PERGAMON

Water Research 36 (2002) 4330–4340

WATER RESEARCH

www.elsevier.com/locate/watres

Application of Landsat imagery to regional-scale assessments of lake clarity

Steven M. Kloiber[a], Patrick L. Brezonik[b,*], Marvin E. Bauer[c]

[a] *Metropolitan Council Environmental Services, 230 East Fifth Street, St. Paul, MN 55101, USA*

[b] *Department of Civil Engineering, Water Resources Center, University of Minnesota, Minneapolis, 1985 Buford Avenue, St. Paul, MN 55108, USA*

[c] *Department of Forest Resources and Remote Sensing Laboratory, University of Minnesota, 1530 N Cleveland Avenue, St. Paul, MN 55108, USA*

Received 1 June 2001; received in revised form 1 March 2002; accepted 26 March 2002

Abstract

A procedure that uses Landsat imagery to estimate Secchi disk transparency (SDT) of lakes was developed and applied to ~ 500 lakes with surface areas > 10 ha in the seven-county metropolitan area of Minneapolis and St. Paul, MN, USA, to assess spatial patterns and temporal trends in lake clarity. Thirteen Landsat MSS and TM images over the period 1973–1998 were used for the analysis. Satellite brightness values from lake surfaces were calibrated against available historical data on SDT ($n = \sim 20-40$) measured nearly contemporaneously with the acquisition date of each image. Calibration regression equations for the late-summer TM images had a range of r^2 from 0.72 to 0.93. Regression analysis for three late-summer MSS images yielded r^2 values ranging from 0.60 to 0.79. Results indicate that a single late-summer image yields a reliable estimate of regional lake clarity and reasonably accurate estimates of SDT for individual lakes. An analysis of seasonal patterns on a large lake water-quality database was used to develop a model that adjusts synoptic satellite SDT estimates from different dates to a common reference, making them more comparable from year-to-year. Analysis of long-term trends shows that in spite of the large land-use changes within the region over the study period, only 49 (about 10%) of assessed lakes in the region showed significant temporal trends in SDT over the period, and more lakes had increasing SDT (34) than decreasing SDT (15).
© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Landsat; Satellite; Secchi disk; Lake; Clarity; Trends

1. Introduction

Broad-based, regional-scale monitoring is critical for making informed lake management decisions. Data from monitoring programs are used frequently to estimate expected ranges in water quality for lakes in a region, examine regional differences, and investigate the relationships between landscape conditions and water quality. This information can be used to develop water-quality criteria and to guide management decisions. A properly designed monitoring program should have adequate representation across both space and time to provide the information needed for scientific or management purposes. However, the design of regional-scale monitoring programs frequently requires a compromise between spatial and temporal detail. Because of cost and logistical problems, ground-based monitoring programs usually sacrifice spatial coverage, monitoring fewer lakes in favor of more frequent sampling.

Relatively few programs monitor many lakes across large area or states to conduct regional-scale assessments. The Great American Secchi Dip-In [1], which

*Corresponding author. Tel.: +1-612-624-9282; fax: +1-612-625-1263.

E-mail address: brezo001@tc.umn.edu (P.L. Brezonik).

0043-1354/02/$ - see front matter© 2002 Elsevier Science Ltd. All rights reserved.
PII: S0043-1354(02)00146-X

uses volunteers to collect Secchi disk transparency (SDT) measurements across many states during an early July index period, is one example. However, this program illustrates a potential weakness of most ground-based monitoring programs, non-random site selection. Lakes are included in the Dip-In based on interests or judgments of volunteers. Data from such sampling schemes cannot be extrapolated reliably to a broader population of lakes, as is frequently the goal in regional-scale assessments. Using data from judgment-based programs for regional analyses may result in biased results and misleading conclusions [2].

Recently, there has been increasing focus on regional analysis (e.g., [3,4]); however, few monitoring programs exist for this purpose. As such, we have had to rely on data from judgment-based sampling programs. One solution to this problem is to use a probability-based sampling design. Another approach would be to sample the entire population of lakes in the region of interest. For many lake quality characteristics, that would be infeasible, but for characteristics related to the optical properties of water, satellite-based measurements may offer a practical alternative.

Satellite technology can provide spatially unbiased information on certain characteristics of lakes across broad regions and has considerable potential to serve as a cost-effective complement to ground-based monitoring programs. Spectral brightness measured by satellite sensors is correlated with several water-quality variables that affect a lake's optical properties. These include SDT, turbidity, suspended solids, humic color, and chlorophyll *a* (chl *a*) concentration. Several studies have shown strong relationships between brightness data collected by satellite systems and these water-quality variables (e.g., [5–10]). However, previous efforts to use satellite data for lake studies have involved short-term measurements on relatively few lakes rather than long-term, regional-scale programs for lake quality analysis.

Previous work by our group [11] developed a standardized method to make synoptic estimates of SDT in lakes of the Twin Cities (Minnesota) Metropolitan Area (TCMA) lakes. In this paper, we describe how synoptic satellite imagery assessments of lake water clarity can be used to inform lake management decisions. We begin with an analysis of ground observation data, including seasonal patterns for trophic variables (especially SDT and chlorophyll), the timing and frequency of SDT minimum, and the relationship between the summer minimum and summer mean SDT. We compare estimates of summer mean SDT from single and multiple image assessments to the summer mean estimated from ground observations. Finally, we summarize results of satellite-based SDT measurements for nearly 500 lakes across the 7700 km^2 TCMA over a 25-year period.

2. Background

Ground-based lake monitoring programs usually use monthly or biweekly sampling. In the TCMA, this results in 7–14 samples per lake during the ice-free season each year. However, TCMA monitoring programs were not designed to collect data for regional analysis. Only a small portion of the lakes in the TCMA are monitored in a given year, and more than half of the lakes have never been monitored. This type of "judgement-based monitoring" yields a spatially incomplete picture of water quality for any given year [2]. Furthermore, because monitoring programs in the TCMA do not include the same lakes every year, the data have limited use for analysis of temporal trends.

Although satellite imagery cannot be used to evaluate the many non-optically active variables of interest to lake managers, a single Landsat image allows simultaneous measurements on all lakes in a given region, potentially providing a comprehensive picture of optically related water-quality characteristics at a given time. A Landsat image covers an area ~ 180 km (110 miles) on a side or 3.24×10^4 km^2. The seven-county TCMA, which encompasses 7700 km^2, thus fits within one Landsat image.

Satellite-based water-quality assessments are unlikely to achieve the same temporal frequency as ground-based monitoring can achieve anytime soon. Landsat 5 and 7 each collect imagery for a given scene every 16 days. With both satellites operating concurrently, an image for a given scene is acquired every 8 days. Landsat 5 has far exceeded its expected life, however, and it is uncertain how much longer it will continue to operate. Furthermore, not every Landsat image is suitable for analysis because of interference from cloud cover. Kloiber et al. [10] found that there was only a 30% chance that any given image of the TCMA would be cloud-free ($<10\%$ cloud cover). With both Landsat satellites in operation, there is a 76% probability that there will be at least one cloud-free image each month in the TCMA. With only one Landsat system in operation, the probability of a cloud-free image in a given month decreases to about 50%. Landsat data thus cannot be used reliably to monitor short-term trends in water clarity, but satellite systems are well suited to provide broad spatial coverage and long-term (multi-year) trends.

3. Methods

Data used for this study include historic ground observation data of SDT and chlorophyll *a* (chl *a*) concentrations and 10 Landsat TM and three Landsat MSS images of the TCMA. Ground observation data were used to analyze intra-seasonal variability of lake water quality and to calibrate satellite brightness data.

The database used for the former purpose is from a Metropolitan Council program that uses citizen volunteers to measure SDT and collect samples for chl *a* and other analyses. This database includes information on 146 lake basins with 1–10 years of measurements each for a total of 374 lake–years of observations over the period 1986–1997 [12]. Ground observations of SDT used to develop relationships with satellite reflectance data were obtained from the EPA's storage and retrieval (STORET) database, which includes data from the Metropolitan Council's lake monitoring programs. Ground-observation data from within ± 1 day of each Landsat overpass were used to develop satellite response relationships, following the recommendations of Kloiber et al. [11]. Calibration data for each image included 20 or more ground observations and spanned a wide range of ambient water clarity conditions. Included among the 13 satellite images were four Landsat TM images from June through September 1991 that were used for analysis of intra-seasonal trends.

The 13 cloud-free or nearly cloud-free ($<10\%$ cloud cover) images were selected from the period 1973 through 1998. All images were registered to the Universal Transverse Mercator (UTM), Zone 15 geographic projection using the North American Datum 1983 (NAD 83). The registration process used ~ 40 carefully selected, ground control points for each image. The root mean square error (RMSE) for positional accuracy was on the order of 0.25 pixels (7.5 m for TM data). A nearest-neighbor re-sampling scheme was used to preserve the original brightness values of the MSS and TM data. ERDAS Imagine, version 8.3 was used for all image-processing steps [13].

The first step in extracting the image data for lakes used an unsupervised classification method based on a clustering algorithm with 10 clusters specified. The clusters then were aggregated into land and water classes, and the resulting raster map was used as a binary mask to create a water-only image of the original brightness data. Using the image-processing software, we drew a polygon around a cluster of pixels that represented typical open-water conditions to define an area-of-interest (AOI) for each lake. An AOI was created for each of the ~ 500 TCMA lakes >10 ha (25 acres) in surface area. The AOI polygons were developed using the following sources of information: (1) A geographic information system (GIS) coverage of sampling point locations (imported into ERDAS) was used as a guide to locate the AOI. (2) Bathymetric maps obtained from the Minnesota Department of Natural Resources (MDNR) were examined to avoid shallow, littoral areas (depth <5 m) of the lake whenever possible. (3) An unsupervised classification map of the water-only image was used as a guide to avoid areas possibly influenced by of macrophyte vegetation. These areas tend to exhibit high spatial variability in their spectral signatures compared with the relatively well-mixed, open-water portion of the lake. The size of the AOI generally followed the recommendation of Lillesand et al. [6] to select a large number of pixels from the profundal zone (depth >5 m) of the lake. AOIs ranged from just a few pixels (minimum of nine for small lakes) up to about 1000 pixels (large lakes); the median AOI had 38 pixels. Each AOI polygon was tagged with a unique MDNR lake identification number that was used later to join the data to the ground observation database. The AOI polygon coverage was used to generate a statistical report of the brightness values for each lake.

A regression equation was developed that relates ground-measured SDT to spatially averaged brightness in selected Landsat bands for 20–40 lakes per image [11]. We used bands 1 and 3 for Landsat Thematic Mapper data and bands 1 and 2 for Landsat multi-spectral scanner data. Each regression equation was applied to its image to create the database of satellite-estimated SDT for all lakes with AOIs. All statistical analyses were conducted using SPSS (version 10.0).

4. Results

4.1. Seasonal analysis of ground observation data

For reasons described earlier, measurements of water clarity and related indicators of water quality can be conducted only relatively infrequently using satellites. It thus is important to determine how this affects the usefulness of satellite-based assessments for lake management purposes. Stadelmann et al. [12] examined the TCMA lake database to address this issue. They found that when normalized to the lake–year average, SDT and chlorophyll levels in TCMA lakes followed a predictable, seasonal cycle that fit a sine function. In addition, the period from mid-July through early September was found to be the optimal time to measure trophic conditions because: (1) the variance in SDT is at its seasonal minimum, (2) trophic state conditions are typically at their worst, (3) lake use for water-contact recreation is high, and (4) the rate of change in SDT and chlorophyll is smaller during this period (Fig. 1). One to three measurements in this index period could be used to estimate the growing-season maximum and mean chl *a* and the growing-season minimum and mean SDT. Their analysis is extended here to determine whether we can reliably estimate mean or maximum trophic conditions from one or a few satellite measurements.

Because the inter-lake variability of chl *a* and SDT of TCMA lakes was large, the data were normalized. Growing-season averages were calculated for each lake and year, and the observations were converted to percent differences from the "lake–year" means. To

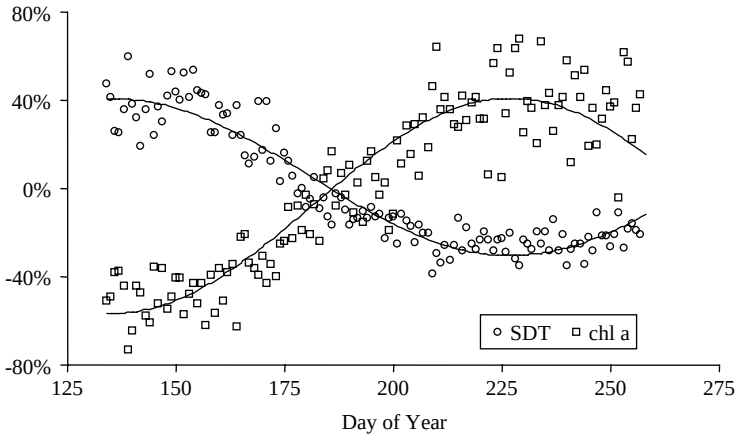

Fig. 1. Seasonal cycle of SDT and chl *a* in TCMA lakes based on data normalized to percent difference from lake–year mean.

examine seasonal patterns of chl *a* and SDT across the region, the data were further aggregated by calculating the average of the normalized values for all lakes on a given day. This step condensed the data such that the daily averages could be plotted over the course of the growing season to show general seasonal patterns (Fig. 1). Chlorophyll levels tended to be lower than the growing-season mean for the first half of the growing season (May through mid-July), increased rapidly in early summer, and stayed above the mean for the second half (mid-July–September). As expected, SDT followed the opposite trend; transparency tended to be greater than the mean during the first half of the growing season and less than the mean for the latter half. During the growing season, SDT and chl *a* values are at or near their mean values for only a short period—a few days in early July. This is also the period with the greatest rate of change in trophic conditions. Maximum trophic indicator conditions occur in late summer (approximately the second half of July to early September), which is the period when the rate of change is at its minimum. This period typically coincided with stronger regression equations between satellite data and ground observations of SDT. Therefore, this period was considered as an index period.

The seasonal trends in chl *a* and SDT data (Fig. 1) fit the following sine functions:

$$\mathrm{SDT}_{\mathrm{rel}} = a[\sin(2\pi(j-90)/182.5)] + b, \quad (1)$$

$$\mathrm{chl}\, a_{\mathrm{rel}} = a[\sin(2p(j-90)/182.5)] + b, \quad (2)$$

where $\mathrm{SDT}_{\mathrm{rel}}$ is the relative percent difference from the lake–year mean for SDT, chl a_{rel} is the relative percent difference from the lake–year mean for chl *a*, *j* is day of the year (between 1 and 365), and *a* and *b* are the fitted slope (amplitude) and intercept (offset). The values for *a* and *b* differ for SDT and chl *a* and were determined by regression analysis. The resulting equation was used to predict growing-season mean SDT and chl *a* using a single measured value and the day of the year for the measurement. Using Eq. (1), we estimated lake–year mean values of SDT for each of the 3223 observed values of SDT and compared these values to lake–year mean values calculated from the observed data (Fig. 2). Based on this analysis, Eq. (1) explains nearly 75% of the variability for the lake–year mean SDT values. Similarly, Eq. (2) explains 60% of the variance in chl *a*. Eq. (1) also can be used to estimate the growing-season minimum SDT from a single measurement because the fitted sine function shows that the minimum SDT is

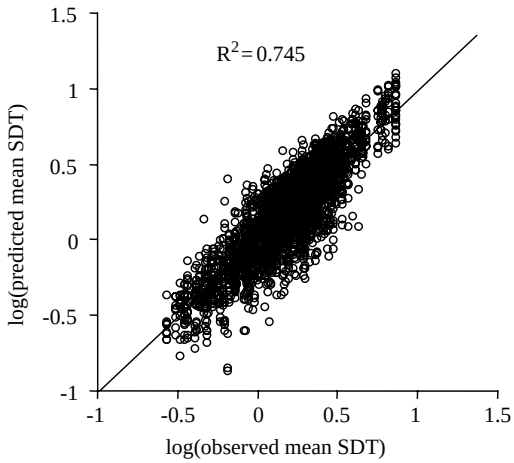

Fig. 2. Comparison of ground-based, growing-season mean SDT and growing-season mean SDT estimated from a single ground observation using sine function model (Eq. (1)).

~30% less than the mean. A regression relationship between growing-season mean and minimum SDT found that the observed minimum is about 75% of the mean (slope of the regression line was 0.75), which agrees well with the seasonal sine function.

For lake managers, the growing-season mean and the worst-case conditions (maximum chl *a* and minimum SDT) are of great interest for lake monitoring, whether by ground-based measurement or satellite imagery. Selecting an image from the late summer allows for a direct estimate of the worst-case condition, and because there is a strong relationship between the minimum and the mean, we also can estimate the growing-season mean. Given limitations on the frequency of satellite data collection, the optimal time for image acquisition also would occur when the temporal variability of trophic variables is at a minimum. The normalized seasonal data in Fig. 1 suggest that SDT variability over time is relatively small during late summer. Analysis of within-lake variability for the TCMA lake database during the period from late July to early September indicates that the relative standard deviation is about 19% (about 0.26 m) for a given lake during this period [12]. The within-lake variability of chl *a* was higher, with a relative standard deviation of 38% (about 14 μg/L).

4.2. Regression equation results

The form of the best regression equations relating Landsat data to water clarity for the TCMA images is given by the following equations [11]:

$$\ln(\mathrm{SDT}) = a(\mathrm{TM1} : \mathrm{TM3}) + b\,\mathrm{TM1} + c, \quad (3)$$

$$\ln(\mathrm{SDT}) = a(\mathrm{MSS1} : \mathrm{MSS2}) + b\,\mathrm{MSS1} + c, \quad (4)$$

where *a*, *b*, and *c* are fitted coefficients from the regression for each individual image. The coefficients varied from image to image because of a variety of factors, including differences in atmospheric conditions, as well as sun angle differences. However, if each image is independently calibrated, the relationship is consistently strong without over-fitting the relationship [11]. Regression equations relating Landsat TM data and SDT for each image (Table 1) produced generally strong relationships, with r^2 ranging from 0.53 to 0.92. Images from the summer index period tended to have higher coefficients of determination than images outside or near the limits of the index period. Images outside (or near the limits) of the index period also tended to have larger standard errors of the estimate (SEE) (Table 1). These trends may be explained at least partially by the relative instability in water clarity outside the index period. In June, the rate of change in water clarity is higher than it is during late summer. By roughly mid-September, increasing algal senescence is an added complication, and some lakes may have begun to de-stratify. Nonetheless, models relating Landsat MSS or TM data and SDT were not always weak on dates outside the index period. For example, the relationship was strong for MSS data and SDT for an image from July 3, 1973 ($r^2 = 0.79$) and weaker for an MSS image from September 6, 1983 ($r^2 = 0.59$).

4.3. Seasonal analysis of satellite-estimated SDT data

Satellite-estimated SDT values of TCMA lakes from four Landsat TM images for the 1991 growing season (June 16, July 18, September 4, and September 20) were compared with historical ground observation data. The number of lakes involved in each comparison was limited to ~20–40 lakes based on the availability of ground observation data. Satellite-estimated SDT values compare favorably with ground observations collected within a window of ±1 day of the satellite overpass

Table 1
Regression equations for prediction of Secchi disk transparency from Landsat TM data

Image date	Equation	r^2	Standard error
7/3/73	ln(SDT) = 2.965(MSS1:MSS2) − 0.0847(MSS1) − 2.99	0.791	0.344
8/7/75	ln(SDT) = 5.230(MSS1:MSS2) − 0.245(MSS1) − 2.37	0.759	0.402
9/6/83	ln(SDT) = 3.029(MSS1:MSS2) − 0.264(MSS1) − 0.384	0.599	0.436
8/21/86	ln(SDT) = 1.324(TM1:TM3) + 0.00777(TM1) − 4.87	0.867	0.250
8/26/88	ln(SDT) = 0.657(TM1:TM3) − 0.0114(TM1) − 2.39	0.929	0.177
6/16/91	ln(SDT) = 0.546(TM1:TM3) + 0.0562(TM1) − 4.43	0.526	0.423
7/18/91	ln(SDT) = 0.882(TM1:TM3) + 0.0754(TM1) − 7.07	0.756	0.369
9/4/91	ln(SDT) = 1.073(TM1:TM3) + 0.0828(TM1) − 8.23	0.816	0.254
9/20/91	ln(SDT) = 1.113(TM1:TM3) + 0.115(TM1) − 10.14	0.608	0.312
8/24/93	ln(SDT) = 1.162(TM1:TM3) + 0.0352(TM1) − 6.05	0.722	0.367
7/29/95	ln(SDT) = 1.262(TM1:TM3) + 0.0376(TM1) − 6.25	0.902	0.208
7/15/96	ln(SDT) = 1.066(TM1:TM3) − 0.0588(TM1) − 0.557	0.771	0.307
9/7/98	ln(SDT) = 0.953(TM1:TM3) − 0.00815(TM1) − 3.18	0.841	0.236

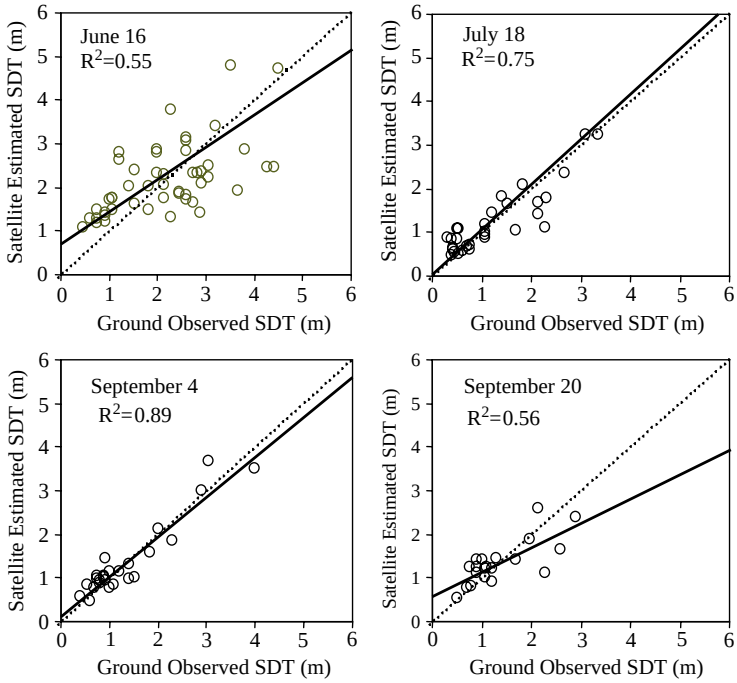

Fig. 3. Comparison of satellite-estimated and ground observed SDT for Landsat TM images from 1991: June 16, July 18, September 4, and September 20.

(Fig. 3), but the correlations are weaker for the two images outside the index period (June 16, September 20) than those for the images within the index period (July 18, September 4). Furthermore, the relationship between satellite-estimated SDT and ground-observed SDT closely matched the 1:1 prediction line for the images inside the index period, but the relationship deviated from the 1:1 line for the images outside the index period.

An analysis of satellite-estimated SDT for 17 lakes with sufficient ground observations found the same seasonal pattern as the ground observation data. When the satellite data for these lakes were normalized to growing-season mean, they followed the same sine function as the historic data (Fig. 4). In addition, statistical distributions of satellite-estimated SDT values for the four 1991 images agree well with distributions of the ground observations.

The equation defining the relative percent difference for SDT can be rearranged to solve for the mean SDT

$$SDT_{mean} = SDT_{obs}/(SDT_{rel}\text{-}1). \quad (5)$$

We substituted Eq. (1) into Eq. (5) and applied it to the satellite-estimated SDT values to derive growing-season mean SDT values. These satellite-estimated, growing-season mean values were compared with the growing-season mean values calculated for 20 lakes with sufficient ground data (at least five SDT observations over the season). The correlation between the growing-season mean SDT from ground observations and the growing-season mean SDT estimated from one satellite image (Fig. 5) was moderate for the June 16 image ($r^2 = 0.54$) and strong for the July 18, September 4 and September 20 images ($r^2 = 0.90$, 0.82, and 0.81). The index period images in this analysis (July 18 and September 4) had strong correlations between ground-based and satellite-based estimates for both synoptic and growing-season mean SDT. The synoptic results show that the results for the index period images closely match a 1:1 line between the observation and the prediction (Fig. 3). The growing-season mean analysis show that the results for the index period images are reasonably close to the 1:1 line (Fig. 5). Landsat images from the summer index period thus provide a reliable estimate of SDT for the date of the image and also provide a reasonable estimate the growing-season mean SDT.

4.4. Spatial–temporal distribution of satellite-estimated SDT

Eq. (1) can be applied to synoptic Landsat images to estimate either the minimum SDT or the growing-season mean SDT. For the purpose of long-term trend analysis, we applied this seasonal adjustment to estimate the growing-season mean SDT for the 13 Landsat images. The results then were examined to determine distributions of lake clarity for the region and identify regional and individual lake trends. Satellite-estimated

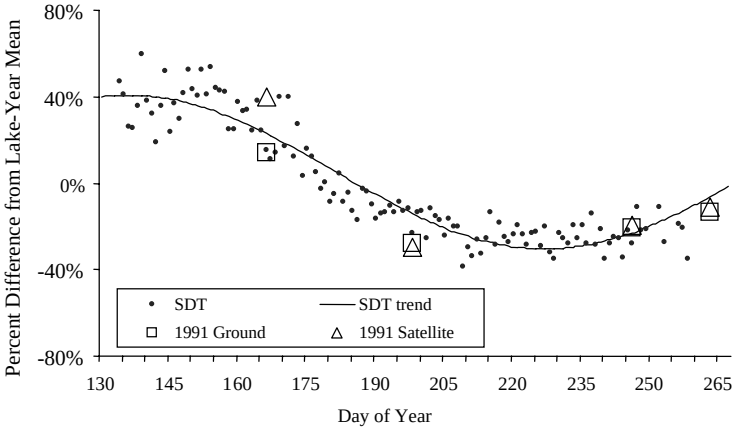

Fig. 4. Normalized ground data for SDT versus Julian day. Closed dots: mean of normalized ground data for a given day for all TCMA lakes measurements in early 1990s, calculated as percent difference from growing-season means (total of 3223 SDT measurements); open squares (□): grand means of normalized data for 17 lakes with sufficient measurements to calculate growing-season means and measurements within 1 day of each 1991 image (values plotted on dates of the four 1991 images); open triangles (△): satellite-estimated SDT for these lakes for each image normalized to lake–year means.

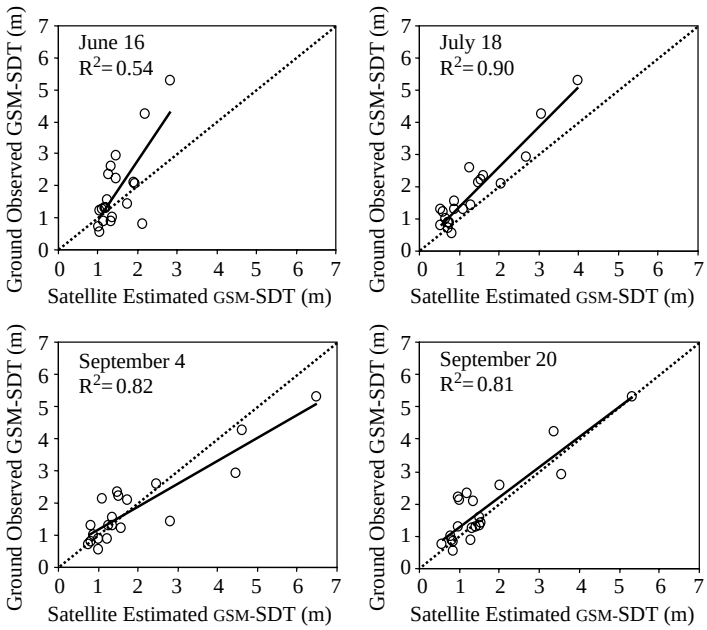

Fig. 5. Comparison of satellite-estimated, growing-season mean SDT and ground-based, growing-season mean SDT for four Landsat TM images from 1991: June 16, July 18, September 4, and September 20.

growing-season mean values of SDT for TCMA lakes >10 hectares in area ($n \sim 500$) ranged from <0.5 m to ~5.0 m (Fig. 6). When the lakes were grouped into clarity classes based on 1 m intervals, the most common lake clarity class for all images was found to be 1.0 m < SDT <2.0 m. The median growing-season mean SDT for all lakes in the region ranged from 1.1 to 1.6 m over the study period with a typical value of 1.4 m.

Excellent agreement between satellite-estimated and ground-observed SDT trends can be achieved, as is demonstrated by the similarity of the trends in SDT (Fig. 7) found by the two methods for Coon Lake (Anoka County). Although some scatter is shown in the individual estimates of growing-season mean SDT, these two methods give virtually identical trends in water clarity.

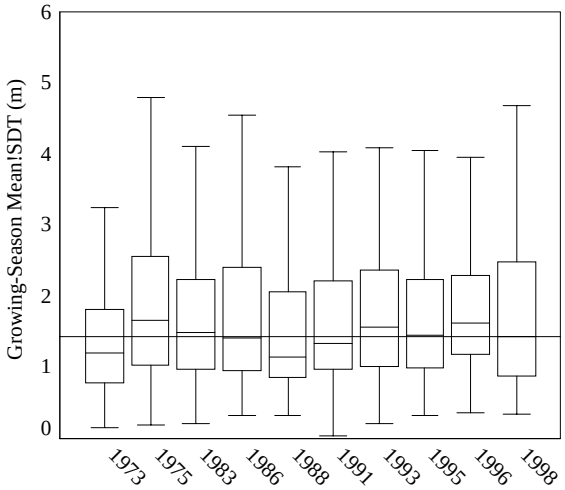

Fig. 6. Box plot summarizing satellite-estimated, growing-season mean SDT for ~500 TCMA lakes over a 25-year period, 1973–1998.

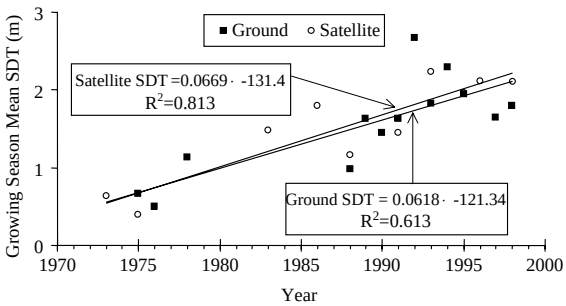

Fig. 7. Comparison of ground-based and satellite-based water clarity trends for Coon Lake, Anoka County, Minnesota.

The Kruskal–Wallis test showed no significant differences in TCMA regional water clarity between the years 1983, 1986, 1991, 1995, and 1998. Water clarity was below normal for the years 1973 and 1988 and above normal for 1975, 1993 and 1996. Although there has been no apparent trend in water clarity for lakes at the scale of the whole TCMA region, a Kendall-Tau-b analysis indicated that 49 lakes did have significant temporal trends. Water clarity has improved in 34 lakes since 1973 and declined in 15 lakes (Table 2). For the lakes with statistically significant increases or decreases in clarity over time, the most common trend (43% of lakes with trends) was an increasing trend of 4–8% per year. Lakes with decreasing water clarity generally are small or quite shallow, which suggests that the morphometry of a lake may predispose it to respond more strongly to changes in natural or anthropogenic stresses.

5. Discussion

5.1. Requirements for estimating mean SDT

When evaluating the suitability of lake water quality for body-contact recreation, lake managers are commonly interested in determining the average condition for the recreation season. Standard practice has been to measure water-quality variables such as phosphorus, chl *a*, and SDT several times during the growing season and then calculate an arithmetic mean. Guidance regarding the frequency of sampling for this type of analysis ranges from weekly to every other week [14,15]. These recommendations are predicated on the need to know the mean value within some specified level of precision, as defined by the standard deviation, confidence interval, or a similar statistical measure of variability. The assumption behind this treatment of the data is that the measurements are independent of each other. This clearly is not the case because the data are serially correlated (Fig. 1). As such, the assumption of independence overstates the variability and inflates the number of samples required to achieve a specified level of precision.

A significant portion of the observed variability of SDT throughout the growing season can be explained using a simple sine function of the day of the year (Fig. 2). When this sine function was used to normalize the SDT data to account for predictable seasonal changes, the standard deviation for a typical lake in the TCMA database was reduced by half (from 1.37 to 0.69 m). According to Gilbert [16], the following formula can be used to estimate the number of samples required for simple random sampling when the data are not serially correlated:

$$n = (Z_{1-\alpha/2}\eta/d_r)2, \tag{6}$$

where n is the number of samples required to estimate the mean to a pre-specified relative error (d_r), $Z_{1-\alpha/2}$ is the standard normal, two-tailed Z-value, and η is the coefficient of variation (i.e. the standard deviation divided by the mean). According to this equation, decreasing the standard deviation by one-half reduces the required number of samples by a factor of four. If we assume that the data are not serially correlated, we would need between 5 and 13 SDT measurements for a reasonably reliable estimate of the growing-season mean [12]. However, the existence of strong serial correlation implies we actually could use between one and three measurements to estimate the growing-season mean with the same level of precision.

Using the satellite-estimated SDT and the sine-function model (Eq. (1)), we are able to estimate the growing-season mean SDT from any single Landsat image with a typical combined standard error of ± 0.5 m. This includes the error from the synoptic

Table 2
Trends in water clarity for TCMA lakes from 1973 through 1998

Lake ID	Lake name	Area (acres)	Tau-b	Count	Change (%/yr)
02000300	Otter	286.6	0.644	10	8.5
02002100	Tamarack	88.3	0.667	9	2.5
02004200	Coon	1493.3	0.733	10	6.0
02004800	South Coon	48.4	0.556	10	5.7
10001000	Tamarack	32.9	0.556	10	6.9
10004500	Steiger	162.5	0.556	10	5.0
19002000	Chub	232.8	−0.600	10	−5.6
19002500	Keller	57.9	0.556	10	6.5
19005600	O'Leary	10.8	0.556	9	10.4
19005700	Fish	31.0	−0.611	9	−9.0
19006300	Schwanz	12.6	0.714	8	6.3
19007900	Pickerel	119.4	−0.556	10	−6.4
19016100	East Thomas	9.5	0.714	7	11.2
27003501	Sweeney	67.4	0.600	10	4.4
27004203	Twin−North	120.6	−0.511	10	−3.5
27006100	Champlin Mill Pond	31.0	0.556	9	6.0
27007600	Red Rock	78.9	−0.556	10	−3.8
27009300	Glen	66.6	0.733	10	8.7
27011900	Cedar Island	86.8	−0.644	10	−3.3
27013303	Carsons Bay	121.1	0.511	10	9.9
27013311	Maxwell Bay	297.2	0.600	10	6.6
27013313	North Arm	316.3	0.556	10	6.1
27013600	Silver	41.5	−0.810	7	−11.5
27014100	Tanager	56.5	0.644	10	5.1
27015400	Katrina	207.0	−0.600	10	−8.4
27016500	Jubert	80.5	−0.511	10	−4.5
27017800	Ox Yoke	101.6	0.556	10	0.6
62000400	Pigs Eye	590.6	0.511	10	1.1
62000600	Kohlman	85.5	0.511	10	2.0
62001000	Keller	72.2	0.556	10	2.5
62005700	Josephine	116.3	0.822	10	6.9
62006200	Charley	35.2	−0.556	10	−5.7
62006900	Pike	35.1	0.600	10	1.1
62007000	Round	110.7	0.600	10	7.6
62007100	Valentine	60.9	0.511	10	4.3
62007200	Karth	16.1	−0.643	8	−7.8
62023400	Little Pig's Eye	41.9	0.556	9	1.8
70001000	Murphy	46.8	0.556	10	2.8
70005200	Cynthia	193.1	0.600	10	2.8
70006900	Fish	169.7	−0.733	10	−5.6
82000900	Cloverdale	38.2	0.644	10	5.6
82001600	Silver	77.0	−0.644	10	−9.4
82005200	Big Marine	1840.6	0.511	10	8.5
82005500	Nielson	40.7	−0.556	10	−4.3
82010300	Olson	90.7	0.556	10	3.3
82012300	Bass	28.0	0.571	8	10.0
82015500	Long	39.5	−0.556	9	−3.4
82016300	Clear	433.6	0.644	10	4.2
82016800	Mud	176.1	0.667	9	7.8

estimate of SDT and the error from the relationship between one measurement and the growing-season mean. Better estimates are obtained if only late-summer images are used (combined standard error decreases to 0.40 m). For example, if all 1991 images are included, the relationship has an R^2 of 0.642 and SEE = 0.527. With just the July 18 and September 4 images, the relationship improves to $R^2 = 0.81$ and SEE = 0.393.

5.2. Use of satellite-estimated SDT

Satellite remote sensing data can be used to estimate water clarity in lakes that would have otherwise gone unmonitored. These data can be used to identify sub-regional differences, assess regional status of lakes, examine temporal trends, and investigate the relationships between the landscape and water quality.

Rohm et al. [17] provide one example of the potential use of regional-scale assessments. They used phosphorus data from over 2600 lakes (about 12% of the total number of lakes) in the northeastern United States to map regional patterns of phosphorus and to define 61 distinct sub-regions. Satellite assessment of SDT could support this type of analysis. In fact, because satellite imagery can be used to sample the entire population of lakes within broad regions, it could be used to establish lake sub-regions with far less effort. This information could then be used to set regionally appropriate lake clarity standards.

Regional-scale data on lake clarity also have potential importance for assessments of recreational suitability. Heiskary and Wilson [18] showed that user perception of swimmability in Minnesota is closely correlated with water clarity. Guidelines from the Minnesota Pollution Control Agency (MPCA) suggest that for TCMA lakes to fully support body-contact recreation, the lakes should have growing-season mean total phosphorus $\leqslant 40\,\mu g/L$ [19]. Based on established relationships between trophic indicator values [20], this concentration corresponds to a growing-season mean SDT of 1.2 m. Using this guideline, we found that only 60% of the TCMA lakes assessed by the 1998 satellite image were fully supportive of swimming; the remaining 40% were either partially supportive or non-supportive of swimming. A MPCA assessment based on ground data for the North Central Hardwood Forest ecoregion, within which the TCMA predominantly lies, shows that 51% of lakes fully support swimming. However, interpretation of the MPCA assessment is difficult because the assessment used data spanning a period of more than two decades. Differences between lakes in such an assessment may reflect annual climate variability or long-term changes in pollutant loading as well as differences in lake and watershed characteristics. With a satellite-based assessment, all lakes can be assessed with a single snapshot in time. Consequently, differences resulting from annual climate variability or long-term water-quality trends are controlled.

Use of non-random (judgment-based) data for regional assessments also may lead to biased results [2]. One way to avoid this is to use a probability-based sampling design. The US Environmental Protection Agency has designed one such a program, the Environmental Monitoring and Assessment Program (EMAP), to evaluate ecological conditions, including lake water quality, across regional and national scales (e.g., [4,21]). EMAP is designed to monitor and assess resources through a probability-based monitoring framework based on a systematic grid. Although results of this program should be unbiased regionally, only a very small fraction of the total population of lakes would be sampled. Thus, for any unmonitored lake, a statement could be made only regarding the probability distribution for the region within which the lake lies. Using a satellite assessment method, clarity estimates can be made for each lake. However, satellite assessment is limited to water-quality variables that are related to optical properties of water, and there are many other variables of interest in lake management that are not directly related to optical properties. For these variables, ground-based sampling is required.

Another important use of regional lake data is to examine the relationships between lake and watershed characteristics and water quality. Although the TCMA has undergone significant change in land use over the 25-year period covered by the satellite images analyzed in this study, no area-wide effects were observed on lake clarity. Only about 10% of the assessed lakes were found to have significant temporal trends in clarity, with 7% improving and 3% declining. The distribution of water clarity and the occurrence of long-term trends appear to be related to lake morphometry. This conclusion is supported by the findings of Day [22], who showed that lake depth had a larger influence on clarity of lakes in central Minnesota than did land use. Further research should be undertaken to quantitatively define the relative influence of landscape and lake characteristics on the spatial and temporal distribution of lake clarity in the TCMA.

6. Conclusions

The results of this research show that Landsat imagery data can be used to perform population-based water clarity assessments over large areas. Using this approach, we were able to evaluate water clarity on a given date for 500 lakes across the entire TCMA, an area of 7700 km^2, from a single Landsat image. (In fact, the study area occupied only about one-fourth of a Landsat image.) An important benefit of a Landsat water clarity assessment is the ability to use this technique in a retrospective analysis. The Landsat imagery archive extends back to 1973. Using 13 Landsat images, we were able to conduct a regional, water clarity assessment over a 25-year period, and through this analysis, we identified statistically significant water clarity trends for 49 lakes. Given the increasing interest in the assessment of status and trends for environmental indicators at regional and national scales, we conclude that satellite-based assessment methods, such as this, can

be an important complement to traditional ground-based monitoring programs.

References

[1] Lee J, Binkley MR, Carlson RE. The Great American Secchi Dip-In—GIS contributes to national snapshot of lake-water transparency. GIS World 1997;10:42–4.

[2] Peterson SA, Urquhart NS, Welch EB. Sample representativeness: a must for reliable regional lake condition estimates. Environ Sci Technol 1999;33:1559–65.

[3] Cohen P, Alley WM, Wilber WG. National water-quality assessment: future directions of the US geological survey. Water Res Bull 1988;24:1147–51.

[4] Overton WS, White D, Stevens Jr. DL. Design report for EMAP—Environmental Monitoring and Assessment Program. EPA/600/3-91/053, US Environmental Protection Agency, Environmental Monitoring and Assessment Program, Washington, DC, 1990, 52pp.

[5] Brown D, Warwick R, Skaggs R. Lake condition in east central Minnesota. Rept. No. 5022, Minnesota Land Management Information System, Center for Urban and Regional Affairs, University of Minnesota, Minneapolis, MN, 1977.

[6] Lillesand TM, Johnson WL, Deuell RL, Lindstrom OM, Meisner DE. Use of Landsat data to predict the trophic state of Minnesota lakes. Photogramm Eng Remote Sensing 1983;49:219–29.

[7] Lathrop RG, Lillesand TM. Utility of Thematic Mapper data to assess water quality in southern Green Bay and west-central Lake Michigan. Photogramm Eng Remote Sensing 1986;52:671–80.

[8] Lathrop RG. Landsat Thematic Mapper monitoring of turbid inland water quality. Photogramm Eng Remote Sensing 1992;58:465–70.

[9] Cox RM, Forsythe RD, Vaughan GE, Olmsted LL. Assessing water quality in the Catawba River reservoirs using Landsat Thematic Mapper satellite data. J Lake Reserv Manage 1998;14:405–16.

[10] Kloiber SM, Anderle TH, Brezonik PL, Olmanson LG, Bauer ME, Brown DA. Trophic state assessment of lakes in the Twin Cities (Minnesota, USA) region by satellite imagery. Arch Hydrobiol Adv Limnol 2000;55:137–51.

[11] Kloiber SM, Brezonik PL, Olmanson LG, Bauer ME. Development of a remote sensing protocol for synoptic, regional water quality assessment. Remote Sensing Environ, 2002, in press.

[12] Stadelmann TH, Brezonik PL, Kloiber SM. Seasonal patterns of chlorophyll-*a* and Secchi disk transparency in lakes of east-central Minnesota: Implications for design of ground- and satellite-based monitoring programs. Lake Reserv Manage 2001;17:299–314.

[13] Olmanson LG, Bauer ME, Brezonik PL, Kloiber SM. Upper Great Lakes RESAC lake water quality image processing protocol. Special Report, Water Resources Center, University of Minnesota, St. Paul, MN, 2000.

[14] Heiskary S, Lindbloom J, Wilson CB. Detecting water quality trends with citizen volunteer data. J Lake Reserv Manage (special issue) 1994;9:4–9.

[15] Osgood RA. An evaluation of lake and stream monitoring programs in the Twin Cities Metropolitan Area. Publ. No. 590-89-128, Metropolitan Council, St. Paul, MN, 1989.

[16] Gilbert RO. Statistical methods for environmental pollution monitoring. New York: Van Nostrand Reinhold, 1987.

[17] Rohm CM, Omernik JM, Kiilsgaard CW. Regional patterns of total phosphorus in lakes of the northeastern United States. Lake Reserv Manage 1995;11:1–14.

[18] Heiskary S, Wilson CB. Minnesota lake water quality assessment report. Minnesota Pollution Control Agency, St. Paul, MN, 1988.

[19] Heiskary S. Proposal for listing lakes on the 303(d) list of impaired waters based on nutrient overenrichment. Minnesota Pollution Control Agency, St. Paul, MN, 2000.

[20] Carlson RE. A trophic state index for lakes. Limnol Oceanogr 1977;22:361–9.

[21] Baker JR, Peck DV, Sutton DW, editors. Environmental monitoring and assessment program surface waters: field operations manual for lakes. EPA/620/R-97/001, US Environmental Protection Agency, Washington, DC, 1997.

[22] Day EE. Landsat 5 and 7 as tools for monitoring water quality. M.S. thesis, University of Minnesota, Minneapolis, 2000.