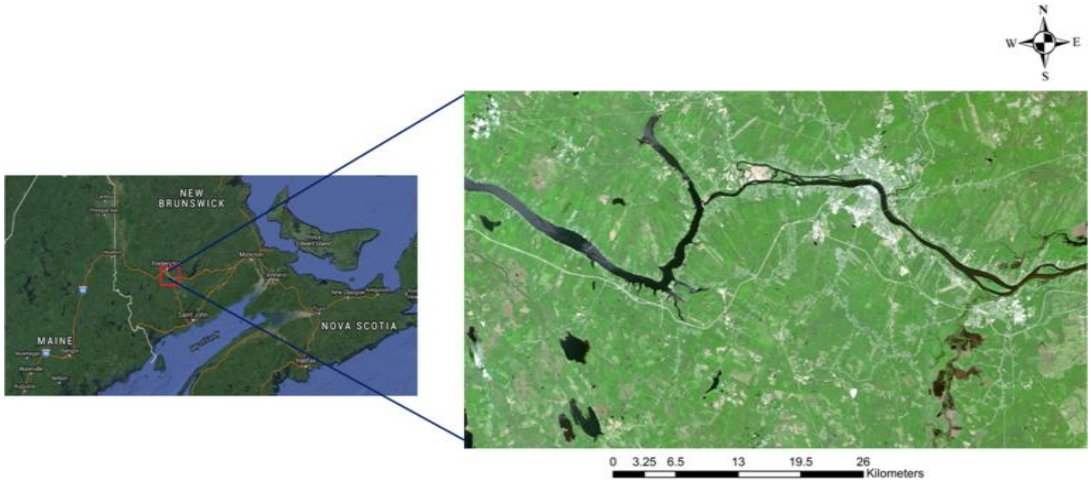


Fig. 2. Selected study site

Landsat8 satellite data

Three L8 sub-scenes, acquired at June 2015 and April and May 2016, were used in this study. These images were geometrically corrected and rectified to the Universal Transverse Mercator projection. Digital numbers of L8 satellite images are stored in 16 bits unsigned integer format. Therefore, they were rescaled to obtain the top of atmospheric (TOA) reflectance using radiometric rescaling coefficients provided in the metadata file. Surface reflectance values were calculated using the Dark Object Subtraction (DOS) method in order to remove the effects of the atmosphere and consequently represent only water-leaving reflectance (Chavez 1988). This method is very efficient in mapping wetland areas and well accepted by the geospatial community to correct light scattering in remote sensing data (Song, et al. 2001). Finally, the normalized difference water index method was utilized to separate water pixels and to delineate the concentrations of optical and non-optical SWQPs over a specific water body (Mcfeeters 1996).

Ground truth data

In this study, 39 water samples were selected and distributed over the study area of the SJR and one sample was excluded due to cloud coverage, as shown in Fig. 3. Sampling events were selected at different months to increase variation between sampling points. At each station, TSS and DO were analyzed according to the standard methods for lab examination of water and wastewater suggested by the American Public Health Association (APHA 2005).

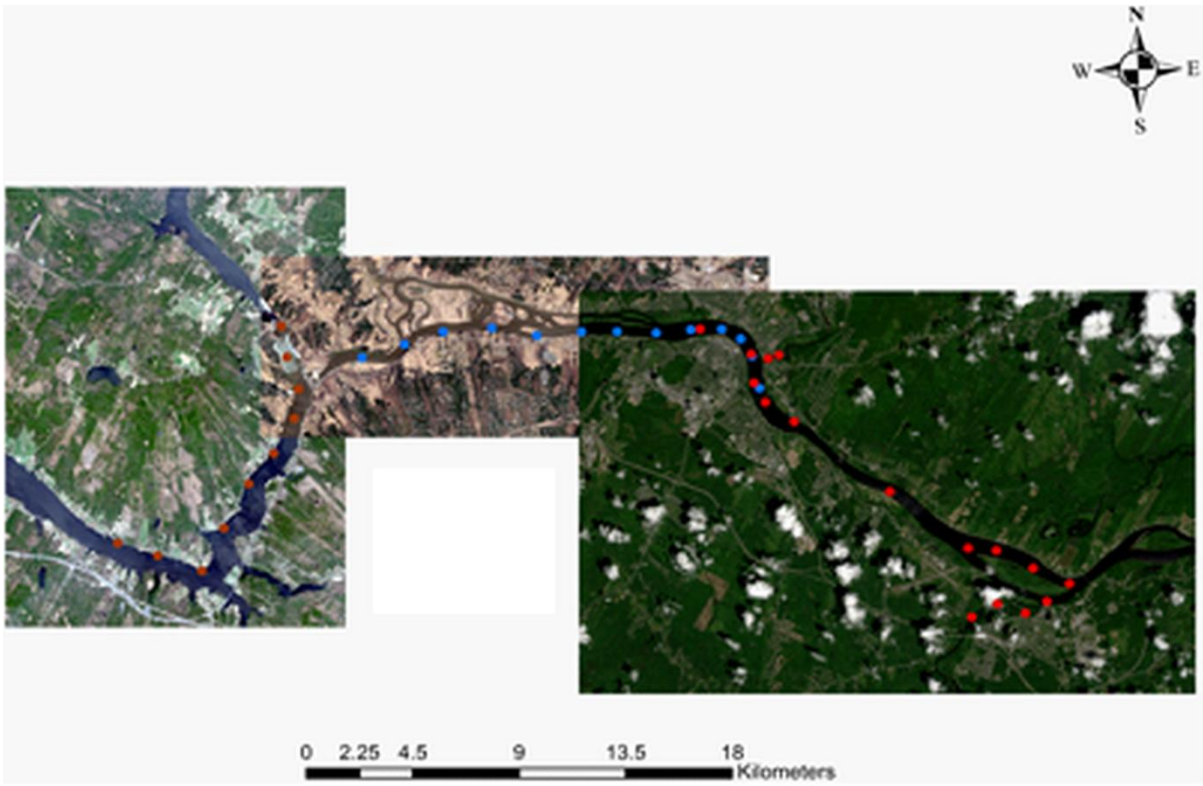


Fig. 3. Water sampling stations

Mapping SWQP concentrations using the L8-based-SWR

The SWR technique was utilized to model the corresponding relationship between L8 reflectance and concentrations of optical and non-optical SWQPs. Sampling points were subdivided into two datasets; calibration (75% of all samples) and validation (25% of all samples) to establish and validate the developed models. The performance of the developed models was evaluated by using scatter plots, determination coefficient (R^2), root mean square error (RMSE), significant value (P-value), and residual prediction deviation (RPD). The RPD can be used as an indication of model stability (Nduwamungu, et al. 2009). Three classes of model were identified based on R2 and RPD values (Chang, et al. 2001) and they are:

- 1st category ($0.80 \leq R^2 \leq 1.00$ and RPD ≥ 2.00) means accurate prediction.
- 2nd category ($0.50 \leq R^2 < 0.80$ and $1.40 \leq$ RPD < 2.00) means satisfactory prediction.
- 3rd category ($R^2 < 0.50$ and RPD < 1.40) means unreliable prediction.

RESULTS AND DISCUSSION

Concentrations of TSS and DO

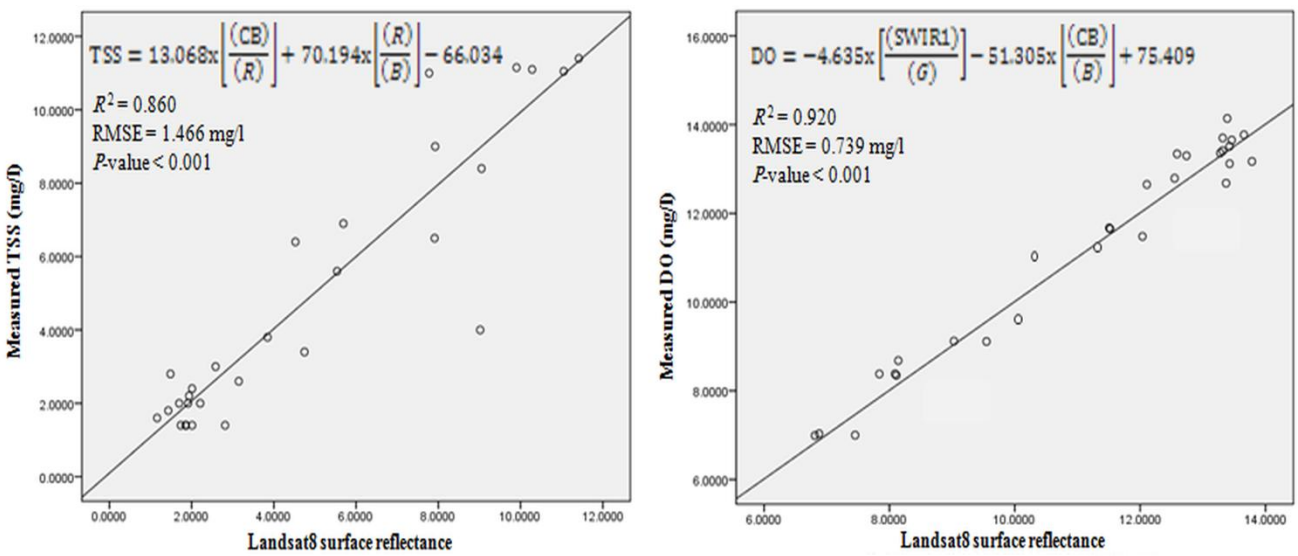
The descriptive statistics, shown in Table 1, were measured for TSS and DO concentrations. For the 39 water samples from the SJR, the concentrations ranged from 1.20 to 11.40 mg/l with an average 4.78 mg/l and 6.99 to 14.14 mg/l with an average 11.06 for TSS and DO respectively.

Table 1. Statistics of concentrations of TSS and DO

SWQPs	Mean	Minimum (Min)	Maximum (Max)	Standard deviation (Std)
TSS (mg/l)	4.78	1.20	11.40	3.61
DO (mg/l)	11.06	6.99	14.14	2.51

Relationship between L8 data and SWQPs

L8 multi-spectral bands, such as blue (B), green (G), red (R), shortwave infrared 1 (SWIR1), and shortwave infrared 2 (SWIR2), significantly contributed to the process of developing accurate models to estimate concentrations of both optical and non-optical SWQPs in the SJR. The new costal blue (CB) band which was added to the L8 multi-spectral bands performed very well in developing TSS and DO estimation models, as shown in Fig.4. These findings confirm the competence of using the L8 OLI imagery in water quality studies.



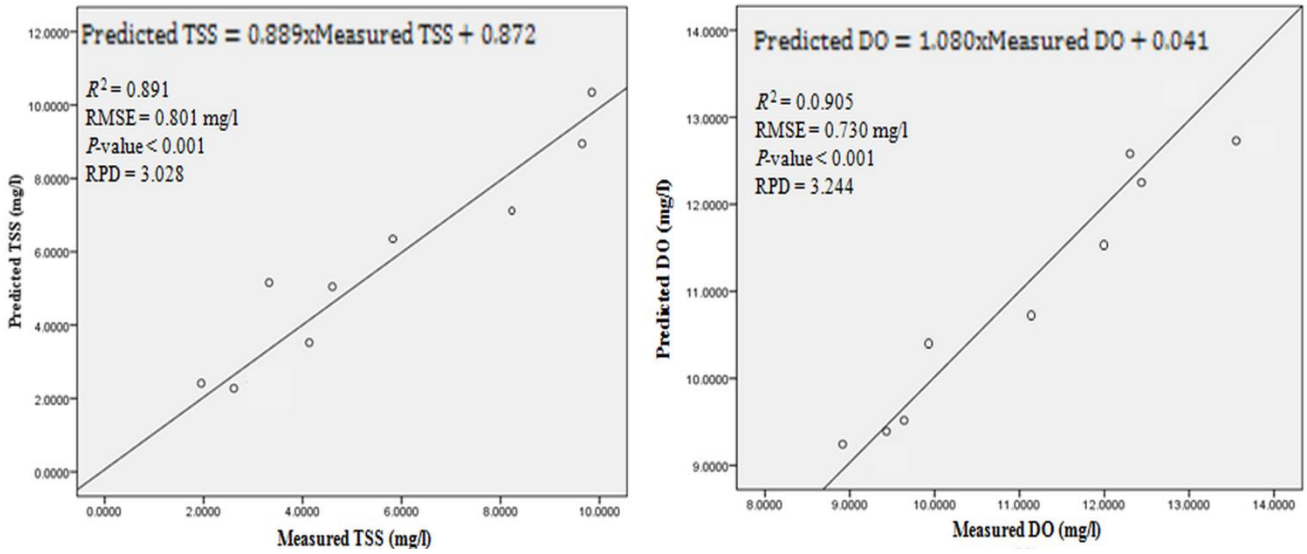
(a) (b)

Fig. 4. Correlation between the L8 data and TSS (a) and DO (b) based on calibration dataset

Estimation and validation of the L8-based-SWR models

The L8-based-SWR estimation models were established based on the calibration dataset and the bands and band ratios that showed the highest correlations were considered in the mathematical model of TSS and DO. To test the reliability of the developed L8-based-SWR models in estimating concentrations of TSS and DO, an independent validation dataset of the remaining water samples was used to validate their performance.

Concentrations of TSS were significantly estimated using the L8-based-SWR approach. The validation models for TSS remained very stable with (R^2 = 0.891, RMSE = 0.801 mg/l, P-value < 0.001, and RPD = 3.028) as shown in Fig. 5. Similarly, the estimation model, which was developed to estimate DO concentrations, was found to be very accurate. Moreover, the validation model for DO was stable with (R^2 = 0.905, RMSE = 0.730 mg/l, P-value < 0.001, and RPD = 3.244).



(a) (b)

Fig. 5. Accuracy measures between measured and predicted concentrations of TSS (a) and DO (b) based on validation dataset

L8-based-SWR spatial distribution maps

The developed L8-based-SWR models were applied to each pixel of the selected study area of the SJR to generate highly accurate spatial concentration maps for TSS and DO. For the entire study area of the SJR, the estimation values of TSS and DO ranged from 1.100 to 14.000 mg/l and 6.500 to 15.000 mg/l respectively, as shown in Fig. 6.

Highly accurate TSS and DO concentration maps were achieved using the developed L8-based-SWR approach because of:

- Band rationing approach was used due to its ability to enhance the spectral contrast between different features.
- The SWR technique was selected as the proposed regression-based technique due to its ability of managing large amounts of independent variables and tuning the model to choose the best independent variables from the available data and its computational speed is mostly faster than other regression techniques.
- The L8 imagery was utilized because its multi-spectral bands were designed to be narrower in order to support coastal applications.
- The L8 surface reflectance data were used to provide only the water-leaving reflectance without introducing any atmospheric distortions.
- Water sampling was performed at the same time of the L8 over pass.

CONCLUSION

Mapping concentrations of optically and non-optically active SWQPs from space is essential to provide both spatial and temporal variability of water quality. In our study, a remote sensing-based-SWR approach was developed to estimate concentrations of TSS and DO using L8 spectral information. Based on the obtained results, the L8 OLI sensor was able to estimate concentrations of optical and non-optical SWQPs in the selected study site. Commonly, regression-based techniques have poor ability to model the complex relationship between remotely sensed data and non-optical SWQPs which do not have direct optical properties and spectral characteristics. However, non-optical SWQPs may be correlated with optical SWQPs such as TSS concentrations which have direct optical properties that can be estimated by remote sensing means.

As a result, in our study, the L8-based-SWR approach was found to be very efficient in establishing and developing highly accurate models to map both optical and non-optical SWQPs with an $R^2 > 0.89$. Finally, to produce more accurate estimation models, especially for non-optical SWQPs, another mapping tool (i.e. artificial intelligence) which is capable of modelling unknown and complex relationship between satellite reflectance and concentrations of SWQPs is needed.

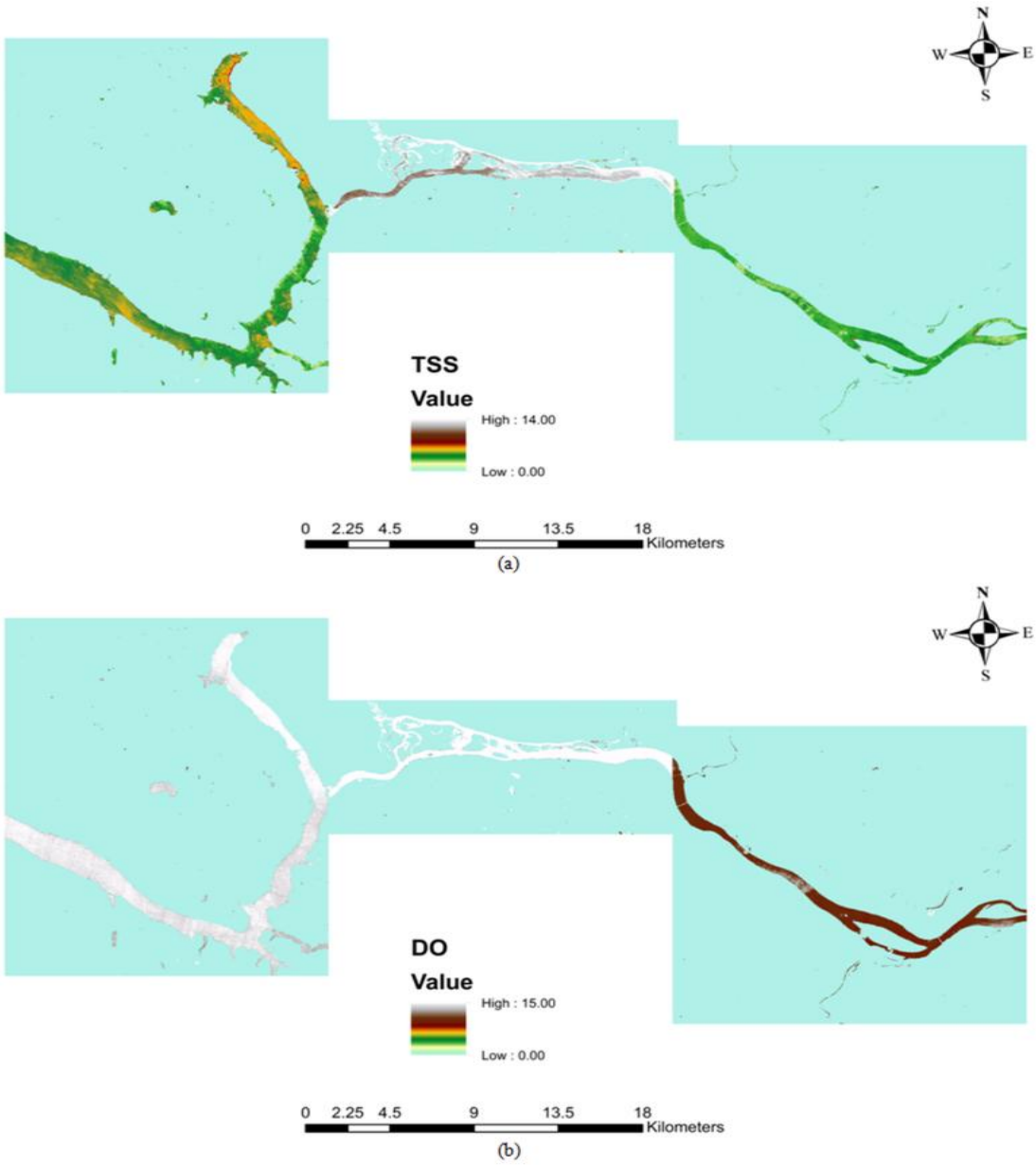


Fig. 6. Spatial distribution maps for TSS (a) and DO (b)

ACKNOWLEDGEMENT

This work was jointly supported by the Egyptian Ministry of Higher Education, the Egyptian Cultural Affairs and Mission Sector, and the Canada Research Chair Program. The authors would like to thank Prof. Dr. Katy Haralampides and Dennis Connor for their participation in the field work. The authors also wish to acknowledge the USGS Landsat Archive Center for the three L8 Level 1T images utilized in this study.

REFERENCES

Alparslan, E., C. Aydöner, V. Tufekci, and H. Tüfekci, 2007. Water Quality Assessment at Ömerli Dam Using Remote Sensing Techniques, Environ. Monit. Assess. 135: pp. 391-398.

APHA, 2005. Standards Methods for the Examination of Water and Wastewater. 21th, American Public Health Association Washington DC, USA.

Arseneault, D., 2008. The Road to Canada - Nomination Document for the St. John River, New Brunswick, The St. John River with the support of the New Brunswick Department of Natural Resources.

Bresciani, M., Stroppiana D., Odermatt D., Morabito G., and Giardino C., 2011. Assessing Remotely Sensed Chlorophyll-A for the Implementation of the Water Framework Directive in European Perialpine Lakes, Sciences Total Environment 409 (17): pp. 3083-3091.

Chang, C. W., Laird D. W., Mausbach M. J., and Hurburgh C. R., 2011. Near-Infrared Reflectance Spectroscopy Principal Components Regression Analyses of Soil Properties, Soil Science Society of America Journal 6: pp. 480-490.

Chavez, P. S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data, Remote Sensing of Environment 24: pp. 459-479.

Derksen, S., and H. J. Keselman, 1992. Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables, British Journal of Mathematical and Statistical Psychology 45: pp. 265-282.

He, W., S. Chen, X. Liu, and J. Chen, 2008. Water Quality Monitoring in Slightly-Polluted Inland Water Body through Remote Sensing-A Case Study in Guanting Reservoir, Beijing, China, Frontiers Environment Sciences Engineering China 2 (2): pp.163-171.

Krista, Alikasa., et al., 2015. Satellite-based products for monitoring optically complex inland waters in support of EU Water Framework Directive, International Journal of Remote sensing 36: pp. 4446-4468.

Liu, D., C. Chin, J. Gong, and D. Fu, 2010. Remote Sensing of Chlorophyll-a Concentrations of the Pearl River Estuary from MODIS Land Bands, Int. J. Remote Sens. 31: pp. 4625-4633.

Mancino, G., A. Nolè, V. Urbano, M. Amato, and A. Ferrara, 2009. Assessing Water Quality by Remote Sensing in Small Lakes: The Case Study of Monticchio Lakes in Southern italy, IForest: pp. 154-161.

Mcfeeters, S. K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features, INT. J. REMOTE SENSING VOL. 17, No.7: pp. 1425-1432.

Murdoch, P. S., J. S. Baron, and T. L. Miller, 2000. Potential Effects of Climate Change on Surface-Water Quality in North America, J. Am. Water Resour. Assoc: pp. 347-366.

Nduwamungu, C., Ziadi N., Parent L. E., Tremblay G. F., and Thuries L., 2009. Opportunities For, and Limitations Of, near Infrared Reflectance Spectroscopy Applications in Soil Analysis: A Review, Canadian Journal of Soil Science 89 (5): pp. 531-541.

Poets, Cf., M. G. Costa, J. C. B. Da Silva, A. M. Silva, and M. Morais, 2010. Remote sensing of water quality parameters over Alqueva Reservoir in the south of Portugal, International Journal of Remote sensing 12: pp. 3373-3388.

Song, C., C. E. Woodcock, K. C. Seto, M. P. Lenney, and S. A. Macomber, 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects, Remote Sensing of Environment 75: pp. 230-244.

Xiang, Yu., Yi. Huapeng, Liu. Xiangyang, Wang. Yebao, Liu. Xin, and Zhang. Hua, 2016. Remote-sensing estimation of dissolved inorganic nitrogen concentration in the Bohai Sea using band combinations derived from MODIS data, International Journal of Remote Sensing 37:2: pp. 327-340.