1-1-2013

TB207: A Manual for Remote Sensing of Maine Lake Clarity

Ian M. McCullough

Cynthia S. Loftin

Steven A. Sader

THE UNIVERSITY OF MAINE

Maine Agricultural and Forest Experiment Station

Technical Bulletin 207 • January 2013

ISSN: 1070-1524

A Manual for Remote Sensing of Maine Lake Clarity

Ian M. McCullough

Cynthia S. Loftin

and

Steven A. Sader

College of Natural Sciences, Forestry, & Agriculture

Maine Agricultural and Forest Experiment Station Technical Bulletin 207 January 2013

A Manual for Remote Sensing of Maine Lake Clarity

Ian M. McCullough

Department of Wildlife Ecology

Cynthia S. Loftin

U.S. Geological Survey, Maine Cooperative Fish and Wildlife Research Unit

Department of Wildlife Ecology

and

Steven A. Sader

School of Forest Resources

College of Natural Sciences, Forestry & Agriculture
5782 Winslow Hall
University of Maine
Orono, ME 04469-5782

ACKNOWLEDGMENTS

This report and the research it contains were made possible through generous financial support from the University of Maine's Department of Wildlife Ecology, the United States Geological Survey (USGS), Maine Cooperative Fish and Wildlife Research Unit, and the Maine Department of Environmental Protection (MDEP). Bathymetric and spatial watershed data were provided by MDEP, and we appreciate the efforts of Linda Bacon, Doug Suitor, and Leslie Latt. Other spatial data were provided by the Maine Office of GIS. Remote estimation of lake clarity depends on field data; Secchi data used in this report were collected by citizen scientists in the Maine Volunteer Lake Monitoring Program in conjunction with MDEP. Comments provided by Aram Calhoun and William Halteman of the University of Maine improved the methods described in this report. Reviews by Manuel Gimond of Colby College and Sarah Nelson of the University of Maine improved this manual. Cover photo was taken by Jonathan Schilling. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the U.S. government.

Contents

Figures

Tables

SUMMARY

The purpose of this manual is to support use of satellite-based remote sensing for statewide lake water-quality monitoring in Maine. We describe step-by-step methods that combine Landsat and MODIS satellite data with field-collected Secchi disk data for statewide assessment of lake water clarity. Landsat can be simultaneously used to assess Maine's more than 1,000 lakes ≥ 8 ha, whereas MODIS can be used to assess a maximum of 364 lakes ≥ 100 ha (250-m image resolution) or 83 lakes ≥ 400 ha (500-m image resolution). It was our intention that the detailed instructions in this manual would assist and expedite implementation of our methods. The methods described here were developed in a master's thesis resulting in several peer-reviewed journal manuscripts (McCullough et al. 2012a, b, in review a, b). We assumed readers had reasonable knowledge of lake ecology, statistics, and geographic information systems (GIS). Although our methods were specifically developed for Maine, other states or non-Maine agencies may find these methods as useful starting points in developing their own protocols for regional remote lake monitoring.

INTRODUCTION

Long-term maintenance or improvement of water quality is essential for the continuation of diverse recreational, economic, and cultural activities associated with lakes. Increased lake water quality is positively correlated with lakefront property value in Maine (Michael et al. 1996; Boyle et al. 1999) and New Hampshire (Gibbs et al. 2002) and enhances user perception of lake health in Minnesota (Heiskary and Walker 1988). Lakes also provide important habitat for a variety of plants and animals, including economically important fishes and waterfowl.

Maine contains more than 5,500 lakes > 1 ha in size and more inland surface waters than any state east of the Great Lakes (Davis et al. 1978). The sheer number of lakes in Maine is both a blessing and a burden. Maine is fortunate to hold so many of these natural resources, but monitoring and maintaining the health of thousands of lakes, many of which are remote and inaccessible, is a difficult and expensive task.

Water clarity (or transparency) is an ideal metric of regional lake-water quality. Often measured in terms of Secchi disk depth (SDD), water clarity is strongly correlated with other measurements of water quality including chlorophyll-a, total phosphorus, and trophic status (Carlson 1977). Unlike these variables, however, water clarity can be easily measured in the field with minimal equipment and no chemicals. Therefore, SDD is arguably the most efficient metric of water quality when attempting to assess a large area. Many of the water-clarity data collected in Maine are gathered by lakeshore residents who volunteer with the Maine Volunteer Lake Monitoring Program (VLMP). Numerous other states have similar organizations, but Maine's is the longest running in the United States. The Maine Department of Environmental Protection (MDEP) initiated statewide monitoring of water clarity in 1971 jointly with the VLMP. Maine continues to rely greatly on volunteers for monitoring water quality. These citizen scientists are not only capable of making substantial contributions to our knowledge of Maine's lakes through the construction of long-term datasets, but they also are important stakeholders in the issue of lake-water quality. The continued involvement of these stakeholders in lake monitoring is an integral component of successful long-term management of lakes.

Owing to the relative ease of gathering data on lake clarity, we have considerably more data covering a greater geographic extent on clarity than other water-quality metrics. Average SDD in Maine has consistently remained 4 to 6 m since 1971, with a historical average of 5.28 m during the period from 1971 to 2011. Although Maine lakes are generally considered to be in good condition, there is concern that field sampling is spatially biased and may not constitute a representative sample of water quality statewide. Assessed lakes are concentrated in accessible areas (particularly in southern Maine) and near roads. Remote lakes are rarely or never sampled, and it is difficult to make definitive judgments about the water quality of Maine's lakes with incomplete data. Probability-based, random field sampling is necessary to avoid false conclusions derived from biased field sampling (Wagner et al. 2007). Remote sensing allows simultaneous assessment of hundreds of lakes and can greatly reduce costs associated with traditional field methods. By combining field data and satellite imagery to model the statistical relationship (e.g., regression) between water clarity and satellite-measured reflectance, we can then estimate SDD of unsampled lakes. These analyses can also be performed retrospectively with archived satellite imagery and historical field data to assess statewide changes in water quality over time.

This manual describes methods for remote monitoring of water clarity with Landsat Thematic Mapper (TM) and Moderate-Resolution Imaging Spectroradiometer (MODIS) satellite imagery. Both satellite platforms have advantages and disadvantages related to spatial resolution (pixel size), image-capture frequency, and amounts of image processing necessary; however, using both Landsat and MODIS data as part of a flexible program of remote monitoring provides benefits offered by both platforms and maximizes remote data collection.

SATELLITE BACKGROUND

Landsat Thematic Mapper (TM)

The Landsat program spans seven U.S. satellites, the first of which was launched in 1972. The extensive image archive makes Landsat an important source of historical data in many areas of monitoring and research in addition to water quality. As of the date of this publication, two Landsat satellites are currently in orbit; however, the likelihood of either providing high-quality imagery in the future remains in doubt. Landsat 5, launched in 1984, experienced an amplifier failure in November 2011 and was suspended for 180 days in an attempt to restore operation. Although the Thematic Mapper (TM) sensor was not revived, the Multispectral Scanner (MSS; 57-m resolution), powered-down since 1995, was turned back on. Despite the revival of the MSS, Landsat 5 has long exceeded its intended lifespan and is not a reliable source of future long-term data. Landsat 7 was launched in 1999 and continues to capture imagery, but image quality has been compromised since 2003 by the failure of the scan-line corrector (SLC), an instrument that corrects for the forward motion of the satellite. As a result, post-2002 Landsat 7 images (SLC-off) contain rows of null values, which complicate estimation of remote lake clarity. Despite these issues, SLC-off imagery can still be used for monitoring remote lake clarity (Olmanson et al. 2008). The expected 2013 launch of the Landsat Data Continuity Mission (LDCM; Landsat 8), if successful, ensures future availability of Landsat data for remote lake monitoring. Landsat 9 is in the preliminary planning stages as of this writing.

Landsat imagery can be freely downloaded from the USGS Global Visualization Viewer (glovis.usgs.gov). Images are indexed by path and row. Path 11 (rows 27–29) and path 12 (rows 27–30) capture most of Maine (Figure 1). Landsat images cover an extent approximately 185 km wide and have a 30-m resolution (Table 1), which is considered moderate among other satellites. Both Landsat satellites contain three visible and four infrared bands; Landsat 7 also includes a 15-m panchromatic band and an additional thermal infrared band. Landsat does not measure UV reflectance. Bands 1 (visible blue) and 3 (visible red) are strongly correlated with lake water clarity (Kloiber et al. 2002; Chipman et al. 2004; Olmanson et al. 2008; McCullough et al. 2012a). Images are captured every 16 days, and this relative infrequency of image capture is one of the greatest limitations of Landsat data when short windows of time (e.g., month of August) are of interest. Landsat images do not receive atmospheric precorrections, which must be performed by the user if desired.

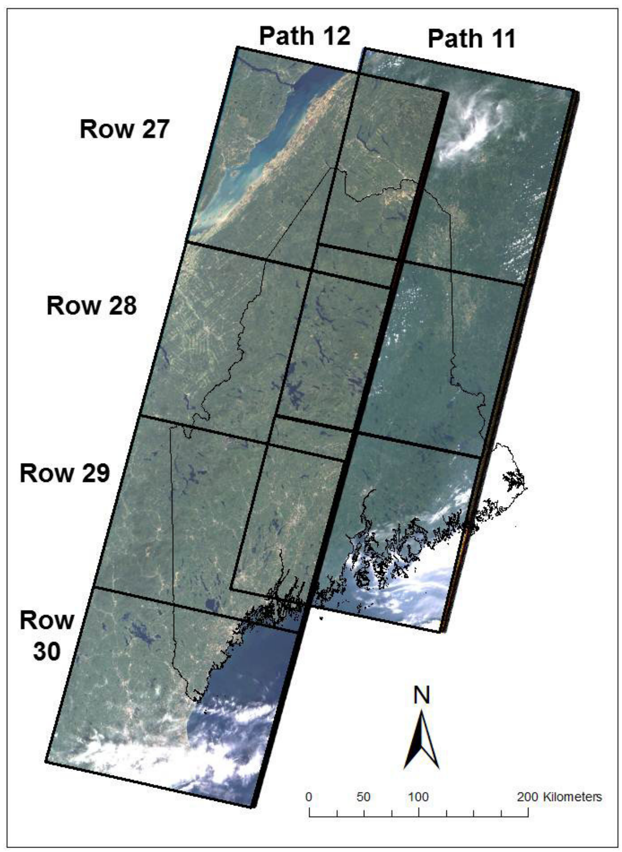

Figure 1. Landsat TM paths 11 and 12 and individual scenes over Maine. Images are true color composite (RGB 3, 2, 1).

Table 1. Comparison of Landsat and MODIS specifications

Specifications	Landsat	MODIS
Spectral resolution	7 bands	36 bands
Pixel size	30 m	250, 500 or 1000 m
Scene width	185 km	2330 km
Image frequency	16 days	Twice daily
Corrections	None	Surface reflectance
Cost	Free	Free

Moderate-Resolution Imaging Spectroradiometer (MODIS)

Moderate-Resolution Imaging Spectroradiometer (MODIS) sits aboard two NASA satellites: Terra, launched in 1999, and Aqua, launched in 2002. Each satellite captures daily images of the entire Earth's surface, yielding two images of a specific location per day. MODIS images cover a large extent (approximately 2,300 km wide); these images are often used by the weather media to illustrate the track of hurricanes and other weather patterns. MODIS contains 36 bands at various resolutions: bands 1–2 (250 m), bands 3–7 (500 m) and bands 8–36 (1,000 m) (Table 1). NASA creates numerous products with MODIS-based images that are useful for a wide variety of applications. Level 1B daily surface reflectance 500-m data from either Aqua (MYD09GA) or Terra (MOD09GA) contain the spectral sensitivity best suited for remote lake monitoring. Although MODIS data are available at 250-m resolution, 250-m data do not contain the visible blue band. MODIS 250- and 500-m products contain a preconversion to surface reflectance, which theoretically precludes additional atmospheric corrections, but these corrections were designed for analysis of land, not aquatic features. Because images are captured twice a day, users have a greater number of images from which to select, increasing the chances of acquiring clear imagery. This high frequency of image acquisition has the additional benefit of reducing image-processing requirements compared to those needed when using Landsat data. The twice-daily image capture is a major advantage of MODIS over Landsat. MODIS images also can be freely downloaded from the USGS Global Visualization Viewer (glovis.usgs.gov).

PART 1—APPLICATION OF LANDSAT TM DATA FOR REMOTE WATER-CLARITY MONITORING

Uncited methods, findings, and explanations in Part 1 are based on McCullough et al. (2012a, in review a).

General Methods

How many lakes in Maine can be assessed?

Landsat TM allows remote monitoring of approximately 1,500 lakes > 8 ha in size. This lake size cutoff was selected based on methods used in a similar study of Minnesota lakes (Olmanson et al. 2008). For accurate remote monitoring, lakes must be sufficiently large to contain several water-only pixels in deep areas where SDD is measured in the field. Path 11 contains 1,121 lakes and path 12 contains 1,090 lakes > 8 ha that are eligible for remote monitoring; the path overlap region contains 570 eligible lakes. Lists of eligible lakes can be found here under supplementary material (www.coopunits.org/Maine/People/Cyndy_Loftin/Publications). Some lakes > 8 ha were eliminated owing to narrow, convoluted shorelines and insufficient quantity of water-only pixels.

Selecting images of sufficient quality

All Landsat images can be previewed online prior to download. Images should be free of clouds or nearly so, but regular monitoring of lakes may necessitate using images with some clouds, as long as the clouds do not cover lakes of interest or ground control points used in radiometric normalization. The online interface (glovis.usgs.gov) (hereafter referred to as GloVis) displays a percentage cloud cover value, but this number is relatively unreliable because thin clouds are inadequately accounted for and clouds covering areas outside the area of interest are considered in this value. In general, images containing < 10% cloud cover are acceptable. We included some examples of images we have used in our analyses (Table 2). Preliminary data assessment includes addressing the following questions:

1. Are there sufficient image data of suitable quality available to justify image processing and model construction?
2. Will clouds cause inaccurate predictions?

Generally, cloudier images yield less lake data. Thick, plainly visible clouds are easily extracted and will not

Table 2. Landsat imagery used for remote estimation of lake clarity

Path[a]	Rows	Acquisition Date	% Clouds	Satellite
12	27–30	8/30/2010	0	Landsat 5
12	27–30	9/14/2004	0	Landsat 5
12	27–30	9/1/1999	0	Landsat 5
12	27–30	9/6/1995	0	Landsat 5
12	27–30	9/8/1990	0	Landsat 5
11	28–29	9/5/2009	6	Landsat 5
11	27–29	8/9/2005	8	Landsat 5
11	27–29	8/9/2002	0	Landsat 7
11	27–29	8/14/1995	2	Landsat 5

[a]Path 11, row 27 scene omitted due to cloud cover on 9/5/2009

compromise model predictions, but they cast dark shadows, which may not be as easily identified or extracted, thereby influencing model results (see Technical Methods). Thin clouds and haze are more difficult to detect and do not necessarily obscure lakes completely. Thin clouds and haze cause atmospheric scattering and may make images appear brighter, yielding underestimates of SDD.

What image dates are preferred?

The eventual goal of remote monitoring of lake water clarity is to detect changes in water clarity over space and time. Therefore, images used in successive analyses should have been captured at roughly the same time of year to avoid error associated with changing lake conditions that reflect seasonal change (e.g., intra-annual algal community development or lake stratification). The late summer stable period of July 15 to September 15, with preference for August, is the optimal window for monitoring remote lake clarity owing to seasonal lake stability and seasonally low clarity conditions (Stadelmann et al. 2001; Olmanson et al. 2008). The preference for August is strongly emphasized. Dimictic lakes in northern Maine may undergo fall turnover (vertical mixing of the water column) as early as late August, though early to mid-September is more common (Davis et al. 1978). Subsequently, mid- to late August imagery (August 10–31) captures annual peaks in algal growth and therefore provides the most direct measurement of lake productivity, whereas July and September images may capture periods before and after this peak. Early September images should be used with caution. We found that images captured 9/8/1990, 9/6/1995, and 9/14/2004, contained lakes that showed evidence of turnover, whereas an image captured 9/5/2009 did not indicate turnover. Turnover dates fluctuate annually, and September 5 is not an absolute cutoff, even though we used data captured on this date in 2009. Satellite-derived or field-collected SDD values that are considerably shallower than in previous summers are strong evidence of lake turnover where algae have been mixed throughout the water column. Climate change may lengthen future growing seasons in lakes, though annual fluctuations in turnover dates likely will continue nonetheless.

Can images from different dates be combined and analyzed together?

Technically, this is possible, but it is impractical and computationally intensive. Combining images across multiple dates would require separate model calibrations for each date because lake and atmospheric conditions will likely vary by date. Availability of calibration data may be limited for individual Landsat scenes, especially north of row 29, owing to the remoteness of these areas. In addition, radiometric normalization (see Part 1: Technical Methods) is difficult if images captured on different dates contain haze. An appealing advantage of using images from only a single date is the ability to capture a one-day snapshot of a large portion of Maine (path 11 or 12). This is only possible for a single Landsat path (11 or 12) because the paths are captured on separate dates.

Radiometric normalization

Although clouds are relatively easy to identify and remove, haze presents a greater problem. Haze is difficult to identify systematically and generally is not uniform throughout a large geographic area. Haze particles in the atmosphere increase Rayleigh scattering (particularly at the shorter wavelengths of TM band 1), which can influence satellite radiometric responses. Radiometric normalization is a standard technique used to minimize haze interference. The idea is to designate a clear set of images as the reference or master images and to scale all other images radiometrically to the reference images. This can be done by selecting a group of large (bright and dark) ground features, known as pseudo-invariant ground targets, that are presumed to be unchanged during the study period. If unchanged ground features appear different in two images, the difference is attributed to haze. Radiometric normalization corrects this difference. Selection of an adequate number of sufficiently large ground features can be difficult in relatively remote areas that lack large developed features (e.g., airstrips, stadium roofs). We created a GIS points layer of ground targets that can be used for future normalizations; however, the point features identified by this layer require cross-referencing with recent, high-resolution aerial imagery (e.g., Google Maps, Bing) to ensure that they are unchanged from the reference images. We used a 9/1/1999 reference image for path 12 and a 8/14/1995 image for path 11. Future analyses may wish to include more recent reference images.

The normalization processes described in this manual are orthogonal regression and principal component analysis (PCA). Most regressions are ordinary least squares in which residuals are measured along the y-axis (dependent variable) from the regression line. In other words, only observation errors for the dependent variable are taken into account. Orthogonal regression (also known as perpendicular regression and total least squares) calculates observational errors along both the x and y axes (dependent and independent variables) and is more appropriate for our purposes because we manually select pseudo-invariant ground features. PCA involves orthogonal transformation and is appropriate for this analysis because analyses involving reference and non-reference images always contain two components and are therefore easily analyzed with PCA.

In the unusual event that an image contains no haze, radiometric normalization is not technically necessary. Unfortunately, haze is difficult to detect visually. It is therefore important to compare pixel values between the reference and non-reference images for pseudo-invariant ground features. This is especially important for TM band 1. If you encounter small or negligible differences in pixel values across the study area, then haze effects are minimal and not a concern; however, normalizing a haze-free or nearly haze-free image will not hurt model performance.

Model calibration data

Model calibration requires the incorporation of field data collected near the date of satellite image capture. Ideally, calibration data should cover a spatially balanced geographic extent and encompass a wide range of SDD values (Nelson et al. 2003). Unfortunately, the lack of spatial balance in existing field monitoring is in part what prompted this initiative on remote monitoring in the first place. SDD data (collected on or near the date of satellite image capture) in conjunction with the lake pixels from the respective satellite image(s) are used to model the relationship between satellite-measured reflectance and water clarity. Generally, insufficient calibration (SDD) data are available on the same date as satellite image capture, and we are forced to use field data collected within a certain number of days from the satellite overpass. Advantages of creating remote SDD-estimation models based on mid- to late-summer imagery are that this is the time of year when field data are most abundant and when lake conditions are the most stable. Olmanson et al. (2008) reported that field data collected ±10 days from satellite image capture were usable in model calibration; however, time windows exceeding ±7 days are rarely, if ever, needed and ±7-day windows (or shorter) can be used to calibrate accurate models (Kloiber et al. 2002). Longer windows of time allow more data to be included in calibrations, but the likelihood increases that lake conditions may change from those captured by the satellite. In general, calibration datasets of ±1 to 3 days are suitable. The smallest calibration dataset we successfully used included 31 data points, but using 50 to 60 points helps ensure wide geographic and numeric variability of SDD values.

Why are average lake depth and watershed wetland area included in models?

Although there are advantages to relying solely on satellite data as the independent variable, adding ancillary variables that reflect conditions of the study landscape improves the accuracy of the model when they are available. We analyzed the effects of several lake and watershed variables on regional water clarity and found average lake depth and the proportion of watersheds covered by wetlands to be consistently significant predictors of SDD. We also tested lake area, lake perimeter, area/perimeter ratio, watershed size, total watershed wetland area, elevation, and maximum lake depth and found these variables to be inconsistently correlated with SDD or redundant with average depth and watershed wetland area. It is possible that some of these variables may be correlated with water clarity of lakes in areas other than Maine. In addition, Maine's lakes are relatively clear and generate weaker satellite reflectance values than more productive lakes (e.g., those in the Great Lakes Region); hence ancillary predictor variables are particularly useful in a clear lake dataset. We found that addition of ancillary data improved model R^2 by 0.03–0.07.

The effect of ancillary variables is subject to geographic variation. In Maine, the proportion of lake watersheds containing wetlands is a strong predictor of SDD in path 11 lakes, but not in path 12 lakes. The likely explanation for this is that the area encompassed by path 12 is relatively mountainous with fewer wetlands, whereas wetland coverage is reflective of eastern Maine geography. Wetlands are a source of dissolved organic carbon (DOC), which negatively affects water clarity (Detenbeck et al. 1993). Additional variables (e.g., portion of watershed devoted to agriculture) may drive regional SDD elsewhere, so including variables that describe local geography may improve the predictive capacity of water-clarity models in other regions. Asking the questions What landscape factors potentially influence water clarity? and How can they be incorporated in remote lake monitoring? will identify variables to include in the predictive models.

What software have we used?

All of the GIS analyses described in this report were designed for use in ArcGIS® version 10.0. ERDAS Imagine® is an alternative software package, but our specific directions would need some modification for use in ERDAS.

We used R version 2.12.0 for all statistical analyses, along with the user interface Rcmdr. Rcmdr provides some relief from the command-line interface of R. The instructions in this manual follow the use of R and Rcmdr, but other statistical software may be used if desired. R can be freely downloaded (cran.r-project.org). R is particularly useful for the principal component analysis portion of radiometric normalization, but users may find that other software packages are easier to use for building regression models.

What projection is used?

GloVis provides Landsat images of Maine in a WGS1984 UTM Zone 19N projected coordinate system. All analyses in this manual, including analyses of MODIS imagery, were performed using this coordinate system. The data layers we reference are also in this projection. If you create any of your own layers, continue to use this projection to maintain spatial accuracy and consistency.

Analysis of spatial and temporal patterns of water clarity

A frequent application of satellite-based remote sensing of water clarity is change detection over space and time. Analyses of detection of multiyear changes require the use of images captured at roughly the same time of year. Therefore, use of August and early September images is ideal, assuming that the intent is to characterize regional water clarity at the seasonal, late summer low just prior to fall turnover. Restricting analyses to August and early September unfortunately is problematic owing to the 16-day temporal resolution and because clear imagery is not reliably available. A maximum of three images for both paths 11 and 12 would be available each year. Clear, usable imagery was available in August and early September in 2009, 2010, and 2011, but no clear imagery was available during this time of year from 1990 to 1994. Remote monitoring must be flexible and cannot adhere to a strict expectation of image availability every certain number of years. Some years will have multiple dates with clear imagery, whereas several years may pass without clear imagery during August and early September.

Analyzing path 11 in one year and path 12 in the next may lead to biased conclusions about statewide water clarity owing to geographic differences between the two paths. An alternative, practical approach is to focus analysis on the overlap region between paths 11 and 12 (Figure 2). This area represents a belt transect, covers a strong north-south gradient, and contains 570 lakes > 8 ha. The primary advantage of analyzing the overlap region is the ability to use images from either path for the purpose of detecting changes, essentially doubling the chances of obtaining a clear image in a given year. Another advantage is the ability to analyze the same lakes in successive years. Focusing study on the overlap region does not result in data loss; whereas only lakes in the overlap region would be used for detecting changes, data still could be collected for all other lakes > 8 ha in the originally analyzed path.

A useful method for detecting potential changes in water clarity across space and time is to divide Maine into regions. Peckham and Lillesand (2006) and Olmanson et al. (2008) used Landsat data to analyze changes in clarity of Wisconsin and Minnesota lakes, respectively,

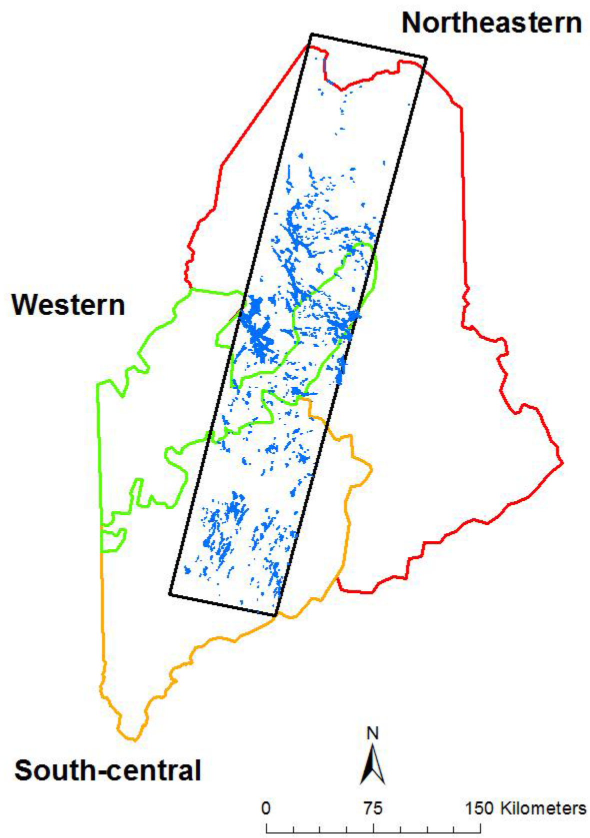

Figure 2. Map of lake regions of Maine and overlap region between Landsat TM paths 11 and 12, containing 570 lakes > 8 ha.

using Omernik's (1987) Level III ecoregions. Only two Level III ecoregions cover substantial portions of Maine, rendering an ecoregion analysis no more effective than a comparison of Landsat paths 11 and 12 (which roughly correspond to the two ecoregion boundaries). The U.S. EPA currently recognizes Level IV ecoregions (19 in Maine), which could potentially be used in fine-scale analyses of estimated SDD. (This link provides maps and data specific to Maine [ftp://ftp.epa.gov/wed/ecoregions/me/].) The most practical, simple way to divide Maine may be to use the three MDEP-recognized "modified ecoregions," or lake regions, based on morphometric and chemical lake variables: northeastern, south-central, and western (Bacon and Bouchard 1997) (Figure 2). All three lake regions are well represented in the overlap area, which includes 227, 256, and 162 sample stations on lakes > 8 ha in each respective region (some lakes contain > 1 station). It is not practical to model each lake region separately because there are generally insufficient calibration data in the relatively remote northeastern and western lake regions.

Technical Methods

Image preprocessing and cloud removal

1. CREATE A WORKING DIRECTORY. Data organization is of critical importance, even for experienced GIS users. Create a folder such as P12_8.17.11 (which in this example is a placeholder for Landsat path 12 and the image acquisition date of 17 August 2011). You will later create a series of subfolders within this main folder.

2. ORDER AND DOWNLOAD IMAGES. Preview and select images (scenes) from GloVis. Inspect all scenes from the desired path (path 12: rows 27–30; path 11: rows 27–29). Before using the service, you must create a free user account and provide some basic information. You may wish to refer to the user guide available on the GloVis website. Some images must be ordered and cannot be immediately downloaded, in which case you will receive an email when the download is ready (may take a few days). Download as GEOTIFF files if you are presented with data format options. The image files will be provided in a *.tar.gz compressed format. You will need a program to unpack downloaded files. We used 7zip (www.7-zip.org), a free Windows-based application to uncompress image files. Create subfolders for each scene in your working directory and extract each image to its respective subfolder. Note: scene files are large and an analysis of a path may require 10 GB. You can save disk space later by deleting unnecessary intermediate files you create during analysis.

3. CREATE COMPOSITE IMAGES. Each scene has seven (eight for Landsat 7) band files ending in B10 for TM band 1, B20 for TM band 2. Band 6 is split into two separate bands: B61 and B62 (low and high gain, respectively) in Landsat 7 images. Load all bands from a single scene (e.g., path 11, row 29) into a blank GIS map document. Navigate to the *Composite Bands* tool (located under *Data Management Tools, Raster,*

Raster Processing). Add all bands to the input list and order them lowest to highest. Save the composite band file in the same folder as the other bands of the respective scene, naming it similarly (e.g., L71011029_02901206 01_ Comp). Repeat this procedure for each scene in your analysis. Save the map document per example: P12_8.17.11 (path number, date) (the underscore is unnecessary, but it is a good file-naming habit when designating path names with anything GIS-related).

4. MOSAIC IMAGES. Add individual scene images (27–30 for path 12, 27–29 for path 11) from TM band 1 to the map. Create a subfolder "Mosaic" in your working directory. In the toolbox, navigate to *Data Management Tools, Raster, Raster Dataset* and select *Mosaic to New Raster*. Add all TM1 images as *Input Rasters*. Specify the Mosaic folder as the *Output Location*. Under *Raster Dataset Name with Extension,* type "TM1_Mos" (if you do not specify a file extension, the program defaults to a GRID file format, which is fine). Enter 1 in the *Number of Bands* field (if mosaicking a composite image, enter 7 for Landsat 5 scenes or 9 for Landsat 7 scenes; B61 and B62 are considered two bands.) Under *Mosaic Operator,* select *MAXIMUM*. The black areas have pixel values of 0, so when scenes overlap, the computer uses the non-zero pixel values to build the mosaic. Repeat these steps for the TM3 and composite images.

   Note: The color scheme may automatically change after mosaicking or extracting; the software automatically recolors images according to the range of values in the image (which you just manipulated by cropping out certain areas).

5. IDENTIFY CLOUDED AREAS. Add the mosaicked composite image to your map. The RGB combination 1, 6, 6 (visible blue and thermal infrared) allows easy visual interpretation of the extent of clouds and cloud shadows. Select *Symbology* from the composite image's *Properties*. Under the drop down menus next to red, green, and blue, select bands 1, 6 and 6, respectively. Red and pink areas will indicate clouds. Another useful RGB combination is bands 3, 2, 1, which represent "true color" and may be a useful reference (Figure 3). (Note: if an image contains no clouds over Maine, ignore steps 6–8).

6. UNSUPERVISED CLASSIFICATION. Add the mosaicked TM1 and TM3 to the map. Enable the *Image Classification* toolbar via the *Customize, Toolbars* pull-down menu. In the *Image Classification* toolbar, select *Iso Cluster Unsupervised Classification*. Add the TM1 mosaicked image to the *Input raster bands* list. Enter 10 under *Number of classes.* Create a subfolder "Unsupervised" in your working directory and save the classification raster as "TM1_US" (there is a 13-character name limit if you save in GRID format). Repeat this process for TM3, naming the output "TM3_US." It is not necessary to classify the composite image.

7. RECLASSIFY CLOUD PIXELS AS NULL. Zoom in on a particularly clouded area, using the RGB 1, 6, 6 image as a reference. In the TM1_US and TM3_US raster layers, change the colors of the pixel classes that show up as clouds (by double clicking on the colored boxes in the table of contents) to a single color (e.g., black). Enable the *Spatial Analyst* extension under *Customize* (pull-down menu), *Extensions, Spatial Analyst*. Use the *Reclassify* tool (located under *Spatial Analyst Tools, Reclass*) to reclassify clouded pixel classes as null values. The input raster should be the unsupervised classification file (e.g., TM1_US). *Reclass field* should default to value (which is what you want). Enter "No Data" in respective boxes for pixel classes that represent clouds in the *New Value* column. Create a subfolder "Reclassified" in your working directory and save the reclassification as "TM1_RC" (RC = reclassified). Repeat this process for TM3.

8. ELIMINATE CLOUD PIXELS. Use the *Extract by Mask* tool (under *Spatial Analyst Tools, Extraction*) to remove cloud pixels from further analysis. Select the original TM1_Mos file as the input raster. Select the respective band reclassification file (TM1_RC) as the feature mask data. Create a subfolder "Extraction" in your working directory and save your file as

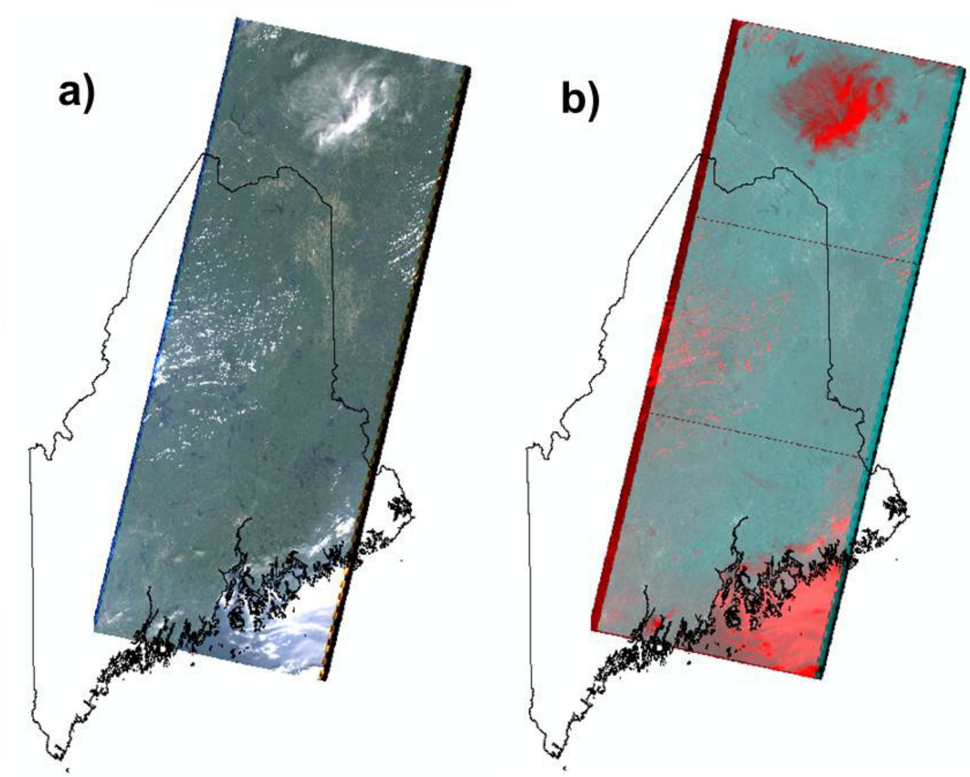

Figure 3. (a) True color composite image (RGB 3, 2, 1) and (b) cloud indicator visible blue/thermal infrared composite image (RGB 1, 6, 6) of path 11 Landsat TM image captured 8-9-05. Images considerably cloudier than this image (8%) are not recommended for use.

"TM1_ebm" (ebm = extract by mask). Repeat this process for TM3.

Radiometric normalization

1. IDENTIFY A REFERENCE IMAGE. The reference image must have minimal haze and be as clear as possible. For paths 11 and 12, we used images captured on 8/14/1995 and 9/1/1999, respectively. For future analyses, though, it may be preferable to use more recent reference imagery. The likelihood of pseudo-invariant ground targets undergoing change increases with longer gaps between reference and non-reference images. Add TM band 1 and 3 files from your reference image to the map document.

2. ADD NORMALIZATION POLYGONS. Pseudo-invariant ground targets used for normalization must be well distributed geographically throughout the study area and span dark and bright features. Add the buffered normalization point file (polygon shapefile) for the appropriate path. The path 11 file contains more polygons because cloud obstruction is more prevalent in this path. Check to see if clouds or fog are obscuring any of your polygons, and if so, select all unaffected points and export these as a new layer in a new subfolder "Normalization" (right-click on layer in table of contents, select *Data,* then *Export Data*). If no polygons are affected, simply export the whole layer anyway to the Normalization subfolder as "Normalization_Pts_8.17.11" (replace 8.17.11 with your project folder date). Select *Yes* when asked to add the layer to the map. Path 12 features were buffered using a 50-m radius, but the path 11 features were buffered using only a 10-m radius because many of the target features are smaller. A 10-m buffer may or may not contain multiple pixels, but this is okay. Single pixels may also be used if

clouds are covering desired parts of features (we used this technique occasionally in path 11 normalization).

3. EXTRACT SATELLITE DATA FROM GROUND TARGETS. Use the *Zonal Statistics* as a *Table* tool (under *Spatial Analyst Tools, Zonal*) to calculate the mean pixel value in each buffered zone. Note that if you are using single pixels, you may choose to use the main identify tool (blue circle with an "i") and manually enter values into a spreadsheet. In *Zonal Statistics as a Table*, set *Feature zone data = buffered normalization points* (e.g., Normalization_Pts_8.17.11), *Zone field = Name, Input value raster* = TM1_ebm (or TM1_Mos if no clouds). Save your table as "TM1_Norm" in the Normalization folder. Uncheck the box next to *Ignore No Data* (if any points were obscured by clouds and you did not delete them earlier, this will discard them automatically in the output table). Repeat this process for TM3 (TM3_ebm or TM3_Mos) and for the reference images for TM1 and TM3. Save the reference image tables as "TM1_Ref" and "TM3_Ref" in the Normalization folder.

4. JOIN TABLES TOGETHER. Join the reference and non-reference tables together as one table. Use *Name* as the join attribute. Export the resulting table to the Normalization folder as a .csv file. Set *Save as Type* to *Text File* and type extension .csv after the table name (e.g., Norm_81711.csv). Note: avoid periods or other punctuation marks when saving .csv files (other than in the file extension itself).

5. PREPARE TABLE FOR ANALYSIS. Open the .csv file containing normalization data (created in Step 4) in Excel. Delete all columns except those containing the names of the ground features and the satellite data. Rename satellite data columns as "TM1_Ref," "TM1_(year)," "TM3_Ref," and "TM3_(year)." Inspect all data columns. If numbers in non-reference columns are vastly different from reference columns, it is possible that specific ground features were affected by undetected clouds. Inspect non-reference images and eliminate features if necessary. Permanent modifications may eliminate usefulness of a particular ground feature. It is also possible that temporary modifications (such as construction) may be responsible for pixel differences. In analysis of a 9-1-2008 image, we found it necessary to eliminate the normalization polygon for the Portland International Jetport owing to temporary construction. Resave the .csv file.

6. NORMALIZATION ANALYSIS. Open R and load the Rcmdr package via the command `library(Rcmdr)`. As with any R package, you may need to install Rcmdr if it is not currently on your system. Rcmdr is a useful package that provides a GUI (graphical user interface) for R and its many functions. Rcmdr is useful in this step only to load the appropriate dataset (hereinafter referred to as "Dataset"). Load the dataset under *Data* (pull-down menu), *Import data, from text file, clipboard or URL*. Select *Commas as the Field Separator.* The R command `princomp` (stat.ethz.ch/R-manual/R-patched/library/stats/html/princomp.html) will be used to conduct the PCA within the R console.

   Once the dataset is loaded into R, run the following command (note that the variable names TM3_2011 and TM3_Ref refer to the column names you defined in Step 5).

```
PC3=princomp(~TM3_2011+TM3_Ref,
cor=FALSE,data=Dataset)
```

   The results of the PCA are stored in the variable PC3. In the next step, you will list the *loadings* and *center* parameters.

```
show(PC3$loadings)
show(PC3$center)
```

   Your output for the loadings parameter might look something like this:

```
Loadings:
            Comp.1     Comp.2
TM3_2011    0.673      -0.739
TM3_Ref     0.739       0.673
```

Your output for the center parameter might look something like this:

```
TM3_2011    TM3_Ref
  58.817     66.725
```

Record the loadings and center values on a sheet of paper or an empty spreadsheet. For each variable (TM3_2011, TM3_Ref), you should have at least two components (Comp.1, Comp.2). Repeat the above steps for TM1.

7. CONSTRUCT THE NORMALIZATION EQUATION. Use the second eigenvector (component 2 in the loadings parameter) of the PCA and the centers (means of the reference and non-reference data) to obtain the normalization equation in y = mx + b format (where *m* represents gain, *b* represents offset, *x* is the original pixel value, and *y* is the normalized pixel value). TM1 and TM3 images must be normalized in separate calculations.

   Continuing with our example in Step 6, the equation for TM3 will take on the following form

   $-0.739(x - 58.817) + 0.673\ (y - 66.725) = 0$

   Using basic algebra, we can solve for y to generate the equation's final form

   $y = 1.098x - 4.336$

   Repeat the same process for TM1, generating a separate equation.

   Note: a gain (coefficient *m*) value of < 1 represents a potential loss in radiometric resolution (shrinking of the range of pixel values in your data). To correct for this, multiply the equation by the inverse of the gain value (i.e., 1/m). If normalizing more than one set of images, multiply all equations by the inverse of the lowest gain value (K. Legaard, University of Maine, pers. comm.). Keep the reference image as is.

8. MAP ALGEBRA. Use the *Raster calculator* to normalize the pixel values (under *Spatial Analyst Tools, Map Algebra*). Build the expression by typing directly into the expression box or by selecting and double-clicking items from map layers and variables and using the operations keypad. Following our sample equation for TM3, our expression for the normalization of TM3 imagery will be

   (1.098 × TM3_ebm) – 4.336.

   Save the output raster in the Normalization subfolder as "TM3_Norm." Follow the same steps to normalize TM1 imagery (TM1_ebm) using the equation generated for TM1 in Step 7. Once you have a normalized rasters, you are ready for model calibration.

   Note: an offset value that results in negative values (more likely to be a problem with TM3 because we are interested in greater TM1 values that are less likely to approach zero) should receive an added constant (some number large enough to make all pixel values of interest > 0) (K. Legaard, pers. comm.). If normalizing more than one set of images, add the constant to all normalized images. Keep in mind that you are interested only in lake pixels, so it is possible to have negative values in your dataset that are not contained in lakes.

Model calibration and development

1. COMPILE FIELD DATA. Open the spreadsheet of field data for the year of interest. Sort all entries by date, oldest to newest. Select all field data collected ±3 days of image capture and copy to a new spreadsheet, along with the column headings. Re-sort this new spreadsheet by SECCBOT, A to Z, and remove all rows containing bottomed out Secchi data (indicated as "B"). Next, re-sort by MIDAS (Maine lake identification number), smallest to greatest, and then by date, oldest to newest. Save the spreadsheet in your working directory as "SecchiData_8.17.11" (or the date pertaining to your project).

2. CREATE NEW CALIBRATION POINTS LAYER. In your working map document, add the Landsat remote sampling points layer (Landsat_SamplingPts_75m). This layer contains circular (75-m radius) remote sampling stations used to extract satellite data from lakes. Attributes contain the name of each lake, sample station number, MIDAS, unique remote sampling station identifier (IAN_ID) and various physical lake and watershed measurements. Add

the shapefile for either path 11 or 12 and use the *Clip* tool (under *Analysis Tools, Extract*) to remove points outside your path of interest. Save the data layer in a new folder "Remote_Sampling" within your working directory as "Calibration_Pts_8.17.11." Remove the original master points layer.

3. CHECK FOR CLOUD SHADOWS. Cloud shadows can make portions of images appear darker than they actually are, and in some cases, appear similar to lakes. Use the RGB 1, 6, 6 band combination to check for lakes potentially affected by cloud shadows and delete respective lake stations from the calibration points layer. The most efficient way to do this is to zoom in on clouded areas and inspect images visually. If you are spending an inordinate amount of time trying to avoid cloud shadows, your image might be too cloudy for analysis in the first place.

4. CHECK FOR FOG. Unfortunately, the unsupervised classification is insensitive to foggy areas, which tend to form over large lakes and the coast. Landsat captures imagery during mid-morning when fog may not have completely dissipated. Fog causes some scattering and the automated classification may identify fog pixels simply as relatively bright lake pixels. Therefore, if fog is undetected, shallow Secchi values may be falsely predicted. Manually zoom in on large lakes throughout the state and inspect for fog. Remove suspect lake stations from the Calibration_Pts_8.17.11 layer.

5. ENTER FIELD DATA INTO GIS. Open the calibration point attribute table and sort entries by MIDAS, smallest to greatest. Make sure the *Editor* toolbar is checked on (under *Customize, Toolbars*). In the *Editor* toolbar, select *Start Editing* and the Calibration_Pts_8.17.11 as the layer to edit. Scroll through the spreadsheet of available field data and enter the Secchi disk depth (m) values in the Secchi column. Also include the date in the date column (enter as 8-17 for August 17). If a lake station was sampled more than once during the 3-day window, use the data from the date closest to image capture. If choosing among data points an equal number of days from image capture, average the values (this also applies to multiple samples taken on the same day) and record either date. Pay attention to lakes with multiple sample stations and avoid substituting data from among different stations (e.g., do not use field data taken at a sample site #1 as calibration data for a sample site #2). Periodically save your edits as you add data to the table (under the *Editor* toolbar).

6. EXTRACT SATELLITE DATA. Use the *Zonal Statistics as a Table* tool (under *Spatial Analyst Tools, Zonal*) to extract the mean pixel value in each 75-m zone of the remote sampling sites. Use the Calibration_Pts_8.17.11 layer as the *feature zone data*. Select IAN_ID as the *zone* field. Select desired satellite band image (e.g., TM1_Norm) as the *Input value raster*. Be sure to use the normalized image file. Create a subfolder "Zonal" in your working directory and save the output table TM1_75m in this folder. Uncheck the box next to *Ignore NoData*. This ensures that any 75-m areas which include cloud pixels are removed from the zonal extraction. Select *MEAN* under *Statistics type* and click *OK*. Select *Yes* if prompted to add the output table to the map. Repeat for TM3.

7. ADD SATELLITE DATA TO CALIBRATION DATA LAYER. Right click on the calibration data layer, select *Joins and Relates* and then *Join*. Select the option to join attributes from a table (usually the default). Select *TM1_75m* as the table to join to this layer. Select *IAN_ID* as the base field in each drop down menu. Select *Keep all records*. If you make an error, keep in mind that joins are completely reversible. Repeat this process for TM3_75m. This will create a second join to the calibration data layer.

8. EXPORT JOINS TO A PERMANENT LAYER. Open the attribute table of the calibration data layer to inspect join results. A successful join should look like the original attribute table of the calibration data layer with extra columns from the TM1 and TM3 zonal statistics tables added to the right end of the table. Right click on the layer in the table of contents, select *Export* and resave the joined layer in the Remote_Sampling subfolder as SecchiTM_8.17.11 (the file name

indicates that this file contains Secchi and satellite data). Indicate *Yes* when asked whether to add this layer to the map, inspect that the export was successful, and remove the original calibration points layer from the table of contents. At this point, you may wish to clean the new layer's attribute table by removing redundant or unnecessary attributes. Delete joined columns except for those containing the TM1 and TM3 pixel values. You can temporarily relabel column headings as aliases in the attribute table; however, you will have to do so again in Excel after exporting the table.

9. EXPORT SPREADSHEET OF CALIBRATION DATA. Open the attribute table of SecchiTM_8.17.11 and sort by date. Select all entries containing a date (some rows may contain no satellite data owing to cloud interference) and export (via the attribute table dropdown menu) the selection as .csv table to the Remote Sampling subfolder as "SecchiTM_81711_3d" (.csv does not support many character types, but underscores are okay). The "3d" designation indicates inclusion of field data captured ±3 days of the satellite overpass. You will have to type the extension .csv in the *Output table name* field box (in place of default .dbf).

10. ORGANIZE CALIBRATION SPREADSHEETS. Open SecchiTM_81711_3d in Excel. Create a new column "lnSecchi" and calculate the cell values using the natural log function

    `ln(Secchi column)`.

    Sort rows by date if necessary. Save these changes. Select rows of data ±1 days from satellite overpass, copy to a new spreadsheet and save as a new .csv file named "SecchiTM_081711_1d" in the Remote_ Sampling subfolder.

11. MODEL CALIBRATION IN R.

    a. In Rcmdr, import your desired calibration dataset per instructions in Step 6. Retain the default name (Dataset), or rename your dataset as desired. Navigate to the Remote_ Sampling folder and select either the ±1- or 3-day data file (depending on which one you want to try first).

    b. From the Rcmdr window, select *Fit models, Linear model* under the *Statistics* menu. lnSecchi is the response (dependent) variable. Select the variables TM1, TM3, AVGDEPTHFT (average lake depth in feet) and PctWet (percentage of wetland cover in a watershed) as the predictor (independent) variables. Click *OK*.

    Note: PctWet is not a consistent, strong predictor of water clarity in path 12 models owing to the lack of large wetlands in western Maine.

    c. Figuring out the best model can be a multistep process. Nonsignificant variables can sometimes be eliminated, but keep in mind that path 11 and path 12 models generally adhere to the following:

    Path 11: lnSecchi = TM1 – TM3 + AvgDepth – PctWet + intercept

    Path 12: lnSecchi = TM1 – TM3 + AvgDepth + intercept

    TM1 may contribute little to model performance in turbid waters. The short wavelengths of TM1 do not penetrate these waters well, and this phenomenon may be reflected by its nonsignificance, provided the calibration data accurately represent the greater population of Maine lakes. Additionally, AvgDepth may contribute little to the model if fall turnover has occurred.

    R^2 is a good cursory measurement of model performance, but outliers may be affecting R^2. Under the *Models* menu, select *Graphs and basic diagnostic plots*. The top two graphs are particularly useful. The residuals vs fitted values plot indicates whether there are any data points with particularly large residuals. The normal Q–Q plot shows potential departures from assumed normality in error (as well as large residual values).

    Steps d–f are best conducted together. Regression assumptions of normal error and constant variance should be verified in conjunction with potential outlier removal.

d. Use the Bonferonni outlier test to check for potential outliers. Type `outlierTest(Model name)` in the command prompt. Any data points with a Bonferonni $p < 0.05$ will be displayed. Attempt removal of each data point one by one and observe changes in R^2 (use the *Edit data set* button to remove data values; note that eliminating the lnSecchi value eliminates the entire data entry from analysis). Reinspect diagnostic plots with each new model.

e. Verification of constant variance assumption: Select *Models, Numerical diagnostics and Breusch-Pagan test for heteroscedasticity.* Default settings for this test are fine. A p value < 0.05 means the assumption of constant variance likely has been violated.

f. Use the Shapiro-Wilk normality test to verify the normal error assumption: Type: `shapiro.test(rstudent(Model name))`. A p value < 0.05 means the normality assumption likely has been violated. If you are able to justify eliminating data points with large residuals, doing so can improve the Shapiro-Wilk p value.

g. Note, or copy and paste all final model output information into a spreadsheet or word-processing file. You may still wish to repeat these steps using calibration windows of different lengths to compare results, particularly if calibration datasets contain < 30 data points. With smaller calibration datasets, there is some risk of fitting models that fit the calibration data but not actual conditions throughout the landscape. In summer, calibration windows of ±3 days usually contain enough data if ±1 day windows do not.

h. Add residuals and fitted values to the data table under *Models, Add observation statistics to data.* Make sure you have the correct model selected as the active model (beneath the toolbar).

i. Export resulting data table via *Data, Active data set, Export active data set.* Check *write variable names* and *write row names.* Select commas as the file-delimiter option. Click *OK*. Navigate to the Remote_Sampling folder and save as .csv as "Export_1d_81711" (if using a 1-day window). You can then use this spreadsheet to graph and calculate differences between observed and model-estimated SDD values.

j. Not all lakes have bathymetric data. Therefore, it is necessary to fit alternative models using the same methods described above without depth as a variable, even though it is a significant predictor. Alternative models are used only to estimate clarity of lakes with unknown depth. Use the same calibration data, but record the alternate model in your working spreadsheet or word processing document. A few lakes may also be missing wetland data, in which case you may need another alternate model (path 11 only).

12. ESTIMATE REGIONAL WATER CLARITY. In your map document, export the attribute table of SecchiTM_81711 as "Remote_Secchi_81711.csv." This file already contains satellite data extracted from all eligible lakes > 8 ha in size in the path study area unaffected by clouds. Open the exported table in Excel and use the final calibrated model to estimate water clarity for the unsampled lakes in the path. Remember that the model estimates ln(Secchi), so the actual Secchi disk depth = `2.71828 ^ (model-estimated lnSecchi)`. It is helpful to sort rows by average depth to make applying the alternate model in the absence of depth data more efficient.

13. ANALYSIS OF TROPHIC STATES. It is often useful to classify and analyze lakes by trophic status. SDD < 4 m = eutrophic, 4–7 m = mesotrophic and > 7 m = oligotrophic (Maine PEARL 2011).

Model validation

Model validation is necessary to determine if models suitably represent the greater population of lakes and not just the calibration data. Calibration datasets that are small or poorly distributed numerically or geographically are more likely to produce models that fail to pass validation. Model validation may not be strictly necessary if the primary interests are the bottom-line SDD estimates (i.e., there are no plans to publish a scientific paper); however, model validation provides an indication of the repeatability of the modeled results. If you are at all concerned that your model is a relic of your calibration dataset, validation can verify or refute this suspicion. It is the decision of the user whether to perform validation.

The models described in this manual were validated using calibration datasets of 31–119 data points. If the calibration dataset contains < 50 points, use leave-one-out jackknifing. If it contains ≥ 50 points, use subsamples of a random 25% of the calibration dataset. If primary models (containing depth) have been validated, validation of alternative models is unnecessary because alternative models contain the same calibration data.

Leave-one-out jackknifing

This method involves running models using the same input parameters with calibration datasets consisting of all but one data point. The idea is to determine if single points have disproportionately large influences on the rest of the model. Outliers have already been removed, but influential data entries may still remain. Influential data entries are not necessarily bad; the purpose of jackknifing is not to suggest removal of certain data points, but rather to determine if the model was fit the way it was because of specific data points. To reiterate, you want a model that represents the greater population of lakes, and if certain models demonstrate considerable change in predictive capacity as a result of slight changes in the calibration dataset, you may not have a truly representative model.

The easiest way to perform jackknifing is to reload the calibration dataset used in the final model in R. Delete the first data entry (i.e., the first record in the dataset) and run the model using the same variables defined in the original model. Create an Excel spreadsheet to record the coefficients of all input variables, intercepts as well as R^2 (if desired for reference). Add the first data entry back into the dataset, delete the second entry and rerun the model. Add the second entry back into the dataset, delete the third entry and so on. Once all jackknifed models have been run, average the variable coefficients and intercepts and compare to those of the original model. If the numbers are all reasonably close, then the model has passed validation. "Reasonably close" is subjective, but if you are unsure what constitutes reasonably close, try applying the jackknifed model and the original model for SDD estimation and compare the results. If they are acceptably similar, validation has been successful. Save the spreadsheet of jackknifing results in a new folder "Validation" in your working directory. Use the original model for SDD estimation, not the jackknifed model (only used for validation purposes).

The following is the sole jackknifed model we used in our manuscript (we generally used > 50 data points in our model calibrations). The coefficients are more than reasonably close, and no single calibration point was particularly influential on the model.

Original model: -0.4270 (TM3) + 0.0045 (AvgDepth) + 6.202

Jackknifed model: -0.4274 (TM3) + 0.0045 (AvgDepth) + 6.203

Subsampling

Subsampling is not practical with small calibration datasets because there are too few data. A cutoff of 50 calibration points ensures adequately large subsampled datasets consisting of 25% of the original entries (W. Halteman, University of Maine, pers. comm.). The general procedure for subsampling is to create 10 random subsets of 25% of the original calibration dataset, run the original model using these subsets and then compare the SSE (sum of squared error) to the PRESS (predicted residual sums of squares).

Ideally, we would resample field data or use subsamples of data not used in model calibration (reserve data). Because we are working with historical data and use all available calibration data within a specified time frame, neither of these options is practical.

1. Excel can be used to select a random 25% of the calibration dataset. For example, a calibration dataset of 60 data points should produce random subsets of 15 data points each. Open the final calibration dataset in Excel. Insert two blank columns to the left of column A. Enter `rand()` in the top cell of the new column A and copy/drag into all data rows. Copy these cells and paste as values (right click on destination cells, then click on the *123* icon; use *Paste Special* in older versions of Excel) into column B.
2. Select all data in the spreadsheet and sort by column B, smallest to largest. Copy the desired number of rows to a new Excel spreadsheet.

Table 3. Example of model validation by subsampling.

Subsample	n	R^2	SSE	PRESS
1	30	0.8088	1.2170	2.0012
2	30	0.7861	1.0840	1.8614
3	30	0.8389	0.9224	1.4333
4	30	0.7649	1.2628	2.6664
5	30	0.8063	1.0603	1.4220
6	30	0.8776	0.8424	1.3705
7	30	0.7403	1.8045	2.5664
8	30	0.8362	0.9939	2.0296
9	30	0.8072	1.2922	1.6862
10	30	0.8618	0.8699	1.2611

Delete columns A and B. Save each spreadsheet as a .csv as "Subsample1_2011." Repeat steps 1 and 2 nine times to create 10 validation datasets. Save these subsets in a subfolder named "Validation" in your working directory.

3. Open R, Rcmdr and load Subsample1_2011 in R. Create a model using the same input variables as the original model.
4. Create an Excel spreadsheet and save it as "Validation_08.17.2011." Create columns to record SSE and PRESS statistics.
5. Retrieve SSE from the analysis of variance table associated with the linear model. Type `anova(model name)`. The *Sum Sq* value next to *residuals* is SSE.
6. Get the PRESS statistic using the command `sum((model name$residuals/(1-hatvalues(model name)))^2)`.
7. Repeat for the remaining nine subsamples. Compare PRESS statistics to respective SSE values for "reasonable closeness," which is a subjective evaluation. Table 3 contains the results of a validation of a 1995 Landsat-based model in which differences between SSE and PRESS statistics were deemed reasonably close.

Mapping water clarity

Creating a water-clarity map is quick and easy. Join a table of estimated SDD values to the lakes GIS layer, using MIDAS as the join attribute, keeping only matching values. Export the join as a new layer, if desired. You can edit the layer symbology to display lakes according to different categories of water clarity, such as trophic states. Be aware that a join will join based on the first match of MIDAS numbers in each table (i.e., if a lake has more than one sample station, the join will pick the first one listed in the table). You can always edit the new layer manually, if desired.

Analyzing spatial and temporal patterns of water clarity

In the final output spreadsheet containing SDD estimates, the Ecoreg attribute can be used to identify lake stations in each lake region and analyze regional lake clarity (e.g., 1 = northeastern, 2 = south-central, and 3 = western). Although some lakes occur in multiple regions, each 75-m sample station occurs in just one region.

Basic statistics

Statistics are used to make inferences about larger populations based on sampling, so if we are comfortable assuming that the overlap region adequately represents Maine, then there is little need for statistics beyond basic comparisons of means, medians, ranges. These can easily be calculated in Excel.

Pairwise t-tests

Pairwise t-tests allow determination of statistically significant differences between SDD means in two separate years. Pairwise t-tests do not require equal sample sizes, which you likely will not have owing to cloud cover (some lakes will inevitably be obscured by clouds). Excel does not compute t-tests with our desired requirements, but doing so with the command prompt in R is not difficult with the command pairwise.t.test (stat.ethz.ch/R-manual/R-devel/library/stats/html/pairwise.t.test.html). This link explains the meaning of the various input parameters.

1. Let data file = Dataset, with columns for SECCHI (estimated SDD) and Year (e.g., 2011)
2. Convert Year to factors rather than a continuous variable. In Rcmdr, select *Data, Manage variables in active data set* and *Convert numeric variables to factors*. Next, select *Year, use numbers* and click *OK*.

3. Conduct pairwise t-tests using the R command prompt. Type:

```
pairwise.t.test(Dataset$SECCHI,
Dataset$Year,p.adjust.method="none"
,pool.sd=FALSE,paired=FALSE,altern
ative="two.sided",var.equal=TRUE)
```

4. Copy/paste or transcribe the p-values into another document. A *p* value < 0.05 indicates a statistically significant difference (0.05 is the standard cutoff in statistics and indicates a 5% chance that a detected difference is not really a difference and instead a result of chance).

Confidence intervals

Confidence intervals provide useful estimates of the true mean based on available data. For example, there is a 95% probability that the true population mean falls within the lower and upper limits of a 95% confidence interval (a 90% confidence interval would be narrower). Confidence intervals take into account error (standard deviation) and sample size and can be calculated in Excel.

1. Open the exported table of Secchi estimates uniquely identified by IAN_ID in Excel.
2. Calculate standard deviation using the command `STDEV.P(Secchi value in all rows)`.
3. Use the command `CONFIDENCE(alpha, standard deviation,n)`. Alpha = 0.05 for a 95% confidence interval, standard deviation is the value in the cell just calculated with `STDEV.P` and n = number of rows with a Secchi value. The number produced by `CONFIDENCE` is the absolute value of the number to subtract/add to the overall mean to obtain the confidence interval.

PART 2—APPLICATION OF MODIS DATA FOR REMOTE WATER-CLARITY MONITORING

Uncited methods, findings and explanations in Part 2 are based on McCullough et al. (2012b, in review b).

General Methods

What are the advantages of MODIS over Landsat?

Landsat is the primary data source for monitoring remote water clarity in Maine because more than 1,000 lakes can be simultaneously assessed in either path 11 or 12; however, clear Landsat imagery is available irregularly. The twice-daily MODIS image capture is an advantage over Landsat, and MODIS 250- and 500-m imagery can be used to estimate water clarity of 364 (≥100 ha) and 83 large lakes (≥400 ha), respectively in Maine (Figure 4). For purposes of estimating water clarity, the advantage of 500-m over 250-m data is that 500-m imagery contains both the visible blue and red bands, whereas 250-m imagery contains only the red band. Although the large pixel size restricts analysis to large lakes, the frequent image capture allows monitoring to occur annually or intra-annually during spring and summer (May–September), whereas Landsat analyses are more sporadic. In May, however, adequate calibration data often are lacking and some lakes conceivably could still be frozen, so remote monitoring in May is relatively unreliable. It also is unlikely that sufficient calibration data would exist to create an October model; our three attempts were unsuccessful.

Why consider 250-m instead of 500-m MODIS imagery?

Although MODIS 500-m imagery enables remote monitoring of 83 lakes during May through September, these 83 lakes represent a restricted sample of large lakes ≥ 400 ha. MODIS 250-m imagery permits remote monitoring of 364 lakes ≥ 100 ha (Figure 4), more than quadrupling the number of eligible lakes; however, 250-m imagery does not contain the visible blue band, an important predictor of SDD. Without the blue band, accurate remote monitoring of lake clarity is unreliable during early to mid-summer or periods when algal communities are not well developed. This is because the short wavelength of the visible blue band penetrates

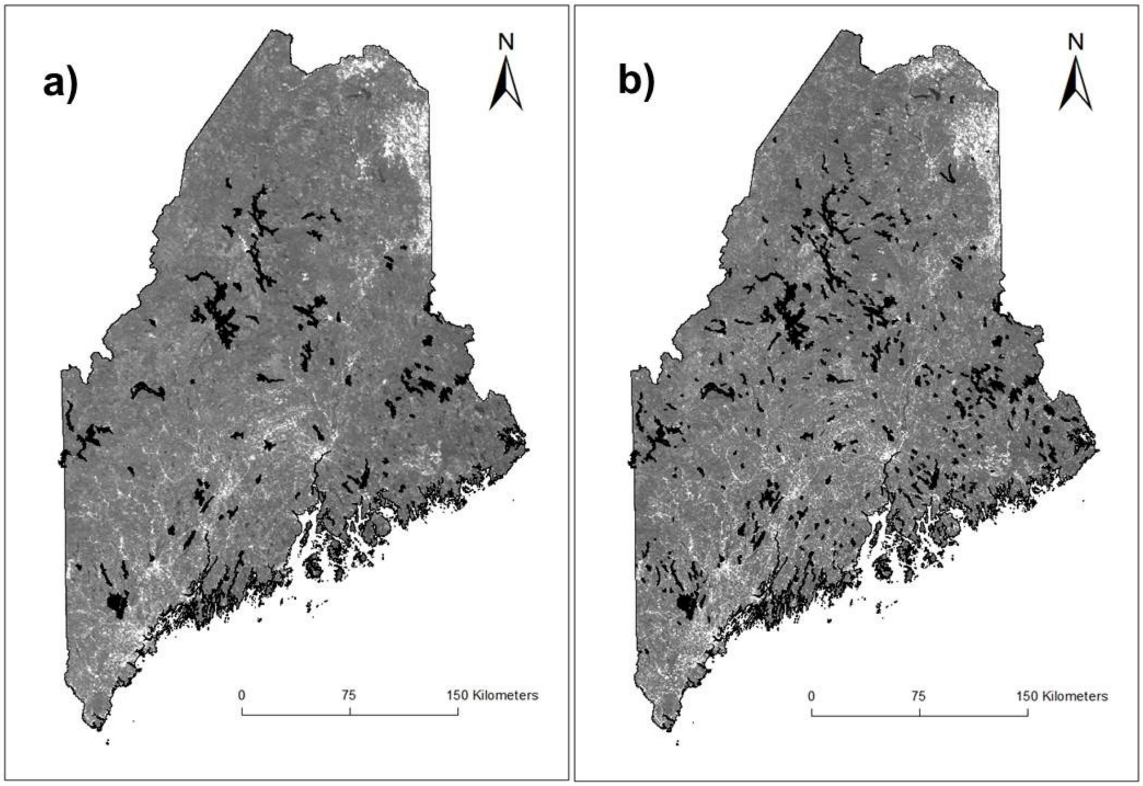

Figure 4. (a) Map of Maine's 83 lakes eligible for monitoring with MODIS 500-m imagery and (b) map of 364 Maine lakes eligible for monitoring with MODIS 250-m imagery.

only relatively clear water, conditions more likely to occur early in the growing season. Subsequently, during productive periods (e.g., late summer), the blue band is less useful than the red band in monitoring remote lakes and may even be expendable, enabling use of 250-m imagery during these periods. MODIS 250-m imagery also contains a near infrared band, but this band is not useful for remote sensing of lake clarity.

Lake eligibility requirements

Eligible lakes must be sufficiently large to contain three to five contiguous, water-only pixels. Some lakes as large as 500 ha with particularly jagged, convoluted shorelines are still unsuitable for remote monitoring with MODIS data owing to lack of water-only pixels. Of the 364 Maine lakes eligible for monitoring with 250-m imagery, some lakes are as small as 100 ha, representing 73% of Maine lakes ≥ 100 ha. Olmanson et al. (2011) reported that 125 ha was the eligibility cutoff, at which > 50% of Minnesota lakes were eligible for monitoring with 250-m data. The 83 lakes eligible for monitoring with 500-m imagery represent 49% of Maine lakes ≥ 400 ha.

Image selection

Level 1B Daily Surface Reflectance products (Aqua: MYD09; Terra: MOD09) are appropriate for this application. Imagery can be freely downloaded from the USGS Global Visualization Viewer (glovis.usgs.gov). If using 250-m imagery, select clear or mostly clear imagery of scenes captured during late summer (i.e., August to early September) owing to the absence of the blue band. This is the same time period targeted for Landsat-based change detection analyses (Part 1). Owing to loss of the blue band, lakes require well-developed algal communities that represent peaks or near-peaks of lake primary productivity prior to fall turnover. We previously used imagery captured during August 7 to September 1 for years 2000 to 2011. Imagery captured 8-1-2001 did not contain representative late summer lake conditions, but future imagery captured on or around this date may not necessarily yield similar results. If using 500-m imagery, images captured during May through September are usable, though May is less reliable owing to sparse calibration data.

Is MODIS a viable alternative to Landsat?

The inherent size differences between Landsat and MODIS-eligible lakes, the larger spatial scale of MODIS-based analyses and the fact that there are at most 364 MODIS-eligible lakes all render comparisons between MODIS and Landsat somewhat difficult (a Landsat path can assess more than 1,000 Maine lakes). MODIS models are calibrated for all of Maine, whereas Landsat models are calibrated only for Landsat path 11 or 12. When available, Landsat data are a considerably better data source than MODIS based on spatial resolution, but given limitations of Landsat image availability and uncertainty surrounding current and future Landsat satellites, successful development of an alternative, cost-effective approach for regional remote lake monitoring is potentially significant. Because 364 Maine lakes can be assessed with 250-m imagery, this data source is likely a better alternative to Landsat than 500-m imagery.

We cannot consider MODIS 250-m data a practical alternative to Landsat unless both produce similar results. We compared SDD estimates derived from two sets of concurrent Landsat and MODIS imagery (8-26-2000, 8-17-2011). The Landsat image from 8-17-2011 contained considerable fog and initially caused disagreement between Landsat and MODIS, but removal of foggy lakes established strong agreement ($t = 0.6891$, $df = 209$, $p = 0.492$). The Landsat image from 8-26-2000 contained only minor fog, but we nonetheless encountered strong disagreement between MODIS- and Landsat-derived SDD estimates. We suspected that the coarse resolution of MODIS imagery was failing to detect small patches of algal growth, so we resampled the Landsat red band (30 m) to 250 m, after which we found strong agreement between MODIS and Landsat ($t = -0.3696$, $df = 283$, $p = 0.713$). To test this hypothesis further, we resampled the Landsat red band in the 2011 image and also found strong agreement, even when including foggy lakes ($t = 0.2074$, $df = 277$, $p = 0.837$). Our findings of agreement in both years after upscaling suggest that relatively coarse MODIS 250-m data are less sensitive than Landsat data to localized areas of algal growth and fog. This reduced sensitivity results in deeper SDD predictions derived from MODIS than from Landsat. Average MODIS-estimated SDD exceeded Landsat-estimated SDD by 0.35 m (2000) and 0.49 m (2011) prior to resampling, whereas differences were 0.02 m (2000) and 0.04 m (2011) afterward.

MODIS 250-m imagery can be used for remote lake monitoring during late summer (or during times of algal abundance) with the caveat that MODIS-based estimates may potentially overpredict SDD. Furthermore, MODIS-based analyses are inherently biased toward large lakes, so assessing statewide lake water quality with MODIS data alone requires caution. Random samples of lakes would include numerous lakes < 100 ha. A study of Wisconsin lakes concluded that assessments of regional water quality characteristics are influenced by inclusion of small lakes in samples (Hanson et al. 2007).

Although we might expect similar disagreement with 500-m imagery, we actually found no statistically significant differences between Landsat and MODIS-based SDD estimates derived from four sets of concurrent imagery. Although annual means in statewide clarity based on remote SDD estimates differed 0.01 to 0.33 m, estimates on individual lakes were quite variable. If common calibration datasets could be used for both Landsat and MODIS models, resulting models and respective estimates may be more consistently similar. The smaller sample size in these analyses could explain the inconsistencies between 250-m and 500-m imagery. Additionally, 500-m imagery fortunately contained no fog, another potential source of MODIS-Landsat disagreement. Despite our findings of agreement, the fact that only 83 lakes can be assessed with MODIS 500-m imagery renders these data an impractical alternative to Landsat.

Methodological differences from Landsat

The methods used to calibrate remote SDD estimation models with MODIS data are similar to Landsat methods, though with fewer steps. Steps that are largely the same are not repeated here and readers are referred to Part 1. Although MODIS imagery must be reprojected to WGS1984 UTM Zone 19N, the freely downloadable MODIS Reprojection Tool simplifies this step. The Level 1B daily surface reflectance product that we use in this manual does not require atmospheric corrections (e.g., radiometric normalization), thus considerably decreasing image processing time. The reduced processing time of MODIS data is another notable advantage over Landsat. Clouds can be removed by unsupervised classification as with Landsat imagery, but the twice-daily image capture allows users to be especially selective and choose completely or nearly cloud-free imagery. Models are calibrated in the same manner as Landsat models, although time windows of ±3 days or more usually are necessary when using 500-m imagery, given that there are only 83 eligible calibration lakes and many of these rarely are assessed in the field. As a result, 500-m models are generally calibrated with relatively small calibration datasets and there is increased risk of over-fitting models. An over-fit model is a model that closely fits the calibration data, but not the overall population of lakes. The simplest way to determine if a model is over-fit is to assess the feasibility of the SDD estimates it produces (ask the question: do they make sense?). The jackknifing validation procedure also guards against over-fit models. When using 250-m imagery, windows of ± 1 to 3 days are sufficient and there is relatively less risk of over-fitting.

Technical Methods

Initial steps

1. CREATE A WORKING DIRECTORY. If analyzing multiple dates of imagery within a single year, create subfolders for each date. Save all files associated with each date in each respective subfolder. Create a separate map file for each date.
2. ACQUIRE IMAGERY. Download desired Level 1B daily surface reflectance imagery (Terra: MOD09GA or Aqua: MYD09GA) from the USGS Global Visualization Viewer (glovis.usgs.gov). Scenes H/V 13-4 and 12-4 cover Maine. Because images from Aqua and Terra are captured on the same day, there is generally little difference between images from the same day. Aqua imagery, however, is captured during the afternoon, after which morning clouds/fog may have dissipated. The twice-daily image capture frequency allows selection of cloud-free or nearly cloud-free imagery. You will receive an email when images are ready for download. Save images according to date in your working directory.
3. DOWNLOAD AND INSTALL THE MODIS REPROJECTION TOOL. (https://lpdaac.usgs.gov/tools/modis_reprojection_tool).
4. REPROJECT IMAGERY TO WGS1984 UTM ZONE 19N. Open the MODIS Reprojection Tool. Use default settings other than those directed here (Figure 5). Click *Open Input File* and navigate to the downloaded image. All bands will be automatically placed under *Selected Bands;* remove all except band 1 (sur_refl_b01_1; visible red). Click *Specify Output File* and create a folder in your working directory called "Reprojected." Name the file B1_81711. Set *Output File Type* to GEOTIFF. Set *Resampling Type* to Nearest Neighbor. Set *Output Projection Type* to UTM. Click *Edit Projection Parameters,* select WGS84 as the datum, type 19 in the UTM Zone box and click *OK*. By leaving *Output Pixel Size* blank, the reprojection will default to *500* m (there is no point in making a smaller pixel size; the resolution of the data will not change and you will only increase file size). Click *Run* to process the reprojection. If analyzing 500-m imagery, repeat this process for band 3 (sur_refl_b03_1; visible blue) and then for the other image of Maine, saving as slightly different file names.

Image processing

1. MOSAIC IMAGES. Complete for bands 1 and 3 (1 only if using 250-m imagery). Follow step 4 under "Part 1: Technical Methods: Image pre-processing and cloud removal."
2. CLIP OUT MAINE. Use the *Extract by Mask* tool (under *Spatial Analyst Tools, Extraction*) to remove areas outside Maine from analysis. Use the band 1 mosaicked surface reflectance image as the Input raster and the state shapefile of Maine as the mask. Create a subfolder "Cut," and save extracted band 1 files as "B1_81711_cut." Repeat for band 3, if necessary.
3. CLOUD REMOVAL. Because we download only the best-quality MODIS imagery, cloud removal may not be necessary. If small clouds exist, they can be identified by using the *Iso Cluster Unsupervised Classification* tool, reclassified as null values (Reclassify tool) and removed from analysis using the Extract by Mask tool. Refer to steps 6-8 from "Part 1: Technical Methods: Image pre-processing and cloud removal" for greater detail. There is no thermal image (RGB 1, 6, 6) as with Landsat in this case.

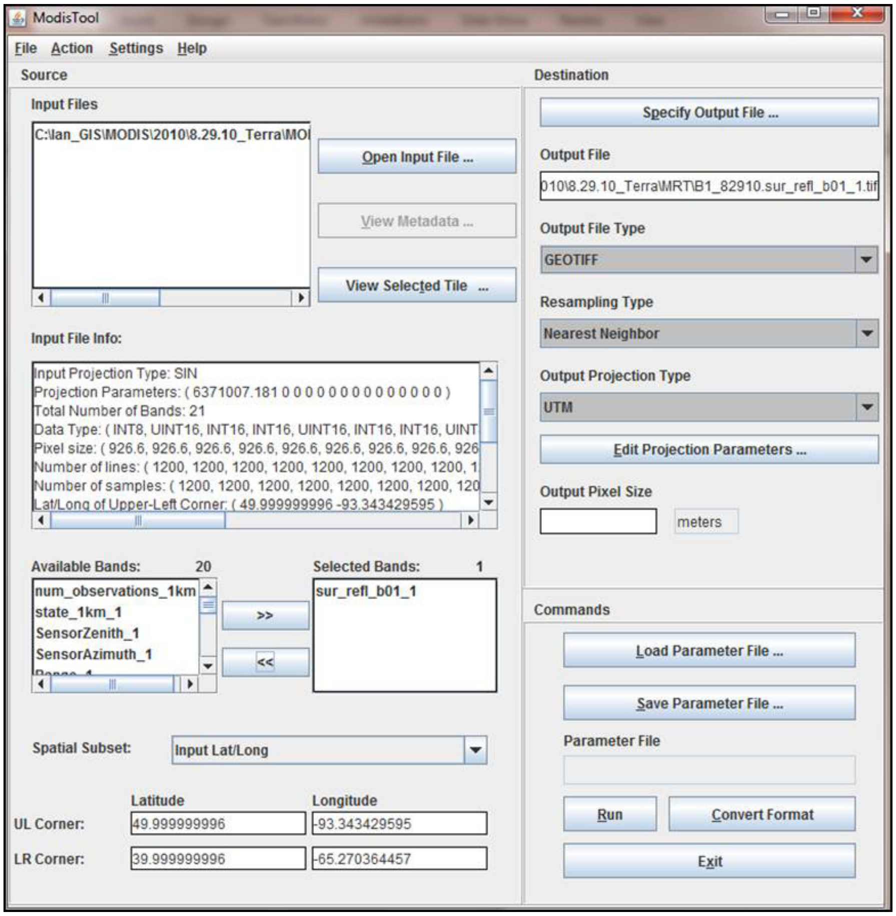

Figure 5. Screenshot of MODIS Reprojection Tool being used to reproject the visible red band of a MODIS image to WGS 1984 UTM Zone 19N.

4. ATMOSPHERIC CORRECTION. Corrections are included in Level 1B daily surface reflectance products.

Model development and execution

5. MODEL CALIBRATION. Refer to "Part 1: Technical Methods: Model calibration and development" section. The only difference is to use the MODIS_500m_Lakes (or MODIS_250m_Lakes) and the MODIS_500m_SamplingPts (or 250 m) layers in place of the Landsat layers. In the Landsat methods, we caution against substituting different sample stations within the same lake for each other in model calibration. With MODIS 500-m data, many existing sample stations are located too close to the shoreline for accurate remote sampling. Owing to the loss of sample sites and the coarseness of 500-m data, it is acceptable to cautiously substitute sample stations within lakes at this resolution if you are comfortable assuming that the bathymetry of substituted stations is roughly equal. When calculating zonal statistics, use MOD_ID as the *Zone* field. If using 250-m imagery, use IAN_ID as the *Zone* field. These IAN_ID values correspond to those in Landsat analyses.

6. ESTIMATE REGIONAL WATER CLARITY. Refer to step 12 of the "Part 1: Technical Methods: Model calibration and development" section. Analyze by trophic states or lake regions if desired.

7. MODEL VALIDATION. Model validation is important for MODIS-based models owing to their relatively small calibration datasets. Refer to section of the same name in "Part 1: Technical

Methods: Model validation." Use leave-one-out jackknifing if your calibration dataset contains < 50 data points.

QUALITY CONTROL MEASURES FOR SUCCESSFUL IMPLEMENTATION

1. Use both Landsat and MODIS, recognizing that each has separate applications in lake monitoring. Landsat can monitor more than 1000 lakes simultaneously, making it the primary satellite data source given the small size of most lakes in Maine. Landsat imagery is acceptable through 2011 and the successful launch of the Landsat Data Continuity Mission (LDCM) in 2013 would ensure future availability of Landsat data with potentially few gaps. Landsat data generally are available at least every 3 to 4 years for both paths 11 and 12, but the overlap region can be used if quality imagery from both paths is unavailable. MODIS 500-m imagery is not really a substitute for Landsat imagery, given that only 83 lakes can be reliably monitored using MODIS 500-m imagery; however, MODIS data are useful for within-year, seasonal monitoring of clarity of large lakes throughout Maine. MODIS 250-m imagery is a better substitute when Landsat data are unavailable, but declines in SDD are more difficult to detect.
2. The Landsat Multispectral Scanner (MSS) was revived to serve as a stopgap until the launch of the LDCM. The MSS can be used for regional water quality monitoring (Lillesand et al. 1983), but Landsat 5 is nearing the end of its life, LDCM is coming soon, and other alternatives exist during the interim (e.g., Landsat 7 and MODIS 250, 500-m imagery). The 57-m MSS resolution is coarser than that of Landsat 5 or 7 and would reduce the number of lakes eligible for monitoring. Development of a monitoring protocol for remote lakes using MSS data is not a priority.
3. Continue to publish a list of Landsat overpass dates for members of the VLMP and any others who collect water-clarity field data, emphasizing the importance of August and early September sampling. Target field-sampling dates as close as possible to the satellite overpass, though field sampling within 1 to 3 days of satellite overpass is acceptable.
4. Accurate models for estimating remote lake clarity require a numerically and geographically well-distributed set of calibration lakes. If possible, field sample lakes spanning a wide variety of SDD values over a large spatial extent in Maine.
5. If possible, survey more lakes for bathymetry. Depth is acceptably consistent year to year on a landscape scale such that future reassessment is not necessary. Depth also is useful for interpreting predicted SDD. For example, a SDD estimate of 2 m has different implications in lakes with average depths of 3 and 10 m.
6. Although remote analyses may be conducted months after actual image capture date, valuable information may be obtained from follow-up field visits to lakes remotely identified as undergoing eutrophication. These occasional field trips could serve to validate model findings and identify water-clarity drivers of individual lakes and watersheds missed by the coarse remote-sensing methods.
7. Beware of fog. Manually inspect images for foggy areas, which may not be eliminated by the unsupervised classification (cloud removal) procedure. Fog is more common over large lakes and along the coast. Undetected fog results in artificially shallow SDD estimates. Be especially careful when using Landsat 7 images because the scan lines may obscure fog.

LITERATURE CITED

Bacon, L., and R. Bouchard. 1997. Geographic analysis and categorization of Maine lakes: A trial of the draft bioassessment and biocriteria technical guidance. Maine Department of Environmental Protection, Augusta.

Boyle, K.J., P.J. Poor, and L.O. Taylor. 1999. Estimating the demand for protecting freshwater lakes from eutrophication. *American Journal of Agricultural Economics* 81(5): 1118–1122.

Carlson, R.E. 1977. A trophic state index for lakes. *Limnology and Oceanography* 22(2): 361–369.

Davis, R.B., J.H. Bailey, M. Scott, G. Hunt, and S.A. Norton. 1978. Descriptive and comparative studies of Maine lakes. Life Sciences and Agricultural Experiment Station Technical Bulletin 88.

Detenbeck, N.E., C.A. Johnston, and G.J. Niemi. 1993. Wetland effects on lake water quality in the Minneapolis/St. Paul metropolitan area. *Landscape Ecology* 8(1): 39–61.

Gibbs, J.P., J.M. Halstead, K.J. Boyle, and J. Huang. 2002. An hedonic analysis of the effects of lake water clarity on New Hampshire lakefront properties. *Agricultural and Resource Economics Review* 31(1): 39–46.

Hanson, P.C., S.R. Carpenter, J.A. Cardille, M.T. Coe, and L.A. Winslow. 2007. Small lakes dominate a random sample of regional lake characteristics. *Freshwater Biology* 52:814–822.

Heiskary, S.A., and W.W. Walker. 1988. Developing phosphorus criteria for Minnesota lakes. *Lake and Reservoir Management* 4(1): 1–9.

Kloiber, S.M., P.L. Brezonik, L.G. Olmanson, and M.E. Bauer. 2002. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sensing of Environment* 82:38–47.

Lillesand, T.M., W.L. Johnson, R.L. Deuell, O.M. Lindstrom, and D.E. Meisner. 1983. Use of Landsat data to predict the trophic state of Minnesota lakes. *Photogrammetric Engineering and Remote Sensing* 49(2): 219–229.

Maine PEARL. 2011. Lakes guide. Senator George J. Mitchell Center for Environmental Research, University of Maine, Orono. www.umaine.edu/waterresearch/pearl/ [Accessed 1/18/11]

McCullough, I.M., C.S. Loftin, and S.A. Sader. 2012a. Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity. Remote Sensing of Environment 123:109–115.

McCullough, I.M., C.S. Loftin, and S.A. Sader. 2012b. High-frequency remote monitoring of large lakes with MODIS 500 m imagery. *Remote Sensing of Environment* 124:234–241.

McCullough, I.M., C.S. Loftin, and S.A. Sader. In review a. Application of Landsat TM imagery reveals declining clarity of Maine's lakes during 1995–2010. *Freshwater Science*.

McCullough, I.M., C.S. Loftin, and S.A. Sader. In review b. Lakes without Landsat? An alternative approach to remote lake monitoring with MODIS 250 m imagery. *Lake and Reservoir Management*.

Michael, H.J., K.J. Boyle, and R. Bouchard. 1996. Water quality affects property prices: a case study of selected Maine lakes. Maine Agricultural and Forest Experiment Station Miscellaneous Report 398.

Nelson, S.A.C., P.A. Soranno, K.S. Cheruvelil, S.A. Batzli, and D.L. Skole. 2003. Regional assessment of lake water clarity using satellite remote sensing. *Journal of Limnology* 62:27-32.

Olmanson, L.G., M.E. Bauer, and P.L. Brezonik. 2008. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sensing of Environment* 112:4086–4097.

Olmanson, L.G., P.L. Brezonik, and M.E. Bauer. 2011. Evaluation of medium to low resolution satellite imagery for regional lake water quality assessments. *Water Resources Research* 47 W09515, doi 10.1029/2011WR011005.

Peckham, S.D., and T.M. Lillesand. 2006. Detection of spatial and temporal trends in Wisconsin lake water clarity using Landsat-derived estimates of secchi depth. *Lake and Reservoir Management* 22(4): 331–341.

Stadelmann, T.H., P.L. Brezonik, and S.M. Kloiber. 2001. 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 and Reservoir Management* 17(4): 299–314.

Wagner, T., P.A. Soranno, K.S. Cheruvelil, W.H. Renwick, K.E. Webster, P. Vaux, and R.J.F. Abbitt. 2007. Quantifying sample biases of inland lake sampling programs in relation to lake surface area and land use/cover. *Environmental Monitoring and Assessment* 141(1–3): 131–147.