

Regional assessment of lake water clarity using satellite remote sensing

Stacy A.C. NELSON*, Patricia A. SORANNO¹⁾, Kendra Spence CHERUVELIL¹⁾, Sam A. BATZLI²⁾ and David L. SKOLE³⁾

Center for Earth Observation, Department of Forestry, North Carolina State University, Raleigh, N.C. 27695, USA

¹⁾ Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI 48824, USA

²⁾ Environmental Remote Sensing Center and Nelson Institute for Environmental Studies University of Wisconsin, Madison, WI 53706, USA

³⁾ Department of Geography and Basic Science and Remote Sensing Initiative, Michigan State University, East Lansing, MI 48823-5243, USA

*e-mail corresponding author: sanelso2@ncsu.edu

ABSTRACT

Lake water clarity as measured by Secchi disk transparency (SDT) is a cost-effective measure of water quality. However, in regions where there are thousands of lakes, sampling even a small proportion of those lakes for SDT year after year is cost prohibitive. Remote sensing has the potential to be a powerful tool for assessing lake clarity over large spatial scales. The overall objective of our study was to examine whether Landsat-7 ETM+ could be used to measure water clarity across a large range of lakes. Our specific objectives were to: 1) develop a regression model to estimate SDT from Landsat data calibrated using 93 lakes in Michigan, U.S.A., and to 2) examine how the distribution of SDT across the 93 calibration lakes influenced the model. Our calibration dataset included a large number of lakes with a wide range of SDT values that captured the summer statewide distribution of SDT values in Michigan. Our regression model had a much lower r^2 value than previously published studies conducted on smaller datasets. To examine the importance of the distribution of calibration data, we simulated a calibration dataset with a different SDT distribution by sub-sampling the original dataset to match the distribution of previous studies. The sub-sampled dataset had a much higher percentage of lakes with shallow water clarity, and the resulting regression model had a much higher r^2 value than our original model. Our study shows that the use of Landsat to measure water clarity is sensitive to the distribution of water clarity used in the calibration set.

Key words: Landsat, Secchi depth, water clarity, Michigan

1. INTRODUCTION

Inland lakes provide an important recreational, commercial and aesthetic resource to the public. Thus, it is important to monitor the quality of these lakes. This monitoring often includes measuring water clarity using an inexpensive Secchi disk. Although more sophisticated measurement techniques exist, Secchi disk transparency (SDT) remains an effective coarse measure of water clarity.

Despite the ease of obtaining SDT data, collecting it on large numbers of lakes can be costly and challenging for monitoring agencies. Within the post-glaciated northern states of the United States of America, several states contain a large number of lakes, often on the order of thousands. These lakes are important to the economy of the states as these resources are a valuable source of water, commercial and sports fisheries, and recreation. Routine lake monitoring is important to measure changes in water quality so effective management decisions can be made.

To expand monitoring programs for lakes, many state agencies have developed citizen-volunteer lake monitoring programs for collecting SDT. These pro-

grams have allowed data collection over larger regions than is possible through local and state agencies alone. Many studies have shown that the quality of volunteer collected data is comparable to data collected by professional monitoring personnel with no statistical difference in summer averages of SDT between volunteer and professional measurements (Heiskary *et al.* 1994; Kerr *et al.* 1994; Obrecht *et al.* 1998; Canfield *et al.* 2002). However, extending volunteer collection programs to thousands of lakes, on a statewide level, is still costly and logistically prohibitive. Thus, new approaches are still required to collect SDT on large numbers of lakes.

Satellite remote sensing using Landsat Thematic Mapper (TM) has been explored in several studies as a method of reducing the cost and labor of sampling water clarity in the field (Khorram & Cheshire 1985; Lathrop 1992; Kloiber *et al.* 2000; Dewider & Khedr 2001). One of the strengths of remote sensing techniques is in its ability to provide spatial measurements of features across large areas that are typically not possible with *in situ* sampling. In addition, sensors such as Landsat provide an excellent tool to examine changes over time because this imagery has been collected over large areas of the earth's surface for several decades. To date, Landsat has been used extensively for land-based studies

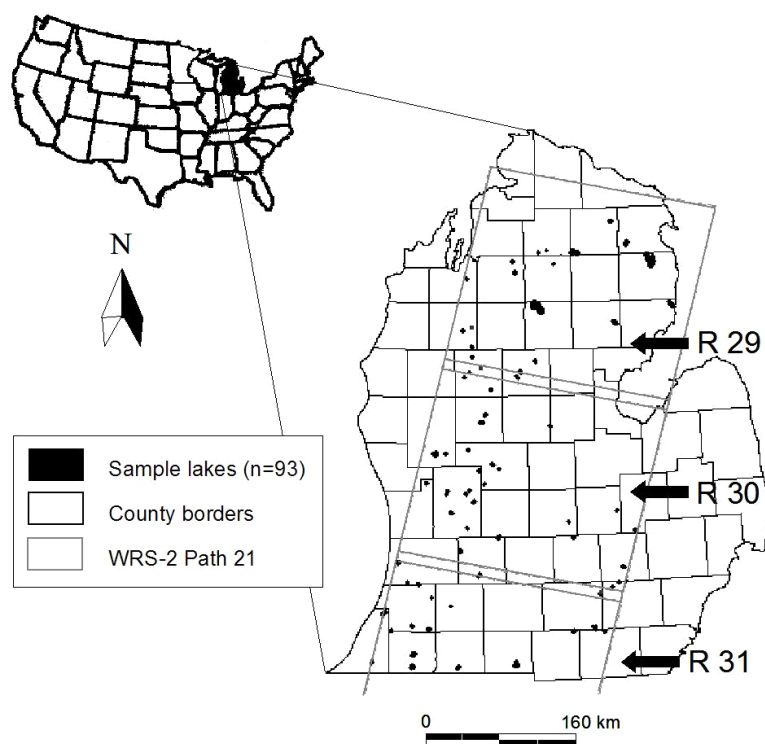


Fig. 1. The study area within Path 21 of the Landsat Worldwide Reference System-2 (WRS) in Michigan, USA. The three individual scene rows within Path 21 used in this study are designated as R29, R30, and R31. All images are from the same acquisition date, 28 August 2001.

(Wilkie & Finn 1996), and less so for inland lake water quality studies, which is primarily due to the spectral band resolutions of the Landsat Thematic Mapper (TM) sensors. Shorter wavelengths are scattered by the atmosphere and longer wavelengths are absorbed by clear water, thus reducing the amount of spectral energy returned to the sensor for aquatic feature detection. Despite these potential limitations, Landsat has been used successfully in aquatic studies for the remote sensing of SDT, chlorophyll-*a*, and suspended sediments (Khorram & Cheshire 1985; Lathrop 1992; Lathrop & Lillesand 1986).

For inland lakes, the relationship between remotely sensed images and SDT has been examined through linear regression analysis of in-situ measurements of SDT and spectral values from the sensor (Lillesand *et al.* 1983; Lathrop & Lillesand 1986; Kloiber & 2002). Regression models developed for single water bodies using a single image have resulted in high r^2 values (0.80 – 0.90, $p < 0.05$) (Lathrop & Lillesand 1986; Lavery *et al.* 1993; Giardino *et al.* 2001). These strong relationships suggest that remote sensing may be useful for measuring SDT in a large number of lakes across larger spatial scales. However, to apply these approaches to the regional scale, such as for the lakes of Michigan U.S.A., this approach must be tested where the spectral values are obtained from multiple Landsat images and for a wider diversity of lakes than is typically found in a single image. The objectives of our study were to deter-

mine if Landsat-7 could be used to measure lake water clarity across a large region of lakes using three Landsat images, and to determine how the distribution of water clarity across the calibration dataset affected the regression models.

2. METHODS

The state of Michigan has approximately 3,500 inland lakes >10 ha in surface area and many thousands of smaller lakes. This study included 93 lakes in the lower peninsula of Michigan (Fig. 1). The lakes were distributed throughout an area of 94,350 km², which made up approximately 80% of the lower peninsula of the state. The lakes ranged in surface area from 12 – 4,125 ha with a mean Secchi depth of 3.1 m between July and August 2001.

Field observations of SDT data were obtained for the 93 lakes from three sampling programs: (1) the Michigan Citizens Lake Monitoring Program (CLMP), (2) the Michigan Department of Environmental Quality's (MDEQ) Lake Water Quality Assessment (LWQA) Monitoring Program, and (3) Michigan State University (MSU). Lake were selected that: (1) were sampled from late July through August to ensure samples were taken during the summer stratified period; (2) had a surface area of >10 ha within the lower peninsula of Michigan; and (3) were sampled within +/-7 days of the satellite imagery.

Three Landsat-7 Enhanced Thematic Mapper Plus (ETM+) scenes from August 28, 2001 from ground track Path 21 (Worldwide Reference System-2) that covered the lower peninsula of Michigan, U.S.A. (Fig. 1) were used in this study. Spectral digital number (DN) values were extracted within the pelagic region of each lake, which was identified by creating an area of interest (AOI) using ERDAS Imagine version 8.4 image processing software. The pelagic region of each lake was defined as the area >4.5 m in depth. Bathymetric maps were used to identify the 4.5 m depth contour within the 80 sample lakes for which maps were available. For the 13 lakes without bathymetric maps, AOIs were created for each lake from groups of pixels within the center of each lake, avoiding any shoreline or shallow areas to ensure that each AOI would be free of reflections from the lake bottom or submersed macrophytes. From each lake AOI, the mean DN value was extracted for each spectral band to use in all subsequent analyses. The AOI sizes ranged from 8 to 1,012 pixels.

All statistical analyses were done using SYSTAT (SPSS Software, Inc., 2001). Probability distribution plots of SDT indicated that a natural log-transformation was necessary for SDT. A linear regression model was developed using the band ratio of ETM1/ETM3 as the independent variable and the natural log of SDT as the dependent variable. Previous studies found these regression variables to be the best predictor of SDT (Lathrop 1992; Pattiaratchi *et al.* 1994; Kloiber *et al.* 2000). Outliers and lakes having large leverages within the SDT and Landsat datasets were removed from the final regression analysis, reducing the number of lakes for the final analysis from 96 to 93.

To examine the role of lake SDT distribution in the regression model, a second regression model was developed with a different SDT distribution based on a subsample of the complete dataset. We used a manual selection technique to create the subsampled dataset with a similar SDT distribution as a previous study that found a strong relationship between multiple lake SDT values and Landsat spectral values (Kloiber *et al.* 2000). To compare the regressions of the complete and subsampled datasets, a slope heterogeneity test and analysis of covariance (ANCOVA) was used. An F-test was also performed on the residual values to determine if there was a significant difference between the variances of the residuals of each dataset.

To compare the distribution of our study lakes to the distribution of lake across the entire state of Michigan, we compiled summer SDT data for Michigan lakes from the US EPA Storage and Retrieval (STORET) database. These data included 675 lakes (>20 ha) sampled during the 1970's and 1980's. If multiple years of data were available for a single lake, then the data were averaged across sample years. Otherwise, the data represent average summer SDT values from biweekly samples in June, July, and August.

3. RESULTS

The regression of SDT and ETM1/ETM3 produced an r^2 of 0.43 ($p < 0.001$; Fig. 2), which seemed low relative to similar studies. The 93 lakes used in our study had a broad distribution of SDT, from 1 to 8 m. Thus to examine the effects of SDT distribution on the model results, we adjusted the SDT sample distribution to have a mean Secchi depth of 1.8 meters and 47% of the lakes having a SDT of <1.5 m. A similar sample distribution yielded strong relationships between SDT and Landsat data in Minnesota, U.S.A. ($r^2 = 0.83$; Kloiber *et al.* 2000). The regression model for our subsampled dataset had an r^2 value of 0.82 ($p < 0.001$; Fig. 3).

To compare our two regression models, we compared the slope of the regression using the complete dataset (n=93) to the slope of the regression using the subsampled dataset (n=17). The slopes were not significantly different (ANCOVA $p = 0.247$). However, examination of the residuals against their predicted values shows that there is less variability in the residuals about the mean for the subsampled dataset than in the complete dataset (Fig. 4). An F-test shows that there is a significant probability that the residual variances in the two datasets are statistically different ($F = 2.86$, $p = 0.01$).

4. DISCUSSION AND CONCLUSIONS

Our results show that the use of Landsat to measure water clarity is sensitive to the distribution of water clarity in the calibration dataset. Several previous studies have developed strong statistical relationships between a single Landsat image and water quality of a single lake or water body. However, to date only Kloiber *et al.* (2000) have examined this relationship on a large scale, using 47 lakes within a 7,700 km² area. The results from Kloiber *et al.* (2000) show promise in developing larger scale satellite-based water clarity monitoring protocols using current sensors. However, our study expands on this work by examining almost twice the number of lakes, with a much wider range of SDT in the calibration dataset. Our low r^2 values, which were lower than the Kloiber *et al.* (2000) study and single-lake studies (Lathrop 1992; Pattiarachi *et al.* 1994; Cox *et al.* 1998), were likely a result of the SDT distribution of our calibration dataset.

Our original calibration dataset included a broad range of SDT values, including deep SDT lakes. Lakes with deeper SDT values return less signal reflectance from algal turbidity in the water column to a remote sensor and, thus, may not be well detected using current sensors. In contrast, the Kloiber *et al.* (2000) study was based on a cluster of lakes within the metropolitan area of Minneapolis and St. Paul, Minnesota, USA, where over 60% of the calibration lakes were eutrophic (Forsberg & Ryding 1980). Similarly, our subsampled dataset included a large percentage of eutrophic lakes and

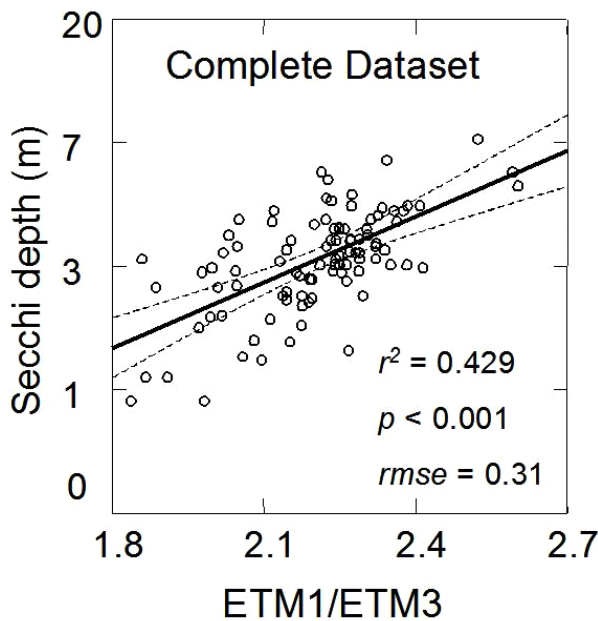


Fig. 2. Regression of Landsat-7 ETM+ data (ratio of two bands, ETM1 and ETM3) and field measured Secchi depth transparency (SDT; meters) for the 93 study lakes. The regression model is $\ln(\text{SDT}) = 1.830 (\text{ETM1}/\text{ETM3}) - 2.976$. The *rmse* residual value (standard error of estimate) has the same units as the predicted ETM values and represents the standard deviation. The dashed lines represent the 95% confidence interval.

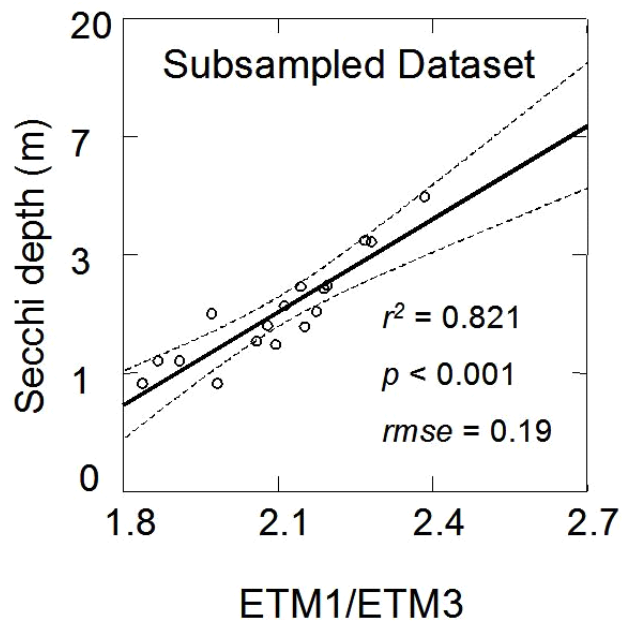


Fig. 3. Regression of Landsat-7 ETM+ data (ratio of two bands, ETM1 and ETM3) and field measured SDT (meters) for our subsampled dataset. The regression model is $\ln(\text{SDT}) = 2.630 (\text{ETM1}/\text{ETM3}) - 5.009$. The *rmse* residual value (standard error of estimate) has the same units as the predicted ETM values and represents the standard deviation. The dashed lines represent the 95% confidence interval.

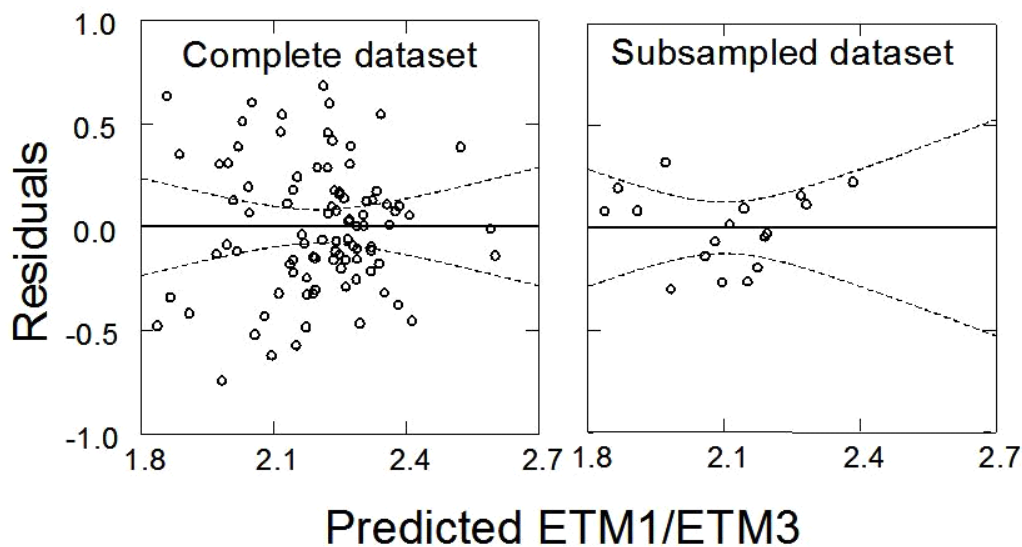


Fig. 4. Scatter plots of the residuals for the complete and subsampled datasets against the predicted value. The dashed lines represent the 95% confidence intervals.

yielded a comparable r^2 value to the Kloiber *et al.* (2000) study.

Our results suggest that to assess the usefulness of Landsat for remote sensing of SDT, one must consider the purpose of the model. If the purpose is to focus on monitoring eutrophic lakes, for example, those lakes in watersheds that have agricultural or urban development, then Landsat can be a very effective moni-

toring tool. However, if the goal is to examine SDT at a statewide level, across thousands of lakes, a broader range of inter-lake variability must be included in the calibration dataset. Unfortunately, it also has to be recognized that the regression model will be more poorly fitted to the data. Our complete dataset is representative of the kind of model that would be needed to monitor all the lakes of Michigan. The range of the SDT data in-

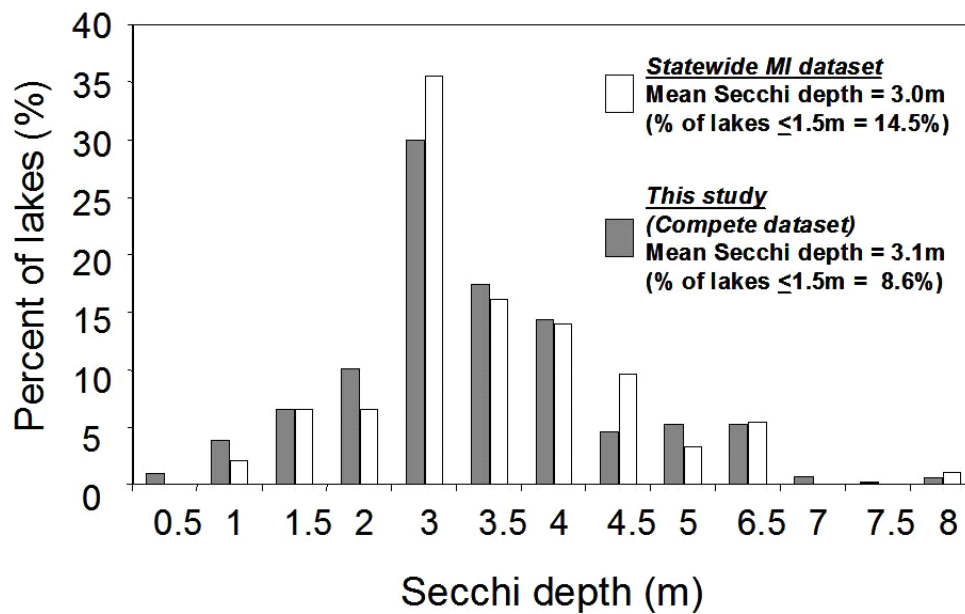


Fig. 5. Histogram of the percent of lakes in each 0.5 m Secchi depth category from a statewide lake database ($n = 675$) and for our 93 study lakes (complete dataset).

cluded in our complete dataset model is comparable to the range of SDT collected from a survey of 675 Michigan lakes (Fig. 5). Secchi disk transparency in Michigan lakes ranges from 0.5 – 8.0 m, with a mean of 3.0 m. The complete dataset used in our study captures this range, from 0.9 – 7.6 m with a mean of 3.1 m.

There were several complications that arose during our study. To use remote sensing for statewide assessments of water clarity, there still must be investment in some statewide sampling efforts that coincide with satellite overflights. Although ground observations were included in this study from a number of reliable sources, available data on a large number of lakes were still hard to find. For example, during this study, Michigan experienced a large number of days with high cloud cover that resulted in few available cloud free scenes. This made it difficult to match field SDT data across the state to the satellite over flight of ± 7 days. In addition, we experienced difficulties in creating a single regression relationship that included images from multiple paths. Further research is necessary to explore the effect of path to path atmospheric variations. For example, within the three contiguous Landsat scenes used in this study, atmospheric corrections were not necessary for sampling SDT because the images were free from atmospheric haze and all of the images were taken on a single date. However, atmospheric correction would have been needed for multiple paths or multiple dates.

In conclusion, although a standard equation relating Landsat data and water clarity over a large area can be used to predict SDT, the sensitivity of the regression model to the distribution of lake calibration data must be taken into consideration. Further research is required

to determine the optimal ranges in which Secchi depth can be predicted from Landsat data, the reliability of predictive models in shallow as well as deep SDT lakes, and the best regression estimates for regions with a wide range of Secchi depth values. Remote sensing can play a vital role in reducing the cost, labor, and time required to monitor inland lake water clarity. It has the potential to be used as a tool to develop statewide assessments that are currently impossible using traditional field operations. However, caution should be used in assessing water quality based solely on remotely sensed data. Current methods and sensors may be able to effectively measure shallow SDT lakes, but may not do as well for lakes with deeper SDT. Until such methods can be improved or better sensors are developed for aquatic applications, inland lake remote sensing will be most useful as a supplement to existing volunteer and agency monitoring programs, rather than as a replacement of these programs.

ACKNOWLEDGMENTS

This research was funded in part by the Michigan State University Minority Competitive Doctoral Fellowship. Landsat-7 ETM+ images were provided by the Basic Science and Remote Sensing Initiative (BSRSI, now called the Center for Global Change and Earth Observation) and through the NASA-funded Upper Midwest RESAC project (Regional Earth Science Applications Center). Additional funding was provided through a USGS/DEQ grant to BSRSI. The CLMP field observation data were provided by Ralph Bednarz, MDEQ. We would like to thank Jiaguo Qi for image processing support and Michael Belligan for database assistance.

We would additionally like to thank Jonathan Chipman, Narumon Wiangwang, Marc Linderman, and Chris Barber for technical support, and Daniel Hayes for statistical advice. Finally, we thank William "Bill" W. Taylor for his never-ending support.

REFERENCES

- Bukata, R.P., J.H. Jerome & J.E. Burton. 1988. Relationships among Secchi disk depth, beam attenuation coefficient, and irradiance attenuation coefficient for Great Lakes waters. *Journal of Great Lakes Research*, 14: 347-355.
- Canfield, D.E., Jr., C.D. Brown, R.W. Bachmann & M.V. Hoyer. 2002. Volunteer lake monitoring: testing the reliability of data collected by the Florida LAKEWATCH program. *Lake and Reservoir Management*, 18: 1-9.
- Cox, R.M., R.D. Forsythe, G.E. Vaughan & L.L. Olmstead. 1998. Assessing water quality in the Catawba River reservoirs using Landsat Thematic Mapper satellite data. *Lake and Reservoir Management*, 14: 405-416.
- Dekker, A.G. & S.W.M. Peters. 1993. The use of the Thematic Mapper for the analysis of eutrophic lakes: a case study in the Netherlands. *International Journal of Remote Sensing*, 14: 799-821.
- Dewider, K. & A. Khedr. 2001. Water quality assessment with simultaneous Landsat-5 TM at Manzala Lagoon, Egypt. *Hydrobiologia*, 457: 49-58.
- Forsberg, C. & S.O. Ryding. 1980. Eutrophication parameters and trophic state indices in 30 Swedish waste-receiving lakes. *Archiv für Hydrobiologie*, 88: 189-207.
- Giardino, C., M. Pepe, P.A. Brivio, P. Ghezzi & E. Zilioli. 2001. Detecting chlorophyll, Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *Science of the Total Environment*, 268: 19-29.
- Heiskary, S., J. Lindbloom & C.B. Wilson. 1994. Detecting water quality trends with citizen volunteer data. *Journal of Lake and Reservoir Management*, 9: 4-9.
- Kerr, M., E. Ely, V. Lee, & A. Mayo. 1994. A profile of volunteer environmental monitoring: National survey results. *Journal of Lake and Reservoir Management*, 9: 1-4.
- Khorram, S. & H.M. Cheshire. 1985. Remote sensing of water quality in the Neuse River Estuary, North Carolina. *Photogrammetric Engineering and Remote Sensing*, 51: 329-341.
- Kloiber, S.M., T.H. Anderle, P.L. Brezonik, L. Olmanson, M.E. Bauer & D.A. Brown. 2000. Trophic state assessment of lakes in the Twin Cities (Minnesota, USA) region by satellite imagery. *Archive Hydrobiologie Special Issues Advances in Limnology*, 55: 137-151.
- Kloiber, S.M., P.L. Brezonik, L.G. Olmanson & M.E. Bauer. 2002. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sensing of Environment*, 82: 38-47.
- Lathrop, R.G. 1992. Landsat Thematic Mapper monitoring of turbid inland water quality. *Photogrammetric Engineering and Remote Sensing*, 58: 465-470.
- Lathrop, R.G., & T.M. Lillesand. 1986. Utility of Thematic Mapper data to assess water quality. *Photogrammetric Engineering and Remote Sensing*, 52: 671-680.
- Lavery, P., C. Pattiaratchi, A. Wyllie & P. Hick. 1993. Water quality monitoring in estuarine waters using the Landsat Thematic Mapper. *Remote Sensing of Environment*, 46: 268-280.
- Lillesand, T.M., W.L. Johnson, R.L. Deuell, O.M. Linstrom, & D.E. Meisner. 1983. Use of Landsat data to predict the trophic state of Minnesota lakes. *Photogrammetric Engineering and Remote Sensing*, 49: 219-229.
- Obrecht, D.V., M. Milanick, B.D. Perkins, D. Ready & J.R. Jones. 1998. Evaluation of data generated from lake samples collected by volunteers. *Journal of Lake and Reservoir Management*, 14:21-27.
- Packard, G.L. & W.J. Emery. 1982. *Descriptive physical oceanography: an introduction*. 4th Ed. Pergamon Press. New York: 47-58.
- Pattiaratchi, C., P. Lavery, A. Wyllie, P. Hick. 1994. Estimates of water quality in coastal waters using multi-date Landsat Thematic Mapper data. *International Journal of Remote Sensing*, 15: 1571 -1584.
- Song, C., C.E. Woodcock, K.C. Seto, M.P. Lenney, & S.A. Macomber. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? *Remote Sensing of Environment*. 75: 230-244.
- Van Belle, G. 2002. *Statistical rules of thumb*. Wiley and Sons. New York: 16-19.
- Wallin, M.L., & L. Hakanson. 1992. Morphometry and sedimentation as regulating factors for nutrient recycling and trophic level in coastal waters. *Hydrobiologia*, 235: 33-45.
- Zilioli, E. 2001. Lake water monitoring in Europe by means of remote sensing. *Science of the Total Environment*. 268: 1-2.