UC Irvine
UC Irvine Previously Published Works

Title
Measurement of oak tree density with Landsat TM data for estimating biogenic isoprene emissions in Tennessee, USA

Permalink
https://escholarship.org/uc/item/56z7m999

Journal
International Journal of Remote Sensing, 22(14)

ISSN
0143-1161

Authors
Baugh, W
Klinger, L
Guenther, A
et al.

Publication Date
2010-11-25

DOI
10.1080/01431160117906

License
CC BY 4.0

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

INT. J. REMOTE SENSING, 2001, vol. 22, no. 14, 2793–2810

Measurement of oak tree density with Landsat TM data for estimating biogenic isoprene emissions in Tennessee, USA

W. BAUGH[1], L. KLINGER[1], A. GUENTHER[1] and C. GERON[2]

[1] Atmospheric Chemistry Division, National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307, USA
[2] US Environmental Protection Agency, Office of Research and Development, National Risk Management Research Laboratory, Research Triangle Park, NC 27711, USA

(*Received 20 September 1999, in final form 5 June 2000*)

Abstract. Isoprene emissions from oak trees in the eastern USA play an important role in tropospheric ozone pollution. Oak trees (*Quercus*) emit an order of magnitude more isoprene than most other emitting tree species, and are by far the largest source of biogenic isoprene in the eastern US. We used Landsat TM data to measure oak tree abundance near Oak Ridge, Tennessee, to estimate fluxes of isoprene. The Landsat classification was performed using multi-date data, supervised classification techniques, and an iterative approach. Training sites were selected based on transect data, and ten vegetation classes were mapped. A supervised classification algorithm called the Spectral Angle Mapper was used to classify the data. Empirical vegetation emission data were used to estimate the isoprene flux from each of the vegetation classes. The resultant isoprene flux maps were compared with concentrations measured in the field, and a good correspondence was observed. We also compare the Landsat classification with three other landcover schemes including the USGS's Global Landcover Classification, which is based on AVHRR data. Results from these landcover classifications are used as input for models that predict tropospheric ozone production and are used to investigate ozone control strategies.

1. Introduction

One of the most persistent air pollution problems in the USA is tropospheric ozone, a photochemical oxidant and an important component of smog. Elevated ozone concentrations in urban and rural air were recognized in the early 1950s (Haagen-Smit and Fox 1954); and massive and costly cleanup efforts have been implemented since the 1960s. Controlling ozone pollution is difficult because ozone is not emitted directly into the atmosphere. Instead, ozone is the product of precursor gas emissions and atmospheric chemical reactions. Under certain meteorological conditions, ambient ozone concentrations in and downwind of some urban areas can reach levels higher than the concentration considered acceptable for public health by the US Environmental Protection Agency (EPA) (National Research Council 1991).

Ozone (O_3) is a reactive oxidant gas that is produced naturally in the Earth's atmosphere. Most of the atmospheric ozone occurs in the stratosphere (figure 1).

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2001 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160010006944

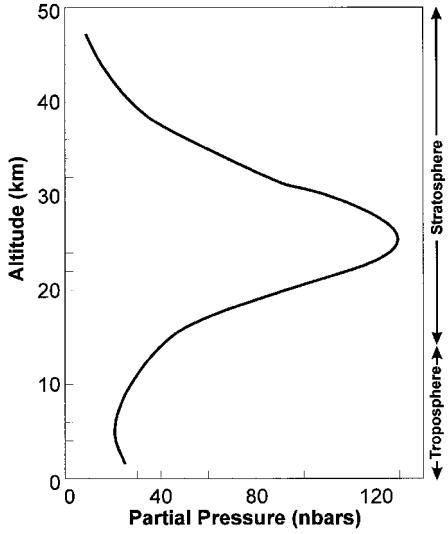

Figure 1. Generalized atmospheric ozone profile (after National Research Council 1991).

Stratospheric ozone plays a critical role in absorbing damaging ultraviolet solar radiation. A smaller portion of naturally occurring ozone is found in the troposphere (figure 1). Tropospheric ozone accounts for about 10–15% of the atmospheric total. Its concentration is measured in orders of 10 parts per billion (ppb), while stratospheric ozone concentrations can exceed 10 000 ppb (National Research Council 1991).

Ozone is formed by a series of atmospheric photochemical reactions. The reactions require two types of gaseous compounds: nitrogen oxides (NO_x) and volatile organic compounds (VOCs). NO_x and VOC sources are both anthropogenic (of human origin) and natural. Most NO_x is produced from atmospheric nitrogen during high temperature combustion. Examples of NO_x sources are automobiles, fossil-fuel power plants, lightning, forest fires, and natural soil processes. Anthropogenic VOC sources include fossil fuel combustion products, and evaporation of fuels, solvents, and paints. Natural (biogenic) VOCs are emitted primarily by forests.

The presence of naturally emitted biogenic VOCs increases the complexity of ozone control. On regional scales, natural sources of these precursor components are often greater than anthropogenic sources (Lamb *et al.* 1993, Guenther *et al.* 1995). Yet, until recently, ozone control strategies were based on modelling studies that did not consider biogenic emissions.

Biogenic VOCs include many organic compounds, and the amount emitted varies considerably among plant species (table 1). One of the compounds emitted by certain deciduous trees is isoprene (C_5H_8); conifers commonly emit monoterpenes such as α- and β-pinene. These few compounds account for about 50% of the biogenic VOC emissions; with the remainder, a wide variety of other compounds (Guenther *et al.* 1994). In the USA, forests are believed to emit about 90% of the biogenic VOCs, with agricultural and scrub lands contributing the rest (Lamb *et al.* 1987).

In the south-eastern USA (where ozone pollution problems are particularly persistent), most of the biogenic VOC emissions are from relatively few tree species. Regional models that incorporate tree emission rates and total biomass of each

Table 1. Isoprene and terpene emission rates for the dominant tree genera in Oak Ridge, TN, study area (from Guenther *et al.* 1994).

Genus	Common name	% Leaf biomass	Isoprene emission ($\mu g\ C\ g^{-1}\ h^{-1}$)	Terpene emission ($\mu g\ C\ g^{-1}\ h^{-1}$)
Quercus	Oak	29.9	70	0.2
Pinus	Pine	12.6	<0.1	3.0
Acer	Maple	6.1	<0.1	1.6
Ulmus	Elm	5.3	<0.1	<0.1
Oxydendrum	Sourwood	5.1	<0.1	0.6
Liriodendron	Tulip Poplar	4.4	<0.1	0.2
Juniperus	Juniper	4.0	<0.1	0.6
Carya	Hickory	3.8	<0.1	1.6
Nyssa	Blackgum	2.8	14	0.6
Liquidambar	Sweetgum	2.5	70	3.0
Fagus	Beech	2.4	<0.1	0.6

species have shown that isoprene emitted by oak trees dominates the daytime biogenic VOC flux in the eastern USA (Geron *et al.* 1994, Guenther *et al.* 1995).

In response to a need for accurate distribution maps of oak trees, the use of Landsat Thematic Mapper (TM) data for mapping oak tree densities in the Oak Ridge, TN, area was investigated. Species level emissions of biogenic VOCs are well known in this area, but distributions of oak trees are not well known. This classification was compared with the US Geological Survey's (USGS) Seasonal Land Cover Regions (SLCR) landcover classification, which is based on Advanced Very High Resolution Radiometer (AVHRR) data; the US Forest Service's (USFS) Forest Inventory and Analysis (FIA) database; and the EPA's Biogenic Emission Landcover Database (BELD). Both satellite landcover maps were converted to isoprene emission maps by assigning field-measured leaf-level emission rates to the landcover classes. Field-measured isoprene concentrations were compared with the isoprene emissions predicted by the remotely sensed maps. These emission maps are used as input to regional ozone production models and therefore represent a critical link between leaf-level measurements and regional isoprene flux estimates.

2. Field study methods

2.1. *Study area*

This study was conducted in and around the Walker Branch Watershed ($35^\circ\ 57'\ 30''$, $-84^\circ\ 17'\ 30''$), which is located within the Oak Ridge National Laboratory (ORNL), a US Department of Energy reservation near Oak Ridge, TN (figure 2). The Oak Ridge area (and the south-eastern USA) commonly exceeds EPA ozone standards during the summer months. Field studies were conducted in the summers of 1992 and 1995; the latter was timed to coincide with the 1995 Southern Oxidant Study (SOS) field campaign. During this campaign other institutions conducted experiments ranging from soil sampling to aircraft air sampling.

The field area lies on the western edge of the Appalachian Mountains, and is characterized by forest-covered, northeast trending ridges and valleys with a relief of about 125 m. The forests are primarily hardwood with some conifer stands. The dominant tree genera are: oak (*Quercus*), maple (*Acer*), hickory (*Carya*), pine (*Pinus*), sweetgum (*Liquidambar*), sourwood (*Oxydendrum*), and blackgum (*Nyssa*) (table 1). A 44 m walkup tower provided a platform for leaf and branch enclosure experiments,

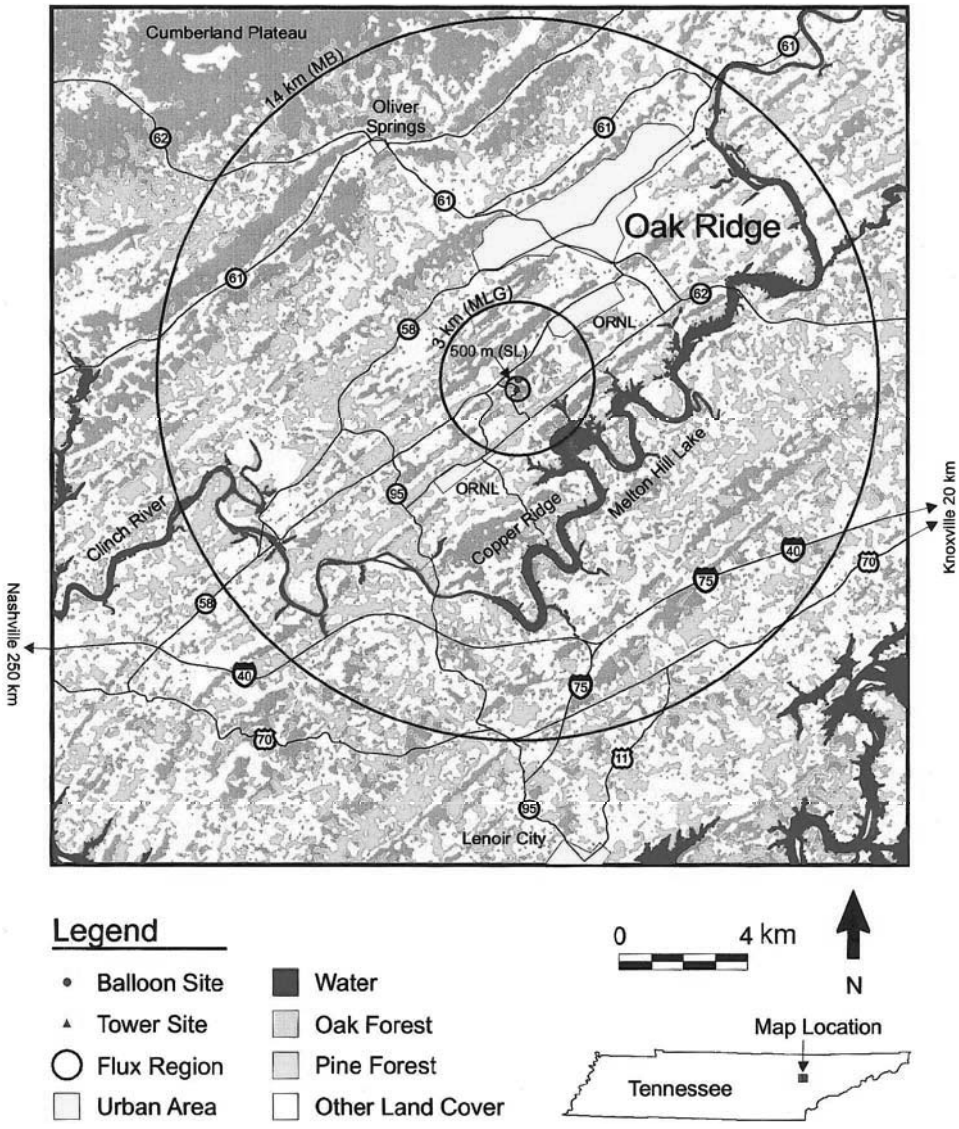

Figure 2. Index map of Oak Ridge, TN. Land classes are based on Landsat TM data from this study. 'Oak forest' includes both high and medium oak classes; 'pine forest' includes all conifer classes. The circles show the fetch areas for various sampling methods and are described in §4.3.

and for above-canopy atmospheric measurements. A tethered balloon atmospheric sampling system was deployed from a clearing about 400 m north of the tower. Further information about the site can be found in Johnson and Van Hook (1989).

2.2. *Ground transects*

In 1992 three transects with a total of thirty-two 10 m × 10 m (100 m^2) individual plots were sampled. Each transect was approximately 1 km long. Transects began at benchmarks with known locations and were measured with tape and compass. In

1995 eight 200 m long belt transects were sampled radially outward from the 44 m tower in the following directions: N, NE, E, SE, S, SW, W, and NW (figure 3). Each belt transect consists of adjacent 10 m × 10 m plots. In all transects each plot was sampled for: slope, aspect, species composition, diameter at breast height (dbh) [~1.5 m], tree height, sapling count, and percentage cover of the dominant growth forms. Trees are defined as individuals greater than 4 cm dbh; saplings are individuals less than 4 cm dbh and taller than 1.5 m.

Transect data were processed to estimate the crown-area percentage of each tree species in the canopy (Guenther *et al.* 1996a). The plots were classified using these percentages and the same classification scheme that was applied to the TM data (§3.2). A subset of the classified TM image surrounding the tower and the classified transect map are compared in figure 3.

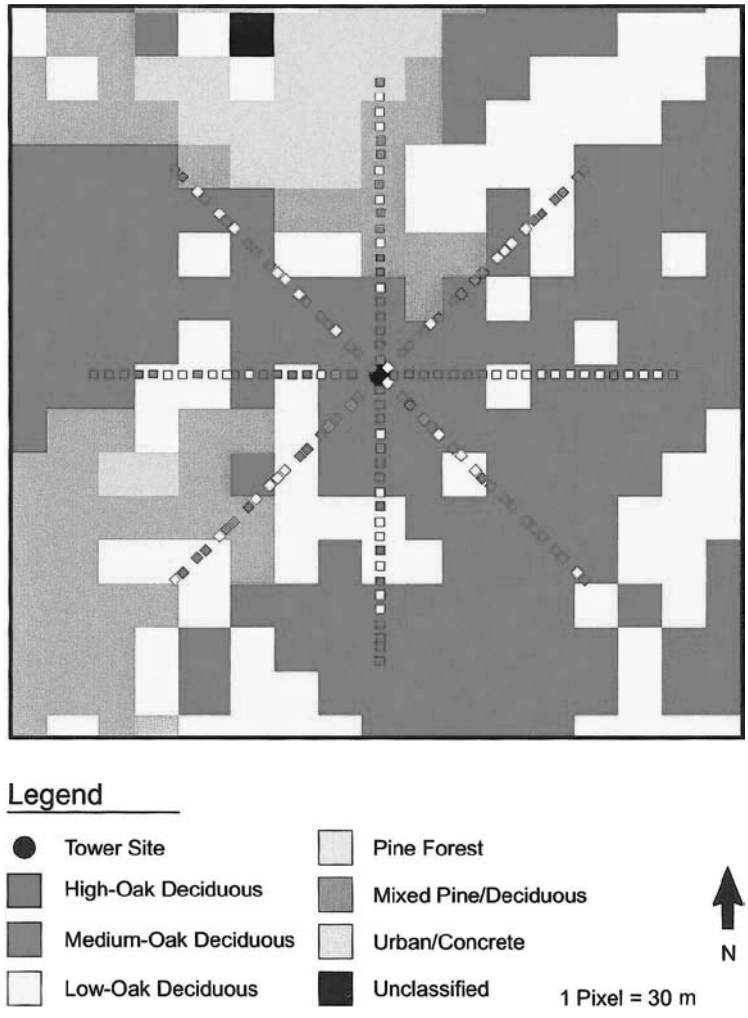

Figure 3. Comparison of tower transects and Landsat classification. Larger background pixels are from the Landsat classification; smaller squares are tower transect plots. The transect plots are shown as 5 m squares for clarity; the actual plots are 10 m squares. See discussion in §2.2 and §4.1.

2.3. *Field observations*

With only one day available to work outside the training area, reconnaissance-style observations were made to help assess the accuracy of the TM classification. Walking transects were conducted at two sites in the ORNL, and at two sites on the Cumberland Plateau, about 20 km to the north-east of the tower site. Walking transects consist of careful map reading; locating plots about every 30 m and recording tree species, diameter, and canopy cover; and comparison with the TM image.

2.4. *Gas flux measurements*

Biogenic isoprene fluxes were measured at three scales in the Walker Branch Watershed (Guenther *et al.* 1996a). At the leaf-level scale, leaf and branch emissions were studied with enclosure systems. Clear Teflon bags were used for branch enclosure measurements, while leaf cuvettes were used to sample foliar emissions in a controlled environment. Emissions are estimated with a mass balance approach, in which constituents of the air flowing to and from the sample chamber are compared. On the canopy level, a Relaxed Eddy Accumulation (REA) system and a micrometeorological vertical gradient system were used to estimate fluxes in the atmospheric surface layer above the forest (Guenther *et al.* 1996a). These systems were mounted on the 44 m tower and use ambient isoprene concentrations to estimate the flux. At the landscape level, fluxes were estimated with data collected from a tethered balloon profiling system at elevations up to 2 km above ground level. These sampling and flux estimation methods are described in more detail in Guenther *et al.* (1996a, b), and the results are compared with TM-based flux maps in §4.3.

3. Image processing methods

3.1. *General*

The desired product of the image classification was a landscape characterization showing the relative density of oak trees in the forest. The classification was performed using multi-date Landsat TM data, supervised classification techniques, and a multi-layered (iterative) approach. An existing 1 km landcover classification by the USGS was also evaluated as a lower cost alternative to Landsat TM data. The two classifications were co-registered and compared for correspondence of similar classes. Both classifications were also converted to isoprene emission maps using empirical field measurements. Isoprene flux estimates based on the remotely sensed data were compared with ambient isoprene concentrations measured during the 1992 field campaign. Image processing was performed with the 'Environment for Visualizing Images' (ENVI) [RSI, Boulder] software on a Power Macintosh 8100/80 computer. Additional data manipulation and iterative classification were done with programs written in the 'Interactive Data Language' (IDL) [RSI, Boulder].

3.2. *Landsat TM processing*

The Oak Ridge data set consists of two standard seven-band digital Landsat TM images of path 019, row 035 that were acquired on 17 December 1991 (winter) and 30 September 1992 (early fall). From these images 10 land cover classes were identified: high oak/other deciduous (greater than 50% oak in deciduous forest); medium oak/other deciduous (25–50% oak); low oak/other deciduous (less than 25% oak); conifer; mixed conifer/deciduous; conifer plantation; shrubs, grasses, and agriculture (non-forest vegetation); water; bare soil; and urban/concrete. Classification parameters for the Landsat TM data are summarized in table 2.

Table 2. Landsat TM image classification parameters.

Class	Subclass	Number of training pixels	Method of plot location	Classification priority	Season
Conifers	Loblolly	195	Field observations	Highest	Winter
	Mixed	129	Field observations	Highest	Winter
	Plantation	151	Aerial photographs	Highest	Winter
Deciduous	High oak	11	Transect plots	2nd	Autumn
	Medium oak	10	Transect plots	2nd	Autumn
	Low oak	9	Transect plots	2nd	Autumn
Nonforest vegetation	Shrub/grass	130	Field observations	3rd	Autumn
	Agriculture	149	Field observations	3rd	Autumn
Soil		196	Field observations	3rd	Autumn
Urban		259	TM image	Lowest	Winter
Water		241	TM image	Lowest	Winter

The following processing was applied to both TM images before classification. TM band six (thermal infrared) was removed from the data. A regression adjustment (dark object subtraction) was used to correct for atmospheric and sensor effects (Jensen 1986, Chavez 1988). Subsets (1024 pixels by 1024 pixels) covering the ORNL and Oak Ridge city areas were extracted from the larger image to accelerate iterative processing on a slow computer. The subsets were registered to the Universal Transverse Mercator (UTM) coordinate system (zone 16) with a pixel size of 30 m using nearest neighbour resampling and second order polynomial equations. A root mean squared (rms) error of 0.49 pixel was achieved for the fit between each image and the 1:24 000 scale USGS topographic maps using 24 evenly distributed ground control points.

Field measurements made in 1992 were used to choose training sites for each class. The field data consist of detailed vegetation inventories in transects with well-located 10 m $\times$ 10 m plots. Transect plot locations were measured by tape and compass, and linked to ORNL benchmarks. The transects, unfortunately, were not measured for the purpose of Landsat TM training. The 10 m square plots are significantly smaller than the 30 m TM pixels, and we must assume that the transect plots represent the vegetation in the corresponding TM pixel. The risks due to this assumption are described in §4.1.

Averaged reference spectra were extracted from the training site pixels. The reference spectra were used as input for the Spectral Angle Mapper (SAM) classification algorithm, and the results were combined manually with an iterative analysis. As pixels were classified in the original Landsat data, the corresponding pixels on a 'blank slate' image were assigned a value indicating the class (e.g. pine = 10). Once a pixel was assigned to a class, it was considered permanent and not re-assigned (Wolter *et al.* 1995).

The SAM is a multi-spectral supervised classification algorithm developed at the University of Colorado (Boardman unpublished data, Kruse *et al.* 1993). This method treats image spectra as *n*-dimensional vectors, where n is the number of bands. A test spectrum (from a pixel) is compared with a reference spectrum (from the desired

class); a smaller angle between the spectral vectors indicates a better match. A threshold angle (chosen by inspection) is used to determine whether a spectrum is classified as a specific class. The power of the SAM method is that it is very sensitive to small spectral variations, yet it is resistant to illumination differences.

We began the classification process by mapping the coniferous trees. These evergreen trees are distinct in the winter image because the surrounding deciduous trees have lost their leaves. Individual stands of loblolly pine and mixed deciduous/pine forest were identified during the 1992 field study and indicated on aerial photographs. The corresponding areas were located on the winter Landsat image, and training spectra were extracted and averaged (table 2). The averaged spectra were used by the SAM to classify the TM winter subset.

Inspection of the results showed that some obvious evergreen stands were not classified as loblolly or mixed pine. Interpretation of stereo aerial photographs showed that young conifer trees planted in regular rows densely cover these areas. These areas are assumed to be cultivated conifer trees (plantation). Training pixels were selected from these unclassified evergreen areas, and the winter image was reclassified with the three conifer classes.

Because the conifer trees are easily detected with the winter image, the three conifer classes were added to the final result image with the highest priority. These classes would not be changed during further manipulations. The SAM output consists of three bands (images) of floating point data, one for each conifer class. The value of each pixel is the spectral angle, in radians, between the class' training spectrum and the pixel's winter TM spectrum. Pixels with lower values have a better match. A program was written to select the band with the lowest value for each pixel. If this value was below an arbitrary threshold of 0.1 radians then that pixel was assigned to the corresponding class. The ENVI software can be used to automatically evaluate the SAM results and return a final classified image. In this case, however, the evaluation was done manually so that the individual classes could be added one layer at a time, with weighted priorities, to the final classification image.

The second iteration of classification focused on the deciduous and oak classes, and used the early fall TM image. Oak trees commonly lose their leaves (senesce) 1 to 2 weeks later than other deciduous trees. Using this brief period of time when oak trees are the only deciduous trees with significant remaining leaves has been quite effective in mapping oak trees in the eastern USA (Baugh unpublished data, Wolter, *et al.* 1995). The early fall image used in this study, however, is about 1 month too early to utilize this effect and the oak mapping relies only on apparent spectral differences.

Training spectra for high, medium, and low oak classes were derived from pixels that correspond with specific ground transect plots from the 1992 survey. Detailed vegetation characteristics were measured at 32 plots in three transects. The plots were classified as high, medium, or low-oak deciduous based on the vegetation surveys. All plots with more than 30% pine were excluded (mixed pine class already classified). Remaining forest plots with less than 25% oak were considered low oak deciduous; 25–50% oak were medium oak deciduous; and greater than 50% oak were high oak deciduous.

Plot locations were mapped with tape and compass, and marked on stereo aerial photos. The points were carefully transferred to 7.5 minute topographic maps and the UTM coordinates measured. Based on the UTM coordinates, the corresponding TM pixels were identified, and average TM spectra were extracted (table 2). The

image was classified with the SAM, and the results were added to the final classification image with the same method that was applied to the pine classes. The relatively small number of pixels available for deciduous class training is of some concern, and we had to assume that the detailed transects provided representative training spectra.

Areas of shrubs, grasses, agriculture, and bare soil were identified based on aerial photographs that were annotated during the 1992 field work. The early fall image showed the best differentiation for these classes. Classification was done with the method described above. The SAM results were added to the final classification image with a lower priority than the coniferous or deciduous classes.

Finally, average spectra for the remaining classes (urban and water) were derived from regions that were identified by inspection of the Landsat images, and classification was done using the winter image. These results were added (with the lowest priority) to the final classification image. To aid interpretation, the classification image was colour coded and roads and highways were added from USGS 1:100 000-scale digital line graphs. When we were comfortable with the results of the classification for the subset image we repeated the appropriate steps to classify the complete 90 km × 90 km Landsat TM quad (figure 4).

3.3. *Seasonal Land Cover Regions (SLCR) processing*

The USGS's Seasonal Land Cover Regions (SLCR) satellite classification, a part of the global land cover characterization data base, was compared with the Landsat TM classification results and evaluated as a low cost alternative to classifying Landsat data. The SLCR classification is a vegetation classification with 1 km pixels that covers most of the world's continents. It is based on seasonal vegetation changes in 10-day averages of AVHRR Normalized Difference Vegetation Index (NDVI) data spanning April 1992–March 1993. The classification contains 205 vegetation and land cover classes (varies regionally) and is available globally. These data are distributed on the Internet at the EROS Data Center website: http://edcwww.cr.usgs.gov/landdaac/glcc/glcc.html.

According to the USGS web site, SLCR classification units are not based on precisely defined mapping units in a predefined land cover classification scheme. Instead the classes serve as summary units for both descriptive and quantitative attributes, and delineate areas of similar characteristics. The regions (classes) are composed of relatively homogeneous land cover associations (e.g. similar floristic and physiognomic characteristics) which exhibit distinctive phenology (e.g. onset, peak, and seasonal duration of greenness), and have similar levels of primary production. Further information can be found at the USGS website: http://edcwww.cr.usgs.gov/landdaac/glcc/globdoc1_2.html.

For biogenic emission land cover mapping, the SLCR database could provide a useful data source if it is suitably accurate, or if it can be locally calibrated. Using the TM classification of the Oak Ridge area, we evaluated the usefulness of the SLCR database for predicting biogenic emissions. To make the comparison between the Landsat TM classification and the SLCR database, the following steps were taken. A subset of the North American SLCR image, slightly larger than the Oak Ridge Landsat quad, was extracted. The subset was reprojected from the original Lambert Azimuthal Equal Area projection to UTM coordinates using the original projection units and equations from Snyder (1987). The 1 km SLCR pixels were resampled to 30 m pixels using nearest neighbour resampling. A subset was extracted

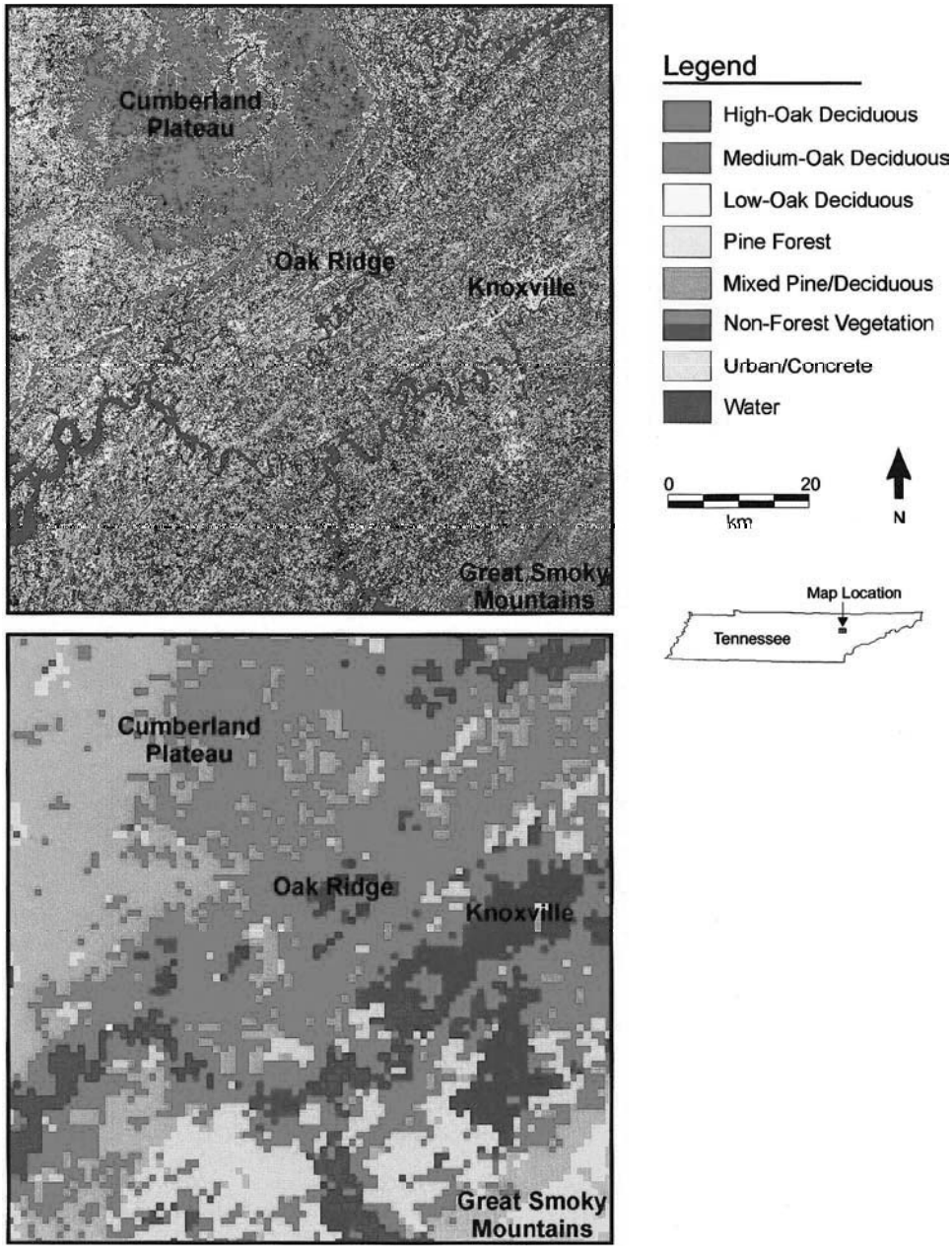

Figure 4. Landsat TM classification results (upper) and the corresponding AVHRR SLCR image (lower). Oak Ridge and the study area are located approximately in the center of the images. The high oak area in the upper part of the Landsat image is the Cumberland Plateau, and the pine and high oak area in the lower right corner is the edge of the Great Smoky Mountains.

from the resampled SLCR image that had the same dimensions and map locations as the TM quad classification image. The SLCR and TM images were compared pixel by pixel to look for relationships between the classes (§4.2).

3.4. *Isoprene emission maps*

Isoprene emission maps were created from the TM and SLCR classifications. Each TM class was converted to isoprene flux using empirical emission factors that are based on vegetation enclosure measurements and a simple canopy isoprene emission model, table 1, (Guenther *et al.* 1996a). SLCR classes were converted to isoprene flux using emission factors recommended by Guenther *et al.* (1994). The results are images where each pixel value consists of a biogenic emission capacity. The emission predictions can be resampled to other scales and used as input for ozone production models. A comparison between the two remotely sensed isoprene flux estimates and field isoprene flux measurements is presented in §4.2.

4. Results and discussion

4.1. *Error analysis*

This study lacks a direct quantitative error analysis of the image classification. Verification consists of three somewhat indirect methods that are not individually conclusive, but together suggest good agreement. The three methods are: comparisons with two existing land cover databases—the USFS's Forest Inventory and Analysis (FIA) database and the EPA's Biogenic Emission Landcover Database (BELD); comparison with tower transects; and qualitative field observations.

4.1.1. *FIA and BELD comparisons*

FIA data used in this study were collected in 1988 and 1989. Results of the comparisons are presented in table 3. FIA data are collected by the USFS on approximately ten-year intervals to monitor forest conditions. The process consists of air photo interpretation and field surveys (Hansen *et al.* 1992). FIA field plots

Table 3. Comparison of the Landsat TM classification (described in this paper), the Forest Inventory and Analysis (FIA) database (Hansen *et al.*1992), and the Biogenic Emissions Landcover Database (BELD), (Kinnee *et al.* 1997).

Land use types	FIA (% of crown area)	Landsat (% of forest classes)
Oak trees	37.7	30.0 (18.6–41.3)
Non-oak trees	47.9	46.7 (35.4–58.1)
Conifer	14.4	23.3
Other evergreen	0.02	NA
Total area (ha)	NA	653 672
Land use type	**BELD (% coverage)**	**Landsat (% of all classes)**
Total forest	60.6	70.2
Oak forest	21.0	15.4 (9.7–21.1)
Non-forest vegetation	14.5	14.9
Urban	5.4	6.1
Barren	NA	0.8
Water	1.3	3.1
Other	18.2	4.9
Total area (ha)	997 266	997 265

Note that for the FIA comparison only the forested pixels of the Landsat image are considered, and in the BELD comparison the oak forest category is a subset of the forest category.

are sampled primarily by prism cruising, with fixed radius subplots to characterize seedling/sapling populations. Recorded observations include tree diameter at breast height, species, and other measurements that enable the prediction of the trees' volume.

FIA inventories are extensive inventories that provide reliable estimates for large sampling areas. As data are subdivided into smaller areas, such as a county, the sampling errors increase and the reliability decreases. According to Hansen *et al.* (1992), the eastern FIA inventories are designed to have a maximum 3% error per 1 million acres of timberland at the 67% confidence limit (one standard error) for basal area and volume. The Landsat image (quad) covers about 2.4 million acres; therefore, the corresponding FIA data should have a 1.9% maximum sampling error at the 67% level (standard error of the mean).

About 240 FIA plots are evenly distributed throughout the Landsat image, and it was originally hoped that these plots could be used to create a contingency table of error for the TM classification. But, because the plots' map coordinates describe generalized polygon centres rather than specific transect locations and because of the error discussion above, the FIA data were summarized for the entire Landsat image area. Tree basal area values from the FIA plot data were converted to crown area using equations from Geron *et al.* (1994). Crown areas were summed for four categories: oak trees, all other deciduous trees, all conifer trees, and other evergreen trees, such as *Magnolia* and *Ilex*. Values are reported as percentages of total crown area for the four categories.

Because the FIA data pertain only to forested areas, only the forest categories from the TM classification were compared with the FIA data. The TM forest categories are: conifer; conifer plantation; mixed conifer/deciduous; and low, medium, and high oak deciduous. From these categories the areas covered by oak, other deciduous, and conifer trees were estimated. The parameters used to create the TM classes (§3.2) resulted in the following equations for the average area coverages (maximum and minimum values were also calculated for 'oak' and 'other deciduous' and are presented in table 3):

$$\text{Oak} = (\text{LO} \times 0.125) + (\text{MO} \times 0.375) + (\text{HO} \times 0.75) + 0.35(0.332 \times \text{MF}) \quad (1)$$

$$\text{Other Deciduous} = (\text{LO} \times 0.875) + (\text{MO} \times 0.625) + (\text{HO} \times 0.25) + 0.35(0.668 \times \text{MF}) \quad (2)$$

$$\text{Conifer} = \text{Conifer} + \text{Conifer Plantation} + (0.65 \times \text{MF}) \quad (3)$$

LO are low oak pixels, MO are medium oak pixels, HO are high oak pixels, MF are mixed conifer/deciduous forest pixels, and Conifer and Conifer Plantation are the respective number of pixels. In the mixed forest component of the 'oak' and 'other deciduous' equations, 0.35 is the fraction of deciduous in the mixed forest. Within this 35% fraction, 0.332 is the factor that describes the oak component and 0.668 is the factor that describes the other deciduous component. In the mixed forest component of the 'conifer' equation, 0.65 is the fraction of conifer in the mixed forest. The TM conifer value was held constant to simplify the table.

Since the FIA data are limited to forested areas only, BELD data were used to evaluate the remaining categories. This database is a compilation of several data sources that were assembled by Kinnee *et al.* (1997). Non-forest categories were taken from the USGS Land Cover Characteristics database and the US Census Bureau's TIGER database; forest categories were taken from resampled FIA data.

All are generalized to the county level. The BELD comparison used here (table 3) was taken directly from Kinnee *et al.* (1997). Values are reported as percentages of total area for consistency with the FIA comparison. Note that the larger total area for the Landsat image in the BELD comparison (relative to the FIA comparison) is due to all classes being considered—not just the forest classes.

Although the results of these comparisons are not a substitute for a contingency table based on numerous field plots, they do provide feedback about the overall accuracy of the Landsat classification in the absence of such data. The comparisons between the TM classification and the FIA and BELD data are quite good (table 3). In the FIA comparison, the oaks and other deciduous categories match very well. Conifers appear to be over-predicted by the TM data. This could be due to other evergreen trees and shrubs, such as *Magnolia*, *Ilex*, or *Rhododendron*, being misclassified as conifer trees. The low occurrence of non-conifer evergreen trees in the FIA data (0.02%) suggests that evergreen trees are not a significant component of the TM conifer category. But, *Rhododendron* shrubs, which are not included in the FIA database, could be contributing to this over-classification.

In the BELD comparison, the total forest, oak forest, non-forest vegetation, and urban comparisons are very good. The TM classification shows significantly more water. This is probably due to the small TM pixel identifying water bodies that the coarser BELD classification misses. There is no 'barren' category in the BELD, and the 'other' category includes the unclassified TM pixels.

4.1.2. *Tower transects*

The next method of verification, a qualitative comparison between the TM classification and the 1995 radial tower transects, shows a general match of forest patterns (figure 3). Note that the high oak area surrounding the tower is well characterized, as is the mixed deciduous area north of the tower. Other correspondences can be found by inspection. Note that the transects were measured to characterize the forest near the tower, and not as a ground-truth check. The sampling strategy was therefore not optimized for ground-truth purposes. The difference in scales between the 900 m^2 satellite pixels and 100 m^2 transect plots hinders the comparison. A feature that dominates a transect plot may be an insignificant part of a larger satellite pixel. Also, the number of transect plots available (46) is too small for a statistically valid comparison (Jensen 1986).

4.1.3. *Field observations*

The third method of verification, qualitative field observations, is also not suitable for statistical error analysis. It does, however, provide excellent feedback about the accuracy of the classification when it is combined with knowledge of the field area. For example, it was known from experience that most oak trees grow on ridge tops because they thrive on well-drained soils. The classification showed high concentrations of oak trees on the ridge tops, and walking transects in the field confirmed this observation.

Two sites in the Copper Ridge area of the ORNL, comprising high to medium and high to low oak classes, were visited (figure 2). At the first site, vegetation was observed along a 200 m walking transect that crossed from medium to high oak pixels. Observed vegetation corresponded well with the classified image. In the medium oak zone, oak trees were about 30–50 cm dbh and about 15 m tall. Hickory (*Carya*) and maple (*Acer*) trees were intermixed, and the canopy was relatively dense. About 75 m into the transect, the density and size of oaks notably increased. There

were fewer hickory and maple trees, the canopy became nearly closed, and the oaks were larger with a dbh of 50–90 cm and a height of about 25 m.

The second walking transect was a circle about 250 m in diameter and included low, medium, and high oak pixels. Based on observations in this transect, the low-oak pixels were under-classified. Oak trees covered about 35–40% of the area in the low-oak zone (a medium concentration). The trees were tall and spindly, however, and likely exposed much of the highly reflective limestone bedrock to the satellite. The bedrock exposure may account for the under-classification.

The high and medium oak pixels on the second walking transect were classified more accurately than on the first. The canopy was mostly closed, and the bedrock was covered by soil. Tree characteristics were similar to those described for the first walking transect.

At the two Cumberland Plateau sites, trees were identified with binoculars from an automobile. The Cumberland Plateau is of particular interest because the classification shows consistent, widespread high oak densities at the higher elevations (NNW portion of figure 4). The first site was about 11 km north of Oak Ridge at the top of Walden Ridge. About 80% hardwood and 20% conifer cover the densely forested ridge. Oak species comprise about 40–60% of the total canopy, with the remaining hardwoods being maple (*Acer*), tulip poplar (*Liriodendron*), walnut (*Juglans*), and hickory (*Carya*). The pine trees are concentrated in bands and patches along the ridge. This is consistent with the classification image, although the oaks may be slightly over-classified.

The second site is characterized by rural residential use, and the forest has been heavily disturbed by human activity. Several hundred meters separate homes; and yards contain lawns, gardens, non-oak hardwood trees, and pine trees. High concentrations of oak trees were observed, however, on hills behind the residential zone. The classification shows scattered high- to low-oak deciduous, and conifer pixels for this area, consistent with the heterogeneous rural land use patterns.

The final site overlooks a ridge to the southwest of State Route 116. The ridge is forested with a high density of oak trees and correctly classified as homogeneous high oak.

4.2. *Landsat—SLCR comparison*

We found that in the Oak Ridge area there is a poor relationship between the SLCR class names and the corresponding TM classes. The SLCR classes also did a poor job of distinguishing the land cover types (oak trees) that are important for characterizing biogenic emissions. These results are to be expected, however, when using 1 km AVHRR data to try to identify specific tree species in a mixed forest. The high-oak tree stands are too small to be distinguished and therefore are lumped into generalized deciduous and mixed deciduous SLCR classes. Despite these differences, SLCR data were able to model isoprene emissions to within a factor of 3 of measured values (§4.3). This is only a fair result for a well-studied site like Oak Ridge, but it would be a good result for a more remote site with few previous studies.

To compare the TM and SLCR images, the fraction of each TM class that occurs in each SLCR class was calculated as a percentage (i.e. SLCR classes were described in terms of TM categories). The results are presented in table 4. The columns of this table are SLCR classes. The rows are the composition of each class in terms of TM (classification) pixels. The numbers indicate the percentage of the TM class that occurs in the SLCR class.

Table 4. Percentage of AVHRR landscape classification pixels occurring in Landsat classification categories for Oak Ridge, TN, area.

	AVHRR classification[1]															
	Cropland Classes						Grassland		Oak woodland	Mixed low oak	Mixed high oak	Degraded forest	Mixed conifer forest	Mixed deciduous forest		Water
	13	19	23	39	43	44	63	64	94	104	105	113	154	168	171	205
Unclassified	8	11	8	8	7	8	4	2	7	6	4	12	6	5	3	7
Loblolly	<1	2	2	0	<1	3	0	0	2	<1	3	13	8	1	6	2
Mixed Pine/ Deciduous	7	16	25	3	10	17	1	2	16	2	21	18	32	13	24	10
Conifer Plantation	<1	<1	<1	0	0	<1	0	0	<1	0	<1	25	1	0	1	<1
High Oak	1	2	3	<1	0	4	0	0	3	64	18	7	9	<1	21	7
Medium Oak	1	1	2	<1	<1	2	0	<1	2	9	7	3	4	1	8	5
Low Oak	6	5	11	3	4	13	<1	1	15	10	22	18	13	12	24	14
Non-forest Vegetation	9	12	19	6	10	13	1	4	13	7	16	4	21	4	10	4
Water	7	17	5	0	<1	18	2	0	25	0	2	0	<1	63	<1	46
Barren	2	3	2	1	6	2	3	3	1	<1	1	<1	1	0	<1	1
Urban/ Suburban	59	30	23	79	62	21	89	89	16	2	6	<1	6	3	2	4

[1] 13-44, Cropland; 63-64, Grassland; 94, Oak Woodland; 104, Mixed Deciduous Forest (low oak); 105, Mixed Deciduous Forest (high oak); 113, Degraded Forest; 154, Mixed Conifer/Deciduous Forest (loblolly); 168, Mixed Deciduous Conifer Forest (low oak, pine); 171, Mixed Deciduous/ Coniferous Forest (medium oak, pine); 205, Water.

Inspection of the table shows that there is very little correlation between the two classification methods. General patterns do emerge, however. The crop and grassland SLCR classes primarily correspond with urban and suburban areas. The mixed deciduous forest (SLCR 104) class is well correlated with the TM high oak areas. There are relatively few SLCR 104 pixels in the image, however, and the bulk of the high oak areas are accounted for by SLCR class 105 (mixed deciduous). SLCR 105 also accounts for most of the medium- and low-oak areas as well as much of the mixed deciduous/conifer areas. The general SLCR mixed forest classes do correspond well with the TM mixed forest classes.

4.3. *Emission modelling*

Biogenic isoprene flux was modelled using oak tree (and other vegetation) distributions from the TM and SLCR classifications. The TM results match field-measured isoprene fluxes to within a factor of 2. We consider this to be reasonably good agreement since a factor of 2 random difference due to heterogeneity of sources within this region can be expected (Guenther *et al.* 1996c). The flux modelling technique is described in Guenther *et al.* (1996a), and uses estimated foliar densities and area-averaged estimates of leaf-level parameters. In this case foliar density was not used in the flux estimate. It could be incorporated to a limited degree, however, using a derivative of NDVI from the remotely sensed data sets to estimate relative foliar density.

Fluxes were calculated for circular fetch areas with radii of 0.5, 3, and 14 km from the tower/balloon sites. The 0.5 and 3 km footprints correspond with different methods of tower measurements, and the 14 km footprint corresponds with balloon measurements (Guenther *et al.* 1996a). The footprint is the upwind ground area responsible for emissions received at the sampling site. The estimates presented here could be further refined by using the wind direction recorded at the time of sampling and modelling the footprint as a wedge-shaped area upwind from the sampling site (rather than a complete circle around the sampling site). Table 5 compares the modelled fluxes with the measured fluxes. The agreement is surprisingly good and supports the validity of estimating biogenic VOC fluxes with remotely sensed data.

The SLCR-based flux estimates are about two to three times greater than the measured fluxes. This is still a fair agreement, and it is encouraging for the future use of SLCR data in isoprene emission studies. Also, the correspondence between the TM and SLCR flux estimates begins to converge over larger areas (TM quad scale). It is likely that over larger regions the differences between TM and SLCR classifications will diminish.

Table 5. Comparison between measured and modelled isoprene flux (after Guenther *et al.* 1996a).

Footprint radius	Measured flux ($mg\ C\ m^{-2}\ h^{-1}$)	TM modelled flux ($mg\ C\ m^{-2} h^{-1}$)	SLCR modelled flux ($mg\ C\ m^{-2} h^{-1}$)
0.5 km	4.8 ± 0.6	4.6	10.5
3 km	5.0 ± 1.7	3.4	10.4
14 km	2.4 ± 0.3	2.5	9.7
Landsat 'Quad'	4.8 ± 0.6	6.1	8.6

4.4. *Forest classification and ozone modelling*

The overall utility of this method, of estimating isoprene fluxes from satellite vegetation maps, is strongly supported by the agreement between the estimated isoprene fluxes and the measured fluxes, even though the accuracy of this TM classification is not validated on a pixel-per-pixel basis. This evaluation has significant uncertainties due to the limited available transect data. Future efforts should use transects designed specifically for the training of the intended remotely sensed data set.

Regional ozone models use landcover data at a much coarser resolution than that of Landsat TM. They use grid cells that range from hundreds to thousands of square kilometres. The purpose of using higher resolution satellite data is to maximize the accuracy of the oak tree density mapping. With TM data, the pixel size approaches the crown size of large oak trees or small tree clusters. With fairly pure pixels (unmixed), the spectral mapping is able to separate oak trees from other deciduous trees. The 1 km pixels of the SLCR database are too large to effectively separate oak trees from the deciduous forest. Yet coarser data, such as SLCR or the anticipated MODIS, are much more convenient for mapping at the global scales of ozone models. A balance must be found between the accuracy of small pixel data and larger pixel coverage. This balance will probably take the form of locally calibrating coarse data with higher resolution imagery and field work.

Future work with this method will include a study of the effect of changing scales and classification accuracy. The scale of classification necessary to provide meaningful biogenic input for local and regional ozone models has yet to be determined. We will continue to work on locally calibrating the SLCR classes with higher resolution satellite imagery and field studies. We also eagerly await the availability of MODIS data. The higher spectral resolution, more precise bandwidths, and (some) higher spatial resolution channels (250, 500 m) may prove very useful for this application.

5. Conclusions

The multi-temporal Landsat TM classification was successful in measuring oak density in the Oak Ridge, Tennessee, area. The classification was achieved using detailed ground transects to provide training sites for the supervised classification. Information was extracted from the early fall and winter TM scenes and added to the final classification as layers. Field observations and comparisons with the FIA and BELD databases support the accuracy of the classification. The isoprene flux maps created from the TM classification predicted the isoprene fluxes observed in the field to within a factor of 2.

Existing landcover data (SLCR) with 1 km pixels are available from the USGS via the World Wide Web. The correspondence between the TM and SLCR classifications is fairly poor. Isoprene emission estimates from the SLCR data are two to three times greater than the measured (and TM estimated) fluxes. This difference is small enough, however, to encourage continued work in calibrating SLCR data with field and other satellite landcover measurements. Due to its lower cost and smaller volume, SLCR data should prove to be more practical than TM data for compiling isoprene flux estimates required by regional and global chemistry models.

Acknowledgments

This research was partially supported by the US Environmental Protection Agency, Research Triangle Park, NC, under Interagency Agreement Grant No.

DW49934973-01-0, and by the Southern Oxidants Study (SOS) under Assistance Agreement No. CR 817766 to North Carolina State University. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

References

CHAVEZ, P. S., 1988, An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, **24**, 459–479.

GERON, C., GUENTHER, A., and PIERCE, T., 1994, An improved model for estimating emissions of volatile organic compounds from forests in the eastern United States. *Journal of Geophysical Research*, **99**, 12 773–12 791.

GUENTHER, A., ZIMMERMAN, P., and WILDERMUTH, M., 1994, Natural volatile organic compound emission rate estimates for US woodland landscapes. *Atmospheric Environment*, **28**, 1197–1210.

GUENTHER, A., HEWITT, C., ERICKSON, D., FALL, D., GERON, C., GRAEDEL, P., HARLEY, L., KLINGER, L., LERDAU, M., MCKAY, W., PIERCE, T., SCHOLES, B., STEINBRECHER, R., TALLAMRAJU, R., TAYLOR, J., and ZIMMERMAN, P., 1995, A global model of natural volatile organic compound emissions. *Journal of Geophysical Research*, **100**, 8873–8892.

GUENTHER, A., BAUGH, W., DAVIS, K., HAMPTON, G., HARLEY, P., KLINGER, L., VIERLING, L., ZIMMERMAN, P., ALLWINE, E., DILTS, S., LAMB, B., WESTBERG, H., BALDOCCHI, D., GERON, C., and PIERCE, T., 1996a, Isoprene fluxes measured by enclosure, relaxed eddy accumulation, surface layer gradient, mixed layer gradient, and mixed layer mass balance techniques. *Journal of Geophysical Research*, **101**, 18 555–18 567.

GUENTHER, A., GREENBERG, J., HARLEY, P., HELMIG, D., KLINGER, L., VIERLING, L., ZIMMERMAN, P., and GERON, C., 1996b, Leaf, branch, stand and landscape scale measurements of volatile organic compound fluxes from US woodlands. *Tree Physiology*, **16**, 17–24.

GUENTHER, A., ZIMMERMAN, P., KLINGER, L., GREENBERG, J., ENNIS, C., DAVIS, K., POLLOCK, W., WESTBERG, H., ALLWINE, G., and GERON, C., 1996c, Estimates of regional natural volatile organic compound fluxes from enclosure and ambient measurements. *Journal of Geophysical Research*, **101**, 1345–1359.

HAAGEN-SMIT, A. J., and FOX, M. M., 1954, Photochemical ozone formation with hydrocarbons and automobile exhaust. *Journal of the Air Pollution Control Association*, **4**, 105–109.

HANSEN, M. H., FRIESWYCK, T., GLOVER, J. F., and KELLY, J. F., 1992, *The Eastwide Forest Inventory Database: Users Manual*, General Technical Report, NC-1512 (St. Paul, MN: Forest Service, Department of Agriculture).

JENSEN, J. R., 1986, *Introductory digital image processing, a remote sensing perspective* (Englewood Cliffs, NJ: Prentice Hall).

JOHNSON, D. W., and VAN HOOK, R. I., (editors), 1989, *Analysis of biogeochemical cycling processes in Walker Branch Watershed* (New York, NY: Springer-Verlag).

KINNEE, E., GERON, C., and PIERCE, T., 1997, United States land use inventory for estimating biogenic ozone precursor emissions. *Ecological Applications*, **7**, 46–58.

KRUSE, F. A., LEFKOFF, A. B., BOARDMAN, J. W., HEIDEBRECHT, K. B., SHAPIRO, A. T., BARLOON, P. J., and GOETZ, A. F. H., 1993, The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, **44**, 145–163.

LAMB, B., GUENTHER, A., GAY, D., and WESTBERG, H., 1987, A national inventory of biogenic hydrocarbon emissions. *Atmospheric Environment*, **21**, 1695–1705.

LAMB, B., GAY, D., WESTBERG, H., and PIERCE, T., 1993, A biogenic hydrocarbon emission inventory for the U.S. using a simple forest canopy model. *Atmospheric Environment*, **27**, 1673–1690.

NATIONAL RESEARCH COUNCIL, 1991, *Rethinking the Ozone Problem in Urban and Regional Air Pollution* (Washington, D.C.: National Academy Press).

SNYDER, J., 1987, *Map projections; a working manual*. US Geological Survey professional paper 1395, Washington D.C.

WOLTER, P. T., MLADENOFF, D. J., HOST, G. E., and CROW, T. R., 1995, Improved forest classification in the northern lake states using multi-temporal Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, **61**, 129–143.