A procedure for regional lake water clarity assessment using Landsat multispectral data

Steven M. Kloiber a, b, Patrick L. Brezonik b, *, Leif G. Olmanson c, Marvin E. Bauer c

a Metropolitan Council, St. Paul, MN 55101, USA
b Department of Civil Engineering, University of Minnesota, Minneapolis, MN 55455, USA
c Department Forest Resources, University of Minnesota, St. Paul, MN 55018, USA

Received 27 December 2000; received in revised form 20 June 2001; accepted 20 November 2001

Abstract

Although previous investigations have demonstrated reliable empirical relationships between satellite data and nearly contemporaneous ground observations, satellite imagery has not been incorporated into routine lake monitoring programs. This paper focuses on key issues involved in applying satellite imagery to the regional assessment of lake clarity. Ten Landsat Thematic Mapper (TM) images and four Multispectral Scanner (MSS) images of the Twin Cities Metropolitan Area (TCMA) spanning 25 years (1973–1998) were analyzed. Based on this analysis, recommendations are made for a Landsat-based procedure for water clarity assessment. Closeness of fit ($r^2$ values) of regression models between satellite brightness data and measured Secchi disk transparency (SDT) decreased with increasing size of the time window between image collection and ground observation of SDT. Use of SDT data collected within ±1 day of the image date is recommended, but where SDT data are limited, windows up to ±7 days yield reasonable results, especially in late summer when water clarity is relatively constant. Average brightness data from at least nine pixels in the deep open area of a lake should be used to predict lake clarity, but the accuracy of predicted SDT did not improve much as the number of pixels in the area of interest (AOI) increased above this value. A three-coefficient regression model using the TM1/TM3 ratio and TM1 was a consistent and reliable predictor of SDT ($r^2$ values of .7–.8). A similar relationship involving the MSS1/MSS2 ratio and MSS1 was a reasonable predictor of SDT for MSS data. Efforts to produce a standard prediction equation for SDT applicable to images collected on different dates were not successful, but a simple regression procedure to account for differences in atmospheric conditions among image collection dates substantially decreased the range of coefficients in the regression model. © 2002 Published by Elsevier Science Inc.

1. Introduction

Satellite imagery holds significant potential for enhancing regional monitoring and assessment of lake water quality and trophic conditions. Although several satellite systems have been used for water quality assessment, the combination of temporal coverage, spatial resolution, and data availability makes the Landsat system particularly useful for assessment of inland lakes. Several investigations have demonstrated that reliable empirical relationships can be developed between Landsat Multispectral Scanner (MSS) or Thematic Mapper (TM) data and ground observations of water clarity and chlorophyll (Brown, Warwick, & Skaggs, 1977; Cox, Forsythe, Vaughan, & Olmsted, 1998; Dekker & Peters, 1993; Lathrop, 1992; Lathrop & Lillesand, 1986; Lillesand, Johnson, Deuell, Lindstrom, & Meisner, 1983). Recently, Pulliainen et al. (2001) demonstrated that the estimation of water quality from remote sensing data for numerous lakes within a region could be achieved using ground observation data for only a few representative lakes from the region. Although satellite remote sensing has significant potential for regional assessment of lakes, its use has moved slowly from proof of concept to routine application. Possible reasons for this include a lack of familiarity among inland aquatic scientists with remote sensing technology and a lingering perception that the data are expensive and difficult to process. With the return to public sector distribution in 1999, the costs of Landsat data acquisition dropped significantly. Along with the advent of today’s powerful desktop computers and sophisticated software, processing and analysis of satellite imagery has become relatively inexpensive and easy to perform.
The overall objective of this study was to develop a method for using satellite remote sensing data to estimate key variables related to lake management issues, such as trophic state condition and water clarity. The three variables most commonly used to indicate trophic state are total phosphorus (TP), chlorophyll a (Chl), and Secchi disk transparency (SDT). Measurements of these variables, along with various transformations such as the trophic state indices of Carlson (1977), are widely used by lake management agencies and organizations. Of these variables, SDT and Chl have promise for estimation from satellite imagery data (Lillesand et al., 1983). In this study, we focus on SDT because ground observations for this variable are more available. SDT has been widely incorporated into volunteer monitoring programs and is used frequently to identify trends in lake conditions (Heiskary, Lindbloom, & Wilson, 1994). It has been shown to correlate with user perceptions of water quality (Heiskary & Walker, 1988). As such, it is important for defining aesthetic and recreational impairment of lakes and helps to quantify the “fishable and swimmable” goals of the Clean Water Act in a lake management context.

The general approach used in previous applications of Landsat imagery to lake clarity estimation has been regression analysis of Landsat data and nearly contemporaneous ground observation data. A combination of atmospheric and lake factors unique to a specific scene is imbedded in the coefficients of the resulting equation. Ideally, a single (standard) relatively simple equation with constant coefficient values would be used to calculate a satellite-based water clarity index or a phytoplankton abundance index. The use of a single standard equation would avoid the need for calibration to contemporaneous ground data. This would enhance the usefulness of remote sensing in the broader context of regional assessment or retrospective trend analysis because the lack of ground (lake) data in some areas and for some time periods is a limitation. While the practicality of developing a standard equation for remotely sensed water clarity is still uncertain, three refinements to the general approach for satellite-based assessment of lake clarity are presented here that enhance its usefulness for routine applications.

One factor influencing the strength and reliability of ground–satellite relationships for water clarity is the length of the “time window” between ground and satellite observations. The effects of varying the width of this window were evaluated. A wide time window would provide significant advantages. A monitoring program with a narrow time window may be difficult to maintain routinely, especially in more remote areas, and may be unnecessary if short-term, within-lake variability of water clarity is small compared with seasonal, annual, or between-lake variability. A wide time window between collection of ground and satellite data would allow more flexibility, possibly allowing the use of existing monitoring programs or volunteer monitoring programs to collect calibration data. In addition, a wide time window would allow the use of more historical ground observation data to conduct retrospective analyses with archived satellite imagery. Second, the effect of changing the size of the area of interest (AOI) for a given lake was examined. The optimal AOI is the group of pixels whose spectral characteristics are most closely correlated with the conditions measured at the ground observation (lake surface) sites. It is commonly assumed that a single site in the center of the lake provides a representative sample of the entire lake. Satellite imagery provides data across the entire lake surface, including such areas as shallow bays and littoral areas with emergent vegetation that may have water quality that is significantly different from the deeper open water area. Therefore, when extracting spectral data from a satellite image, care must be taken to include only the pixels from areas that have similar water quality to the ground observation site. Third, rather than using regression equations where the independent variables are different for each image, we examined the feasibility of using a consistent equation form to relate ground observations and satellite data. Use of a consistent equation form is preferable because it allows for easier comparison of the results from different images. It also would be preferable for the independent variables to have a physical basis for their inclusion. We describe work below that promotes the development of a standardized method for lake clarity assessment using satellite remote sensing.

2. Methods

The study area was Twin Cities Metropolitan Area (TCMA) in east–central Minnesota. This area consists of a seven-county area of about 7700 km² of rural and suburban land surrounding the urban core of Minneapolis and St. Paul, MN. The regional landscape is characterized by glacial landforms, such as moraines and outwash plains, resulting in topography that ranges from nearly level to gently rolling hills and bluffs. Most lakes in the region were formed when large ice blocks were left behind at the retreat of the last glaciation. Glacial and postglacial activity resulted in the formation of four physiographic regions in the TCMA and accounts for the distribution and character of the lake basins (Osgood, 1989; Wright, 1989). The TCMA includes more than 500 lakes > 10 ha and many more lakes and ponds less than 10 ha. Most of the lakes are considered moderately eutrophic, and water clarity in the lakes generally is controlled by the abundance of algae (Stadelmann, Brezonik, & Kloiber, 2001). The largest lake, Lake Minnetonka, is a multibasin lake with a surface area of 5666 ha. Nearly 80% of the TCMA lakes have surface areas < 40 ha. Maximum depths for TCMA lakes range from < 1 to 43 m. Of the 330 lakes that have been surveyed for depth, the median depth is 4.3 m (unpublished data from the Minnesota Department of Natural Resources and the Metropolitan Council).

Only images with less than 10% cloud cover were used for analysis. This criterion significantly reduced the pool of images suitable for analysis, but most years had at least one
late summer image that met the criterion (Kloiber et al., 2000). From the pool of suitable images, we selected 10 Landsat TM images and four MSS images spanning a 25-year period (1973–1998) (Table 1). 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 approximately 40 carefully selected ground control points for each image. The root mean square error (RMSE) for positional accuracy was generally \( \sim 0.25 \) pixels (7.5 m for TM data). A nearest-neighbor resampling scheme was used to preserve the original brightness values of the image. ERDAS Imagine version 8.3 was used for all image processing.

The procedure used here followed the procedure described in Olmanson, Kloiber, Bauer, and Brezonik (2001) and briefly summarized here. The first step in extracting image data for lakes used an unsupervised classification method based on a clustering algorithm with 10 clusters that were then aggregated to land and water classes. 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, a polygon was drawn around a cluster of pixels that represented typical open water conditions to define an AOI for each lake. An AOI was created for each of about 500 lakes. The size of the AOI generally followed the recommendation of Lillesand et al. (1983) to select a large number of pixels from the profundal zone (depth >5 m) of the lake. AOIs ranged from just a few pixels for small lakes to about 1000 pixels for large lakes, with a median of 38. Auxiliary information used to guide AOI selection included a geographic information system (GIS) coverage of sampling point locations and bathymetric maps. In addition, an unsupervised classification map of the water-only image was used as a guide to avoid areas possibly influenced by macrophytic 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. Each AOI polygon was assigned a unique identification number that was used to join the satellite data to the ground observation database.

Available SDT and Chl data were obtained from the US EPA’s Storage and Retrieval (STORET) database for lakes within the project area and for time periods within \( \pm 7 \) days of each satellite image used for the study. These data included results from the Metropolitan Council’s lake monitoring programs. In general, SDT data were more readily available than Chl, although in recent years Chl monitoring has become more common (Table 1). A data set that included 20 or more ground observations per image, spanning a wide range of ambient conditions, was considered minimally acceptable for this effort. SDT data were sufficient to meet this criterion for all the images (number of observations ranging from 30 to 184), but Chl data frequently were unable to meet the criterion. Six images had insufficient corrected Chl and nine images had insufficient uncorrected Chl. As such, the historical analysis could be completed only for SDT.

The average spectral brightness value for each AOI was paired with ground observations of SDT made within 7 days of each satellite image acquisition date. Pearson correlation coefficients and step-wise multiple regression analysis were used to identify the Landsat spectral bands most correlated with SDT for several satellite images. Regression assumptions were tested and a sensitivity analysis was conducted to evaluate the effects of model forms.

### 3. Procedure development

#### 3.1. Selection of ground data

To address the issue of the time window, we conducted regression analyses for a Landsat TM image from July 29, 1995. Regression statistics were calculated for the relationship between the natural log of SDT and TM3/TM1, successively increasing the time window between satellite overpass and ground observation from same-day ground observations to \( \pm 7 \) days (Fig. 1). The TM3/TM1 ratio was chosen because previous investigators found it to be a strong predictor of SDT (Cox et al., 1998; Lathrop, 1992), and this was confirmed by our analysis. As expected, increasing the time window decreased \( r^2 \) (.86, .74, and .72 for \( \pm 1 \), \( \pm 3 \), and \( \pm 7 \) days, respectively) and increased the standard error of estimate (SEE) (.245, .358, and .361). The loss of strength in the correlation was partially offset by cumulative increases in the number of ground observations as the window was enlarged. The analysis was repeated with random selection of a constant number of ground observations (n = 20) using various time windows. Keeping the number of samples constant emphasized the loss of correla-

### Table 1

<table>
<thead>
<tr>
<th>Image date</th>
<th>Number of ground observations of SDT</th>
<th>Number of ground observations of Chl Corrected*</th>
<th>Number of ground observations of Chl Uncorrected*</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/03/73</td>
<td>30</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8/07/75</td>
<td>64</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>9/06/83</td>
<td>64</td>
<td>5</td>
<td>13</td>
</tr>
<tr>
<td>8/21/86</td>
<td>93</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>8/26/88</td>
<td>82</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>6/16/91</td>
<td>109</td>
<td>22</td>
<td>7</td>
</tr>
<tr>
<td>7/18/91</td>
<td>132</td>
<td>48</td>
<td>8</td>
</tr>
<tr>
<td>9/04/91</td>
<td>110</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>9/20/91</td>
<td>90</td>
<td>26</td>
<td>15</td>
</tr>
<tr>
<td>8/24/93</td>
<td>138</td>
<td>28</td>
<td>31</td>
</tr>
<tr>
<td>7/29/95</td>
<td>170</td>
<td>33</td>
<td>40</td>
</tr>
<tr>
<td>7/15/96</td>
<td>184</td>
<td>41</td>
<td>46</td>
</tr>
<tr>
<td>9/07/98</td>
<td>134</td>
<td>13</td>
<td>52</td>
</tr>
</tbody>
</table>

*a Corrected chlorophyll (total chlorophyll minus pheophytin, a degradation byproduct of chlorophyll) is a better representation of living algal biomass, but given the large bandwidths of satellite sensors, uncorrected chlorophyll probably is more closely correlated with satellite data.
tion strength, with $r^2$ decreasing from .85 to .57 and SEE increasing from 0.267 to 0.388. Based on this analysis, we concluded that a window of ±1 day between satellite and ground observations was optimal for our data set. Limiting the calibration data to ±1 day of image acquisition provided the highest correlation and still provided at least 20 ground observations for calibration. For regions with fewer ground observations, a wider window may be necessary.

3.2. Extraction of image brightness data

Another factor that may influence the strength of the relationship between satellite and ground observation data is the number of pixels included in the AOI for each lake. It could be argued that if ground observations were strictly contemporaneous with the satellite image, a small cluster of pixels containing the location of the ground observation would provide the best correlation between satellite- and ground-based measurements. While it is possible to use GPS to further increase the accuracy of the ground station, this is usually impractical for citizen volunteer monitoring programs and impossible for historical studies. The procedure followed for this study (described in Section 2) was to use several pixels from the profundal zone of each lake, similar to the method recommended by Lillesand et al. (1983). This approach was tested in two ways: (1) by analysis of prediction residuals and (2) by experimenting with the size of the AOIs for a consistent subset of lakes. Analysis of residuals from the regression for each image showed that there was no correlation between the residual and the number of pixels in any given AOI. This suggests that SDT is predicted equally well for lakes having at least nine pixels (the minimum criteria for this study) as for lakes with more pixels. In a separate analysis, the AOIs were varied from 1 to 250 pixels for a set of 20 large lakes. Successive regression analyses were performed on each of these data sets. Correlation strength, as measured by $r^2$ and SEE, improved significantly as the AOI was increased from one to nine pixels, but the benefits of increasing the size of the AOI were marginal beyond 25 pixels.

3.3. Regression model development

Because the general approach used for modeling relied on regression analysis, it was important to test the basic assumptions of this technique: (1) regression residuals have constant variance, especially with respect to the independent variable(s); (2) residuals are independent; and (3) residuals are normally distributed. Analysis of the residuals showed that these assumptions were met by regressing log-transformed SDT versus untransformed TM3/TM1. This was found to be generally true for correlations between SDT and other TM bands as well. A few previous studies have used nonlinear power models ($y = ax^b$) to address the curvilinear behavior of this relationship (e.g., Cox et al., 1998; Lathrop, 1992). Although a power model provided a strong correlation, residuals from it were not normally distributed. In contrast, a semilog equation met the model assumptions, as also found by Pattiaratchi, Lavery, Wyllie, and Hick (1994). Therefore, subsequent regression analyses used log-transformed SDT.

A Pearson correlation matrix was developed to examine the relative strength of correlation between SDT and various Landsat TM bands and band ratios. Previous investigations (Cox et al., 1998; Lathrop, 1992; Lavery et al., 1993) have suggested that band ratios provide useful relationships, and several ratios were included in the correlation matrix. A summary of absolute Pearson correlation coefficients for all 10 TM images (Fig. 2) shows that highest coefficients generally were found for TM3 and TM3/TM1. Other variables also were strongly correlated with SDT: TM4, TM2, TM2/TM1, and TM3/TM2. TM1 was moderately correlated to SDT, but TM5, TM6, TM7, and TM4/TM3 had little or no correlation. SDT thus was strongly correlated with bright-
ness in the visible and near infrared regions of the spectrum (TM1, TM2, TM3, and TM4), as expected, and not well correlated with brightness in the middle to far infrared regions (TM5, TM6, and TM7).

Backward, step-wise multiple regression was performed using all seven Landsat TM bands on each image. The threshold for factor removal was set at $P > .10$. Results indicate that the most frequently significant factors were TM1, TM2, TM3, and TM4 (Table 2). As with the analysis of Pearson correlation coefficients, TM5, TM6, and TM7 were not significant factors and were dropped from further analysis. With two exceptions, these results agree with findings of the pair-wise correlation analysis. First, TM1 was a more important factor in the multiple regressions than suggested by the pair-wise correlation. Although TM1 is not highly correlated to SDT by itself, it is a significant component in multiple regression models. Second, TM2 was less important in the multiple regressions than suggested by the pair-wise correlations. This probably reflects the fact that much of the variance (i.e., information) in TM2 is contained in either TM1 or TM3. The Pearson correlation coefficient and step-wise regression analyses were used to select a subset of the Landsat bands for subsequent use in a multiple regression model.

Further multiple regression analysis focused on various combinations of TM1, TM3, and TM4 along with TM3/TM1. The analysis was repeated for each of the 10 study images (Fig. 3). TM3 has the strongest single-band relationship with SDT (average adjusted $r^2 = .67$). Use of TM3/TM1 improved the mean adjusted $r^2$ to .73. Adding a term for TM1 to the TM3/TM1 ratio improved the mean adjusted $r^2$ to .75, and adding a term for TM4 to that relationship increased the mean adjusted $r^2$ to .77. The mean SEE of the estimate decreased with increasing mean $r^2$ from .339 to .280 (Fig. 3). Additional analysis of the regression equation with the combination of TM3/TM1 and TM1 showed that this model had a slight tendency to underpredict SDT at high SDT values but that an inverse transformation to TM1/TM3 corrected this deficiency.

The regression analysis performed on the Landsat TM images was repeated for four Landsat MSS images (Table 3). By including MSS imagery in the study, we were able to extend our retrospective trend analysis back to the early 1970s, while TM data extend only to the early 1980s. As with the TM image data, a semilog relationship between log-transformed SDT and MSS brightness was found to meet the

Table 2

<table>
<thead>
<tr>
<th>Image Date</th>
<th>n</th>
<th>$r^2$</th>
<th>Two-tailed significance level ($P$ value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8/21/86</td>
<td>21</td>
<td>.86</td>
<td>** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>8/26/88</td>
<td>22</td>
<td>.95</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>6/16/91</td>
<td>53</td>
<td>.76</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>7/18/91</td>
<td>31</td>
<td>.82</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>9/04/91</td>
<td>26</td>
<td>.83</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>9/20/91</td>
<td>20</td>
<td>.83</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>8/24/93</td>
<td>38</td>
<td>.77</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>7/29/95</td>
<td>43</td>
<td>.87</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>7/15/96</td>
<td>50</td>
<td>.88</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
<tr>
<td>9/07/98</td>
<td>30</td>
<td>.82</td>
<td>** *** *** *** *** *** *** *** *** ***</td>
</tr>
</tbody>
</table>

* Time window of ±1 day between satellite overpass and ground observations.
** Significance level of $P < .10$.
*** Significance level of $P < .05$.
**** Significance level of $P < .01$.

![Fig. 2. Box–Whisker plot of Pearson correlation coefficients between ln(SDT) and various Landsat TM bands and band ratios.](image_url)

![Fig. 3. Mean and 95% confidence intervals for $r^2$ and SEE for regression equations between ln(SDT) and combinations of TM bands and band ratios.](image_url)
3.4. Radiometric calibration

The brightness values of the pixels in a satellite image are affected by Sun angle, atmospheric interference, changes in detector response, and numerous other factors. If radiometric correction techniques can account for these factors, then the coefficients for Eqs. (1) and (2) should be more consistent, and one set of coefficients may apply for different images across time and space. However, radiometric calibration may not correct for all factors affecting these coefficients.

An example of a nonatmospheric phenomenon that might affect the coefficient values is wind-induced wave action, which affects surface reflection. Koponen et al. (2001) discussed the effect of varying atmospheric effects and surface reflection on the degradation of accuracy in remote estimation of lake water quality. They concluded that atmospheric correction and the use of a bidirectional scattering model reduces the bias in estimating Chl concentration. In the case where the images cannot be suitably corrected, each image needs to be calibrated to in situ measurements made near concurrently with the collection of the remote sensing data.

The most rigorous method of radiometric calibration involves the use of radiative transfer models to produce an absolute correction. However, some data required to perform such a calibration are unavailable for these historic images. We tested simple radiometric correction techniques such as dark pixel subtraction and Sun angle correction and found these insufficient. This conclusion agrees with a similar analysis by MacFarlane and Robinson (1984).

We also used relative (image-to-image) normalization by applying regression models to pseudoinvariant ground targets (targets that have stable reflectance over time) to normalize multitemporal images to a single reference scene (Coppin & Bauer, 1994). According to Eckhardt, Verdin, and Lyford (1990), suitable normalization targets should have (1) similar elevation to the rest of the scene, (2) minimal vegetation, (3) a relatively flat surface, and (4) constant pattern or general appearance over time. In addition, they should be sufficiently large that a multiple-pixel sample, free of mixed pixels, can be identified. Furthermore, normalization targets should be selected to represent a broad range of spectral-radiometric responses. Based on these criteria, we identified seven normalization targets in the TCMA: Hubert H. Humphrey Metrodome, tarmac of the Minneapolis-St. Paul International Airport, parking lot of the former Met Center, Gillette Building, and three lakes with high transparency (thus very low reflectance)—Square Lake, White

### Table 3

Summary of step-wise multiple regressions between natural log of SDT and Landsat MSS bands

<table>
<thead>
<tr>
<th>Date</th>
<th>n</th>
<th>$r^2$</th>
<th>Significance level (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/3/73</td>
<td>21</td>
<td>0.83</td>
<td>***</td>
</tr>
<tr>
<td>8/7/73</td>
<td>49</td>
<td>0.75</td>
<td>***</td>
</tr>
<tr>
<td>9/6/83</td>
<td>46</td>
<td>0.72</td>
<td>***</td>
</tr>
<tr>
<td>9/4/91</td>
<td>25</td>
<td>0.83</td>
<td>***</td>
</tr>
</tbody>
</table>

* Time window of ±1 day between satellite overpass and ground observations for all images except July 3, 1973, which used a 4-day window to allow for sufficient ground data.

** Significance level of $P<.05$.

*** Significant level of $P<.01$.

...
Bear Lake, and Lower Lake Minnetonka. The brightness levels from these targets typically range 51–254 for TM1 and 10–211 for TM3 out of a possible range of 0–255 (Fig. 5). Average digital brightness levels for each normalization target were computed for each of the TM bands. Target responses from nine images were paired with the response from a reference image (August 24, 1993). Correlations for the calibration relationships were very strong ($r^2$ ranging from .96 to 1.00). Gain correction was relatively small, ranging from $-23.3\%$ to $+13.5\%$. Offset corrections were most significant for the TM1 band, with four images requiring corrections of more than 10 digital brightness levels. This was expected, as TM1 is strongly affected by atmospheric scattering. The calibration relationships were used to rescale the original digital brightness levels. The raw brightness data yielded a wide range of regression lines for the TM images (Fig. 6a), but the regression lines partially coalesced when the calibrated image data were used (Fig. 6b). This results in a more consistent relationship between satellite and ground observations. However, even with this calibration effort, the relative variabilities for the coefficients $a$, $b$, and $c$ in Eq. (1) for the 10 TM images were $\pm 14\%$, $\pm 180\%$, and $\pm 200\%$, respectively. The variability in the coefficient for the TM3/TM1 ratio is small compared with the variability for TM1, suggesting that this ratio is more consistent over time and for different atmospheric conditions.

Corrected brightness data for multiple Landsat TM images were used to derive a single standard relationship using multiple regression analysis. Including only near-anniversary images reduced the influence of changing Sun angle. Using only images with acquisition dates from August 21 to September 7, we found an $r^2$ of .55 and SEE of 0.465. This relationship is weaker than those for the individual images, which had $r^2>.7$. However, most of the observations that deviate significantly from the composite relationship occurred during 1988. When considered alone, the correlation for the 1988 image was strong, but its slope is different from all the other years. In addition, 1988 was an unusually hot, dry year with frequent occurrence of brush and forest fires in the region and likely had significantly different atmospheric conditions. Water clarity was lower than usual for most lakes in 1988, but how this might affect the slope of the relationship is unclear. When this image was excluded from the composite relationship, the strength of the composite relationship increased, resulting in an $r^2$ of .68 and SEE of 0.356. Although this approaches the range of strengths of the individual image relationships, the image-to-image calibration is insufficient to bring the regression models for the individual images into alignment as a standard model. We suggest two possible explanations for this. First, we selected lakes as dark calibration targets. These targets may be relatively invariant with respect to the entire

![Fig. 5](image.png)

Fig. 5. Range of brightness for TM1 and TM3 for radiometric calibration targets expressed as a percent of total possible range for eight-bit Landsat TM data (0–255).

![Fig. 6](image.png)

Fig. 6. Comparison of relationships between digital brightness values for Landsat TM3 and ln(SDT) before (a) and after (b) relative atmospheric correction for 10 images.
radiometric range, but they still vary. The brightness values for lakes occur at the lower end of the radiometric scale. Thus, even subtle shifts in the correction gain and bias may result in significant errors. Second, the correction was applied as a constant over the entire image. Pixel by pixel correction might provide a significant enhancement. However, the data needed to correct historic imagery on a pixel by pixel basis are not available. Improved atmospheric correction techniques are likely to be important in the effort to develop a reliable standard equation to estimate lake water clarity from satellite data. However, even without a standard model, this procedure still provides a wealth of information about lake conditions at relatively low cost.

3.5. Procedure application

Although a standard relationship between SDT and satellite measured spectral-radiometric data still is an elusive goal, the procedures described in this paper can be applied today as long as each image is calibrated to contemporaneous ground observations. Regression models for each Landsat TM image within the late summer index period had $r^2$ values ranging from .71 to .92 when Eq. (1) was applied. The models were based on ground observations from only a small subset of TCMA lakes, but they can be applied to estimate SDT for all lakes within the study area for which we delineated an AOI. This results in a much more complete regional spatial sampling of lake clarity, thus allowing for mapping and analysis of spatial trends (Fig. 7) as well as statistical analyses (Fig. 8).

Satellite-estimated SDT values for all study lakes and all 13 image dates were seasonally adjusted according to a procedure described in Kloiber et al. (2001). Figure 7. Distribution of lakes in TCMA with inset image of satellite-based SDT for lakes in northern Ramsey and Washington counties on September 7, 1998. Inset area contains 130 lakes >10 ha with water clarity ranging from 0.3 to 4.2 m on this date.

Satellite-estimated SDT values for all study lakes and all 13 image dates were seasonally adjusted according to a procedure described in Kloiber et al. (2001). Figure 8. Frequency distribution of SDT for TCMA lakes for September 7, 1998 compared with threshold values for user-perceived swimming impairment (MPCA 1988). The median surveyed user-perceived swimming impairment at SDT = 1.4 m.
estimated growing season mean values of SDT for TCMA lakes greater than 10 ha in area (n ≈ 500) ranged from <0.5 to about 5.0 m. Dividing the lakes into classes based on 1-m SDT intervals, we found that the most common clarity class is 1.0–2.0 m. The typical growing season mean SDT for the region ranged from 1.2 to 1.7 m over the study period (1973–1998).

4. Discussion

An important consideration when gauging the potential to develop a single standard equation is consistency of the form of the model from place to place and across time. Do we need different combinations of bands and band ratios to obtain predictive relationships for each image or will one set work over a range of time and place? Results of this study show that the ratio TM1/TM3 plus TM1 or the ratio MSS1/ MSS2 plus MSS1 (Eqs. (1) and (2)) provided strong predictive relationships for multiple images over a 25-year period. Several other investigators had success with similar relationships. Lathrop (1992) found that the TM3/TM1 ratio provided a strong relationship to SDT based on an analysis of two Wyoming lakes (r²=.83) and Green Bay of Lake Michigan (r²=.94). Lavery et al. (1993) used TM3 plus TM1/TM3 to predict SDT (r²=.81) for an estuary near Perth, Australia. Pattiaratchi et al. (1994) used TM3 to predict the natural logarithm of SDT (r²=.78) in Cockburn Sound, which is also near Perth. Cox et al. (1998) used TM3/ TM1 to predict SDT (r²=.82) for several reservoirs in North and South Carolina. Although these studies confirm the consistency of the relationship between SDT and Landsat TM data, each study involved only a few surface waters. In a regional analysis of 34 lakes near Alexandria, MN using the methods described here, Day (2000) found that Eq. (1) yielded an r²=.78.

Although the spectral bands of Landsat TM (and by extension MSS) are broad enough to encompass a mixture of spectrally opposing absorption and scattering features (Dekker, Malthus, Wijnen, & Seyhan, 1992), we still were able to develop a consistent and reliable relationship between satellite imagery and water clarity. The fact that several investigators have developed similar relationships between SDT and Landsat TM imagery across a broad geographic range (Australia, Wyoming, Minnesota, Wisconsin, and the southeastern US) and for such widely varying surface waters (estuaries, reservoirs, and inland lakes) is more than coincidence. There is an underlying physical basis for this result. Using spectroradiometers to collect hyperspectral reflectance data, several investigators have identified consistent spectral features of lakes related to optically active substances, including photosynthetic pigments and inorganic suspended sediment (Gitelson et al., 2000; Han, Rundquist, Liu, Fraser, & Schalles, 1994; Mittenzwey, Ullrich, Gitelson, & Kondratiev, 1992). Gitelson et al. (2000) described several important spectral features within the range corresponding to the visible and near infrared bands of Landsat TM and MSS. These features include a broad peak in the green range (near 550–600 nm) corresponding to scattering from algae and suspended sediment, an absorbance minimum near 670 nm due to Chl, and a peak at 700 nm that corresponds to scattering by suspended particles. TM3 (630–690 nm) encompasses the Chl absorbance minimum as well as a significant part of the scattering peaks on either side of this minimum. Although these apparently contradictory features all may influence the value of TM3, the net effect is clear: TM3 brightness increases as water clarity decreases. Undoubtedly, this is because the effect of increased scattering by suspended particles has a broad effect across much of the visible and near infrared portion of the spectrum from about 500 to about 850 nm. In fact, brightness in TM2, TM3, and TM4 all increase as water clarity decreases. This scattering effect overwhelms the subtler influence of other features such as the Chl minimum. Although we clearly can establish a relationship between water clarity and Landsat data, this should not be construed to imply that such relationships can be developed for other related water quality variables such as chlorophyll.

The use of a consistent combination of bands and band ratios helps make the analysis of different images comparable and is an important step towards standardizing this method. The major remaining impediment to developing a standard equation for assessing SDT based on satellite imagery seems to be the lack of adequate and readily applicable methods for atmospheric correction. The regression method we used for atmospheric correction was easy to implement but yielded only partial success. Atmospheric models may provide better results, but they are complicated to implement and require input data that are not always available. Jensen (1989) suggested that for remote sensing of biophysical parameters, such as SDT, in situ calibration may be necessary. The advent of new sensors such as MODIS, which have detectors in spectral regions that may provide direct measurements of atmospheric scattering and absorption (Kaufman, Herring, Ranson, & Collatz, 1998), holds potential that better methods for atmospheric correction can be developed.

Even without a standard model, the procedure described in this paper provides a wealth of information on lake conditions at relatively low cost. Through this effort, SDT was estimated for about 500 lakes in each of 13 image dates over a 25-year period. These data can support an array of important analyses. For example, combining these data with user perception survey data (Heiskary & Wilson, 1988) can aid in assessing the degree of lake impairment for the region. SDT estimates for the most recent image (September 7, 1998) show that the typical lake user likely would consider only half the TCMA lakes as supporting swimming (Fig. 8). The results also can be used to create thematic maps of regional lake clarity (Fig. 7). In addition, routine application of the procedure on an annual basis may enhance our ability
to detect temporal trends at the regional, subregional, and individual lake levels (Kloiber, Brezonik, & Bauer (2002)).

Acknowledgements

This work was funded by the (Twin Cities) Metropolitan Council through the Twin Cities Water Quality Initiative grant program, with additional support from the University of Minnesota Water Resource Center and Agricultural Experiment Station. We gratefully acknowledge the assistance of Erin Day who helped process the image data.

References


