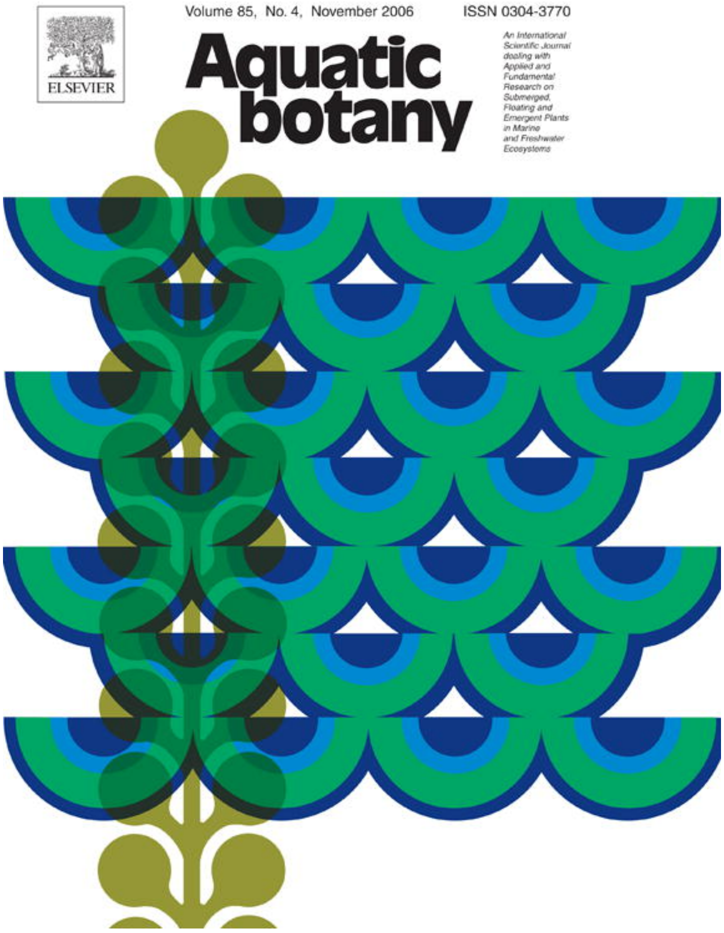
Volume 85, No. 4, November 2006
ISSN 0304-3770
ELSEVIER
Aquatic botany
An International Scientific Journal dealing with Applied and Fundamental Research on Submerged, Floating and Emergent Plants in Marine and Freshwater Ecosystems

Available online at www.sciencedirect.com
ScienceDirect
Aquatic Botany 85 (2006) 289–298
Aquatic botany
www.elsevier.com/locate/aquabot

Satellite remote sensing of freshwater macrophytes and the influence of water clarity

Stacy A.C. Nelson[a,*], Kendra Spence Cheruvelil[b,c], Patricia A. Soranno[c]

[a] *Center for Earth Observation and the Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA*
[b] *Lyman Briggs School of Science, Michigan State University, East Lansing, MI 48824, USA*
[c] *Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI 48824, USA*

Received 13 September 2004; received in revised form 8 May 2006; accepted 1 June 2006

Abstract

In regions with thousands of lakes, large scale regional macrophyte surveys are rarely done due to logistical difficulties and high costs. We examined whether remote sensing can be used for regional monitoring of macrophytes in inland lakes using a field study of 13 lakes in Michigan, USA (nine model development lakes and four model testing lakes). Our objectives were: (1) to determine if different levels of macrophyte cover, different growth forms or specific species could be detected using the Landsat-5 TM sensor, and (2) to determine if we could improve predictions of macrophyte abundance and distribution in lakes by including sediment type or measures of water clarity (Secchi disk transparency, chlorophyll *a*, phytoplankton biovolume, or water color) in our models. Using binomial and multinomial logistic regression models, we found statistically significant relationships between most macrophyte measures and Landsat-5 TM values in the nine model development lakes (percent concordant values: 58–97%). Additionally, we found significant correlations between three lake characteristics and the TM values within lake pelagic zones, despite the inability of these variables to improve model predictions. However, model validation using four lakes was generally low, suggesting caution in applying these models to other lakes. Although the initial model development results suggest that remote sensing is a potentially promising tool for regionally assessing macrophytes, more research is necessary to refine the models in order for them to be applied to unsampled lakes.

Keywords: Remote sensing; Landsat TM; Macrophyte cover; Regional scale; Eurasian watermilfoil; Littoral zone

1. Introduction

Large scale regional macrophyte surveys are rarely possible, even though effective macrophyte management depends in part on understanding the coverage and abundance of macrophytes, the growth forms present or the species present. This lack of survey data is largely due to the expense and challenges associated with sampling macrophytes. Assessments are further complicated in regions containing thousands of lakes. For example, the state of Michigan has approximately 3500 inland lakes >10 ha in surface area and many thousands of smaller lakes. Currently, it is not possible to inventory this many lakes using traditional field sampling techniques such as sampling along transects, within quadrants, or subsampling randomly-stratified lake points. Although these techniques can give good estimates of local macrophyte biomass and species composition at selected sites within a lake, these methods cannot capture whole-lake plant biomass/cover or the patchy distribution of aquatic macrophytes in an entire lake (Zhang, 1998). Remote sensing has the potential to be an important tool to obtain survey information on macrophytes within large geographic areas (Valley et al., 2005; Vis et al., 2003).

Typically, remote sensing has been used to measure macrophyte cover by the labor-intensive process of mapping macrophyte areal distributions along coastal margins using visual interpretations of aerial photographs (Orth and Moore, 1983; Marshall and Lee, 1994). Unfortunately, this approach has limited applicability for assessing macrophyte distributions in large regions with many water bodies. Therefore, an approach that can accommodate many lakes is needed for regional lake monitoring of macrophytes. One approach is to

* Corresponding author. Tel.: +1 919 513 7162; fax: +1 919 515 3439.
E-mail address: sanelso2@ncsu.edu (S.A.C. Nelson).

0304-3770/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.aquabot.2006.06.003

use Landsat Thematic Mapper (TM) satellite images that include many lakes within a single image. Although Landsat was primarily designed for detecting land features, recent improvements provide better spatial and spectral resolutions that may be applicable for aquatic studies (Zilioli, 2001). However, satellite remote sensing of aquatic macrophytes, especially submersed macrophytes, has been less studied than terrestrial vegetation because of the difficulties inherent in interpreting reflectance values of water (Penuelas et al., 1993; Lehmann and Lachavanne, 1997). For example, clear water provides little atmospheric reflectance and either absorbs or transmits the majority of incoming radiation (Lillesand and Kiefer, 1994; Verbyla, 1995). As a result, researchers have used remotely sensed data to detect primarily emergent vegetation or dense homogenous clusters of submersed vegetation, often only in single water bodies (Ackleson and Klemas, 1987; Armstrong, 1993). Despite the potential limitations of using current sensors such as Landsat to detect submersed aquatic macrophytes, more research is clearly needed to determine whether this sensor can be used to assess macrophyte abundance and distribution in large numbers of lakes across a large geographical region.

Additionally, lake characteristics may need to be taken into consideration when attempting to remotely sense macrophytes. For example, several studies have shown that remotely sensed images can measure lake characteristics such as chlorophyll, Secchi disk transparency, and suspended sediments (Lathrop and Lillesand, 1986; Jensen et al., 1993; Narumalani et al., 1997; Lillesand et al., 1983; Khorram and Cheshire, 1985; Dekker and Peters, 1993; Kloiber et al., 2000), all of which may influence the detection of macrophytes, especially submersed macrophytes. Lakes within a region can vary widely in several of the above characteristics, which can also influence how aquatic macrophytes are remotely sensed. For example, because water color and lake water depth may influence the sensor's ability to detect macrophytes, it may be necessary to incorporate such factors into predictive models of aquatic macrophytes. Although water depth has been incorporated into models to detect submersed macrophytes using sensors such as Landsat in individual water bodies (Raitala and Lampinen, 1985; Ackleson and Klemas, 1987; Armstrong, 1993; Narumalani et al., 1997), no studies have linked the influence of this feature or other physical and chemical lake features with the ability to detect submersed macrophytes in multiple lakes at a regional scale.

Our objectives in this study were: (1) to determine if different levels of aquatic plant cover and plant types (cover of: overall littoral macrophytes, emergent plants, floating leaf plants, submersed vegetation, and submersed *M. spicatum*) could be detected using the Landsat-5 TM sensor and (2) to determine if we could improve predictions of lake macrophyte abundance and distribution by including lake characteristics (Secchi disk depth, chlorophyll *a*, phytoplankton biovolume, water color, sediment type, and water depth) in the models. We hypothesize that including these additional lake characteristics will strengthen relationships between macrophyte cover and Landsat-5 TM spectral values.

2. Materials and methods

2.1. Study area

Using a random stratified design, we selected study lakes distributed throughout an area of approximately 70,000 km^2, located in the lower peninsula of Michigan, USA (Fig. 1). The lakes were stratified by surface area (20–140 ha), mean depth (2–10 m), and Secchi disk depth (0.6–8.2 m). A total of 36 lakes were selected and sampled. However, due to cloudy weather during late summer 2001, the final analyses were conducted on only 13 lakes for which cloud-free imagery could be obtained. Four of our 13 sample lakes were randomly selected and withheld from the model development for use as a validation dataset (Table 1). Therefore, nine lakes were used for model development.

2.2. Macrophyte sampling

The study lakes were sampled during the summer-stratified season and peak plant biomass (July 17–September 12, 2001). Macrophytes were sampled using a modification of the point intercept method (Madsen, 1999). Each lake was mapped using a geographic information system (GIS) and then overlain with a grid of points which represented the macrophyte survey points. The sample points were located in the field with a global positioning system (GPS), and plant cover was assessed at each point for an area of 40 m × 40 m or 50 m × 50 m depending on lake area, resulting in 138–467 points per lake. At each point, we measured water depth and assessed plant composition by recording plant presence in each of four categories: emergent, floating leaf, total submersed, and exclusively submersed,

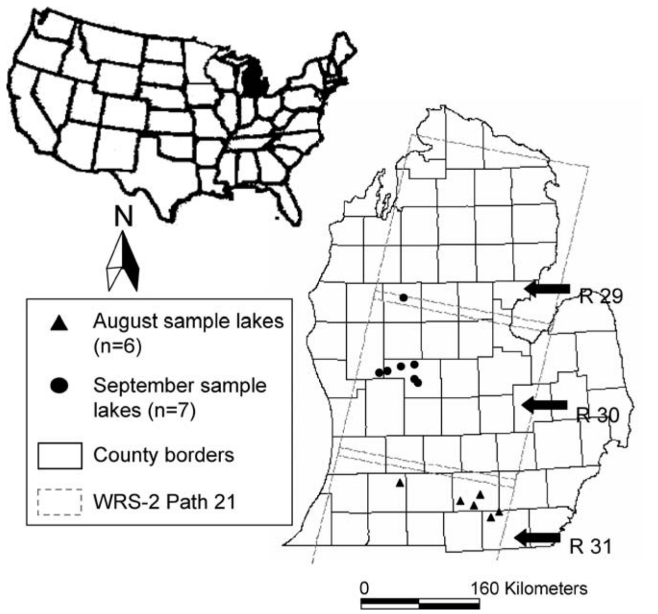

Fig. 1. The study area and lakes within Path 21 of Landsat (Worldwide Reference System-2) in Michigan, USA. The three individual scene rows within Path 21 used in this study are designated as R29, R30, and R31. Images R29 and R30 were acquired September 5, 2001. Image R31 was acquired August 4, 2001.

Table 1
Physical characteristics and image information for the study lakes

Lake	County	Image date	Plant sampling date–image date	# Littoral zone sample points	Surface area (ha)	Mean depth (m)	Sediment type[a]
Deep	Lenawee	8-04-01	−10	44	26.1	7.3	Marl
Eagle	Kalamazoo	8-04-01	−5	38	29.1	5.2	Sand
Round[b]	Lenawee	8-04-01	−9	70	25.8	5.3	Marl
Round	Jackson	8-04-01	−18	191	62.4	2.4	Marl
Swains	Jackson	8-04-01	−12	93	30.8	4.3	Marl
Vandercook[b]	Jackson	8-04-01	−11	116	57.9	6.8	Rocks/cobble/gravel
Baptist	Newaygo	9-05-01	−7	72	34.9	6.9	Sand
Dickerson	Montcalm	9-05-01	−9	114	60.2	7.4	Rocks/cobble/gravel
Horseshoe	Montcalm	9-05-01	−11	47	37.9	5.1	Sand
Little Whitefish	Montcalm	9-05-01	−16	122	72.5	5.5	Marl
Nevins	Montcalm	9-05-01	−9	45	21.0	7.2	Marl
Sunrise[b]	Osceola	9-05-01	+7	62	82.4	8.4	Rocks/cobble/gravel
Winfield[b]	Montcalm	9-05-01	−11	137	47.3	4.5	Sand

[a] Littoral zone sediment type estimates were based on an average of six to ten sample sites per lake which were qualitatively assigned a category: sand, rocks/cobble/gravel, marl, or silt/muck/peat.
[b] Lakes used for model validation, thus were not used for model development.

non-native *Myriophyllum Spicatum*. Examples of the plant species included in the emergent category were *Typha latifolia*, *Pontederia cordata*, and *Scirpus* spp. Representative species of the floating leaf category included *Nuphar advena*, *Nymphaea odorata*, and *Brasenia schreberi*. Examples of macrophytes in the total submersed category included *Chara* spp., *Potamogeton* spp., as well as *M. spicatum*. Plant cover at each site was assessed by qualitatively assigning a 'plant cover level' for each category, which was done by visual inspection in shallow/clear water, or by throwing a two-sided rake in deeper/turbid water. Plant cover levels were: 0 (0–20% plant cover), 1 (21–40% plant cover), 2 (41–80% plant cover), and 3 (81–100% plant cover). An additional binomial category of total littoral zone plant cover was developed by combining the four levels recorded for each plant category at each point. This category captures littoral plant presence or absence at each point by assigning each site either a 0 (0–20% plant cover) or a 1 (21–100% plant cover).

We calculated littoral percent plant cover for each lake as the number of points sampled with any plant category greater than level 0 (i.e., >20% cover at an individual site), divided by the total number of points in the littoral zone. The littoral zone was defined as <4.5 m water depth. Sites with a depth of >4.5 m were regarded as pelagic where we assumed that reflectance of the water column would dominate the reflectance spectra necessary for submersed plant detection by Landsat (Lillesand et al., 1983; Kloiber et al., 2000).

2.3. Lake characteristics

We estimated lake water clarity with Secchi disk depth averaged across two measurements taken over the shady side of the boat. Pelagic water samples were taken from the deepest area of each lake for nutrients, chlorophyll *a*, total alkalinity, and total phytoplankton biovolume (Table 2). For each lake, the depth of the epilimnion was calculated with a temperature

Table 2
Limnological characteristics of the study lakes, including the measurements of water transparency included in predictive models

Lake	Image date	Secchi depth (m)	Water color[a] (Co-Pt units)	Chlorophyll *a* (μg/L)	Total phosphorus (μg/L)	Total nitrogen (μg/L)	Phytoplankton biovolume (μm^3/mL)
Deep	8-04-01	2.6	14	1.8	8	297	1,467,500
Eagle	8-04-01	2.0	5	3.5	25	182	1,415,600
Round Lenawee[b]	8-04-01	3.3	9	2.2	16	486	5,110,100
Round Jackson	8-04-01	2.0	7	1.6	15	462	498,300
Swains	8-04-01	1.6	10	3.0	20	448	7,318,100
Vandercook[b]	8-04-01	2.5	29	3.2	13	329	5,838,000
Baptist	9-05-01	4.5	5	3.3	23	275	10,591,300
Dickerson	9-05-01	2.7	20	2.5	19	385	4,695,700
Horseshoe	9-05-01	1.0	11	13.7	27	429	11,642,100
Little Whitefish	9-05-01	4.1	7	4.1	14	238	3,731,000
Nevins	9-05-01	2.8	3	2.0	6	277	3,189,400
Sunrise[b]	9-05-01	3.1	10	5.0	8	248	7,094,000
Winfield[b]	9-05-01	1.9	30	7.7	18	485	3,226,700

[a] Water color values were collected by MI-DEQ 1974-1981 (STORET), and were measured on the cobalt-platinum scale (Co-Pt units). This scale ranges from 0 in very clear lakes to 300 Co-Pt units in heavily stained bog waters, which have high concentrations of humic substances.
[b] Lakes used for model validation, thus were not used for model development.

profile, and a tube sampler was used to take an integrated epilimnetic sample. For chlorophyll analysis, water was filtered on site through a glass fiber filter (Whatman GF-C) and stored in dark containers until being returned to the lab and frozen. Chlorophyll *a* concentrations were determined fluorometrically with phaeopigment correction following 24 h extraction in methanol (Nusch, 1980). Total nitrogen was determined using a persulfate digestion followed by second derivative spectroscopy (Crumpton et al., 1992). Total phosphorus was determined using a persulfate digestion (Menzel and Corwin, 1965) followed by standard colorimetry (Murphy and Riley, 1962). Secchi disk depth was measured for each lake over the shaded side of the boat. Phytoplankton were preserved with 1% Lugol's solution and stored in the dark until further laboratory analysis occurred. For each lake, a 50 mL phytoplankton sample was settled for 5 days except for 3 lakes (Nevins, Little Whitefish, and Round (Jackson County)), which were only settled for 48 h. For these 3 lakes, we extrapolated to a 5-day count based on comparisons of the 2 settling lengths for 6 lakes that were counted after being settled for both 2 and 5 days. For all lakes, phytoplankton samples were counted and identified using an inverted microscope. For each lake, more than 400 individuals were counted using multiple magnifications (40, 100, 400, and 1000×), and phytoplankton densities (μm^3/mL) were then calculated. Finally, we developed sediment type estimates in the littoral zone by visually assessing 6–10 sites per lake. We categorized lake sediments into four different types that represent different 'colors' of sediment that may reflect light differently: sand, gravel/rock/cobble, marl, or silt/mulch/peat.

2.4. Satellite imagery

We used three Landsat-5 TM scenes: one from August 4, 2001 and two from September 5, 2001. All three scenes were from the Landsat ground track Path 21 (Worldwide Reference System-2; Fig. 1). Image rectification and geoprocessing was conducted using ERDAS Imagine 8.4 image processing software. Atmospheric corrections were not necessary for the two contiguous Landsat scenes (September) used in this study because the images were taken on a single date and showed no signs of being effected by atmospheric haze (i.e., uniformed elevation in spectral pixel values throughout each scene). Spectral pixel values within the entire August image also appeared similar in the two September images. This similarity in pixel reflectance values between all three images suggested no evidence of differential reflectance throughout the August image as a result of atmospheric haze conditions. However, spectral values extracted from the August and September images were independently statistically evaluated before combining the two individual data sets to develop a combined model data set.

The spectral pixel values or digital number (DN) values for all single pixels located at the same position as each lake sample point (using the field-recorded GPS coordinates within each lake) were extracted using the ArcInfo GRID module of ArcGIS 8.1 (ESRI) from the imagery. Spectral DN values for the pelagic region of each lake (sample points with depth measurements >4.5 m) were also extracted to analyze the relationship between pelagic zone lake characteristics and spectral values (see below). Because the pelagic zone is more homogeneous than the littoral zone, we averaged the spectral DN values for all pixels within the pelagic zone resulting in one pelagic spectral value.

2.5. Statistical analysis

We analyzed the satellite imagery DN values and macrophyte data using binomial and multinomial logistic regression (logit models) in SAS/STAT software (SAS Institute Inc., 1995). We included spectral DN values for TM bands 2 and 4 (TM2 and TM4) in the models as the independent variables based on numerous attempts to fit several individual and combined bands to the data using a series of stepwise regression techniques. In all models, we also included lake depth at each sampling point as an interaction term with each main effect variable (TM2 and TM4).

We included sample points from multiple lakes in all logit models, unless noted otherwise. However, because our images were from two different dates, we developed three separate models to assess whether including results from different dates (and potentially different atmospheric levels, i.e., haze) would change model predictions. Thus, we grouped the nine sample lakes into three datasets; two datasets according to the August and September image dates, and one dataset that combined both August and September lakes. The three datasets resulted in 366, 400, and 766 sample points for each logit model, respectively.

Each of the five macrophyte categories (overall littoral plant cover, emergent plants, floating leaf plants, submersed plants, and submersed *M. spicatum*) served as individual response variables for each of the three models. The logit model uses the explanatory and interaction covariates to predict the probability that the response variable will take on a given value (SAS Institute Inc., 1995). For binomial logistic regression, the logit model indicates how the explanatory variable (TM values) affects the probability of the event (macrophytes) being observed versus not being observed. For the multinomial logistic regression, probable outcomes of observations are calculated by analyzing a series of binomial submodels that represent the overall model's ability to predict each of the plant cover response variables. For all logit model analyses, we used the descending option to select the highest plant category level as the response variable reference (level 3 for emergent, floating leaf, total submersed, and submersed *M. spicatum* and level 1 for littoral plant cover). This selection ensured that the results were based on the probabilities of modeling an event (macrophytes present), rather than a non-event (macrophytes not present).

Model fit was determined by examining the percent concordant values and the Wald test statistic. The percent concordant values provide an indication of overall model quality through the association of predicted probabilities and observed responses. These values are based on the maximum likelihood estimation of the percent of paired observations of which values differ from the response variable (Kleinbaum, 1994). Thus, the higher the predicted event probability of the larger response variable (based on the highest plant category level), the greater the percent concordant value will be. The Chi-square level of significance for the Wald test statistic tests the hypothesis that the coefficients of the independent variables are significantly different from zero by fitting the model using the intercept terms (Kleinbaum, 1994; Pampel, 2000). The Wald test is regarded as being more accurate than other test statistics when large sample sizes, such as ours, are used (Kleinbaum, 1994).

To examine whether there were significant differences between data from the two image dates and whether it was valid to include all lakes in a combined model, we developed individual-lake logit models. We then compared the model output and model coefficients from lakes in the August image to lakes in September image. First, we used a two sample t-test to test for differences between the means of the model coefficients (log transformed) of the September lakes and the August lakes. Resultant p-values are for the paired variance and significance was determined at the 0.1 level. Insignificant results from these tests suggest that the means of the nine individual lakes showed no significant difference. Therefore, atmospheric effects may not significantly affect the datasets and it is valid to include lakes in a combined model across image dates. Second, we compared the means of the percent concordant values from the individual models for the September lakes to those from the August lakes. In this analysis, the absence of large differences between the September and August percent concordant values also supports the validity of including all lakes in a combined model across image dates.

Using the logit model for individual lakes, we also examined whether various lake characteristics can help improve predictions of plant cover using Landsat imagery. Using ordinary least squares regression, we regressed each of the four model coefficients (TM2, TM4 and the interaction terms with lake depth for both) from the individual lake logit models against each of the measured lake characteristics individually: Secchi disk depth, water color, chlorophyll a, total phytoplankton biovolume, and sediment type.

2.6. Logit model validation

We validated the results of the logit models by using field-collected data from four lakes not included in model development. The validation was developed by calculating the logit values for each sample point in the four validation lakes from the logistic (for plant cover) and bionomial (for all other plant categories) regression equations for both the September and combined models. The August model was not used for validation comparisons because the plant sample date of these validation lakes more closely matched the September dataset. The logit values represented the cumulative probability of each sample point being each plant cover level (0, 1, 2, and 3) within each plant category. The cumulative probability value of the logit was then used to calculate the actual probability of each sample point being each plant cover level. The actual probabilities were then averaged to determine the overall probability of sample points belonging in each plant cover level and plant category.

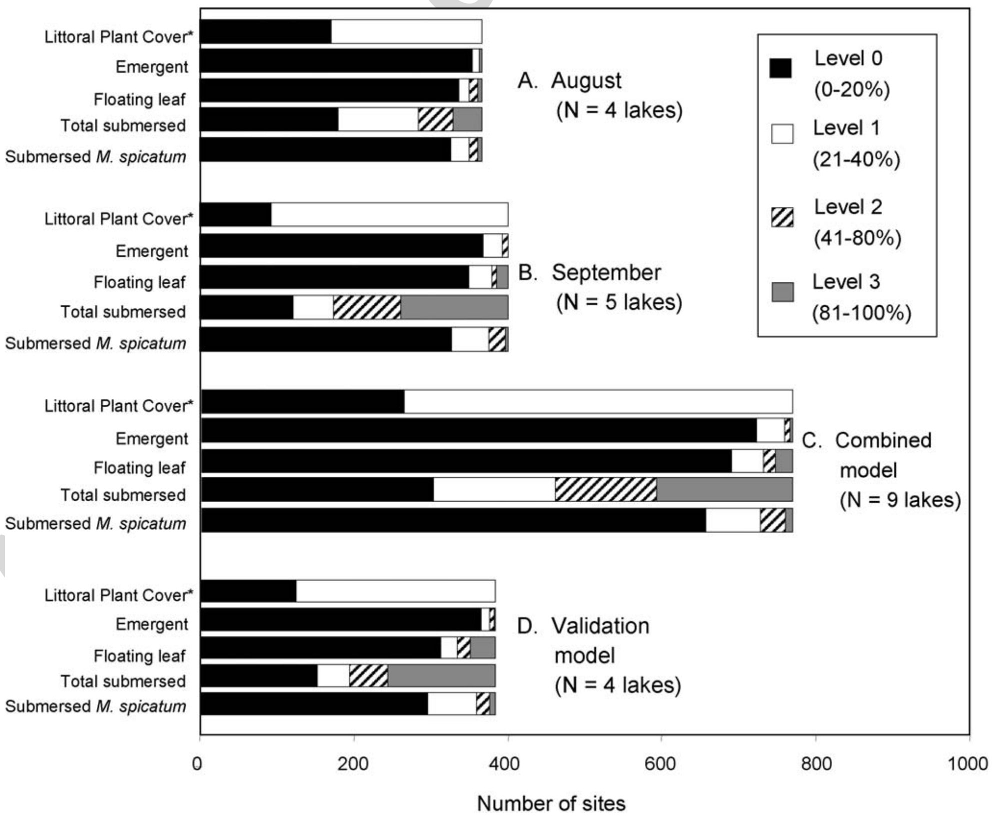

Fig. 2. The distribution of plant categories and plant cover levels within the littoral zone of the lakes comprising our three logit model datasets and validation lakes.

3. Results

3.1. Macrophyte cover in the study lakes

The 13 study lakes had sites with each of the four plant categories, although there was variation across all lakes as well as among categories within each lake. In particular, few lakes had very dense amounts of emergent and floating leaf vegetation as compared to submersed macrophytes. In addition, although *M. spicatum* was present in relatively low amounts as compared to total submersed macrophytes, it was present in all lakes. The overall percent cover of macrophytes in the study lakes ranged from 5 to 42%, and the percent cover of macrophytes in the littoral zone ranged from 10 to 90%. Four lakes had ≤50% plant cover in the littoral zone; in these lakes, the littoral zone was dominated by exposed lake sediments rather than macrophytes. Although the overall presence of macrophytes was quite common in all lakes, plant density at each sample point was quite low (Fig. 2). The majority of sample points had 0–20% plant cover (level 0) at a site, which led to skewed distributions of plant levels across the four categories. The total submersed plant category was the most evenly distributed category across all levels.

3.2. Detecting macrophyte cover with satellite imagery

The highest concordance values from the logit models of plant cover categories and Landsat DN values were for the emergent and floating leaf (91–98%) plant categories (Table 3).

Table 4
T-test *p*-values from testing the difference between the means of the beta coefficients (log transformed) for the 9 individual lakes used in the August and September logit models

Plant category	TM2	TM4	TM2 × depth	TM4 × depth
Littoral plant cover	0.31	0.23	0.42	0.16
Emergent	0.24	0.22	0.16	0.39
Floating leaf	0.21	0.34	0.33	0.32
Total submersed	0.28	0.42	0.64	0.54
Submersed *M. spicatum*	0.23	0.35	0.78	0.42

Lower concordance values were found for the two submersed categories, total submersed (69–73%) and submersed *M. spicatum* (58–61%). These results suggest that it may be easier to predict plant categories that are above the water surface than below the water surface. When all plant categories were aggregated into a single category (littoral plant cover), the logit models produced moderate percent concordance values (66–73%). The *p*-values for the Wald test were highly significant ($p < 0.001$) for all categories and tests.

To assess the validity of grouping lakes into a combined August/September model, we developed individual logit models for each image. We found no significant differences in any model coefficients between the August and September lakes (*t*-test $p = 0.16$–0.78 for all main effects and interactions terms), suggesting that the relationships between the means of the spectral DN values and plant cover estimates were similar across different images, despite any atmospheric differences (Table 4). Additionally, comparisons of the percent concordant

Table 3
Logit regression results for the three models

Plant category	Statistical test	August model	September model	Combined model
Littoral plant cover	Percent concordance[a]	66	73	71
	Wald test[b]	<0.001	<0.001	<0.001
	Number of lakes	4	5	9
	Number of sample points	366	400	766
Emergent	Percent concordance[a]	98	97	91
	Wald test[b]	<0.001	<0.001	<0.001
	Number of lakes	4	5	8
	Number of sample points	366	400	766
Floating leaf	Percent concordance[a]	97	95	96
	Wald test[b]	<0.001	<0.001	<0.001
	Number of lakes	4	5	9
	Number of sample points	366	400	766
Total submersed	Percent concordance[a]	71	69	73
	Wald test[b]	<0.001	<0.001	<0.001
	Number of lakes	4	5	9
	Number of sample points	366	400	766
Submersed *M. spicatum*	Percent concordance[a]	67	58	61
	Wald test[b]	<0.001	<0.001	<0.001
	Number of lakes	4	5	9
	Number of sample points	366	400	766

[a] Percent concordance assesses the overall model quality based on the maximum likelihood estimation of all pairs of observations with different values of the response variable (plant category). For example, because the models predict the probability of detecting a plant category, if the larger response value (>0 for a level of plant cover) has a higher predicted event probability than the smaller response value (0, non-event or <20% plant cover), then the observation pair is concordant. Pairs are calculated from the number of observations within each category level submodel, which makes up the overall logit regression for each plant category.
[b] The Wald test is the Chi-square statistic for overall model fit. It tests the hypothesis that the coefficient of an independent variable is significantly different from zero.

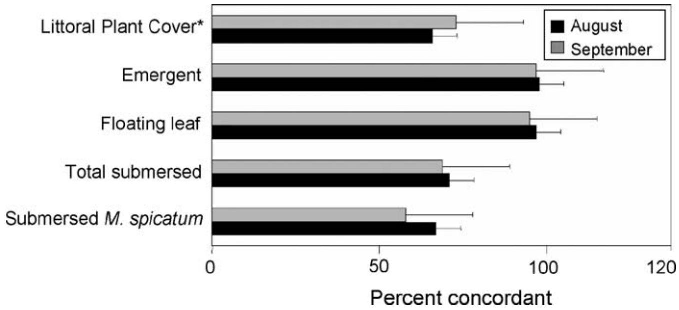

Fig. 3. The percent concordance for individual lake models averaged by image date.

values from the individual lake logit models suggest that the August and September models produced similar model fits to the data (Fig. 3). Given that there were no statistically significant differences between the two images, we concluded it was valid to combine all lakes into a single 'combined' model.

3.3. Improving predictions of macrophyte cover by including lake characteristics

The ordinary least squares regressions of the logit model coefficients estimated for the nine individual lake datasets with lake characteristics produced low r^2 values and insignificant p-values for the majority of lake characteristics. Total phytoplankton biovolume produced the highest r^2 values and only four phytoplankton biovolume-plant category combinations were significant at the 0.10 level (littoral plant cover, emergent, total submersed, submersed *M. spicatum*). This result suggests that knowing phytoplankton biovolume may help improve predictions of littoral plant cover using Landsat. However, although some relationships were significant, this result was not consistent for all independent variables tested. Because we found it surprising that additional lake characteristics did not appear to improve model predictions, we examined how well we could predict these lake characteristics from Landsat DN pelagic zone values. We found significant correlations between all variables and spectral DN values of the pelagic zone except for water color (Table 5). Thus, our regression results suggest that differences in the significant variables (Secchi depth, chlorophyll *a*, and phytoplankton biovolume) should be detectable within the nine lakes sampled.

Table 5
Regression results of Landsat TM2 and TM4 digital values vs. selected pelagic lake characteristics using the combined dataset of all lakes ($N = 9$)

Lake characteristics	r^2	p
Secchi depth	0.72	0.01
Water color	0.02	0.87
Chlorophyll *a*	0.65	0.01
Phytoplankton biovolume	0.54	0.03

See Table 2 for lake characteristic units.

Table 6
Results of the model validation using the 4 lakes not used in model development (see Table 1 for lake identities and characteristics)

Plant category	Cover level	September model: average probability of correctly classifying each level	Combined model: average probability of correctly classifying each level	No. of observed sites
Littoral plant cover	0	0.62	0.69	124
	1	0.26	0.24	261
	Average	0.44	0.47	
Emergent	0	0.05	0.01	367
	1	0.26	0.20	10
	2	0.69	0.32	7
	3	–	0.92	1
	Average	0.33	0.36	
Floating leaf	0	0.04	0.03	313
	1	0.04	0.08	23
	2	0.17	0.16	16
	3	0.46	0.45	33
	Average	0.18	0.18	
Total submersed	0	0.22	0.24	152
	1	0.23	0.22	42
	2	0.12	0.21	50
	3	0.29	0.20	141
	Average	0.22	0.22	
Submersed *M. spicatum*	0	0.00	0.01	297
	1	0.11	0.09	63
	2	0.65	0.67	17
	3	0.00	0.00	8
	Average	0.25	0.26	

3.4. Model validation

We validated the September and the combined models using field-collected data from four lakes. The distributions of overall plant categories and plant cover levels in our four validation lakes were similar to the distributions found in the nine lakes used to develop the models (Fig. 2). Similar to the nine model lakes, the majority of sample points had 0–20% plant cover at most site sampled, resulting in skewed distributions of plant levels across the four categories in our four validation lakes. The total submersed plant category was the most evenly distributed category across all levels.

Evaluation of the validation results showed very little difference between the September and the combined model (Table 6). We found relatively low, yet fairly similar probabilities of correctly classifying each multilevel plant cover category (emergent, floating leaf, total submersed, and submersed *M. spicatum*) and a somewhat higher probability for the aggregated category of littoral plant cover (plant cover presence or absence). This result may be influenced by the fact that we correctly classified level 0 at a higher probability level than we were able to correctly classify level 1. However, within two of the multilevel categories (emergent and floating leaf), we obtained the best probability of correct classification from the higher levels of plant cover (level 2 and 3). High classification levels were also seen for the level 2 submersed *M. spicatum* category, which was the highest level for this category (level 3 was of very low occurrence for this category in our validation lakes). The most evenly distributed range of probabilities occurred in the total submersed category; probabilities of classification ranged from 0.12 to 0.29 for the September model and 0.20–0.24 for the combined model.

4. Discussion

We found relatively strong relationships between the five plant cover categories and Landsat spectral TM values. For each plant cover category, we obtained significant model fits and reasonably high percent concordant values. Not surprisingly, the highest percent concordant values were for models of the plant categories with plant components above water: floating leaf and emergent macrophytes. Lower percent concordant values were seen for the submersed categories (total submersed and submersed *M. spicatum*). This result is expected considering the difficulties inherent in remotely sensing the aquatic environment (Lillesand and Kiefer, 1994; Verbyla, 1995). The fact that the models resulted in percent concordant values for the two submersed categories ranging from 58 to 73%, suggests that remote sensing using the Landsat sensor may prove to be a valuable tool for measuring plant abundance and distribution, including the spread of the non-native *M. spicatum*, in multiple lakes across a large region.

We did not find a consistent effect of lake characteristics on the ability to detect aquatic macrophytes in lakes. We found this result surprising given the wide range of water clarity that was present in the study lakes (Secchi depth range from 1.0 to 4.5 m). However, it is possible that with only nine lakes in the combined model (and 4–5 lakes in the single-date models), the sample size was too low to detect the effects of lake characteristics on predicting plant cover. Additionally, although other studies have found Landsat to be significantly correlated with a variety of water clarity and water quality parameters (Lathrop, 1992; Cox et al., 1998; Kloiber et al., 2000; Nelson et al., 2003), there may be sufficient variation around these relationships to make it difficult to quantify the effect that these lake parameters have on model predictions of aquatic macrophytes across a range of lakes.

An interesting finding of our study is the similar results obtained using the individual date models (August and September) and the combined date models. It is commonly accepted that values sampled from images collected on different dates should not be combined into a single dataset because they are affected by different atmospheric effects (i.e., haze) specific to the date of image acquisition (Song et al., 2001). However, we did not find this to be the case. In fact, atmospheric differences resulted in no statistically significant differences in model predictions between the datasets used to build or validate the predictive models. These results are important because they allowed us to include all lakes into a combined model, thus increasing our sample size, which also allowed us to better examine the effect of lake characteristics on the models.

The model validation step adds additional insights into the use of remote sensing to detect aquatic vegetation. The validation resulted in relatively low probabilities (0.44–0.47) of correctly classifying macrophytes in all plant categories and across all cover levels. This result was surprising because the models produced high overall percent concordant values and significant Wald test statistics, suggesting a good model fit with the data. Based on the distribution of the categories and levels of plant cover, the randomly selected validation lakes appear to be very similar to the model lakes. We can only speculate as to why the validation and the model statistics resulted in different interpretations. First, the four lakes used in the validation may have differed for some other important feature as compared to the model lakes. However, there was no statistical difference in the means between the lakes of the two groups for all of the key limnological characteristics measured in this study (see Table 2 for the list of variables; Nelson, unpublished data). Second, the models themselves, although significant, may be sensitive to the distributions of macrophytes within each category. Therefore, if the plant-cover level distributions of both the model datasets and the validation datasets had been more equally balanced among levels, higher probabilities in the validation may have resulted. In fact, except in the case of littoral plant cover, higher validation probabilities seem to be related to higher plant cover levels (Table 6), of which we had relatively few sample points. Thus, this potential statistical issue demonstrates that new analytical techniques to analyze this type of data may be needed. Finally, our lower concordance values may also have been related to variables within the macrophyte bed structure that could not be easily accounted for in our predictive models, such as variability within the plant cover categories. For example, biomass densities (biomass per unit volume) may differ according to growth forms within each category, such that species within our floating-leaf and canopy-forming submergent beds may have produced plants with variable biomass densities among stands of the same species or combination of species category (White, 1985; Lonsdale and Watkinson, 1983; White, 1981). Thus abundances could be highly variable if estimates are based on canopy cover, height, and differential shading effects as opposed to the degree of packing within the macrophyte stand (Duarte and Kalff, 1990).

Below we explore some possible limitations of our study. First, it is possible that we had low positional accuracy with our hand-held GPS unit. Horizontal accuracy of commercial-grade hand-held GPS units still varies from 7 to 15 m depending on environmental and satellite signal reception conditions. Second, there was the potential for sample point/image pixel misalignment, which sampling at a finer grid resolution in the field may have alleviated. However, it may not be feasible to sample additional points on multiple lakes. In addition, although we sampled a total of 36 stratified-randomly selected lakes for this study, cloud-free images were available for only 13 of these lakes. These lakes may not necessarily be representative of the overall 36 lakes. For example, nine of the 36 lakes for which we had no clear images had overall lake plant cover >42% (Cheruvelil, 2004). Therefore, additional clear images would have increased the number of lakes included in the models and resulted in more evenly distributed plant cover. The final unknown in this study is to what extent we reached the detection limits of the Landsat sensor to detect aquatic macrophytes. Unfortunately, we were not able to quantify any of the above sources of error in our study. However, despite these study limitations, our significant results warrant further examination of this important research area.

Our results demonstrate that the use of remote sensing in freshwater lake studies can play a vital role in reducing the cost, labor, and time required to monitor these systems over a large geographic area. Remote sensing also has the potential to be used as a tool for statewide assessments that are currently impossible using traditional field approaches. Here, we have provided a general method for detecting littoral zone plant cover in inland freshwater lakes using satellite imagery. However, the results of our validation suggest that attempts to predict macrophytes in unsampled lakes may be difficult without further research incorporating more lakes with even more varied levels of plant cover. Additionally, we offer the following suggestions. First, more sample lakes should be included in the model calibration to provide representative distributions of regional plant cover within the lakes modeled. Second, until newer sensors are developed for aquatic applications, inland lake remote sensing may be most useful as a supplement to existing volunteer and agency monitoring programs, rather than a replacement of these programs. Newer sensors, such as Space Imaging's IKONOS sensor and Digital Globe's Quickbird, show some promise by offering comparable spectral bands to the Landsat platform but at higher spatial resolutions (4 m and 2.44 m, respectfully). However, the trade off for regional remote sensing comes at a steep price, as data obtained from high resolution sensors are currently very costly as compared to available Landsat data. Additionally, because of the higher resolution, the total spatial coverage is more limited than with a single Landsat image, thus requiring more images. Finally, these high resolution data are not routinely collected and require additional programming or tasking fees to acquire data over the area of interest, making the use of this data fairly cost prohibitive for many inventory, management, or regulatory agencies forced to operate on limited budgets. Clearly, more research needs to be done to determine the optimal spatial resolution of sensors for aquatic macrophyte detection.

Acknowledgements

This research was funded in part by a Michigan State University Minority Competitive Doctoral Fellowship to S.A.C.N. and an EPA STAR Fellowship U-91595801-2 to K.S.C. Landsat-5 ETM+ imagery was provided by the Michigan State University Center for Global Change and Earth Observation and through the NASA-funded Upper Midwest RESAC project (Regional Earth Science Applications Center). We thank Kevin Hoffman, John Whitlock, Tamara Brunger Lipsey, Kristy Rogers, Sarah Wills, and Konrad Kulacki for field data collection and processing. We are also grateful to David Skole, Jiaguo Qi, Jianguo Liu, David Lusch, Samuel Batzli and Walter Chowmentowski for their contributions to this research. We also thank Michael Belligan, Marc Linderman, and Chris Barber for technical support and database assistance and John Madsen, Chetta Owens, Linda Nelson and Mike Stewart for sharing their endless macrophyte expertise. Additionally, we thank Daniel Hayes, Orlando Sarnelle, Ken Frank, Jim Bence, and Emily Szali for statistical advice. Finally we thank Jan Vermaat, and several anonymous reviewers whose comments have greatly improved this manuscript.

References

Ackleson, S.G., Klemas, V., 1987. Remote-sensing of submerged aquatic vegetation in lower Chesapeake Bay: a comparison of Landsat MSS to TM imagery. Remote Sens. Environ. 22, 235–248.

Armstrong, R., 1993. Remote sensing of submerged vegetation canopies for biomass estimation. Int. J. Remote Sens. 14, 621–627.

Cheruvelil, K.S., 2004. Examining lakes at multiple spatial scales: predicting fish growth, macrophyte cover and lake physio-chemical variables. Ph.D. Dissertation. Michigan State University.

Cox, R.M., Forsythe, R.D., Vaughan, G.E., Olmstead, L.L., 1998. Assessing water quality in the Catawba River reservoirs using Landsat Thematic Mapper satellite data. Lake Reserve. Manage. 14, 405–416.

Crumpton, W.G., Isenhart, T.M., Mitchell, P.D., 1992. Nitrate and organic N analysis with 2nd-derivative spectroscopy. Limnol. Oceanogr. 37, 907–913.

Dekker, A.G., Peters, S.W.M., 1993. The use of the Thematic Mapper for the analysis of eutrophic lakes: a case study in the Netherlands. Int. J. Remote Sens. 14, 799–821.

Duarte, C.M., Kalff, J., 1990. Biomass density and the relationship between submerged macrophyte biomass and plant growth form. Hydrobiologia 196, 17–23.

Jensen, J.R., Narumanlani, S., Weatherbee, O., Mackey Jr., H.E., 1993. Measurement of seasonal and yearly cattail and waterlily changes using multidate SPOT panchromatic data. Photogramm. Eng. Remote Sens. 52, 31–36.

Khorram, S., Cheshire, H.M., 1985. Remote sensing of water quality in the Neuse River Estuary North Carolina. Photogramm. Eng. Remote Sens. 51, 329–341.

Kleinbaum, D.G., 1994. Logistic Regression A Self-learning Text. Springer-Verlag Publishing, New York.
Kloiber, S.M., Anderle, T.H., Brezonik, P.L., Olmanson, L., Bauer, M.E., Brown, D.A., 2000. Trophic state assessment of lakes in the Twin Cities (Minnesota USA) region by satellite imagery. Archiv. Hydrobiologie Special Issues Adv. Limnol. 55, 137–151.
Lathrop, R.G., 1992. Landsat Thematic Mapper monitoring of turbid inland water quality. Photogramm. Eng. Remote Sens. 58, 465–470.
Lathrop, R.G., Lillesand, T.M., 1986. Utility of Thematic Mapper data to assess water quality. Photogramm. Eng. Remote Sens. 52, 671–680.
Lehmann, A., Lachavanne, J.B., 1997. Geographic information systems and remote sensing in aquatic botany. Aquat. Bot. 58, 195–207.
Lillesand, T.M., Johnson, W.L., Deuell, R.L., Linstrom, O.M., Meisner, D.E., 1983. Use of Landsat data to predict the trophic state of Minnesota lakes. Photogramm. Eng. Remote Sens. 49, 219–229.
Lillesand, T.M., Kiefer, R.W., 1994. Remote Sensing and Image Interpretation. John Wiley and Sons, Inc., New York.
Lonsdale, W.M., Watkinson, A.R., 1983. Plant geometry and self thinning. J. Ecol. 71, 285–297.
Madsen, J.D., 1999. Point intercept and line intercept methods for aquatic plant management. Army Corps of Engineers Waterways Experiment Station Technical Note MI-02.
Marshall, T.R., Lee, P.F., 1994. Mapping aquatic macrophytes through digital image analysis of aerial photographs: assessment. J. Aquat. Plant Manage. 32, 61–66.
Menzel, D.W., Corwin, N., 1965. The measurement of total phosphorus in seawater based on the liberation of organically bound fractions by persulfate oxidation. Limnol. Oceanogr. 10, 280–282.
Murphy, J., Riley, L.P., 1962. A modified single solution method for the determination of phosphate in natural waters. Anal. Chimica Acta 27, 31–36.
Narumalani, S., Jensen, J.R., Althausen, J.D., Burkhalter, S.G., Mackey Jr., H.E., 1997. Aquatic macrophyte modeling using GIS and logistic multiple regression. Photogramm. Eng. Remote Sens. 63, 41–49.
Nelson, S.A.C., Soranno, P.A., Cheruvelil, K.S., Batzli, S.A., Skole, D.L., 2003. Regional assessment of lake water clarity using satellite remote sensing. J. Limnol. 62 (Suppl. 1), 27–32.
Nusch, E.A., 1980. Comparison of different methods for chlorophyll and phaeopigment determination. Arch. Hydrobiolia. 14, 14–36.
Orth, R.J., Moore, K.A., 1983. Submerged vascular plants: techniques for analyzing their distribution and abundance. Mar. Technol. Soc. J. 17, 38–52.
Penuelas, J., Gamon, J.A., Griffen, K.L., Field, C.B., 1993. Assessing coummunity type, plant biomass, pigment composition, and photosyntheic efficiency of aquatic vegetation on spectral reflectance. Remote Sens. Environ. 46, 110–118.
Pampel, F.C., 2000. Logistic Regression: A Primer. Sage University Papers Series Quantitative Applications in the Social Sciences, series no. 07-132. Sage Publications, CA.
Raitala, J., Lampinen, J., 1985. A Landsat study of the aquatic vegetation of the lake Luodonjarvi reservoir Western Finland. Aquat. Bot. 21, 321–346.
SAS Institute Inc., 1995. Logistic Regression Examples Using SAS System. Version 6, first ed. SAS Institute Inc., NC.
Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., Macomber, S.A., 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sens. Environ. 75, 230–244.
Valley, R.D., Drake, M.T., Anderson, C.S., 2005. Evaluation of alternative interpolation techniques for the mapping of remotely-sensed submersed vegetation abundance. Aquat. Bot. 81, 13–25.
Verbyla, D.L., 1995. Satellite Remote Sensing of Natural Resources. CRC Lewis Publishers, New York.
Vis, C., Hudon, C., Carignan, R., 2003. An evaluation of approaches used to determine the distribution and biomass of emergent and submerged aquatic macrophytes over large spatial scales. Aquat. Bot. 77, 187–201.
White, J., 1981. The allometric interpretation of the self-thinning rule. J. Theoret. Biol. 89, 475–500.
White, J., 1985. The thinning rile and its application to mixtures of plant populations. In: White, J. (Ed.), Studies on Plant Demography. Academic Press, London, pp. 291–309.
Zhang, X., 1998. On the estimation of biomass of submerged vegetation using Landsat Thematic Mapper (TM) imagery: a case study of the Honghu Lake, PR China. Int. J. Remote Sens. 19, 11–20.
Zilioli, E., 2001. Lake water monitoring in Europe by means of remote sensing. Sci. Total Environ. 268, 1–2.